INTERNATIONAL TELECOMMUNICATION UNION

CCITT

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

BLUE BOOK

VOLUME II – FASCICLE II.3

TELEPHONE NETWORK AND ISDN

QUALITY OF SERVICE, NETWORK MANAGEMENT AND TRAFFIC ENGINEERING

RECOMMENDATIONS E.401-E.880

IXTH PLENARY ASSEMBLY
MELBOURNE, 14-25 NOVEMBER 1988

Geneva 1989

INTERNATIONAL TELECOMMUNICATION UNION

CCITT

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

BLUE BOOK

VOLUME II – FASCICLE II.3

TELEPHONE NETWORK AND ISDN

QUALITY OF SERVICE, NETWORK MANAGEMENT AND TRAFFIC ENGINEERING

RECOMMENDATIONS E.401-E.880

IXTH PLENARY ASSEMBLY
MELBOURNE, 14-25 NOVEMBER 1988

Geneva 1989

ISBN 92-61-03271-0

CONTENTS OF THE CCITT BOOK
APPLICABLE AFTER THE NINTH PLENARY ASSEMBLY (1988)

BLUE BOOK

FASCICLE III.6 – Line transmission of non-telephone signals. Transmission of sound-programme and television signals. Series H and J Recommendations (Study Group XV).

FASCICLE III.7 – Integrated Services Digital Network (ISDN) – General structure and service capabilities. Recommendations I.110-I.257 (Study Group XVIII).

FASCICLE III.8 – Integrated Services Digital Network (ISDN) – Overall network aspects and functions, ISDN user-network interfaces. Recommendations I.310-I.470 (Study Group XVIII).

FASCICLE III.9 – Integrated Services Digital Network (ISDN) – Internetwork interfaces and maintenance principles. Recommendations I.500-I.605 (Study Group XVIII).

Volume IV

FASCICLE IV.1 – General maintenance principles: maintenance of international transmission systems and telephone circuits. Recommendations M.10-M.782 (Study Group IV).

FASCICLE IV.2 – Maintenance of international telegraph, phototelegraph and leased circuits. Maintenance of the international public telephone network. Maintenance of maritime satellite and data transmission systems. Recommendations M.800-M.1375 (Study Group IV).

FASCICLE IV.3 – Maintenance of international sound-programme and television transmission circuits. Series N Recommendations (Study Group IV).

FASCICLE IV.4 – Specifications for measuring equipment. Series O Recommendations (Study Group IV).

Volume V – Telephone transmission quality. Series P Recommendations (Study Group XII).

Volume VI

FASCICLE VI.1 – General Recommendations on telephone switching and signalling. Functions and information flows for services in the ISDN. Supplements. Recommendations Q.1-Q.118 *bis* (Study Group XI).

FASCICLE VI.2 – Specifications of Signalling Systems Nos. 4 and 5. Recommendations Q.120-Q.180 (Study Group XI).

FASCICLE VI.3 – Specifications of Signalling System No. 6. Recommendations Q.251-Q.300 (Study Group XI).

FASCICLE VI.4 – Specifications of Signalling Systems R1 and R2. Recommendations Q.310-Q.490 (Study Group XI).

FASCICLE VI.5 – Digital local, transit, combined and international exchanges in integrated digital networks and mixed analogue-digital networks. Supplements. Recommendations Q.500-Q.554 (Study Group XI).

FASCICLE VI.6 – Interworking of signalling systems. Recommendations Q.601-Q.699 (Study Group XI).

FASCICLE VI.7 – Specifications of Signalling System No. 7. Recommendations Q.700-Q.716 (Study Group XI).

FASCICLE VI.8 – Specifications of Signalling System No. 7. Recommendations Q.721-Q.766 (Study Group XI).

FASCICLE VI.9 – Specifications of Signalling System No. 7. Recommendations Q.771-Q.795 (Study Group XI).

FASCICLE VI.10 – Digital subscriber signalling system No. 1 (DSS 1), data link layer. Recommendations Q.920-Q.921 (Study Group XI).

FASCICLE VI.11 – Digital subscriber signalling system No. 1 (DSS 1), network layer, user-network management. Recommendations Q.930-Q.940 (Study Group XI).

FASCICLE VI.12 – Public land mobile network. Interworking with ISDN and PSTN. Recommendations Q.1000-Q.1032 (Study Group XI).

FASCICLE VI.13 – Public land mobile network. Mobile application part and interfaces. Recommendations Q.1051-Q.1063 (Study Group XI).

FASCICLE VI.14 – Interworking with satellite mobile systems. Recommendations Q.1100-Q.1152 (Study Group XI).

Volume VII

FASCICLE VII.1 – Telegraph transmission. Series R Recommendations. Telegraph services terminal equipment. Series S Recommendations (Study Group IX).

FASCICLE VII.2 – Telegraph switching. Series U Recommendations (Study Group IX).

FASCICLE VII.3 – Terminal equipment and protocols for telematic services. Recommendations T.0-T.63 (Study Group VIII).

FASCICLE VII.4 – Conformance testing procedures for the Teletex Recommendations. Recommendation T.64 (Study Group VIII).

FASCICLE VII.5 – Terminal equipment and protocols for telematic services. Recommendations T.65-T.101, T.150-T.390 (Study Group VIII).

FASCICLE VII.6 – Terminal equipment and protocols for telematic services. Recommendations T.400-T.418 (Study Group VIII).

FASCICLE VII.7 – Terminal equipment and protocols for telematic services. Recommendations T.431-T.564 (Study Group VIII).

Volume VIII

FASCICLE VIII.1 – Data communication over the telephone network. Series V Recommendations (Study Group XVII).

FASCICLE VIII.2 – Data communication networks: services and facilities, interfaces. Recommendations X.1-X.32 (Study Group VII).

FASCICLE VIII.3 – Data communication networks: transmission, signalling and switching, network aspects, maintenance and administrative arrangements. Recommendations X.40-X.181 (Study Group VII).

FASCICLE VIII.4 – Data communication networks: Open Systems Interconnection (OSI) – Model and notation, service definition. Recommendations X.200-X.219 (Study Group VII).

FASCICLE VIII.5 – Data communication networks: Open Systems Interconnection (OSI) – Protocol specifications, conformance testing. Recommendations X.220-X.290 (Study Group VII).

FASCICLE VIII.6 – Data communication networks: interworking between networks, mobile data transmission systems, internetwork management. Recommendations X.300-X.370 (Study Group VII).

FASCICLE VIII.7 – Data communication networks: message handling systems. Recommendations X.400-X.420 (Study Group VII).

FASCICLE VIII.8 – Data communication networks: directory. Recommendations X.500-X.521 (Study Group VII).

Volume IX – Protection against interference. Series K Recommendations (Study Group V). Construction, installation and protection of cable and other elements of outside plant. Series L Recommendations (Study Group VI).

Volume X

FASCICLE X.1 – Functional Specification and Description Language (SDL). Criteria for using Formal Description Techniques (FDTs). Recommendation Z.100 and Annexes A, B, C and E, Recommendation Z.110 (Study Group X).

FASCICLE X.2 – Annex D to Recommendation Z.100: SDL user guidelines (Study Group X).

FASCICLE X.3 – Annex F.1 to Recommendation Z.100: SDL formal definition. Introduction (Study Group X).

FASCICLE X.4 – Annex F.2 to Recommendation Z.100: SDL formal definition. Static semantics (Study Group X).

FASCICLE X.5 – Annex F.3 to Recommendation Z.100: SDL formal definition. Dynamic semantics (Study Group X).

FASCICLE X.6 – CCITT High Level Language (CHILL). Recommendation Z.200 (Study Group X).

FASCICLE X.7 – Man-Machine Language (MML). Recommendations Z.301-Z.341 (Study Group X).

CONTENTS OF FASCICLE II.3 OF THE BLUE BOOK

Part I – Recommendations E.401 to E.428

International telephone network management and checking of service quality

Part II – Recommendations E.500 to E.600

Traffic engineering

Part III – Recommendations E.800 to E.880

Quality of services; concepts, models, Objectives, dependability planning

Part IV – Supplements to the Series E Recommendations relating to telephone network management and traffic engineering

MODIFICATIONS TO THE SERIES E RECOMMENDATIONS

1 *Fascicle II.2*

1.1 The following new Recommendations and Supplement did not appear in Fascicle II.2 of the *Red Book*, and are new Recommendations developed during the 1985-1988 Study Period:

Recommendations

E.118	E.184
E.133	E.214
E.152	E.215
E.165	E.216
E.166	E.301
E.167	E.330
E.170[1]	Supplement No. 7
E.172	

1.2 The following Recommendations in Fascicle II.2 of the *Red Book* were revised, in many cases extensively, during the 1985-1988 Study Period:

Recommendations

E.100	E.160
E.110	E.161
E.112	E.163
E.115	E.164
E.116	E.171
E.121	E.180
E.123	E.182
E.141	

1.3 The text of the following Recommendation and Supplement no longer appears in the *Blue Book*; however a reference to this text contained in Fascicle II.2 of the *Red Book*, is given in the current edition:

E.125	Supplement No. 5

2 *Fascicle II.3*

2.1 The following Recommendations and Supplement did not appear in Fascicle II.3 of the *Red Book*, and are new Recommendations developed during the 1985-1988 Study Period:

Recommendations

E.412	E.710
E.428	E.711
E.503	E.713
E.504	E.720
E.507	E.721
E.508	E.855
E.524	E.862
E.525	E.880
E.700	Supplément No. 6
E.701	

[1] This Recommendation has been totally modified from the equivalently numbered Recommendation contained in the *Red Book*.

2.2 The following Recommendations which appeared in Fascicle II.3 of the *Red Book* were revised, in many cases extensively, during the 1985-1988 Study Period:

Recommendations

E.410	E.550
E.411	E.600
E.413 (formerly numbered E.412)	E.800 (formerly Rec. G.106)
E.414 (formerly numbered E.413)	
E.420	
E.424	E.810 (formerly Rec. G.107)
E.500	E.830 (formerly Rec. G.108)
E.501	E.845 (formerly Rec. G.180)
E.502	E.850 (formerly Rec. G.181)
E.506	

2.3 The following Supplements which appeared in Fascicle II.3 of the *Red Book*, have been deleted from the *Blue Book*:

Supplements

No. 6[2)]

No. 8 (updated and converted to Rec. 301, Fascicle II.2)

NOTES

1 The Questions entrusted to each Study Group for the Study Period 1989-1992 can be found in Contribution No. 1 to that Study Group.

2 In this fascicle, the expression "Administration" is used for shortness to indicate both a telecommunication Administration and a recognized private operating agency.

---

2) This material has been superseded by information contained in Recommendations E.411 and Q.297.

PART I

Recommendations E.401 to E.428

INTERNATIONAL TELEPHONE NETWORK MANAGEMENT AND CHECKING OF SERVICE QUALITY

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

SECTION 1

INTERNATIONAL SERVICE STATISTICS

Recommendation E.401

STATISTICS FOR THE INTERNATIONAL TELEPHONE SERVICE (NUMBER OF CIRCUITS IN OPERATION AND VOLUME OF TRAFFIC)

(Statistics exchanged by Administrations)

Administrations exchange each year, *in February*, statistics showing the number of circuits used and the volume of traffic monitored in the preceding year, as well as estimates of the number of circuits which will be required three years and five years later. These statistics shall be drawn up in the form indicated below.

A copy of the statistics shall be sent to the CCITT Secretariat for information.

ANNEX A

(to Recommendation E.401)

How to fill in the table on international telephone traffic statistics

Column 1 — Designation of the connection by giving the name of the outgoing exchange first and then the name of the incoming exchange. Two-way connections will be shown in alphabetical order.

Columns 2 and 3 — Number of circuits in operation as on *31 December* of the year of the statistics.
The number will be shown in column 2 when it refers to outgoing circuits and in column 3 when it refers to both-way circuits.

Columns 4 and 5 — Number of circuits which would have been required during the year of the statistics.

Column 6 — Method of operation.

The following abbreviations will be used:

- A for automatic,
- SA for semiautomatic,
- M for manual,
- A + SA for automatic and semiautomatic.

Column 7 — Destination of traffic.

Each relation will be shown in this column on a separate line.

In the example given, the traffic routed over the Zürich-København circuits is destined for Denmark (terminal), Sweden, Norway and Finland (transit). In this case, the data for each destination will be shown in columns, 8, 9, 10 and 11. The total traffic figure, however, should not be omitted. These data will be bracketed together. If the connection handles traffic only to the country in which the incoming exchange is situated, only the word "terminal" will appear in column 7.

Columns 8 and 9	Busy-hour traffic, expressed in *erlangs.* (See Recommendation E.600.) The traffic measured during the busiest month of the year of the statistics is given in column 9. For two-way circuit groups the total amount of incoming and outgoing traffic should be given. In column 8 the month of the year during which the traffic was measured should be indicated in roman numerals.
Column 10	Busy hour (UTC). This refers to the busy hour as defined in Recommendation E.600.
Column 11	Annual increase, in %. Each Administration should insert in this column the annual traffic increase rate with respect to the previous year.
Columns 12 and 13	Columns 12 and 13 should show the estimated number of circuits required to route traffic in three and five years' time, respectively. For example, if the statistics relating to 1982 are drawn up in February 1983, column 12 will give the estimated number of circuits required in 1986 and column 13 those required in 1988.

International telephone traffic statistics

Year: ....

Circuits	Number of circuits in service		Number of circuits required		Method of operation	Destination of traffic	Busy-hour traffic		Start of busy-hour (UTC)	Annual traffic increase	Estimated number of circuits		Observations
	Outgoing	Both-way	Outgoing	Both-way			Month	Erlangs			In three years	In five years	
1	2	3	4	5	6	7	8	9	10	11	12	13	14
(Examples)													
Zürich-Kobenhavn	24	—	20	—	SA	Terminal	X	8	10.00	15%			
						Sweden [a]	X	4	10.15	12%			
						Norway	X	2	09.45	13%			[a] Overflow traffic on Zürich-Stockholm connection
						Finland	X	1	10.30	7%			
						Total	X	15	10.00	14%	28	32	
Zürich-Stockholm	12	—	11	—	SA	Terminal	IX	5.5	10.15	12%	13	15	

SECTION 2

INTERNATIONAL NETWORK MANAGEMENT

Recommendation E.410

INTERNATIONAL NETWORK MANAGEMENT – GENERAL INFORMATION

1 Introduction

The demand for international telephone service continues to increase substantially. This increasing demand has been met by advances in both technology and operational techniques. The growth of traffic has also required the development of larger transmission systems and exchanges to provide the capacity to meet the required grade of service. With the continued growth of the international automatic service, direct supervision and control over traffic has decreased since operators are no longer involved in establishing most calls.

In addition to the above, the introduction of larger digital transmission and switching systems, along with common channel signalling, has resulted in an international telephone network which is highly interconnected and interactive, and which has become increasingly vulnerable to overload and congestion. This overload and congestion can occur with little or no advance warning.

A number of events may arise which can have a serious effect on the international telephone service. Among these events are:

- failures of international or national transmission systems;
- failures of international or national exchanges;
- planned outages of transmission systems and exchanges;
- abnormal increases in traffic demand. The events which give rise to such traffic demand may be foreseen (e.g., national or religious holidays, international sporting events) or unforeseen (e.g., natural disasters, political crises);
- focussed overloads, and in particular, mass-calling;
- difficulties in meeting the requirements of international traffic resulting (for example) from delays in the provision of additional circuits or equipment;
- congestion in connected networks.

These events can lead to congestion which, if uncontrolled, may spread and thus seriously degrade the service in other parts of the international network. Considerable benefits can be derived for the international network as a whole if prompt action is taken to control the effect on service of such events.

In addition, as the telephone network migrates toward ISDN, interworking with other networks will develop. With interworking, failure or congestion in one network, or in the interface between networks, can have an adverse impact on the performance of the connected network(s).

The above considerations have led to the development of "international network management", which encompasses all the activities necessary to reduce the effect on service of any situation affecting unfavourably the international telephone network, and in the future, the ISDN.

Note – Much of the guidance on international network management may be applicable in national networks.

2 Definition of international network management

international network management is the function of supervising the international network and taking action when necessary to control the flow of traffic.

Network management requires real-time monitoring and measurement of current network status and performance, and the ability to take prompt action to control the flow of traffic.

3 Objective of network management

The objective of network management is to enable as many calls as possible to be successfully completed. This objective is met by maximizing the use of all available equipment and facilities in any situation through the application of the principles given below.

4 Principles of international network management

4.1 *Utilize all available circuits*

There are periods when, due to changing traffic patterns, the demand for service cannot be met by the available circuits in the normal routing. At the same time, many circuits to other locations may be idle due to differences in calling patterns caused by time zones, local calling habits, or busy season variations. After negotiation and agreement amongst the Administrations affected, some or all of the unusually heavy traffic can be redirected to this idle capacity for completion.

4.2 *Keep all available circuits filled with traffic which has a high probability of resulting in effective calls*

The telephone network is generally circuit-limited; therefore the number of simultaneous effective calls is strongly influenced by the number of available circuits. However, ineffective calls can occupy circuit capacity which would otherwise be available for effective calls. Therefore identifying those call attempts which are likely to be ineffective because of a situation in the network (e.g., a failure), and reducing them as close to their source as possible, will allow circuit capacity to be available for call attempts which have a higher probability of being effective.

4.3 *When all available circuits are in use, give priority to calls requiring a minimum number of circuits to form a connection*

When telephone networks are designed using automatic alternate routing of calls, efficient operation occurs when traffic loads are at or below engineered values. However, as traffic loads increase above the engineered value, the ability of the network to carry effective calls decreases since an increased number of calls require two or more circuits to form a connection. Such calls increase the possibility of one multi-link call blocking several potential calls.

Thus automatic alternate routing should be restricted to give preference to direct routed traffic during periods of abnormally high demand.

4.4 *Inhibit switching congestion and prevent its spread*

A large increase in switching attempts can result in switching congestion when the switching capacity of an exchange is exceeded. If the switching congestion is left uncontrolled, it can spread to connected exchanges or networks and cause a further degradation of network performance. Controls should be applied which inhibit switching congestion by removing attempts from the congested exchange which have a low chance of resulting in a succesful call.

Note – Network management assumes that the network is adequately engineered to meet the normal levels of traffic, the requirement for which is described in Recommendations E.171, E.510, E.520, E.522, E.540 and E.541.

5 Benefits derived from international network management

Among the benefits to be derived from international network management are:

5.1 Increased revenue which is derived from an increase in successful calls.

5.2 Improved service to the customer. This can lead, in turn, to:

- improved customer relations;
- stimulation of customer calling rate;
- increased customer acceptance of new services.

5.3 More efficient use of the network. This can result in:

- an increased return on the capital invested in the network;
- an improvement in the ratio of effective to ineffective calls.

5.4 Greater awareness of the actual status and performance of the network. Such awareness can lead to:

- a basis by which network management and maintenance priorities can be established;
- improved network planning information;
- improved information on which future capital investment in the network can be decided;
- improved public relations.

5.5 Protection of revenue and important services, particularly during severe network situations.

6 Network management functions

Network management encompasses all of the activities necessary to identify conditions which may adversely affect network performance and service to the customer, and the application of network controls to minimize their impact. This includes the following functions:

a) monitoring the status and performance of the network on a real-time basis, which includes collecting and analyzing relevant data;

b) detecting abnormal network conditions;

c) investigating and identifying the reasons for abnormal network conditions;

d) initiating corrective action and/or control;

e) cooperating and coordinating actions with other network management centres, both domestic and international, on matters concerned with international network management and service restoration;

f) cooperating and coordinating with other work areas (e.g., maintenance, operator services or planning) on matters which affect service;

g) issuing reports of abnormal network situations, actions taken and results obtained to higher authority and other involved departments and Administrations, as required;

h) providing advance planning for known or predictable network situations.

7 Cooperation and coordination

Effective network management depends on the prompt availability of information indicating when and where a problem is occurring, and a trained group working in cooperation with all parts of the telecommunications organization. Just as there is a need for coordination in planning and building the network, there also is a need for coordination in managing it. The network is such that equipment malfunctions or overloads frequently produce unacceptable performance at a distance from the physical location of the problem. Therefore, those who monitor and manage the network, both nationally and internationally, must cooperate to ensure satisfactory service.

Network management is highly technical in nature, and depends on the skill and creativity of those who share an understanding of network management philosophy, objectives, terminology, tools and techniques. These items are specified in Recommendations E.410 through E.414, and provide a basis for the cooperation and coordination which are a vital part of network management.

8 Further Recommendations on network management

8.1 Recommendation E.411 provides operational guidance for network management including:
- status and performance parameters;
- expansive and protective traffic controls;
- criteria for application of controls.

8.2 Recommendation E.412 provides information on network management controls:
- traffic to be controlled;
- exchange controls;
- automatic controls;
- status of controls;
- operator controls.

8.3 Recommendation E.413 provides guidance on planning for events such as:
- peak calling days;
- failures of transmission systems;
- failures of exchanges;
- failures of common channel signalling systems;
- mass-calling situations;
- disasters;
- introduction of new services.

8.4 Recommendation E.414 provides guidance on the functional elements of a network management organization which need to be identified internationally as contact points. These comprise:
- planning and liaison;
- implementation and control;
- development.

8.5 It is emphasized that it is not necessary to meet the full scope of these Recommendations to achieve some benefit from the application of network management, particularly when getting started. However, the Recommendations do provide detailed information over a wide range of techniques, some of which can be implemented readily, whilst others may require considerable planning and design effort. Additional information may also be found in the handbook on Quality of service, network management and maintenance [1].

Reference

[1] CCITT Manual *Quality of service, network management and maintenance*, ITU, Geneva, 1984.

Recommendation E.411

INTERNATIONAL NETWORK MANAGEMENT – OPERATIONAL GUIDANCE

1 Introduction

Network management requires real-time monitoring of current network status and performance and the ability to take prompt action to control the flow of traffic when necessary (see Recommendation E.410). Operational guidance to meet these requirements, including a description of status and performance parameters, traffic controls and the criteria for their application are included in this Recommendation. It should be noted that the complete range of parameters and traffic controls are not necessary for the introduction of a limited network management capability, however a comprehensive selection will bring substantial benefit (see Recommendation E.410, § 5). In addition, some guidance on beginning network management is provided, along with information on developing a network management centre and the use of common channel signalling for network management purposes.

2 Information requirements

2.1 Network management requires information of where and why difficulties are occurring or are likely to occur in the network. This information is essential to identify the source and effect of a difficulty at the earliest possible time, and will form the basis for any network management action which is taken.

2.2 The information relating to current difficulties can be obtained from:

a) real-time surveillance of the status and performance of the network;

b) information from telephone operators as to where they are experiencing difficulties; or where they are receiving customer complaints of difficulties;

c) transmission system failure and planned outage reports (these reports need not relate only to the network local to one Administration, but should reflect the whole international network);

d) international or national exchange failures and planned outage reports;

e) news media reports detailing unforeseen events which stimulate traffic (for example, natural disasters).

2.3 The information relating to difficulties which are likely to occur in the future will be obtained from:

a) reports of future planned outages of transmission systems;

b) reports of future planned outages of international or national exchanges;

c) knowledge of special events (for example, international sporting events, political elections);

d) knowledge of national holidays and festivals (e.g., Christmas Day, New Year's Day);

e) an analysis of past network performance.

2.4 The system availability information point, defined in Recommendation M.721, will provide a ready source for much of the information indicated above.

3 Network status and performance data

3.1 In order to identify where and when difficulties are occurring in the network, or are likely to occur, data will be required which will indicate the status and measure the performance of the network. Such data will require real-time collection and processing, and may require the use of thresholds (see § 5.1).

3.2 Data may be collected in various ways which include counters in electromechanical exchanges which can be read manually when required (e.g., during periods of heavy traffic or special events), data output reports from SPC exchanges, or computerized network management operations systems which can collect and process data from a large number of exchanges.

3.3 Network status information includes information on the status of exchanges, circuit groups and common channel signalling systems. This status information can be provided by one or more types of displays. These may include printers, video displays, and/or indicators on a display board or network management console. To be useful, network status indicators should be available as rapidly as possible.

3.3.1 Exchange status information includes the following:

Load measurements – These are provided by attempt counts, usage or occupancy data, data on the percent of real-time capacity available (or in use), blocking rates, percentage of equipment in use, counts of second trials, etc.

Congestion measurements – These are provided by measurements of the delay in serving incoming calls, holding times of equipment, average call processing and set-up time, queue lengths for common control equipment (or software queues), and counts of equipment time-outs, etc.

Service availability of exchange equipment – This information will show when major items of equipment are made busy to traffic. This could highlight a cause of difficulty or give advance warning that difficulties could arise if demand increases.

Congestion indicators – In addition to the above, indicators can be provided by SPC exchanges which show the degree of congestion. These indicators can show:

- moderate congestion Level 1;
- serious congestion Level 2;
- unable to process calls Level 3.

Note – While this is desirable, SPC exchanges may not be able to provide a level 3 indicator during catastrophic failures.

The availability of specific exchange status information will depend on the switching technology employed by each Administration. Details of exchange measurements are found in Recommendations E.502 and Q.544.

3.3.2 Circuit group status information relates to the following:

- status of all circuit groups available to a destination;
- status of individual circuit sub-groups in a circuit group;
- status of circuits on each circuit group.

Status indicators can be provided to show when the available network is fully utilized by indicating:

- when all circuits in a circuit group are busy;
- when all circuits in a circuit sub-group are busy;
- when all circuit groups available to a destination are busy.

This would indicate that congestion is present or imminent. Status information can be provided to show the availability of the network for service, by reporting the number or percentage of circuits on each circuit group that are made busy or are available for traffic.

This information could identify the cause of difficulty or give advance warning that difficulties may arise as the demand increases.

3.3.3 Common channel signalling system status provides information that will indicate failure or signalling congestion within the system. It includes such items as:

- receipt of a transfer prohibited signal (Signalling Systems Nos. 6 and 7),
- initiation of an emergency restart procedure (Signalling System No. 6),
- presence of a signalling terminal buffer overflow condition (Signalling System No. 6),
- signal link unavailability (Signalling System No. 7),
- signal route unavailability (Signalling System No. 7),
- destination inaccessible (Signalling System No. 7).

This information may identify the cause of difficulty or give advance warning that difficulties may arise as the demand increases.

3.3.3.1 Network management actions may help to reduce congestion in common channel signalling systems by reducing traffic being offered to common channel signalling circuit groups, or by diverting traffic to conventional signalling circuit groups.

3.4 Network performance data should relate to the following:

- traffic performance on each circuit group;
- traffic performance to each destination;
- effectiveness of network management actions.

It may also be desirable to assemble performance data in terms of circuit group and destination combinations and/or traffic class (for example, operator-dialled, subscriber-dialled, transit). (See Recommendation E.412, § 2.1.)

3.5 Data collection should be based on a system of measurement which is either continuous or of a sufficiently rapid sampling rate to give the required information. For example, for common control switching equipment, the sampling rate may need to be as frequent as every second.

Reports on network status and performance should be provided periodically, for example, on a 3 minute, 5 minute, 15 minute, 30 minute or hourly basis, with the more frequent reports usually being more useful. However, the more frequent reports may produce erratic data due to the peakedness of traffic, especially on small circuit groups. Data reports compiled by a network management operations system take on added value in that a more global view of network performance is provided.

3.6 The network performance data is generally expressed in parameters which help to identify difficulties in the network. Among these parameters are:

3.6.1 **percentage overflow (% OFL)**

% OFL indicates the relationship between the total bids offered to a circuit group or destination, in a specified period of time, and the quantity of bids not finding a free circuit. It will, therefore, give an indication of the overflow from one circuit group to another, or the bids which fail because all circuit groups to a destination are busy.

$$\% \text{ OFL} = \frac{\text{Overflows bids (to another circuit group or to circuit busy signal)}}{\text{Total bids for the circuit group (or all circuit groups to a destination)}} \times 100$$

3.6.2 **bids per circuit per hour (BCH)** [1)]

BCH is an indication of the average number of bids per circuit, in a specified time interval. It will therefore identify the demand and, when measured at each end of a both-way operated circuit group, will identify the direction of greater demand.

$$\text{BCH} = \frac{\text{Bids per hour}}{\text{Quantity of circuits available for service}}$$

It is not necessary to accumulate data for an hour to calculate BCH. However, the calculated BCH must be adjusted when data accumulation is less than hourly. For example, the bids should be doubled if 1/2 hour data is used. The result would be BCH for the data collection period.

3.6.3 **answer seizure ratio (ASR)**

ASR gives the relationship between the number of seizures that result in an answer signal and the total number of seizures. This is a direct measure of the effectiveness of the service being offered onward from the point of measurement and is usually expressed as a percentage as follows:

$$\text{ASR} = \frac{\text{Seizures resulting in answer signal}}{\text{Total seizures}} \times 100$$

Measurement of ASR may be made on a circuit group or on a destination basis.

---

1) International networks contain one-way and both-way operated circuits, and their traffic flow characteristics are inherently different. This difference needs to be taken into account when calculating BCH and SCH either by:

i) multiplying the number of one-way circuits by 2 to derive an equivalent number of both-way circuits or;

ii) dividing the number of both-way circuits by 2 to derive an equivalent number of one-way circuits.

When analyzing BCH and SCH data, and when BCH and SCH data are exchanged between Administrations, it is essential that the method used is understood so that erroneous conclusions may be avoided.

3.6.4 answer bid ratio (ABR)

ABR gives the relationship between the number of bids that result in an answer signal and the total number of bids. ABR may be made on a circuit group or on a destination basis.

$$\text{ABR} = \frac{\text{Bids resulting in answer signal}}{\text{Total bids}} \times 100$$

ABR is expressed as a percentage and is a direct measure of the effectiveness of traffic onward from the point of measurement. It is similar to ASR except that it includes bids that do not result in a seizure.

3.6.5 seizures per circuit per hour (SCH) [2)]

SCH is an indication of the average number of times, in a specified time interval, that each circuit group is seized. When related to the expected values of average call holding times and effective call/seizure rate for the circuit group, it will give an indication of the effectiveness of the service being offered.

$$\text{SCH} = \frac{\text{Seizures per hour}}{\text{Quantity of circuits available for service}}$$

It is not necessary to accumulate data for an hour to compute SCH. (See § 3.6.2 – BCH.)

3.6.6 occupancy

Occupancy can be represented in units (for example, erlangs, hundred-call-seconds (CCS) or as a percentage. It can be measured as a total for a destination or for a circuit group and as an average per circuit on a circuit group. Its use for network management purposes is to show usage and to identify unusual traffic levels.

3.6.7 mean holding time per seizure

This is the total holding time divided by the total number of seizures and can be calculated on a circuit group basis or for switching equipment.

3.6.8 busy-flash seizure ratio (BFSR)

BFSR gives the relationship between the number of seizures that result in a "busy-flash" signal (or its equivalent) and the total number of seizures. Measurement of BFSR is usually made on a circuit group basis.

$$\text{BFSR} = \frac{\text{Seizures resulting in a "busy-flash"}}{\text{Total seizures}} \times 100$$

Note – The source of "busy-flash" signals, or their equivalent, will vary with the signalling system used. Therefore, the calculated BFSR on individual circuit groups may naturally be different, and as a result, caution should be used when comparing BFSR among circuit groups.

---

2) International networks contain one-way and both-way operated circuits, and their traffic flow characteristics are inherently different. This difference needs to be taken into account when calculating BCH and SCH either by:
i) multiplying the number of one-way circuits by 2 to derive an equivalent number of both-way circuits or;
ii) dividing the number of both-way circuits by 2 to derive an equivalent number of one-way circuits.
When analyzing BCH and SCH data, and when BCH and SCH data are exchanged between Administrations, it is essential that the method used is understood so that erroneous conclusions may be avoided.

3.7 The number of parameters possible or necessary for particular Administration purposes will depend upon a variety of factors. These will include:

a) the data available at an exchange;

b) the particular routing arrangements employed (for example, SCH and BCH relate to circuit group performance only; ABR, ASR, and % OFL can relate to circuit group or destination performance);

c) the interrelationships which exist between the parameters (for example, SCH can give similar indications to ASR – see § 3.6.5 above).

4 Interpretation of parameters

The interpretation of parameters on which network management actions are based can most conveniently be made by considering the originating international exchange as the reference point (see Figure 1/E.411).

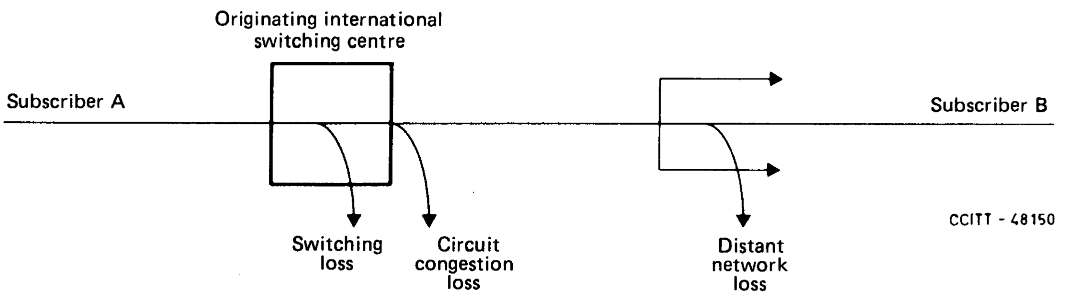


FIGURE 1/E.411

From this reference point, the factors which affect call completion can broadly be divided into three main components:

a) switching loss (near-end loss);

b) circuit congestion loss (near-end loss);

c) distant network loss (far-end loss).

4.1 *Switching loss*

Switching loss may be due to:

1) common equipment or switchblock congestion, or queue overflows or processor overloads;

2) failures in incoming signalling;

3) subscriber/operator dependent errors, such as insufficient or invalid digits, premature call abandonment, etc.;

4) routing errors, such as barred transit access;

5) other technical failures.

Guidance to the identification of switching loss can be obtained from § 3.3.

4.2 *Circuit congestion loss*

This loss will depend on:

1) the number of circuits available for a destination, and:

2) the level of demand for that destination,

3) the traffic performance on the circuit groups to that destination.

Indication that circuit congestion loss may occur can be obtained from the status information detailed in § 3.3.2 above.

Circuit congestion loss can be identified by any of the following:

- percentage overflow (see § 3.6.1),
- a difference between the "bids per circuit per hour" and "seizures per circuit per hour" measurements on the final circuit group (see §§ 3.6.2 and 3.6.5),
- a difference between the "answer bid ratio" and the "answer seizure ratio" (see §§ 3.6.3 and 3.6.4).

It should be noted that for both-way operated circuit groups, excessive demand in the incoming direction may also cause circuit congestion loss. This can be identified by comparing incoming and outgoing bids, seizures or occupancy.

4.3 *Distant network loss*

Distant network loss may be divided into:

1) *technical loss*: due to distant exchange and national circuit faults,
2) *subscriber dependent loss*: due to subscriber B busy, no answer, invalid distant number, number unavailable, etc.,
3) *traffic dependent loss*: these losses are due to lack of distant network capacity to meet traffic demand.

Under normal conditions, and for a large sample measured over a long period, distant network loss can be said to have a fixed or ambient overhead loss (this value depends on destination with some hour-by-hour and day-by-day variations).

Under abnormal situations (heavy demand, failures, etc.) distant network losses can be significantly affected. Variations in distant network loss can be identified by any of the following:

- answer seizure ratio (see § 3.6.3) (this is a direct measurement),
- seizures per circuit per hour (see § 3.6.5) (this is an indirect measurement),
- mean holding time per seizure (see § 3.6.7) (this is an indirect measurement),
- busy-flash seizure ratio (see § 3.6.8) (this is a direct measurement).

5 Criteria for action

5.1 The basis for the decision on whether any network management action should be taken will depend upon real-time information on the status and performance of the network. It is advantageous if the output of this information can be initially restricted to that which is required to identify possible difficulties in the network. This can be achieved by setting threshold values for performance parameters, and for the number or the percentage of circuits and common control equipment which are in service, such that when these threshold values are crossed, network management action can be considered. These threshold values will represent some of the criteria by which decisions are reached.

5.2 Indications that a threshold has been crossed and "all circuits on a circuit group are busy" and "all circuit groups to a destination are busy" may be used to direct attention to the particular area of the network for which detailed performance information will then be required.

5.3 The decision on whether or not to take network management action, and what action will be taken, is the responsibility of the network management personnel. In addition to the criteria mentioned above, this decision will be based on a number of factors, which could include:

- a knowledge of the source of the difficulty;
- detailed performance and status information;
- any predetermined plans that exist (see Recommendation E.413);
- experience with and knowledge of the network;
- routing plan employed;
- local traffic patterns;
- ability to control the flow of traffic (see Recommendation E.412).

This personnel is responsible for ensuring that conventional network management controls, once activated, are not left unsupervised.

6 Network management actions

6.1 *General*

Network management actions fall into two broad categories:

a) "expansive" actions, which are designed to make available lightly loaded parts of the network to traffic experiencing congestion;

b) "protective" actions, which are designed to remove traffic from the network during congestion which has a low probability of resulting in successful calls.

Normally, the first choice response to a network problem would be an expansive action. Protective actions would be used if expansive actions were not available or not effective.

Network management actions may be taken:

– according to plans which have been mutually agreed to between Administrations prior to the event (see Recommendation E.413);

– according to ad hoc arrangements agreed to at the time of an event (see Recommendation E.413);

– by an individual Administration wishing to reduce its traffic entering the international network, or to protect its own network.

6.2 *Expansive actions*

Expansive actions involve the rerouting of traffic from circuit groups experiencing congestion to other parts of the network which are lightly loaded with traffic, for example, due to differences in busy hours.

Examples of expansive actions are:

a) establishing temporary alternative routing arrangements in addition to those normally available;

b) in a country where there is more than one international exchange, temporarily reorganizing the distribution of outgoing (or incoming) international traffic;

c) establishing alternative routings into the national network for incoming international traffic;

d) establishing alternative routings to an international exchange in the national network for originating international traffic.

The protective action of inhibiting one direction of operation of both-way circuits [see § 6.3 a)] can have an expansive effect in the other direction of operation.

6.3 *Protective actions*

Protective actions involve removing traffic from the network during congestion which has a low probability of resulting in successful calls. Such traffic should be removed as close as possible to its origin, thus making more of the network available to traffic which has a higher probability of success.

Examples of protective actions are:

a) Temporary removal of circuits from service (circuit busying). This action may be taken when a distant part of the network is experiencing serious congestion.

   Note – In the case of both-way circuits, it may only be necessary to inhibit one direction of operation. This is called directionalization.

b) Special instructions to operators. For example, such instructions may require that only a limited number of attempts (or none at all) be made to set up a call via a congested circuit group or exchange, or to a particular destination experiencing congestion.

c) Special recorded announcements. Such announcements may be connected at an international or national exchange and, when there is serious congestion within part of the network, would advise customers (and/or operators) to take appropriate action.

d) Inhibiting overflow traffic. This action prevents traffic from overflowing onto circuit groups or into distant exchanges which are already experiencing congestion.

e) Inhibiting direct traffic. This action reduces the traffic accessing a circuit group in order to reduce the loading on the distant network.

f) Inhibiting traffic to a particular destination (code blocking or call gapping). This action may be taken when it is known that a distant part of the network is experiencing congestion.

g) Circuit reservation. This action reserves the last few idle circuits in a circuit group for a particular type of traffic.

6.4 Information on the network management controls (and their method of activation) which can be used for expansive and protective actions is found in Recommendation E.412.

6.5 *Actions during disasters*

6.5.1 Disasters whether man-made or natural can result in damage to the telephone network, they can give rise to heavy calling, or both.

6.5.2 A single point of contact for network-related information should be established to prevent confusion, duplication of effort, and to ensure an orderly process of returning communications to normal. It is recommended that the single point of contact be the network management implementation and control point (see Recommendation E.414, § 4) within the Administration affected by the disaster.

6.5.3 The role of the network management implementation and control point may vary depending on the size or impact of a disaster. However, the following are functions which may be required:

- assess the impact of the disaster on the network (transmission systems, exchanges, circuit groups, destination codes, isolated destinations);
- provide status information, as appropriate, to:
  - i) operator services
  - ii) public relations and media
  - iii) government agencies
  - iv) other network management implementation and control points;
- develop and implement control strategies (expansive and protective);
- assist in determining the need for, and locating, technical equipment to restore communications.

7 **Exchange of information**

7.1 Effective network management requires good communications and cooperation between the various network management elements within an Administration and with similar elements in other Administrations (see Recommendation E.414). This includes the exchange of real-time information as to the status and performance of circuit groups, exchanges and traffic flow in distant locations.

7.2 Such information can be exchanged in a variety of ways, depending on the requirements of the Administrations. Voice communications can be established between or among network management centres using dedicated service circuits or the public telephone network. Certain operational signals, such as switching congestion indicators, may be transported directly by the common channel signalling system. (See Recommendation Q.297 for Signalling System No. 6 and Recommendations Q.722, Q.723, Q.724, Q.762, Q.763 and Q.764 for Signalling System No. 7.) Larger data exchange requirements on a regular basis may be supported by the Telecommunications Management Network (TMN) (see Recommendation M.30) or by use of a packet switched network capability. The transfer of smaller amounts of data on an infrequent basis may be supported by telex or similar media, or by facsimile.

7.3 *Guidance on the use of common channel signalling for network management*

7.3.1 Common channel signalling systems provide a fast and reliable means of transfering network management operational signals between exchanges. An example is the transfer of exchange congestion status signals for the Automatic Congestion Control (ACC) system (see Recommendation E.412, § 3.1). These signals should be given a high priority in common channel signalling flow control. Specific details on the application of network management operational signals in Signalling System No. 6 are found in Recommendation Q.297. In the case of Signalling System No. 7, the details for the Telephone User Part (TUP) are found in Recommendations Q.722, Q.723 and Q.724, and the ISDN User Part (ISUP) are found in Recommendations Q.762, Q.763 and Q.764.

7.3.2 Signalling System No. 7 may also be used to transfer network management data and status information between an exchange and its network management operations system, and between network management operations systems. It should be noted that in these applications, the volume of data to be transferred can be quite large and its frequency of transmission can be as high as every three minutes. When this data is transferred over signalling links which also handle user signalling traffic, stringent safeguards must be adopted to minimize the risk of signalling system overloads during busy periods when both user signalling traffic and network management data transmissions are at their highest levels. These safeguards include the following:

- limiting the amount of network management information to be transferred on signalling links which also carry user signalling messages;
- using dedicated signalling links for network management purposes;
- using the telecommunications management network (TMN), or the Operations and Maintenance Application Part (OMAP) in Signalling System No. 7 (for further study);
- developing appropriate flow control priorities for network management information (for further study);
- equipping the network management operations system in such a way that it can respond to signalling system flow control messages.

8 Beginning network management

The introduction of network management into an existing network should be viewed as a long-term project. This long period is required:

- to gain knowledge and experience of network management;
- to carry out studies on the requirements of an individual network;
- to write specifications for network management requirements in present and future telephone exchanges and to hold discussions with manufacturers;
- to oversee the introduction of facilities and to organize and train suitable network management staff;
- to introduce limited facilities in existing older technology exchanges.

A rational approach would consist in first using existing limited facilities to manage the network, while at the same time developing full network management facilities with the introduction of modern stored program control (SPC) exchanges.

8.1 *Utilizing existing resources and capabilities*

8.1.1 *Responsibility*

As an important first step, the responsibility for network management should be identified and assigned within an organization. This initial organization can then be expanded, as required, in accordance with Recommendation E.414.

8.1.2 *Telephone operators*

Operators are usually aware of problems as they occur in the network, and this information can reveal the need to control traffic. The operators can then be directed to modify their procedures to reduce repeated attempts, or to use alternative routings to a destination. They can also provide special information and/or instructions to customers and distant operators during unusual situations.

8.1.3 *Exchange capabilities*

Exchanges may have been provided with certain features which can be adapted for network management purposes. Data already available for maintenance or traffic engineering purposes could be used for network management, or could be made available through the addition of an interface unit. In addition, manually operated switches or keys can be provided in electro-mechanical exchanges to block certain destination codes or to change alternate routing. They can be provided separately for each item of common control equipment, thereby allowing flexible control of traffic to a destination.

The scope for network management in a telecommunications network may depend on the technology of the exchanges in that network. However, close examination of the manufacturers' specifications for SPC exchanges may reveal that certain network management functions may be available, for example, via a maintenance terminal.

8.1.4 *Circuits*

Both-way circuits can be made busy to one direction of operation to improve the flow of traffic in the other direction. In addition, both-way and one-way circuits can be removed from service, when necessary. Both of these actions may be taken by verbal direction to the responsible maintenance organization.

8.2 *Improving capabilities*

From the experience gained through the use of these simple tools, more sophisticated network management facilities can be specified. In the interest of cost reduction, these up-graded network management capabilities should be planned for introduction as a part of a planned addition or modification to an exchange, and should be specified as a part of the initial installation of new systems. Before purchasing a new exchange, attention should be paid to the ability of the exchange to provide network management requirements as specified in Recommendations Q.542 and Q.544.

In some cases, certain off-line network management information storing and processing needs may be accommodated by the use of personal computers.

9 Considerations for the development of network management

9.1 Network management can be provided on a distributed basis, where network management functions are provided "on-site" at the exchange, or on a centralized basis, where network management functions for a number of exchanges are provided at a single location. Each approach provides certain advantages which should be recognized when deciding which one would be appropriate for an Administration's situation. In general, the decentralized distributed approach may be more appropriate where activity levels are relatively low. It may also be an appropriate way to get started in network management. The centralized approach may be more appropriate in networks where activity levels are high. In some networks, a combination of these approaches may be most effective.

9.2 *Advantages of the decentralized (distributed) approach*

The decentralized (distributed) approach provides certain advantages, which include the following:

- locally available features and capabilities can be developed and used (see § 8.1.3);
- a more detailed analysis and assessment of localized problems are possible;
- survivability of network management functions is improved, since a problem or outage at one location will not usually result in the loss of all network management capabilities;
- network management functions may be assigned to existing staff, eliminating the need to develop a dedicated, specialized staff;
- it may provide an interim capability while a long-term plan is being developed and deployed.

9.3 *Advantages of the centralized approach*

A centralized network management centre provides a number of operational benefits when compared with a distributed approach, where network management functions are provided "on-site" at the exchange. These include:

- more effective network management operations. A centralized approach is inherently more effective in dealing with complex, interrelated network problems in the SPC-common channel signalling environment, and will become more so during the transition to ISDN. In many cases, the most effective response to a problem in the international network might be to take action in the national network, and vice-versa. A centralized approach simplifies the problem of coordination of activities in these cases;

- a more "global" view of network performance. This, in turn, will permit faster and more accurate problem identification, and the development of more effective control strategies which can be implemented with less delay;
- a central point of contact for network management, both internally and with other Administrations (see Recommendation E.414);
- more efficient network management operations. The cost of staffing and training is reduced, and staff expertise is enhanced through specialization.

9.4 *Network management operations systems*

A computer-based network management operations system can provide considerable benefits to a network management centre due to its ability to process large volumes of information and to present that information in a common format. The functions of a network management operations system include the following:

- collecting alarms, status information and network management traffic data from exchanges (see § 3 and Recommendation E.502);
- processing the collected data and calculating network management parameters (see § 3 and Recommendation E.502);
- providing performance reports (see § 9.4.1);
- comparing network management parameters with thresholds to identify unusual conditions;
- applying controls in exchanges based on input commands;
- calculating hard-to-reach status of destinations and providing this information to exchanges;
- interfacing with network management centre visual displays, and work station terminals and printers;
- preparing administrative reports;
- maintaining a database of network statistics and information.

Note – Many of these functions can also be provided to the Network Management Centre by each SPC exchange, however, the provision of these functions in a network management operations system may reduce the requirements placed on the exchanges.

9.4.1 *Performance reports*

Performance reports can be provided in the following ways:

i) *automatic data* – this data is provided automatically as specified in the operations system software, and cannot be readily changed by the network manager;

ii) *scheduled data* – this data is provided according to a schedule established by the network manager;

iii) *demand data* – this data is provided only in response to a specific request by the network manager. In addition to performance data, demand data includes reference data, such as the number of circuits provided or available for service, routing information, assigned threshold values, numbers of installed switching system components, etc.;

iv) *exception data* – this data is provided when a data count or calculation crosses a threshold established by the network manager.

Data reports can be provided on a regular basis, for example, every 3 minutes, 5 minutes, 15 minutes, 30 minutes, or hour. The specific interval for any data report will be determined by the network manager. Historic data relating to at least the previous two or three periods should also be available.

9.4.2 *Other considerations*

It should be noted that shorter collection intervals increase the usefulness of the data to the network manager, but also increase the size and cost of the operations system and may increase the volatility of the data.

It should also be noted that it is important that network management controls should not become completely unavailable due to the failure or malfunction of the network management operations system or of its communications links with exchanges. Therefore, network management operations systems should be planned with a high degree of reliability, survivability and security. This could be achieved through the provision of certain essential capabilities (such as controls and automatic routing protection mechanisms) on-site in the exchange, or by redundancy in computers and data links, or through the provision of alternative stand-by centres.

The failure of a network management operations system should not have an adverse impact on normal traffic flow in the network.

ANNEX A

(to Recommendation E.411)

Terminology for network management

A.1 circuit

A circuit connects two exchanges. A national circuit connects two exchanges in the same country. An international circuit connects two international exchanges situated in different countries. (Based on Recommendation D.150 and Recommendation F.68.)

A.2 circuit group

The set of all switched circuits which directly interconnect one exchange with another.

A.3 circuit sub-group

A set of circuits within a circuit group which are uniquely identifiable for operational or technical reasons. A circuit group may consist of one or more circuit sub-groups.

A.4 destination

A country in which the called subscriber is located or an area or other location that may be specified within that country. A destination can be identified by the digits used for routing the call.

A.5 bid

An attempt to obtain a circuit in a circuit group or to a destination. A bid may be successful or unsuccessful in seizing a circuit in that circuit group or to that destination.

A.6 seizure

A seizure is a bid for a circuit in a circuit group which succeeds in obtaining a circuit in that circuit group.

A.7 answer signal

A signal sent in the backward direction indicating that the call is answered. (Based on Recommendation Q.254.)

A.8 holding time

The time interval between seizure and release of a circuit or switching equipment.

A.9 busy-flash signal (sent in the backward direction)

This signal is sent to the outgoing international exchange to show that either the circuit group, or the called subscriber, is busy (Signalling Systems No. 4 and No. 5, see Recommendations Q.120 and Q.140).

Note – In Signalling Systems No. 6 and No. 7, there is no busy-flash signal. However, the equivalent of busy-flash can be roughly approximated through the aggregation of specific backward failure signals such as circuit group congestion, national network congestion and subscriber busy.

NETWORK MANAGEMENT CONTROLS

1 Introduction

1.1 Network management controls provide the means to alter the flow of traffic in the network in support of the network management objectives given in Recommendation E.410. Most network management controls are taken by or in the exchange (see Recommendation Q.542), but certain actions can be taken external to the exchange. This Recommendation provides specific information on network management controls and gives guidance concerning their application. However, it should be noted that the suggested use for each network management control is given only for the purpose of illustration. Other controls, separately or in combination, may be more appropriate in any given situation.

1.2 The application or removal of network management controls should be based on network performance data which indicates that action is required in accordance with the network management principles in Recommendation E.410, § 4. Performance data will also measure the effect of any network management control taken, and will indicate when a network management control should be modified or removed (see Recommendations E.411 and E.502).

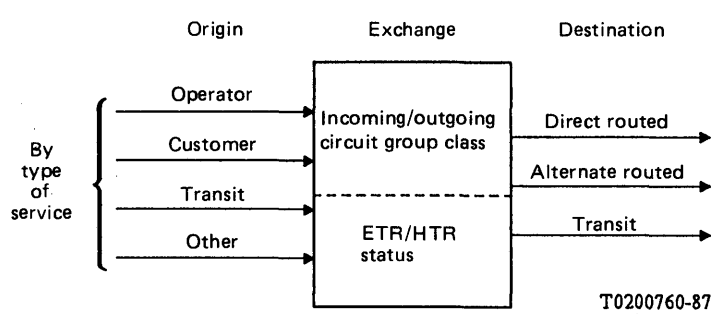
1.3 Controls can be activated or removed in an exchange by input from a network management operations system or by direct input from a terminal. In some cases, controls can be activated automatically either by external or internal stimulus, or when a parameter threshold has been exceeded. [The automatic congestion control (ACC) system is an example (see § 4.1).] When automatic control operation is provided, means for human override should also be provided.

2 Traffic to be controlled

2.1 *Type of traffic*

Exchanges should be capable of applying a range of network management controls (see Recommendation Q.542). For increased flexibility and precision, there is considerable advantage when the effect of a control can be limited to a particular specified traffic element.

The operating parameters of a control can be defined by a set of traffic attributes. As shown in Figure 1/E.412, these parameters include distinctions based on the origin of the traffic, for example customer-dialled, operator-dialled, transit or other such classification as may be specified by the Administration. These can be further classified by type of service, particularly for ISDN.



ETR/HTR Easy to reach/Hard to reach

FIGURE 1/E.412

Traffic distinctions for controls

Additional attributes can be specified based on information which may be available in the exchange. For example, incoming/outgoing circuit group class, or hard-to-reach status of destinations (see § 2.2) can be used. Further distinctions can be based on the outgoing traffic type, for example direct routed, alternate routed or transit.

In general, the more attributes that can be specified for a control, the more precise will be its effect.

Note – Precision is of vital importance, particularly in the case of protective controls.

2.2 *Hard-to-reach (HTR) process*

2.2.1 A hard-to-reach process for network management will enable exchanges to automatically make more efficient use of network resources during periods of network congestion by improving the performance of network management controls. This improved performance is derived from the ability to distinguish between destinations that are easy to reach (ETR) and destinations that are hard-to-reach (HTR), (e.g., destinations with a low answer bid ratio) and applying heavier controls to HTR traffic. This distinction can be based on:

i) internal performance measurements within the exchange and/or the network management operations system;

ii) similar information gathered and reported by other exchanges;

iii) historical and current observations of network performance by network managers.

The network manager should have the ability to set the threshold for HTR determination in the exchange or network management operations system, and to assign a destination as HTR regardless of its actual status.

2.2.2 *Controlling traffic based on HTR status*

When a call to a destination that is on the HTR list is being routed and a network management control on HTR traffic is encountered, the call should be controlled according to the relevant parameters. If a destination is considered HTR, it normally should be HTR for all outgoing circuit groups.

Additional details of the hard-to-reach process can be found in Recommendation Q.542.

2.3 *Methods for specifying the amount of traffic to be controlled*

2.3.1 *Call percentage control*

There is considerable advantage when exchange controls can be activated to affect a variable percentage of traffic (for example 10%, 25%, 50%, 75% or 100%).

2.3.2 *Call rate control*

An ability to set an upper limit on the maximum number of calls to be allowed to access the network during a specified period of time is of particular advantage.

3 Exchange controls

Network management controls may be applied in exchanges to control traffic volume or to control the routing of traffic. The resulting effect on traffic of these controls may be expansive or protective, depending on the control used, its point of application and the traffic element selected for control.

3.1 *Traffic volume controls*

Traffic volume controls generally serve to control the volume of traffic offered to a circuit group or destination. These include the following:

3.1.1 *Destination controls*

3.1.1.1 *Code blocking*

This control bars routing for a specific destination on a percentage basis. Code blocking can be done on a country code, an area code, an exchange identifying code or an individual line number. The last of these is the most selective control available.

Typical application: Used for immediate control of focussed overloads or mass-calling situations.

3.1.1.2 *Call-gapping*

This control sets an upper limit on the number of call attempts allowed to be routed to the specified destination in a particular period of time (for example, no more than 5 call attempts per minute). Thus, the number of call attempts that are routed can never exceed the specified amount.

Typical application: Used for the control of focussed overloads, particularly mass-calling to an individual line number. A detailed analysis may be required to determine the proper call-rate parameters.

3.1.2 *Cancellation of direct routing*

This control blocks the amount of direct routed traffic accessing a circuit group.

Typical application: Used to reduce traffic to congested circuit groups or exchanges where there is no alternate routed traffic.

3.1.3 *Circuit directionalization*

This control changes both-way operated circuits to incoming operated circuits, either on a percentage basis or by a specified number of circuits. At the end of the circuit group for which access is inhibited, this is a protective action, whereas at the other end of the circuit group (where access is still available), it is an expansive action.

Typical application: To enhance the flow of traffic outward from a disaster area while inhibiting incoming traffic. To have an effect, it is recommended that the minimum amount of directionalization be at least 50%.

3.1.4 *Circuit turndown/busying/blocking*

This control removes one-way and/or both-way operated circuits from service, either on a percentage basis or by a specified number of circuits.

Typical application: Used to control exchange congestion when no other control action is available.

3.1.5 *Specialized volume controls*

Both the automatic congestion control (ACC) system and the selective circuit reservation control (SCR) are volume controls, but due to their specialized nature, they are described separately in § 4.1 and § 4.2.

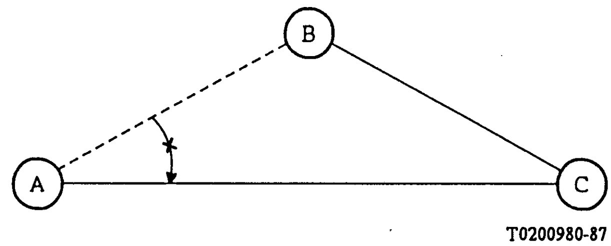
3.2 *Routing control*

Routing controls are used to control the routing of traffic to a destination, or to or from a circuit group. However, it should be noted that in some cases a routing control may also affect the volume of traffic. Controls which are applied to circuit groups may also be applied to circuit sub-groups, when appropriate.

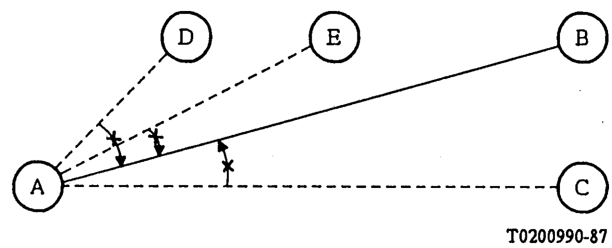
3.2.1 *Cancellation of alternative routing*

Two versions of this control are possible. One version prevents traffic from overflowing *FROM* the controlled circuit group: alternative routing from (ARF). The other version prevents overflow traffic from all sources from having access *TO* the controlled circuit group: alternative routing to (ART). See Figure 2/E.412.

Typical application: There are many uses for this control. These include controlling alternative routing in a congested network to limit multi-link connections, or to reduce alternative routed attempts on a congested exchange.



a) "Alternative Routing From" (ARF) cancellation on A-B circuit group



b) "Alternative Routing To" (ART) cancellation on A-B circuit group

FIGURE 2/E.412

Examples of alternative routing cancellation

3.2.2 *Skip*

This control allows traffic to bypass a specified circuit group and advance instead to the next circuit group in its normal routing pattern.

Typical application: Used to bypass a congested circuit group or distant exchange when the next circuit group can deliver the call attempts to the destination without involving the congested circuit group or exchange. Application is usually limited to networks with extensive alternative routing. When used on both-way circuit groups it has an expansive effect on traffic flow in the opposite direction.

3.2.3 *Temporary alternative routing*

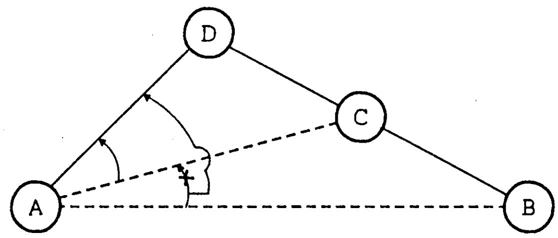
This control redirects traffic from congested circuit groups to other circuit groups not normally available which have idle capacity at the time.

Typical application: To increase the number of successful calls during periods of circuit group congestion and to improve the grade of service to subscribers.

3.2.4 *Special recorded announcements*

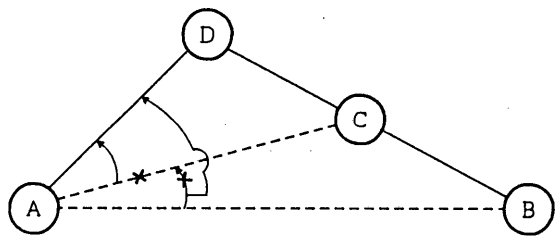
These are recorded announcements which give special information to operators and/or subscribers, such as to defer their call to a later time.

Typical application: Used to notify customers of unusual network conditions, and to modify the calling behavior of customers and operators when unusual network conditions are present. Calls that are blocked by other network management controls can also be routed to a recorded announcement.



T0201000-87

a) Skip alternative routing traffic on A-C circuit group



T0201010-87

b) Skip direct and alternative routing traffic on A-C circuit group

FIGURE 3/E.412

Examples of skip

4 Automatic exchange controls

Automatic dynamic network management controls represent a significant improvement over conventional controls. These controls, which are preassigned, can quickly respond to conditions internally detected by the exchange, or to status signals from other exchanges, and are promptly removed when no longer required. Automatic control applications should be planned, taking into account the internal overload control strategy provided in the exchange software.

4.1 *Automatic congestion control system*

4.1.1 *Exchange congestion*

When a digital international/transit exchange carries traffic above the engineered level, it can experience an overload that diminishes its total call processing capability. Because of the speed of the onset of such congestion and the critical nature of the condition, it is appropriate that control be automatic. The automatic congestion control (ACC) system consists in the congested exchange sending a congestion indicator to the connected exchange(s) using common channel signalling. The exchange(s) receiving the congestion indication can respond by reducing a certain percentage of the traffic offered to the congested exchange, based on the response action selected for each application.

4.1.2 *Detection and transmission of congestion status*

An exchange should establish a critical operating system benchmark, and when continued levels of nominal performance are not achieved (e.g. due to excessive traffic), a state of congestion is declared. Thresholds should be established so that the two levels of congestion can be identified, with congestion level 2 (CL2) indicating a more severe performance degradation than congestion level 1 (CL1). When either level of congestion occurs, the exchange should have the capability to:

1) code an ACC indication in the appropriate common channel signalling messages, and
2) notify its network management centre and support system of a change in its current congestion status.

4.1.3 *Reception and control*

When an exchange receives a signal that indicates a congestion problem at a connected exchange, the receiving exchange should have the capability to reduce the number of seizures sent to the congested exchange.

An exchange should have the capability of:

1) assigning an ACC response action on an individual circuit group[1] basis, as specified by the network manager, and

2) notifying its network management centre and support system of a change in congestion status received from a distant exchange.

There should be several control categories available in the exchange. Each category would specify the type and amount of traffic to be controlled in response to each of the received ACC indicators. The categories could be structured so as to present a wide range of response options.

For a specific ACC response category, if the received ACC indicator is set to a CL1 condition then the receiving exchange could, for example, control a percentage of the Alternate Routed To (ART) traffic to the affected exchange. The action taken by the control would be to either SKIP or CANCEL the controlled calls, depending on the ACC response action that was assigned to that circuit group. In a similar manner, if a CL2 condition is indicated, then the receiving exchange could control all ART traffic and some percentage of Direct Routed (DR) traffic. Other options could include the ability to control hard-to-reach traffic, or transit traffic. In the future, control categories could be expanded to include service-specific controls. This would be particularly useful in the transition to ISDN.

Note – ACC response categories can be set locally in the exchange or by input from a network management centre, or operations system.

Table 1/E.412 is an example of the flexibility that could be achieved in response to a signal from an exchange that is experiencing congestion. In this example, different control actions would be taken based upon the distinction between ART and DR traffic types. These actions could represent the initial capabilities available with the ACC control. Other alternatives in the future could include the ability to control hard-to-reach traffic (see § 2.2), or transit traffic or to provide other controls such as call-gapping. Additional response categories could also be added to Table 1/E.412 to give greater flexibility and more response options to the ACC control. It could also be possible to exclude priority calls from ACC control.

TABLE 1/E.412

ACC control response

Congestion level	Traffic type	Response category		
		A	B	C
CL1	ART	0	0	100
	DR	0	0	0
CL2	ART	100	100	100
	DR	0	75	75

[1] In this context, the term "circuit group" refers to all of the outgoing and both-way circuit sub-groups which may directly connect the congested exchange and the responding exchange.

4.1.4 Any international application of ACC should be based on negotiation and bilateral agreement among the affected Administrations. This includes an agreement as to whether the controlled calls should be skipped or cancelled. Application within a national network would be a national matter. An exchange that is capable of "ACC receive and control" should not indiscriminately assign ACC to all routes since a distant exchange may be equipped for common channel signalling, but may not yet have an ACC transmit capability. This could result in invalid information in the ACC fields in the signalling messages and the inappropriate application of ACC controls at the receiving exchange. Additional details on the ACC system are in Recommendation Q.542.

4.2 *Selective circuit reservation control*

4.2.1 The selective circuit reservation control enables an exchange to automatically give preference to a specific type (or types) of traffic over others (e.g., direct routed calls over alternate routed calls) at the moment when circuit congestion is present or imminent. The selective circuit reservation control can be provided with one or two thresholds, with the latter being preferred due to its greater selectivity. Specific details on the selective circuit reservation control may be found in Recommendation Q.542.

4.2.2 *General characteristics*

The selective circuit reservation control has the following operating parameters:

- a reservation threshold(s),
- a control response,
- a control action option.

The reservation threshold defines how many circuits or how much circuit capacity should be reserved for those traffic types to be given preferred access to the circuit group. The control response defines which traffic types should be given a lesser preference in accessing the circuit group, and the quantity of each type of traffic to control. The control action option defines how those calls denied access to the circuit group should be handled. The control action options for processing of calls denied access to the circuit group may be SKIP or CANCEL.

When the number of idle circuits or the idle capacity in the given circuit group is less than or equal to the reservation threshold, the exchange would check the specified control response to determine if calls should be controlled. The SKIP response allows a call to alternate-route to the next circuit group in the routing pattern (if any) while the CANCEL response blocks the call.

These parameters should be able to be set locally in the exchange for each selected circuit group or by input from a network management operations system. In addition, the network manager should have the capability to enable and disable the control, and to enable the control but place it in a state where the control does not activate (e.g., by setting the reservation threshold to zero). Further, the network manager should have the ability to set the values for the response categories.

4.2.3 *Single threshold selective circuit reservation control*

In this version of the control, only a single reservation threshold would be available for the specified circuit group.

Table 2/E.412 is an example of the flexibility that could be achieved in the control's response to circuit group congestion. In the future, other distinctions between traffic could be identified that would expand the number of traffic types in Table 2/E.412. An example would be to control hard-to-reach traffic as indicated in § 2.2, or to give preference to priority calls.

4.2.4 *Multi-threshold selective circuit reservation control*

The multi-threshold control provides two reservation thresholds for the specified circuit group. The purpose of multiple reservation thresholds is to allow a gradual increase in the severity of the control response as the number of idle circuits in the circuit group decreases. The only restriction on the assignment of reservation thresholds would be that a reservation threshold associated with a more stringent control must always be less than or equal to the reservation threshold of any less stringent control, in terms of the number of reserved circuits, or circuit capacity.

TABLE 2/E.412

An example of a single threshold selective circuit reservation
Percentage control response table

Circuit group reservation threshold	Traffic type	Response category assigned to circuit group		
		A	B	C
RT1	ART	25	50	100
	DR	0	0	25

Table 3/E.412 is an example of the flexibility that could be achieved in the control's response to circuit group congestion with a two-reservation threshold control. In the future, other distinctions between traffic could be identified that would expand the number of traffic types in Table 3/E.412. An example would be to control hard-to-reach traffic as indicated in § 2.2.

TABLE 3/E.412

An example of a two-threshold selective circuit reservation
Percentage control response table

Circuit group reservation threshold	Traffic type	Response category assigned to circuit group				
		A	B	C	D	E
RT1	ART	25	50	75	100	100
	DR	0	0	0	0	0
RT2	ART	50	75	75	100	100
	DR	0	0	25	50	100

5 Status and availability of network management controls

5.1 The exchange and/or network management operations systems should provide information to the network management centre and/or the exchange staff as to what controls are currently active and whether the controls were activated automatically or by human intervenion. Measurements of calls affected by each control should also be available (see Recommendation E.502).

5.2 To help insure the viability of network management functions during periods of exchange congestion, network management terminals (or exchange interfaces with network management operations systems), and functions such as controls, should be afforded a high priority in the exchange operating software.

6 Operator controls

Traffic operators are usually aware of problems as they occur in the network, and this information can reveal the need to control traffic. The operators can then be directed to modify their normal procedures to reduce repeated attempts (in general, or only to specified destinations), or to use alternative routings to a destination. They can also provide information to customers and distant operators during unusual situations, and can be provided with special call handling procedures for emergency calls.

Recommendation E.413

INTERNATIONAL NETWORK MANAGEMENT – PLANNING

1 Introduction

1.1 Many situations arise which may result in abnormally high or unusually distributed traffic levels in the international network, or loss of network capacity, or both. These situations include the following:

- peak calling days,
- failure of transmission systems (including planned outages),
- failure of exchanges,
- failure of common channel signalling systems,
- mass-calling situations,
- disasters,
- introduction of new services.

Experience has shown that advanced planning for these situations has a beneficial effect on overall network management efficiency and effectiveness. The timely application of planned control strategies can be instrumental in improving network performance.

1.2 For known or predictable events, predetermined network management plans should be developed and agreed between Administrations, bearing in mind the costs involved. The degree of detail of any plan will depend on the type of situation to be covered. For example, a recurring event such as Christmas or New Year's Day may be planned in great detail. The lack of real-time network management facilities in an Administration should not preclude planning activities.

1.3 When unforeseen situations arise for which predetermined plans do not exist, ad hoc arrangements will need to be agreed at the time. Whether network management actions result from a negotiated plan, or an ad hoc arrangement, it is essential that agreement be reached between Administrations concerned before such actions are actually implemented.

1.4 Network management planning is normally performed by the "network management planning and liaison" point (see Recommendation E.414).

1.5 Another aspect of network management planning is long-range planning for the development and introduction of new network management techniques and capabilities for surveillance and control. This includes the development of new or improved controls which may be necessary due to the introduction of new services or the transition to ISDN. These functions are normally performed by the "network management development" point (see Recommendation E.414).

2 Development of plans

2.1 A comprehensive network management plan would include some or all of the following, as appropriate:

- Key indicators or criteria which should be used to decide when a plan should be implemented.
- The identification of destinations or points likely to be affected, along with an assessment as to the likely impact on originating and/or terminating traffic.
- Control actions which may be required or that should be considered locally and in distant locations. This includes the identification of temporary alternative routings which may be available for use, and the modifications to automatic controls which may be necessary.

- Special call handling procedures to be used by operators, and notification requirements.
- Communication requirements. This includes identification of the necessary information flows between the network management centre and other organizations which may be involved or may have information concerning the problem (such as maintenance and operator centres).
- Data requirements. This includes determining what information may be relevant and where it is available.
- Key events or milestones. These are critical elements which can measure the success or progress of a plan, and indicate when certain actions should begin or end.

2.2 Regardless of the format or detail in a plan, it will not be fully effective unless it is readily available and understood by all who may be involved including other Administrations. This requires that network management plans be reviewed on a regular basis. Plans should be reviewed to ensure that they reflect changes or additions that may have taken place in the network since the plan was prepared. This is particularly important for plans which are used infrequently. Attention should be directed to changes in routing, the introduction of new circuit groups, new exchanges or common channel signalling, or the addition of new network management capabilities since the plan was first developed.

2.3 When developing network management plans, it is important that they be flexible and, if possible, contain a number of alternatives. This is necessary because a planned action may not be viable or available at a given time, for example:

- it may be under consideration for the same or another problem,
- it may already be in use for some other purpose,
- a planned transit point may not be available due to congestion or a lack of spare capacity to or from the transit point at the time.

3 Peak day planning

3.1 There are a number of days which give rise to heavy calling in the international network. These usually correspond to certain religious or national holidays. Plans should be developed for those holidays which have resulted, or are expected to result, in unusually heavy traffic.

Peak-day calling can result in significant and sustained blockages in the network. This can be caused by two factors:

- the average length of conversation on a peak day in many cases can be significantly longer than on a normal business day;
- the calling pattern (which is usually residential in nature) may be different than the normal pattern (which is usually business-oriented).

A combination of these factors can result in a network that is highly congested and which requires careful planning and extensive network management controls to optimize service and revenues.

It should be noted that many peak calling days may also be public holidays. As a result, staffing in telephone exchanges and administrative offices may be minimal and some traffic data and service measurements may not be readily available. These factors should also be considered in peak-day planning.

3.2 Peak-day plans may include information on the following, as appropriate:

- Network management staffing requirements and expected hours of operations, and the exchange of such information with other network management centres.
- Provision of temporary additional circuits.
- Directionalization of both-way circuits where appropriate.
- Temporary alternative routings to take advantage of anticipated idle capacity.
- Controls to inhibit alternate routing via transit points that are expected to be congested.
- Identification of anticipated hard-to-reach points and planned controls to reduce attempts to hard-to-reach points.
- Special calling procedures for operators, including the exchange of network status information with operator centres.

- Advance testing of new controls, or those infrequently used (including the testing of the rerouting to ensure proper operation and the ability to complete to a terminating number via the transit point).
- Consideration of limiting installation and maintenance activity just prior to the peak day to only essential work in order to insure that all available circuits and switching equipment are in service.
- Procedures to take into account special situations, such as inter-ISC circuit groups, circuit multiplication systems, etc.

4 Transmission system failure planning

4.1 The impact on service of the failure of an international transmission system will depend on a number of variables:

- the size of the failed transmission system and its relationship to the total network capacity;
- its loading (the number of channels that are assigned for use) (this may change frequently);
- the destinations and/or services assigned to the transmission system and their relationship to their respective totals (this may change frequently);
- the traffic intensity during the period from the onset of a failure until restoration or repair (this can vary significantly);
- the duration of the failure (this is usually unpredictable);
- the availability of restoration capacity (this can vary).

Thus, it can be seen that it is difficult, if not impossible, to predict the precise impact on service of a failure at a given point in time. However, recognizing the increasing size and loading of modern transmission systems, the impact of a failure on service can often be severe, and as a result, significant effort has been expended by Administrations to develop and refine transmission system failure restoration plans.

Experience has shown that network management actions can also play a significant role in minimizing the adverse impact of failures on service. However, it should be noted that these network management actions will usually complement or enhance a transmission failure restoration plan and do not necessarily supplant the need for such plans. For short duration failures, e.g., solar interference on satellites, network management plans may be the only viable solution.

4.2 When an international transmission system fails, network management and transmission restoration activities should proceed in parallel on a coordinated basis.

- The network management centre will become aware of the impact of a failure on service via its network surveillance capacity; in some cases, this will occur before the specific details of the failure are known. The network management centre can identify the affected routes, destinations and/or services. This information will guide the application of network management controls and may also be useful to the restoration control point (Recommendation M.725) in setting priorities for restoration.
- The first response of the network management centre should be to consider the use of temporary alternative routings in order to complete traffic which is being blocked by the failure. In many cases, these actions can begin immediately, before the decision is made to activate a transmission restoration plan.
- If significant congestion continues despite the expansive controls, protective controls should be considered. Emphasis should be placed on the identification of destinations that are hard-to-reach and the selective reduction of traffic to these points so that the remaining network can be used by traffic with a higher probability of success.

4.3 It is recommended that a network management plan for the failure of a major international transmission system should include the following, as appropriate:

- identification of destinations or points affected for originating and terminating traffic,
- temporary alternative routings which may be utilized to bypass the failure, and hours of availability,
- notification lists,
- special call handling procedures for operators,
- controls which may be required in connected networks,
- controls to be requested of distant network management centres,
- actions to be taken after fault correction to restore the network to its normal configuration,
- special recorded announcements to customers, when necessary.

5 International exchange failure planning

5.1 The impact on service of the failure of an international exchange will depend on a number of variables, which include:

- whether there is a single or multiple international exchange(s),
- the routing plan and the distribution of circuit groups among the international exchanges,
- the traffic intensity during the failure,
- the duration of the failure,
- the size (capacity) and the current loading of the failed exchange, and its relationship to the total international switching capacity.

In any case, the failure of an international exchange usually will have a severe impact on service. Network management exchange failure plans can provide considerable benefits during the failure by limiting the spread of congestion to connected exchanges and providing alternative ways of routing traffic to bypass the failed exchange.

5.2 It is recommended that a network management exchange failure plan should include the following information, as appropriate:

- general information about the exchange and its function in the network, including diagrams of the normal network configuration and the reconfigured network during a failure,
- actions to be taken to verify a total failure of an exchange to differentiate it from certain fault recovery actions in SPC exchanges which may, at first, appear similar,
- notification lists,
- initial control actions to be taken upon verification of exchange failure,
- additional control actions to be taken based on the prognosis of the failure,
- controls to be applied within the national network,
- controls to be requested of distant network management centres,
- modifications which may be required to automatic controls,
- sequence of control removal when the exchange is restored to normal operation.

5.3 It is recommended that network management exchange failure plans be reviewed and up-dated whenever a significant change in network configuration occurs, or at least annually. A network management exchange failure plan should be prepared for a new international exchange before it is introduced into the network.

6 Common channel signalling (CCS) failure planning

6.1 When a failure in the common channel signalling system interrupts the flow of traffic, the affected traffic may be diverted by network management controls to other unaffected circuits groups. It is preferable that these actions be planned in advance. These plans should identify the modifications to the automatic CCS flow control responses which may be required in the exchanges to permit the planned actions to be taken [for example, to change the normal programmed response to the receipt of a transfer prohibited signal (TFP)].

6.2 It should be noted that, as more of the international network converts to common channel signalling, the availability of potential alternative routings may become limited, which will increase the need for careful planning.

7 Mass-calling planning

7.1 Uncontrolled mass-calling has the potential to seriously disrupt calling in the network. However, with proper planning, the adverse effects of many mass-calling situations may be minimized. The key to success is advance warning and interdepartmental cooperation and planning.

This requires that the Administration be alert to potential mass-calling situations so that the proposed use of the network can be evaluated in advance to determine the potential for congestion. When congestion appears likely, alternative serving arrangements may be proposed, which may include the use of network management controls.

7.2 With widespread availability of call-gapping controls (see Recommendation E.412), certain mass-calling applications may be provided without harm to the network. The call-gap controls can be set at each exchange to limit outgoing calls to only the amount necessary to keep the called lines filled. It must be noted, however, that no mass-calling control strategy can prevent originating congestion and dial tone delays in local exchanges if a large number of customers simultaneously attempt to dial a service or specific number.

8 Disasters

Disasters can be natural (for example, a typhoon, an earthquake) or man-made (an airplane or railroad accident). These events can result in either damage to network facilities or in an extraordinary number of calls, or both. While it is difficult to predict such a disaster, the effects of a disaster on the telephone network can be predicted with some degree of accuracy and plans developed accordingly. These plans should include:

- contact and notification lists,
- control actions required locally and/or in other Administrations,
- arrangements for additional staffing and extended hours of operation.

(See Recommendation E.411, § 6.5.)

9 Planning for the introduction of new services

The introduction of new services in the network may result in new or unusual traffic flow characteristics, and/or unusual traffic demand, particularly when there is strong initial interest in the new service. Therefore, the potential impact on the network of a new service should be evaluated to identify where congestion or deteriorated service might occur, and to identify what special network management surveillance and control capabilities may be required. It is important that this analysis take place well in advance of the planned service availability date, so that the necessary modifications to the exchange and/or network management operations system software can be completed in a timely manner. This will help to insure that the necessary surveillance and control capabilities will be available when the new service is introduced.

10 Negotiation and coordination

10.1 Administrations should exchange information concerning their network management capabilities as part of the network management planning process. Specific plans should be negotiated in advance on a bilateral or multilateral basis, as appropriate. Negotiation in advance will allow time to fully consider all aspects of a proposed plan and to resolve areas of concern, and will permit prompt activation when needed.

10.2 The use of any network management plan must be coordinated with the involved Administrations at the time of implementation. This will include (as appropriate):

- determining that planned transit exchange(s) have switching capacity to handle the additional traffic,
- determining that there is capacity in the circuit group(s) between the planned transit point and the destination,
- advising the transit Administration(s) that transit traffic will be present in its circuit groups and exchanges,
- arranging for the activation of controls at distant locations,
- arranging for surveillance of the plan while in effect to determine the need to modify the plan.

When the use of a plan is no longer required, all involved Administrations should be notified of its discontinuance, so that the network can be restored to its normal configuration.

INTERNATIONAL NETWORK MANAGEMENT – ORGANIZATION

1 Introduction

The required high degree of cooperation and coordination in international network management can best be achieved by efficient and effective interworking between international network management organizations in the various countries. This Recommendation specifies the organizational elements necessary for this purpose, and outlines the functions and responsibilities of each element.

Only those organizational elements vital to the network management development, planning, implementation and control of the international network are dealt with in the Recommendation. It is recognized that other functions must necessarily be carried out within the network management organization, either in support of the functions specified below or in connection with the management of the national network.

It is also recognized that Administrations may not wish to assign each element to a separate staff or create a separate organization. Administrations are, therefore, afforded the freedom to organize such functions in a manner which best suits their own situation and the level of development of network management.

2 International network management – organization

2.1 As far as international cooperation and coordination are concerned, network management should be based on an organization comprising the following elements, all of which should exist in each country practicing international network management:

a) network management planning and liaison;

b) network management implementation and control;

c) network management development.

Each element represents a set of functions and reponsibilities, and are further defined in §§ 3 to 5.

2.2 At the discretion of the Administration concerned, the elements defined in §§ 3 to 5 below can be grouped together in a single organizational entity, for example, an International Network Management Centre. This is likely to be the most convenient and efficient approach where the level of development and degree of practice of network management is high. Where such an approach is not possible, or is impractical, international network management functions could be carried out at locations where related activities are performed. § 6 offers specific guidance on the relationship between network management and network maintenance, and includes consideration for the possible combining of organizational elements involved in the two fields of activity.

2.3 Irrespective of which arrangement an Administration decides for its international network management organization, it must ensure that the functions and responsibilities of a particular organizational element are not divided between two separate locations. Administrations can then issue a list of contact point information (see § 7 for guidance) which will give telephone, telex numbers, service hours etc. for each element.

3 Network management planning and liaison

3.1 Network management planning and liaison is an element within the international network management organization. It is concerned with liaising with other Administrations to develop plans to cater for unforeseen high traffic levels and any other situation likely to adversely affect the completion of international calls.

3.2 Network management planning and liaison is responsible for the following set of functions:

a) liaising with similar points in other Administrations to determine the actions necessary to overcome unforeseen high traffic levels and other situations adversely affecting the completion of international calls;

b) producing plans to cater for the abnormal traffic levels produced by foreseen national and international events;

c) liaising with the restoration liaison officer (RLO) within the Administration concerning network management plans for failures and planned outages;

d) liaising with similar points in other Administrations to establish the required actions when plans to overcome abnormal situations cannot be implemented;

e) ensuring that the facilities and network management controls required for the rapid implementation of agreed plans are available and ready for use when required.

4 Network management implementation and control

4.1 Network management implementation and control is an element within the international network management organization. It is concerned with monitoring the performance and status of the network in real time, determining the need for network management action, and, when necessary, implementing and controlling such action.

4.2 Network management implementation and control is responsible for the following set of functions:

a) monitoring the status and performance of the network;

b) collecting and analysing network status and performance data;

c) determining the need for the control of traffic as indicated by one or more of the following conditions:

- the failure or planned outage of an international or national transmission system,
- the failure or planned outage of an international or national exchange,
- congestion in an international exchange,
- congestion in a distant network,
- congestion to an international destination,
- heavy traffic caused by an unusual situation;

d) applying or arranging for network management control action, as described in Recommendation E.411, and Recommendation E.412;

e) liaising and cooperating with similar points in other Administrations in the application of network management controls;

f) liaising with the fault report point (network)[1] within the Administration concerning the exchange of information available at either point;

g) liaising with the restoration control point[2] within the Administration concerning failures and planned outages;

h) disseminating information as appropriate within its own Administration concerning network management actions which have been taken.

5 Network management development

5.1 Network management development is an element within the international network management organization. It is concerned with the development and introduction of techniques and facilities for the purpose of network management surveillance and control at the international level, although it may also have similar responsibilities for the national network.

5.2 Network management development is responsible for the following set of functions:

a) developing facilities to enable the application of current network management techniques;

b) long range planning for the coordinated introduction of new network management techniques and improved network surveillance and controls;

c) evaluating the effectiveness of current plans, controls and strategies with a view to identifying the need for improved controls, control strategies and support systems, particularly those which may be required for new services and the ISDN.

---

[1] The fault report point (network) is a functional element in the general maintenance organization (see Recommendation M.716).

[2] The restoration control point is a functional element in the general maintenance organization (see Recommendation M.725).

6 Cooperation and coordination between network management and network maintenance organizations

Considerable benefit may be obtained by close cooperation and coordination between the network management organization identified in this Recommendation and the network maintenance organization identified in the M.700 series of Recommendations. For example, reports of network difficulties received by the fault report point (network) in the maintenance organization may assist the network management implementation and control point in refining its control action. Similarly, difficulties reported to the fault report point (network) may be explained by information already available to the network management implementation and control point. For this reason, and to take into account the particular operating situation and stage of development of network management within an Administration, some of the functional elements identified in this Recommendation may be located with one of the groupings of functional elements of the network maintenance organization as outlined in Recommendation M.710.

Where it is advantageous to create a separate international management centre containing the elements defined above, care should be taken to ensure that suitable liaison and information flows occur between such a centre and the network maintenance organization.

7 Exchange of contact point information

For each of the three organizational elements in §§ 3 to 5 above, Administrations should exchange contact point information. Network management contact points should be exchanged as part of the general exchange of contact point information as specified in Recommendation M.93.

SECTION 3

CHECKING THE QUALITY OF THE INTERNATIONAL TELEPHONE SERVICE

Recommendation E.420

CHECKING THE QUALITY OF THE INTERNATIONAL TELEPHONE SERVICE – GENERAL CONSIDERATIONS

1 Quality of service parameters

An adequate picture of the level of quality of service (QOS) in the network can be defined by a set of parameters which are measured, registered and data processed.

In Recommendation E.800 a set of performance concepts is defined in order to provide a satisfactory description of the quality of service, and the interconnection of those concepts is shown. Each performance concept can be impaired by a number of particular causes. These causes, either singly or in groups, lie behind the failure symptoms observed by the user.

A user views the provided service from outside the network and his perception can be described in observed quality of service parameters. The link between the observed quality of service parameters and the impairment causes can be indicated in the form of tables. [1)]

Five main observed quality of service parameters are derived; they reflect the quality of:

i) providing the customer with the ability to use the desired services;

ii) furnishing a desired level of service for:
- connection establishment,
- connection retention,
- connection quality,
- billing integrity.

These main parameters can be supervised by quality of service indicators (e.g. efficiency rate, call cut-off rate, etc.).

Objectives can be set for these indicators and can be revised at regular intervals.

When a deterioration of these supervision indicators is detected, or when an improvement programme is started, more data must be collected by measurements to permit a more detailed analysis in order to locate the impairment causes which lie behind the observed problem areas.

2 Methods of measuring the quality of service

2.1 The following methods of measuring the quality of service are described:

1) service observations by external means;
2) test call (simulated traffic);
3) customer interviews;
4) internal automatic observations.

1) Such tables can be found in the handbook cited in [1].

2.2 Administrations are recommended to draw up a programme for observations and tests designed for assessment of circuits and equipment, supervision of operators and evaluation of the quality of service given to subscribers. It would be desirable if telephone Administrations were to exchange statistics on quality of service.

2.3 Table 1/E.422 relates to the manual and semi-automatic observations of the quality of international automatic and/or semi-automatic service. It provides in particular a check of the percentage of unsuccessful calls due to technical faults (equipment shortages or failures).

Table 2/E.422 relates the same information as Table 1/E.422 but does not include information which can only be obtained by operators listening in (automatic observation).

Table 1/E.423 relates to observations on traffic set up by operators. It provides, in manual and semi-automatic service, a means of determining the efficiency of international circuits, of assessing the work of operators and the quality of transmission.

Table 2/E.423 summarizes observations of the time-to-answer by operators. The table is compiled by automatic means.

Table 1/E.424 is used to record the results of test calls undertaken especially when the observations shown in Table 1/E.422 make it clear that the percentage of faults is too high.

The use of customer interviews as a method of measuring telephone service quality is the subject of Recommendation E.125 which is particularly concerned with the determination of sources of user difficulty which may arise when making an international automatic telephone call.

Recommendation E.426 contains a general guide to the expected percentage of effective international call attempts.

Table 1/E.427 may be used to supplement the information contained in Table 1/E.422 when the observations shown in that table make it clear that the percentage of faults due to customer difficulties is too high or the outcome of the application of Recommendation E.125 demonstrates the need for additional information.

Recommendation E.425 describes the data that might be taken from the switching centres with respect to quality of service, and the exchange of that data.

2.4 Paying attention to the quality of service of the incoming traffic stream is of major importance, since the incoming Administration is in a better position to improve the situation.

In the past less attention has been paid by several Administrations to the quality of service (QOS) on incoming calls than on outgoing calls. This situation should not persist in the future.

Therefore, in addition to the measurement of QOS of the outgoing traffic stream which is described in this series of Recommendations, Administrations are strongly advised to observe the incoming traffic stream with the aim to improve the QOS.

3 Other sources of information on the quality of service

The following sources are useful to consider when trying to improve the quality of service:

- subscriber complaints (see Annex B);
- other Administrations or organizations such as INTELSAT (SPADE reports);
- operators contacting maintenance staff for direct action;
- operators giving information on QOS: if operator traffic is significant one might consider organizing the flow of this type of information by establishing "trouble codes", e.g. echo, no tone, no answer, etc.;
- reports from "national" switching centres: the QOS as experienced by the subscriber does not only depend on the international network and the network of the country of destination but also on the national network of the country of origin;
- user organizations/large companies: as large companies have much to gain from an improved QOS they might be willing to cooperate with Administrations;
- holding time versus conversation time measurements;
- average conversation time;
- traffic measurements;
- transmission measurements.

ANNEX A

(to Recommendation E.420)

A possible approach to integrate activities measuring the quality of service into an overall problem-investigating process

The flowcharts of the resource allocation process and a typical problem identification procedure are given in Figures A-1/E.420 and A-2/E.420. The numbers 1) through 10) in the figures correspond to the processes described below.

1) The exception threshold is set to detect possible isolated destinations. It is up to the individual Administration to set the value.

2) A destination could be regarded as being under isolated condition when the bid frequency is significant enough to show that there is some demand to the destination (e.g. 20 attempts per day) without or nearly without answer.

3) The most practical way to find out whether improvements may be possible is "consultation with other Administrations".

4) Apply, if possible, network management actions, e.g. alternative routing.

5) The destination priority, P, for each destination is calculated as follows:

$$P = BID^{2)} \times (TABR - MABR)$$

where

$BID^{2)}$ is the number of total bids to the destination during a certain period of time (for example, 1 month);

$TABR$ is the target $ABR^{3)}$ (answer bid ratio) performance which is expected as the result of the service improvement activities;

$MABR$ is the measured $ABR^{3)}$ to the destination during the same period with BID.

The $TABR$ is set for each destination and can be based on the average historic ABR and should be higher than that value.

In order to comply with Recommendation E.426, § 2.2, the $TABR$ to be used in the formula for P given above should not be lower than the $MABR$ experienced one period earlier.

6) In order to comply with § 2.4, it is suggested to consider also the *total* international incoming traffic stream as one of the elements which require QOS improvements. It should be noted that the procedure can be well applied to domestic destinations, for example, on an area code basis, and can be applied on an incoming route basis.

7) Perform detailed analysis: when possible, monitor circuit group performance and do analyses on a destination code basis. It is essential to be aware of "killer trunks" (though observation of the QOS is not directly intended to discover killer trunks).

8) Discuss possible improvements with counterpart.

9) In Recommendation M.710 (General maintenance organization for the international automatic and semi-automatic service) the basic maintenance elements, their functions and the cooperation between the elements are described. Recommendation M.1230 (Assessment of the performance of the international telephone network) gives guidance on the relationship between service quality observations, network performance assessment and maintenance procedures. It should be noted that the QOS very much depends on the proper operation of maintenance elements and maintenance procedures. Therefore, Administrations faced with QOS problems are strongly advised to be attentive to the maintenance Recommendations contained in Volume IV.

10) If this procedure does not lead to a successful conclusion, then an escalation procedure may be required (see Recommendation M.711).

[2)] Paid minutes or the revenues can be used.

[3)] In case ABR cannot be used, ASR (answer seizure ratio) is considered to be an acceptable substitute. Seizures, $TASR$ and $MASR$ are then applicable.

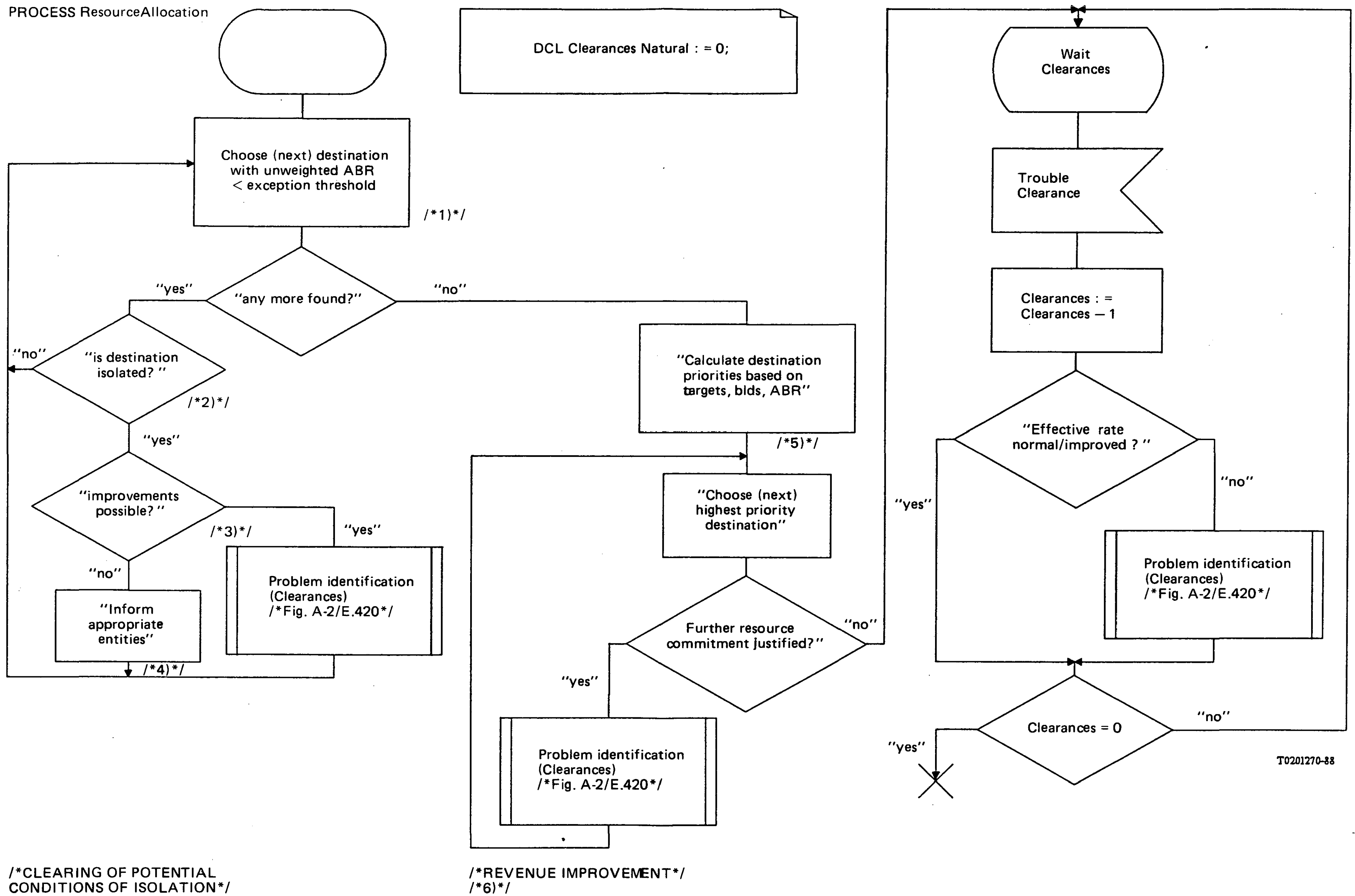


FIGURE A-1/E.420

Process resource allocation

Problemidentification

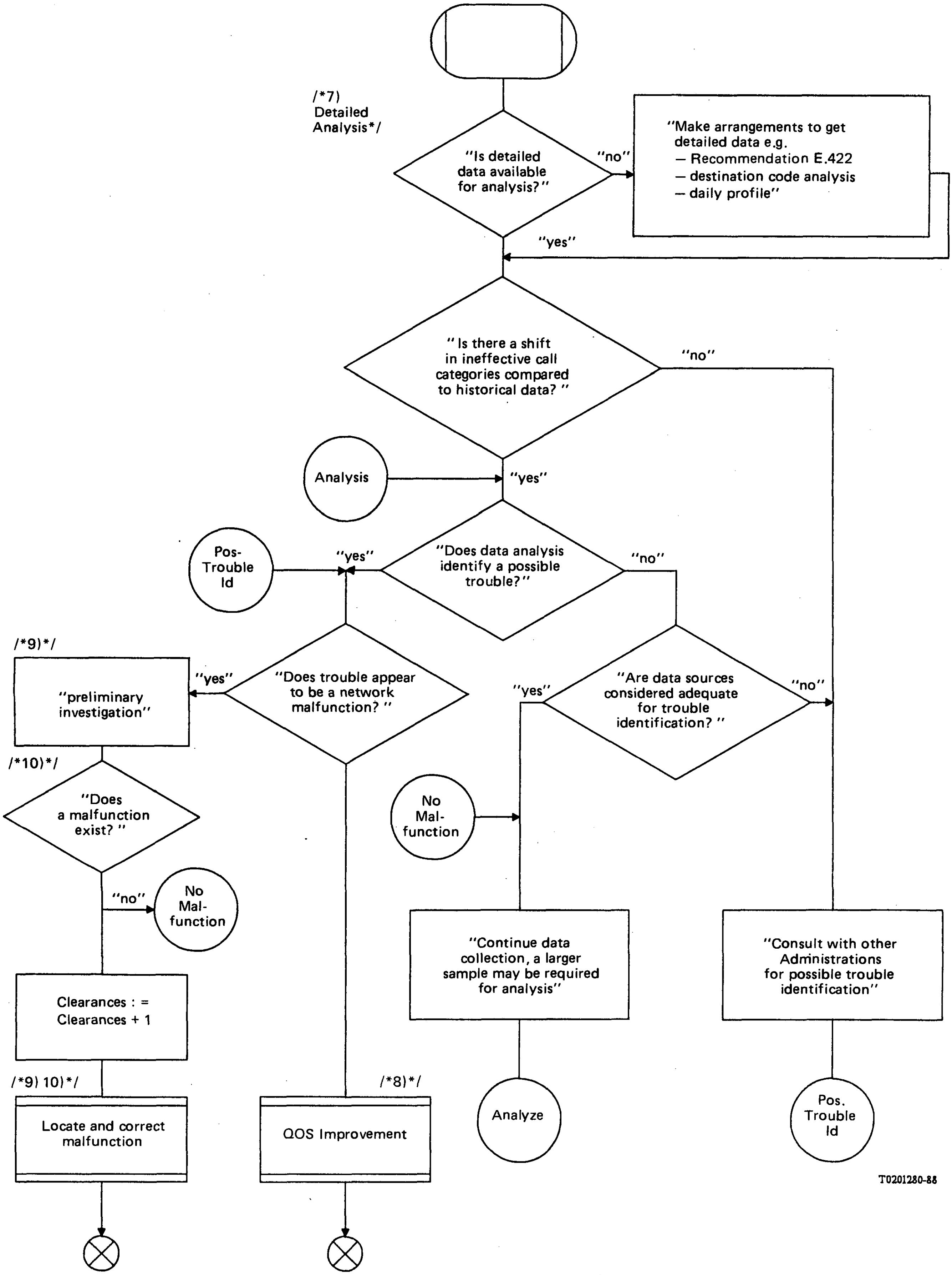


FIGURE A-2/E.420

Problem identification procedure

ANNEX B

(to Recommendation E.420)

Utilization of customer complaints to improve the quality of service for international traffic

It is possible to use subscriber complaints to control processes if the organization of an Administration provides possibilities for centralized collection of these complaints.

The collected data can be processed statistically to provide useful indications for the operations and maintenance staff to correct problems and, in so doing, improve the quality of service.

Three aspects are relevant in the processing of the data:

- the data itself;
- the statistical processes;
- the analysis of complaints.

B.1 *Data to be collected*

The trouble report of an individual customer may be subjective and unqualified since it is usually made by a person, who is not well trained for observation of quality of service. Therefore it will be necessary to make sure that the information about the complaint is as reliable as possible and useful for identifying the possible impairment of the network that caused the complaint.

Examples (see also the handbook cited in [1]):

- data concerning the subscriber numbers involved (route, destination);
- data concerning the observations during the unsuccessful call attempt(s), or the disturbed call;
- time of the observation by the customer.

B.2 *Statistical processes to improve the reliability of the data*

Reliable data is obtained by statistical processing of large numbers of complaints (e.g. an average value during a certain period of time). In order to achieve this, the following methods are considered to be useful:

1) choose complaints whose possible causes seem to relate to *impairments* of the network;
2) *accumulate* complaints for a certain period of time, for example, one month or one week, depending on the number of complaints;
3) calculate the *ratio of complaints statistically* from accumulated data, for example, *complaint-to-completion* ratio (CTCR), for the chosen period of time:

$$\text{CTCR} = \frac{\text{Number of complaints}}{\text{Number of effective call attempts}} \times 100\%$$

It is practical to use the CTCR in combination with one or more classification aspects (see § B.3) such as "per destination".

B.3 *Analysis of complaints*

It is necessary to identify the possible impairment of the network causing the complaint and smoothly clear this impairment in order to actually improve the quality of service. To accomplish this, the complaint needs to be processed into data useful to network maintenance organizations in localizing the possible impairment. The following methods are considered useful:

1) classifying complaints by category of failure;
2) classifying complaints by destination, route (or circuit group) and/or area code;

3) time of day analysis. This may be effective in identifying impairments that may not be apparent when looked at on a total day basis;

4) highlighting relative changes or trends in the statistical data. These changes are likely to reflect a change of the network status and are useful indications along with the values themselves. For example, a rapid increase in the statistical value (e.g. the ratio of complaints) may reflect a new impairment of the network.

Reference

[1] CCITT Manual *Quality of service, network management and network maintenance,* ITU, Geneva, 1984.

Recommendation E.421

SERVICE QUALITY OBSERVATIONS ON A STATISTICAL BASIS

1 Definitions

1.1 service observation

F: observation de la qualité de service

S: observación de la calidad del servicio

Monitoring to obtain a complete or partial assessment of the quality of telephone calls, excluding test calls.

1.2 manual observation

F: observation manuelle

S: observación manual

Monitoring of telephone calls by an observer without using any automatic data-recording machine.

1.3 automatic observation

F: observation automatique

S: observación automática

Monitoring of telephone calls without an observer.

1.4 semi-automatic observation

F: observation semi-automatique

S: observación semiautomática

Monitoring of telephone calls using equipment which records some data automatically. For example, equipment in which information, such as exchange being observed, number dialled by the subscriber, metering pulses and time of call, is recorded automatically on some means suitable for data processing. The observer merely has to key in a code indicating the condition observed.

2 Relative merits of manual, automatic and semi-automatic observations

2.1 The three methods mentioned above in §§ 1.2, 1.3 and 1.4 are not exclusive; for example, automatic observations may be used to supplement observations taken by an operator. It was considered in 1968 that the need for automatic observations would increase in view of the heavy cost associated with manual or semiautomatic observations on the rapidly expanding international network. It was also considered that automatic observations would not entirely supersede observations taken by an observer within the foreseeable future.

The relative merits of the three methods can be assessed as follows:

2.2 *Manual observation*

Provides all the data required in Tables 1/E.422 and 1/E.423.

Observations can be carried out with a minimum of equipment.

Observations can permit the detection of a number of abnormalities which cannot be detected automatically, e.g. very poor speech transmission (item 5.2 of Table 1/E.422), or difficulty with audible tones encountered in the international service (item 6.4 of Table 1/E.422).

2.3 *Semi-automatic observation*

Provides all the data required in Tables 1/E.422 and 1/E.423.

There is a saving in staffing costs compared with manual observation.

Greater accuracy compared with manual observation is possible due to the fact that there is an automatic recording of the number dialled, the time of the call, etc.

It is possible for the observer to devote greater attention to the more critical conditions being checked during observations of calls.

The results are produced in a form suitable for subsequent mechanized analysis.

Owing to the reduction of costs it is possible to obtain a larger sample for the same expenditure.

Semi-automatic equipment may be converted, during certain hours of the day, to automatic operation.

2.4 *Automatic observation*

Operating cost is minimum (staff reduction).

Continuous observation is possible.

It is possible to have a larger sample.

Human error is eliminated.

Automatic processing of data is facilitated.

Conversational privacy is ensured.

Control of the time at which observations are made is facilitated.

Some of the differences between internal and external automatic observations are given below:

2.4.1 Internal automatic observations can be made in the switching centre itself, on the incoming side or the outgoing side or in between, according to the way the switching centre is engineered:

a) Only line signals, such as seizure, answer, etc. can be monitored, and also register signals as long as they do not pass through the exchange in an end-to-end signalling procedure.
b) Signals received are only monitored if the exchange itself operates correctly in that respect.
c) Item b) applies also to outgoing signals. If there is a fault in the exchange it can happen that signals have not been sent in the appropriate way without the exchange being aware of it.

More information on this type of observation technique is given in Recommendation E.425.

2.4.2 External automatic observations are made by means of monitoring equipment which is supervising the traffic on incoming or outgoing lines:

- All signalling signals can be monitored.
- The detection of tones, speech and data is possible if advanced equipment is used.
- This observation technique provides all the data required in Table 2/E.422 and Table 2/E.423.
- The application is very flexible and can be used instead of manual or semi-automatic observation techniques.

3 Time of observations

The results of all observations taken over the whole day should be recorded in Table 1/E.422 or Table 2/E.422.

In the case where observations are not taken over the whole day the observation period is recorded under the heading "Time of observations" and should include the three busiest hours of the day.

4 Observation access points

4.1 Observations for Table 1/E.422 or Table 2/E.422 should be carried out from access points as close as possible to the outgoing international exchange.

The following access points can be considered:

i) outgoing relay set of an international circuit ("exchange" side), i.e. *international circuit access point*[1]);

ii) incoming relay set of a national circuit;

iii) link circuits of the international exchange.

Observations will be made only while the call is being set up, and a few seconds after the called subscriber's reply.

When the circuit access point[1]) is used for observation of international calls it is possible that the service quality of the international exchange may not be checked by either international or national observation programmes.

Preferably, and where technically feasible for the most complete results, observations for Table 1/E.422 should be carried out as close as possible to the international exchange on the national side. This would be more representative of service to the subscriber, and allows observation of call failure at the outgoing international exchange. Where it is not possible to make the distinction between failures in the outgoing international exchange, and failures beyond this exchange, or where there is a meaningful advantage in doing so, observations should be taken on the outgoing side.

It is necessary to state in Table 1/E.422 or Table 2/E.422 the access point where the observations have been made, as observations obtained at each one of the three access points mentioned above are not comparable.

4.2 Observations for Table 1/E.423 must be carried out from access points on the operators' positions.

5 Number of observations

5.1 Service observing programmes should be established in such a manner that statistical results obtained be as reliable as practicable bearing in mind the cost of obtaining large samples.

5.2 According to the studies carried out by the CCITT in 1964-1968, the quantities shown below are considered the *minimum* quantities to provide a general indication of the quality of service.

5.2.1 *Table 1/E.422*

The minimum number of observations per outgoing circuit group for Table 1/E.422 should be 200 per month when more than 20 circuits are included in a group, 200 per quarter when there are between 10 and 20 circuits in a group and 200 per year if there are less than 10 circuits in a group.

5.2.2 *Table 1/E.423*

The minimum number of observations for Table 1/E.423 should be 200 per quarter when there are more than 20 circuits in the group, 200 per semester when there are between 10 and 20 circuits and 200 per year when there are less than 10 circuits in the group.

---

[1]) For definitions of test access points see Recommendation M.700. See also Recommendation M.110.

5.2.3 *Transit traffic*

Where an outgoing circuit group also carries transit traffic it is desirable to obtain data for each destination country reached via this circuit group. In principle, the number of observations for each destination should be obtained as indicated above. To accomplish this, one should use for each destination country its corresponding number of erlangs and derive from these erlangs a theoretical number of circuits.

However, where only a very small amount of traffic is handled, e.g. less than 5 erlangs, each Administration may wish either to make a smaller number of observations or (e.g. in case of no complaints) no observations at all and rely on the information obtained at the transit exchange.

5.3 The number of observations specified above will provide a general indication of results on quality of service in certain broad categories. Administrations may desire more accurate results especially for the individual categories in Table 1/E.422.

Attention is drawn to Table 1/E.421 which gives the number of observations required to obtain a certain degree of accuracy.

TABLE 1/E.421

Expected percentage rate of failure	Number of observations of a random sample required to predict with 95% confidence the true percentage of failure with an accuracy of:					
	± 25%	± 30%	± 35%	± 40%	± 45%	± 50%
2	3136	2178	1600	1225	1030	880
4	1536	1067	784	600	500	440
6	1003	696	512	392	330	290
8	736	511	376	288	245	215
10	576	400	294	225	195	170
12	469	326	239	183	150	132
14	393	273	201	154	128	112
16	336	233	171	131	112	98
18	292	202	149	114	95	80
20	256	178	131	100	85	70
30	149	104	76	60	50	42
40	96	67	50	38	30	24
50	64	44	33	25	20	16

Annex to Table 1/E.421

Examples of use of Table 1/E.421

Example 1 – It is estimated from previous results that a particular type of failure occurs on about 4% of calls. If it is required to confirm, with 95% confidence, that the existing failure rate is between 3% and 5% (i.e. ± 25% of 4%), then observations must be made on a random sample of 1536 calls.

Example 2 – For an expected failure rate of 2%, observations must be made on a random sample of about 1200 calls (1225 in the table) to predict, with 95% confidence, that the true percentage is between 1.2% and 2.8% (i.e. ± 40% of 2%). This means that when 200 observations are taken over a period it is necessary to take the "rolling average" of conditions over six periods. The rate of failure for a number of categories important from the maintenance point of view is expected to be about 2%.

Example 3 – After observations have been taken and the rate of failure in the sample has been calculated, the table may be used in a "backward" direction to give a rough indication of the accuracy of the result.

Suppose that out of a sample of 1000 observations, there were 29 failures due to cause "X" and 15 failures due to cause "Y". The rates of failure in the sample due to X and Y, respectively, are then 2.9% and 1.5%. From the table, it is apparent from this sample of 1000 calls that the true rate of failure due to X has an accuracy of about ± 35% (i.e. is between 1.9% and 3.9%), and that due to Y has an accuracy of about ± 50% (i.e. is between 0.8% and 2.3%).

6 Exchange and analysis of the results of observations

6.1 *Exchange of the results of observations*

The following periodicities are proposed for the exchange of results between Administrations:

Table 1/E.422 or Table 2/E.422 – a monthly exchange is desirable;

Table 1/E.423 or Table 2/E.423 – a quarterly exchange is desirable.

Nevertheless, in the case of small groups of circuits (less than 20 circuits) the information should be exchanged after 200 observations have been made but never later than one year in any case; attention is drawn to the fact that less than 200 observations are of little value.

Results of observations will be reported without delay:

- to the Administrations and the network analysis point of the country where observations are carried out;
- to the Administrations and the network analysis point of the other country (including transit Administrations and their network analysis point when involved).

The benefits to be derived from service observations tend to decrease if there is any increase in the time taken to make the results available to those who can take action to bring about an improvement. The results of service observations according to Tables 1/E.422 and 1/E.423 should therefore be made available to the Administration in the countries of destination as soon as possible after completion of the observation period and in any case within six weeks.

6.2 *Analysis of observation results*

An analysis of the results should be carried out in the country of origin as well as in the country of destination.

Some Administrations have found it useful to distribute to other Administrations concerned, service observation statistics in the form of graphs.

Recommendation E.422

OBSERVATIONS ON INTERNATIONAL OUTGOING TELEPHONE CALLS FOR QUALITY OF SERVICE

1 Objectives concerning Table 1/E.422 and Table 2/E.422

1.1 The purpose of service observation in the international service is to assess the quality of service obtained by the calling subscriber. Consequently, it is essential to have factual or objective recording of observations (i.e., successful and unsuccessful calls), and to present them in the form of a table (see Table 1/E.422 for manual or semi-automatic observations and Table 2/E.422 for automatic observations).

2 Manual or semi-automatic observations (Table 1/E.422)

2.1 Table 1/E.422 should be capable of being completed through the use of a wide range of observation facilities, i.e. from the simple to the sophisticated.

TABLE 1/E.422

Observations of international outgoing telephone calls for quality of service

Country of origin . . . .

Outgoing international exchange . . . .

Group of circuits . . . .

Service { automatic [a)] / semi-automatic [a)]

Period: from . . . . to . . . .

Point of access:

National side . . . .

Link circuits . . . .

Outgoing side . . . .

Time of observations . . . .

Category	Number		Percentage	
	Subtotal	Total	Subtotal	Total
1. Calls successfully put through (see note 1)		...		...
2. Ring tone received but no answer		...		...
3. Unsuccessful calls: *Positive* indication of congestion, including subscriber busy, from beyond the outgoing international exchange. Visual signal, tone or recorded announcement		...		...
3.1 Subscriber busy/congestion indicated by visual signal	...		...	
3.2 Subscriber busy/congestion indicated by busy/congestion tone .	...		...	
3.3 Congestion indicated by a recorded announcement	...		...	
4. Unsuccessful calls: Other visual signals, tones or recorded announcements, not positively identified as category 3 or 8		...		...
4.1 Visual signal received	...		...	
4.2 Tone received	...		...	
4.3 Recorded announcement received	...		...	
5. Unsuccessful calls for other technical reasons		...		...
5.1 Wrong number obtained	...		...	
5.2 Abandoned due to very poor speech transmission	...		...	
5.3 No tone, no answer after waiting ... seconds	...		...	
5.4 Reception of answer signal when the called party does not reply	...		...	
5.5 Other failures of a technical kind	...		...	
6. Unsuccessful calls due to incorrect handling by the calling party . .		...		...
6.1 Wrong number dialled	...		...	
6.2 Incomplete number	...		...	
6.3 Call prematurely abandoned before receipt of signal, tone or announcement (within less than ... seconds)	...		...	
6.4 Call prematurely abandoned after receipt of ring tone (within less than 30 seconds)	...		...	
6.5 Other failures due to incorrect handling	...		...	
7. *Total calls monitored* (categories 1-6)		...		100

TABLE 1/E.422 *(cont.)*

Category	Number		Percentage	
	Subtotal	Total	Subtotal	Total
8. Unsuccessful calls: *Positive* indication of failure from outgoing international exchange		...		
8.1 Congestion on outgoing international circuits	...			
8.2 All other indications	...			
9. Successful calls with defects. These calls are included in category 1		...		
9.1 Non-reception of answer signal on chargeable calls	...			
9.2 Call with impaired intelligibility but not abandoned	...			
9.3 Other calls with defects but not abandoned	...			

[a] Delete whatever is inapplicable.

Note 1 – A successful call is one that reaches the wanted number and allows conversation to proceed. All successful calls are entered in category 1. However, a successful call may or may not have noticeable defects. Successful calls with noticeable defects should also be entered in category 9.

Note 2 – With the exception noted above for categories 1 and 9, the results of one call observation should be entered under one category only, namely the most appropriate one from 1 to 6.

Note 3 – Administrations should periodically exchange necessary information to interpret the observation data recorded under categories 4.1, 4.2 and 4.3.

2.2 Specialized training of observers should be kept to a minimum.

2.3 The table should be self-explanatory so that reference to detailed how-to-complete instructions is unnecessary.

2.4 The major categories should be selected such that:

- they identify the major factors adversely affecting the quality of service;
- they are suitable for the centralized processing of observation results.

2.5 To permit the orderly collection of data for human factors studies to identify sources of difficulty in customer use of the international (automatic) telephone service, Recommendation E.427 contains an additional table to Table 1/E.422.

3 Comments concerning the use of Table 1/E.422

3.1 Table 1/E.422 summarizes observations made on outgoing automatic and/or semi-automatic traffic, on a country of origin to a country of destination basis. A separate form should be used for each country of destination, and if required, for each group of circuits to which traffic to a country of destination has access at the outgoing international exchange (or exchanges). It is not necessary to make observations on both automatic and semi-automatic services. An Administration may select the service to be observed, provided that the service is the majority of the traffic to the country of destination.

3.2 For an explanation of the point of access, see Recommendation E.421, § 4.1.

3.3 The result of each call observed should be entered only under the most appropriate category. In the case of several faults on one call, the most significant cause of failure should be entered.

3.4 In completing Table 1/E.422 reference should be made to the following explanations.

4 How to fill in Table 1/E.422

Category 1 – To ensure objective recording and to avoid producing a biased sample resulting from the exclusion of calls which require subjective assessment, the successful call is defined as a call that reaches the wanted number and allows conversation to proceed. All non-abandoned calls are entered into category 1 and of these calls those which are subjectively adjudged to be defective are also entered into category 9. Thus it is required of the observer to make *two* entries for successful calls with noticeable defects.

Enter in category 1 then, calls successfully put through. This includes answered calls for which a clearback signal is received after some words have been spoken, without knowing for what reason the call is abandoned. If it is observed that the caller has dialled a wrong number, the call will be entered under 6.1. Category 1 will also include calls put through correctly to operator positions, information services, or to machines replying in place of the subscriber or to their equivalents.

Category 2 – Enter in this category calls on which ring tone was heard but the subscriber did not answer before the attempt was abandoned, the caller having waited at least 30 seconds after commencement of ring tone before clearing forward. (See category 6.4 if the call was abandoned *less* than 30 seconds after ring tone commenced.)

Category 3 – Enter in this category all unsuccessful calls in which a *positive* indication of subscriber busy or congestion beyond the outgoing international exchange had been encountered, either by visual signal, tone or recorded announcement. Congestion encountered on common control equipment should be entered in this category as well (e.g. no "proceed-to-send" signal). Where a positive indication of these conditions has *not* been received, enter in category 4.

Categories 3.1, 3.2 and 3.3 are entered for the specific indication received.

When more than one indication is received, e.g. visual signal and audible tone, only one entry should be made. In this case, the preferred order of entry should be tone, announcement, visual signal.

Category 4 – Enter in this category all other indications on unsuccessful calls whether by visual signal, tone or recorded announcement that cannot be positively identified and entered in category 3 or 8.

Categories 4.1, 4.2 and 4.3 are entered for the specific indication received.

When more than one indication is received, e.g. visual signal and audible tone, only one entry should be made. In this case, the preferred order of entry should be tone, announcement, visual signal.

Category 5 – Enter in this category those calls which fail for technical reasons not included in categories 3, 4 and 8. Category 5 subdivides as follows:

Category 5.1 – Calls on which the wrong number was obtained, although the caller dialled correctly.

Category 5.2 – Calls abandoned by the caller because of very poor speech transmission, although the answer signal was received. (See category 9.2 if speech transmission is poor but the call is not abandoned.) In some countries observers may be required to cease listening immediately after conversation is established, thus reducing the number of calls that would be reported in this category.

Category 5.3 – Calls on which the dialling information was correctly and completely sent, but the caller received no signal, tone or announcement before abandoning the call, having waited for at least the specified period before clearing forward.

The value of this time period left open under this category should be filled in by the Administrations of the originating country according to its experience in this matter. The prescribed value may differ depending on the international destination. It is, however, recommended to limit the number of such different quoted periods to a maximum of three values (e.g. 10, 20 or 30 seconds or any other value considered pertinent by the Administrations concerned).

Category 5.4 – Calls on which an answer signal was received, although the called subscriber did not answer.

Category 5.5 – Call failures due to technical reasons which are unable to be entered in categories 5.1 to 5.4. These should be very few, if any, and this category is provided in case they do arise. All possible information about these failures should be supplied as an attachment to the summary of the table. This category includes calls abandoned due to reception of a clear-back signal while connecting with the extension number (PBX).

Category 6 – Enter in this category all unsuccessful calls which have failed due to incorrect handling by the caller (subscriber or operator). Category 6 subdivides as follows:

Category 6.1 – Calls on which it was determined that the number which should have been dialled was different from the number actually dialled.

Category 6.2 – Calls on which it was determined that the number dialled had insufficient digits to be successful.

Category 6.3 – Calls on which the digital information was correctly and completely sent, but the caller abandoned the call without receiving any signal, tone or announcement, and without waiting for at least the specified period.

The value of the time period left open under this category should be filled in by the Administrations of the originating country according to its experience in this matter. The prescribed value may differ depending on the international destination. It is, however, recommended to limit the number of such different quoted periods to a maximum of three values (e.g. 10, 20 or 30 seconds or any other value considered pertinent by the Administration concerned).

The value quoted under category 6 must be the same as that quoted under category 5.

Category 6.4 – Calls prematurely abandoned after receipt of the ringing tone on which the caller disconnected less than 30 seconds after the ringing tone commenced. (See category 2 if the call was abandoned after *more* than 30 seconds had elapsed from the time of commencement of ringing tone.)

Category 6.5 – Calls which failed due to incorrect handling by the caller which cannot be classifed under categories 6.1 to 6.4. All possible information about these failures should be supplied as an attachment to the summary of the table. As in categories 5.5, these should be very few, if any.

Category 7 – Enter in category 7 the number of calls monitored (categories 1-6).

Category 8 – Category 8 will be useful for those Administrations which observe on the national side of the outgoing international exchange. (See Recommendation E.421, § 4.1.) Positive indications of failure, congestion or other, are to be entered here. They are not to be included with categories 1-6, which give the data for calls monitored for category 7.

Thus, when category 8 is viewed with categories 3 and 4 a more complete picture is provided of quality of service received by the caller.

Category 9 – Entries in category 9 are for successful calls (entered in category 1) which encountered defects, but which were not abandoned. They are thus automatically included in the total of category 7.

Category 9.1 – Enter here chargeable calls for which no answer signal was received. If abandonment should be detected on such calls, enter in category 5.5.

Category 9.2 – Enter here calls on which poor speech transmission was observed, but the call was not abandoned. (See category 5.2 if the call was abandoned.) All possible information about these calls should be supplied as an attachment to the summary of the table. Note that in some countries observers may be required to cease listening immediately after conversation was established, thus reducing the number of calls that would be reported under this category.

Category 9.3 – Enter here calls encountering switching, signalling or transmission defects, but which were not abandoned and which cannot be classified under categories 9.1 or 9.2.

5 **Automatic observations** (Table 2/E.422)

Considering the limitation of abilities of automatic observation equipment (for example, automatic observation equipment cannot understand announcements) and the variety of signals used in signalling systems, the table recommended for CCITT Signalling System No. 5 is given below.

TABLE 2/E.422

Automatic observations of international outgoing telephone calls for quality of service

Country of origin . . . .

Outgoing international exchange . . . .

Group of circuits . . . .

Service { automatic [a)] / semi-automatic [a)]

Period: from . . . . to . . . .

Point of access

National side . . . .

Link circuits . . . .

Outgoing side . . . .

Time of observations . . . .

Category	Number		Percentage	
	Subtotal	Total	Subtotal	Total
1. Calls successfully put through		. . .		. . .
2. Ring tone received but no answer		. . .		. . .
3. Unsuccessful calls: *Positive* indication of congestion, including subscriber busy, from beyond the outgoing international exchange. Visual signal or tone		. . .		. . .
3.1 Subscriber busy/congestion indicated by visual signal	. . .		. . .	
3.2 Subscriber busy/congestion indicated by busy/congestion tone	. . .		. . .	
4. Unsuccessful calls: Other tones or recorded announcements, not positively identified as category 3 or 8		. . .		. . .
4.1 Tone received	. . .		. . .	
4.2 Recorded announcement received	. . .		. . .	
5. Unsuccessful calls for other technical reasons		. . .		. . .
5.1 No tone, no answer signals after waiting ..... seconds	. . .		. . .	
5.2 Reception of answer signal when the called party does not reply	. . .		. . .	
5.3 Other failures of a technical kind	. . .		. . .	
6. Unsuccessful calls due to incorrect handling by the calling party		. . .		. . .
6.1 Call prematurely abandoned before receipt of signal, tone or announcement (within less than ..... seconds)	. . .		. . .	
6.2 Call prematurely abandoned after receipt of ring tone (within less than 30 seconds)	. . .		. . .	
6.3 Other failures due to incorrect handling	. . .		. . .	
7. *Total calls monitored* (categories 1-6)		. . .		100
8. Unsuccessful calls: *Positive* indication of failure from outgoing international exchange		. . .	X	X
8.1 Congestion on outgoing international circuits	. . .		X	X
8.2 All other indications	. . .		X	X
9. Successful calls with defects. These calls all included in category 1		. . .	X	X
9.1 Non-reception of answer signal on chargeable calls	. . .		X	X
9.2 Other calls with defects	. . .		X	X

a) Delete whatever is inapplicable.

6 Comments concerning the use of Table 2/E.422

6.1 Table 2/E.422 summarizes observations made on outgoing automatic and semi-automatic traffic, on a country of origin to a country of destination basis. A separate form should be used for each country of destination, and if required, for each group of circuits to which traffic to the country of destination has access at the outgoing international exchange (or exchanges).

6.2 For an explanation of the point of access, see Recommendation E.421, § 4.1.

6.3 The result of each call observed should be entered only under the most appropriate category. In the case of several faults on one call, the most significant cause of failure should be entered.

6.4 As the function of sound analysis by automatic observation equipment is not concerned with the signalling system used and since some signalling systems e.g. Signalling System No. 6 have more information exchanged in the signalling system than those of sound signals, it is expected that the proposed table will be applied to all signalling systems for the present.

6.5 In completing Table 2/E.422 reference should be made to the following explanations.

7 How to fill in Table 2/E.422

Category 1 – The successful call is defined as a call that allows conversation to begin between subscribers, or allows to begin sending facsimile or data. This includes calls put through to operator positions, information services, or to machines replying in place of the subscriber or to their equivalents. In other words, the successful call is such that the automatic observation equipment detected voice on both sending and receiving lines, or that it detected sending tone of facsimiles or data, or that it detected voice on the receiving line after receipt of answer signal.

Category 2 – This category includes those calls for which the automatic observation equipment detected ringing tone, but there was no answer signal and the clear-forward signal was sent 30 seconds after the detection of ringing tone.

Category 3 – Enter in category 3 all unsuccessful calls for which a positive indication of subscriber busy or congestion beyond the outgoing international exchange has been encountered, either by visual signal (busy-flash signal) or by tone (also includes no "proceed-to-send" signal).

Category 4 – Enter in category 4 unsuccessful calls for which the automatic observation equipment detected a tone, but could not classify it, or the equipment detected announcement (that is, it detected voice on receiving line without answer signal).

Category 5 – Enter in category 5 those calls which failed for technical reasons not included in categories 3, 4 and 8. Category 5 subdivides as follows:

Category 5.1 -- Calls on which the dialling information was completely sent, but the automatic observation equipment received no signal, tone or announcement and it received a clear-forward signal after a specified period. The value of this time period left open under this category should be filled in by the Administrations of the originating country according to its experience in this matter. The prescribed value may differ depending on the international destination. It is, however, recommended to limit the number of such different quoted periods to a maximum of three values (e.g. 10, 20 or 30 seconds or any other value considered pertinent by the Administrations concerned).

Category 5.2 – Calls on which an answer signal was received, although the called subscriber did not answer. In other words, calls for which the automatic observation equipment received an answer signal, although it detected no voice on receiving line.

Category 5.3 – Failed calls due to technical reasons which are unable to be entered in categories 5.1 and 5.2. For example, a call for which there was a busy-flash signal after receiving ringing tone.

Category 6 – Enter in category 6 all unsuccessful calls which have failed due to incorrect handling by the caller (subscriber or operator). Category 6 subdivides as follows:

Category 6.1 – Calls on which the dialling information was completely sent, but the automatic observation equipment received no signal, tone or announcement and it received a clear-forward signal within a specified period. (For this period, see category 5.1 above.)

Category 6.2 – Calls prematurely abandoned after receipt of the ringing tone on which a clear-forward signal was received less than 30 seconds after the ringing tone was detected.

Category 6.3 – Calls which failed due to incorrect handling by the caller which cannot be classified under categories 6.1 and 6.2. For example, a call for which the automatic observation equipment received an answer signal after receiving ringing tone, and then the ringing tone stopped, but the equipment could not detect any voice either on the sending line or the receiving line.

Category 7 – Enter in category 7 the number of calls monitored (categories 1-6).

Category 8 – Category 8 will be useful for those Administrations which observe on the national side of the outgoing international exchange. Positive indications of failure, congestion or other, are to be entered here.

Category 9 – Entries in category 9 are for successful calls (entered in category 1) which encountered defects. Category 9 subdivides as follows:

Category 9.1 – Calls on which no answer signal was received, but the conversation was begun.

Category 9.2 – Calls which encountered switching or signalling defects, but on which the conversation was begun.

Recommendation E.423

OBSERVATIONS ON TRAFFIC SET UP BY OPERATORS

1 Comments concerning the use of Table 1/E.423

1.1 This table summarizes observations relating to manual and semi-automatic outgoing traffic originated by operators. These observations will be made, if possible, during the whole call duration. Observations for the categories 1 to 7 may be omitted in case of semi-automatic service, if there is no problem regarding the efficiency of international circuits.

1.2 Administrations should, if possible, make a distinction between the different types of call, e.g. station-to-station, personal and collect calls; they should use a separate column for each under the heading "Type of call".

1.3 For collect calls, the times to be recorded will be those observed in the country where the call request was made.

1.4 It is recommended that these observations be spread over the whole day.

1.5 Each outgoing Administration will select the international circuit groups on which observations should be carried out.

1.6 In completing this table, reference should be made to the following explanations:

2 How to fill in Table 1/E.423 (Traffic observations determined by the operators)

Category 1 – This category should show the mean duration of calls observed which are successful and have been charged for ("effective" calls).

Category 2 – This category will show the mean *chargeable* duration of all effective calls observed.

Category 3 – This category will show, for each type of observed call, the average time per effective call during which the international circuit has been occupied for manoeuvres or for call preparation.

This average should be based on the time during which the international circuit is held:

a) to obtain information concerning the called number;

b) to obtain information about routing and trunk codes;

c) to call operators, in the incoming international exchange;

d) to exchange information on how to set up the call;

e) to (or attempt to) obtain the called number even when it is engaged or does not reply;

f) to (or attempt to) obtain the called person (in personal calls);

g) between replacement of the receiver by the called person and release of the circuit;

h) because the operator is holding the circuit (whether she is on the line or not) and for any other reasons for which the circuit is engaged.

TABLE 1/E.423

Observations on traffic set up by operators

International outgoing exchange: ....

Circuit group: ....

Service { semi-automatic [a]
manual [a]

Period from .... to ....

Category	Type of call [b]			
	Ordinary	Personal		
1. Mean call duration – in seconds				
2. Mean chargeable duration – in seconds				
3. Mean holding time of circuits for manœuvres and preparation of calls – in seconds				
4. Number of effective calls observed				
5. Mean number of times the international circuit was seized per effective call				
6. Mean number of "attempts" per effective call				
7. Percentage of calls set up at the first "attempt"				

8. Time-to-answer by operators	Total number of calls answered and unanswered		Calls answered						Calls unanswered (abandoned calls)			
	Number	Mean waiting time in seconds	under 15 seconds		in 15 to 30 seconds		after 30 seconds		within 30 seconds		after 30 seconds	
Operators:			No.	%	No.	%	No.	%	No.	%	No.	%
— incoming operator (code 11)												
— delay operator (code 12)												
— assistance operator												
— information operator												

9. Quality of transmission from the subscriber's viewpoint:	Number	%	10. Comments
— good			
— defective			
Total		100	

[a] Delete whichever is inapplicable.
[b] In accordance with § 1.2.

The times listed above, which exclude the conversation time, should be added together. This total should be divided by the number of effective calls observed during the period in question to obtain the value to be entered in Table 1/E.423.

Category 4 – The number of effective calls observed considered in category 1.

Category 5 – The mean number of times the international circuit was seized per effective call (see category 3). This number is usually obtained by meter recordings.

Category 6 – The mean number of *attempts* (as specifically defined hereafter from the operating point of view) to set up a call. Should the operator try several times to set up a call while continuously occupied on that call, all these operations must be considered as being one attempt. Similarly, if the operator makes several tries to set up a call and each time encounters a congestion or busy condition and if, after the last try, she informs the caller, only one attempt must be entered. Calls to information services or to obtain routing particulars, and all calls not directly related to the establishment of a call or to information required by the caller, should not be considered as attempts and should not be included.

The total number of attempts during the period of observation should be divided by the number of effective calls observed in the same period to obtain the mean number of attempts per call.

The total number of attempts is usually determined from markings or notations on call tickets.

Category 7 – The data for this category will be taken from all tickets prepared for the relation concerned, during the period of observation or a comparable period.

Category 8 – The mean waiting time for outgoing operators to receive an answer will be indicated in seconds. This average will include both answered and unanswered calls.

An outgoing operator waits on the circuit (waiting time) for the period:

a) until the incoming operator answers, or

b) until she abandons the attempt, should the incoming operator not answer.

Thus while mean waiting time relates to the outgoing operator it is also a measure of the performance of the incoming operators.

Category 9 – It will be difficult to obtain absolutely comparable results from all observers for this category. However, the observer should consider the quality of transmission from the subscribers' viewpoint, taking into account comments made in this respect by subscribers and the number of requests for conversation to be repeated.

Category 10 – This category should include any comments likely to explain the probable cause of difficulties frequently noted during the observations.

3 **Automatic observations of the time-to-answerby operators** (Comments concerning the use of Table 2/E.423)

3.1 This table summarizes observation of the time-to-answer by operators.

3.2 Administrations should make a distinction between the different types of incoming operators if the types of operators are distinguished by the selecting digits.

3.3 It is recommended that these observations be spread over the whole day.

3.4 Each outgoing Administration will select the international circuit groups on which observations should be carried out.

3.5 The time-to-answer of the assistance operator cannot be measured automatically.

3.6 In completing this table, reference should be made to the explanations in § 4.

4 **How to fill in Table 2/E.423** (Automatic observations of the time-to-answer by operators)

The mean waiting time for outgoing operators to receive an answer will be indicated in seconds. This average will include both answered and unanswered calls.

The mean waiting time is defined as the time interval between the instant the outgoing circuit is seized (the seizing signal is sent) and:

a) the instant the incoming operator answers, or

b) the instant the outgoing operator abandons the attempt (a clear-forward signal is sent).

TABLE 2/E.423

Automatic observations of the time-to-answer by operators

International outgoing exchange . . . .

Circuit group . . . .

Service: semi-automatic

Period from . . . . to . . . .

Time-to-answer by operators	Total number of calls answered and unanswered		Calls answered						Calls unanswered (abandoned calls)			
			under 15 seconds		in 15 to 30 seconds		after 30 seconds		within 30 seconds		after 30 seconds	
Operators:	Number	Mean waiting time in seconds	Number	%	Number	%	Number	%	Number	%	Number	%
– incoming operator												
– delay operator												
– information operator												

Recommendation E.424

TEST CALLS

1 General

Test calls carried out manually or automatically to assess the functioning of international circuits of connections are of four types:

a) *Type 1 test call*

A test call conducted between two directly connected international centres to verify that the transmission and signalling on an international circuit of a given group are satisfactory.

b) *Type 2 test call*

A test call conducted between two international centres not directly connected to verify transit operational facilities of an intermediate international centre.

c) *Type 3 test call*

A test call from an international centre to a subscriber type number in the national network of the distant country, generally as a result of a particular kind of fault.

d) *Subscriber-to-subscriber type test call*[1)]

A subscriber-to-subscriber type test call is a test call from a test equipment having the characteristics of an average subscriber line in one national network to a similar equipment in the national network of a distant country.

Test calls types 1, 2, 3 and subscriber-to-subscriber test calls must not interfere with customer traffic. If, however, test calls contributing a significant load on a part of a network are to be made, prior advice should be given to the other Administration(s) concerned. Types 1 and 2 test calls for preventive maintenance should be conducted during light load periods. Types 1 and 2 test calls should be conducted as and when required for the investigation and clearance of faults.

Type 3 test calls should be conducted only after adequate testing has been done by means of type 1 or 2 test calls and after the distant Administration has made the necessary check in its national network. Type 3 test calls should be conducted during light load periods.

In order to find faults in last-choice equipment, circuit multiplication equipment or in-circuit multiplexing equipment, it may be necessary for tests to be carried out at the time when the traffic load approaches the full capacity of the route under test. The agreement of the distant network analysis point will be necessary before this test is carried out.

Subscriber-to-subscriber type test calls can be made by agreement of the network analysis point in the countries concerned.

Normally, unless there is a specific agreement between the Administrations concerned, subscriber-to-subscriber type test calls would be considered for fault location after:

1) verifying that there are no evident faults in the international switching centres involved that would cause the poor quality of service or subscriber complaint being investigated;
2) verifying that type 1 or type 2 test calls have been made on the international circuits that might have been involved;
3) verifying that there are no evident faults in the national network from the outgoing exchange to the international centre in the originating country;
4) verifying that there are no evident faults in the national network in the distant country, from the international centre to the called exchange.

When test calls are undertaken from the international centre to a subscriber number to verify that there are no evident faults in the national network, such calls should be routed through the international centre on the same path as a normal incoming international call. Using the test access facilities in the international centre could route calls via a different path thereby masking a fault.

When subscriber-to-subscriber type test calls are made, the network analysis point in the two countries should consider such factors as:

i) the expected nature of the fault;
ii) international accounting agreements;
iii) the need for making the test calls in the busy hour;
iv) the possibility of causing or aggravating congestion at the time the calls are made.

The responding equipments used for subscriber-to-subscriber type test calls could be those used for maintenance of the national network.

[1)] Recommendation M.1235 describes the use of automatic-to-subscriber test calls in more detail.

2 Results of test calls (see Table 1/E.424)

TABLE 1/E.424

Results of test calls

International outgoing exchange: . . . .

Circuit group: . . . .

Service { semi-automatic [a)] automatic [a)]

Period from . . . . to . . . .

Type of test call
Type 1 [a)]
Type 2 [a)]
Type 3 [a)]
Sub-to-Sub [a)]

Category	Number		Percentage	
	Subtotal	Total	Subtotal	Total
1. Satisfactory tests		. . .		. . .
2. Signalling and charging faults		. . .		. . .
2.1 Wrong number	. . .		. . .	
2.2 No tone, no answer	. . .		. . .	
2.3 Absence of a backward line signal	. . .		. . .	
2.4 Other faults	. . .		. . .	
3. Transmission faults		. . .		. . .
3.1 Conversation impossible	. . .		. . .	
3.2 Call overamplified or underamplified	. . .		. . .	
3.3 Noise	. . .		. . .	
3.4 Fading	. . .		. . .	
3.5 Crosstalk	. . .		. . .	
4. Congestion		. . .		. . .
5. Other faults	. . . . . .	. . .	. . . . . .	. . .
Tests carried out		. . .		100
Tests procedure followed (apparatus used, destination of calls, etc.)				

[a)] Detele whichever is inapplicable.

INTERNAL AUTOMATIC OBSERVATIONS[1]

1 Definitions

1.1 essential information (of internal automatic observations)

The answer seizure ratio (ASR) (see § 1.3) or answer bid ratio (ABR) (see § 1.4), whichever is appropriate in terms of attempts, completed attempts and percentage completed.

1.2 supplementary information (of internal automatic observations)

Information on signalling faults, subscriber behaviour and the network.

1.3 answer seizure ratio (ASR)

ASR gives the relationship between the number of seizures that result in an answer signal and the total number of seizures. This is a direct measure of the effectiveness of the service being offered and is usually expressed as a percentage as follows:

$$\text{ASR} = \frac{\text{Seizures resulting in answer signal}}{\text{Total seizures}} \times 100$$

Measurement of ASR may be made on a route or on a destination code basis.

1.4 answer bid ratio (ABR)

Gives the relationship between the number of bids that result in an answer signal and the total number of bids.

$$\text{ABR} = \frac{\text{Bids resulting in answer signal}}{\text{Total bids}} \times 100$$

ABR is expressed as a percentage and is a direct measure of the effectiveness of traffic from the point of measurement. It is similar to ASR except that it includes bids that do not result in a seizure.

2 Merits of internal automatic observations

The advantage of internal monitoring is that a large volume of records can be collected. The large volume of data obtained from an internal observation system allows day-to-day evaluation of network performance. Daily analysis of this information has proven invaluable in trouble detection, and, coupled with a good maintenance response, is instrumental in providing the best possible quality of service.[2] The disadvantage is that this method does not have the capability of detecting tones or speech and therefore cannot present a complete representation of all call dispositions.

To overcome this disadvantage Administrations are advised to use Recommendation E.422 as well to supplement the data obtained from internal automatic observations.

---

[1] This Recommendation also applies in case external monitoring equipment is used when a route is monitored constantly for all or a large (statistical significance) number of calls. Refer to Recommendation E.421, § 2.4.

[2] Using these techniques one can improve the quality of service even when no distinction can be made between ring no answer, subscriber busy (or congestion indicated by congestion tone) and recorded announcement.

3 Time of observations

The results of all observations taken over the whole day should be recorded.

4 Exchange of the results of observations

4.1 The essential information[3] should be exchanged monthly (preferably by facsimile or telex) to all network analyses points of those Administrations who are interested (the analyses points can then make comparisons between different streams going to the same destination). If information on ASR or ABR can be supplied separately for direct routes and indirect routes via transit countries, this should also be exchanged as being essential information, including the name of the transit country involved.

4.2 With respect to supplementary data such as: signalling faults, failures due to calling subscriber, failures due to called subscriber and failures due to the network, a quarterly exchange of information is appropriate. Because different formats will be required, mail seems the most likely means to be used for exchanging supplementary data.

4.3 Besides the monthly and quarterly exchange of information, a direct contact on all aspects should be made (by telephone) as soon as action is required to prevent a persistent drop in the quality of service.

5 Classes of calls

The distinction between classes of calls (such as operator-operator, subscriber-subscriber and operator-subscriber) is considered useful in identifying problems relating to the quality of service. This can only be done if the language digit[4] and some of the subsequent digits are analyzed.

6 Destination analysis from service observation data

Consideration should be given to include the dialled digits, as observed by the monitoring equipment, in the exchange of information, especially for the sake of destination analyses (see Recommendation E.420, Annex A).

7 Details about supplementary information for CCITT Signalling System No. 5

7.1 *Signalling faults*

- faulty signals;
- time outs, the main item in this category being no proceed-to-send signal;
- busy flash. (Since busy flash is applied in many situations, including failures due to calling and called subscriber and the network, it is considered useful to distinguish between busy flash received within 0-15 seconds, 15-30 seconds and after 30 seconds when making destination analysis.)

7.2 *Ineffective calls associated with the calling subscriber*

Premature release, to distinguish between release before or after having received ringing tone; equipment which can detect audible signals is required.

7.3 *Ineffective calls associated with the called subscriber*

Ringing tone no answer cannot be detected without equipment which can detect audible signals.

---

[3] The Administration supplying the data must indicate whether the ASR or ABR is used.

[4] The language or discrimination digit is inserted automatically, or by the operator, between the country code (Recommendation E.161) and the national (significant) number.

7.4 *Network*

Here only the busy flash can be detected without equipment which can detect audible signals.

8 Equipment impact

8.1 Administrations are recommended to consider inclusion of appropriate facilities in existing and new exchanges to record all or some of the following phases:

a) Calls switched to speech position, then:
   1) answered;
   2) unanswered, but released by calling party;
   3) timed out awaiting answer;
   4) a call failure signal (busy flash or equivalent) received;
   5) timed out after clearback signal;
   6) faulty signal received after answer.

b) Calls failing to switch to speech position:
   1) clear forward signal received;
   2) insufficient digits received;
   3) congestion on international circuits;
   4) faulty signals received into exchange;
   5) signalling fault into next exchange;
   6) time out while signalling to next exchange;
   7) congestion signal received from next exchange;
   8) vacant number received;
   9) busy subscriber signal received;
   10) line out of order signal received;
   11) transferred subscriber signal received.

As a minimum requirement one should be capable of determining the answer seizure ratio (ASR) or the answer bid ratio (ABR). This recording can be done by off-line processing of call records if they contain some more information than the information already required for international accounting.

8.2 Another way to assemble date on the quality of service (QOS) on outgoing circuit groups is through event counters. Five event counters already give a reasonable amount of information, three of them being common to the Signalling Systems No. 5, No. 6 and R2: seizure, answer and busy signals[5].

Signalling System No. 5

The number of:

- seizing signals sent;
- end-of-pulsing (ST) signals sent;
- proceed-to-send signals received;
- busy flash signals received;
- answer signals received.

Signalling System No. 6

The number of:

- initial address messages (IAM) sent;
- congestion (switching-equipment; circuit groups; national network) signals, call-failure signals and confusion signals received;
- address-complete (subscriber-free, charge; subscriber-free, no charge; subscriber-free, coinbox; charge; no charge; coinbox) signals received;
- subscriber busy signals received;
- answer (charge; no charge) signals received.

[5] In case the event counting is used to analyze the quality of service to a particular destination, the counting should be done separately for each signalling system.

The number of:

- seizing signals sent;
- congestion [national network (A4 or B4); international exchange (A15)] signals received;
- address complete (charge; subscriber's line free, charge; subscriber's line free, no charge) signals received;
- subscriber line busy signals received;
- answer signals received.

Recommendation E.426

GENERAL GUIDE TO THE PERCENTAGE OF EFFECTIVE ATTEMPTS WHICH SHOULD BE OBSERVED FOR INTERNATIONAL TELEPHONE CALLS

1 General considerations

1.1 The success of call attempts is fundamental to an automatic international telephone service of high quality.

1.2 The periodic observation of completion ratio [1] and the categorization of failures to destination countries together with the exchange of such information between countries are valuable to establish and/or maintain a high service quality.

1.3 The call completion ratio of the national network of a given country, as manifested through its international switching centre(s), affects the efficiency of operation of all countries routing traffic to that country.

1.4 Call completion ratio information can be provided either internally in an SPC international switching centre or externally at the level of the outgoing international circuits in any international switching centre in which access to the circuits is provided for the purpose of establishing the disposition of call attempts.

1.5 The availability, flexibility and capacity of minicomputers provides an economically attractive method of obtaining call completion ratio information with extreme accuracy. This includes the observation of tones when suitable interfaces with the minicomputer are provided.

2 A guide to the proportion of effective call attempts

2.1 A general guide for the expected percentage of effective call attempts during the mean busy hour and its two immediately adjacent hours, as observed at the originating international switching exchange, is indicated below. An effective call attempt is defined, for this purpose, as one for which an answer is received at the originating international exchange. Faults caused by the originating international exchange shall be excluded to the extent feasible. All attempts which succeed in seizing an international circuit shall be included in the results:

a) low level of effective call attempts: less than 30%;

b) medium level of effective call attempts: 30% to 60%;

c) high level of effective call attempts: more than 60%.

2.2 When an originating country notes a downward change in the level of effective call attempts towards any destination, the originating, destination or transit Administrations should initiate investigations to determine and alleviate the underlying causes (e.g. network provisioning, subscriber behaviour). The objective of this action is to avoid degradation in the level of effective call attempts.

[1] See Recommendation E.600.

COLLECTION AND STATISTICAL ANALYSIS OF SPECIAL QUALITY OF SERVICE OBSERVATION DATA FOR MEASUREMENTS OF CUSTOMER DIFFICULTIES IN THE INTERNATIONAL AUTOMATIC SERVICE

This Recommendation is provided to permit the orderly collection of data required for special studies to identify sources of difficulty in customer use of the international automatic telephone service.

When calls are made to points outside a customer's home country, many different sets of ringing and busy tones are encountered. In order to measure the effect of unusual sounding ringing tones and busy tones on customer behaviour, it has been decided to collect data on how long customers listen to such foreign tones as well as to their national tones in order to compare them.

The data are to be collected in the same manner as those required for the completion of Table 1/E.422. These data are an extension of those collected for Table 1/E.422, and, as an aid to subsequent analysis, a copy of the current version of that table should be used with the table of this Recommendation.

Table 1/E.427 contains questions numbered 1-9. Their relationship to the questions of Table 1/E.422 is shown in parentheses.

A preferred set of analyses for identifying the statistical significance of differences between data collected from subscribers when setting up national calls and the corresponding data collected from subscribers when setting up international calls is given below.

1 Determine the percentage change in any measure by use of the formula:

$$\text{Change } (C_i) = \left[\frac{f_{ij}}{N_j} - \frac{f_{iH}}{N_H}\right] \times 100 \qquad j = \text{A, B, C} \quad i = 0\text{-}2, 2\text{-}5 \ldots, > 30$$

where

f_{ij} is the observed frequency of calls of category i in the country j,

N_j is the total number of observations in the country j sample,

f_{iH} is the observed frequency of calls of category i in the home country H, and

N_H is the total number of observations in the home country sample.

2 Compare the central location of the distributions by use of the Kruskal-Wallis One-Way Analysis of Variance [1].

3 Compare the "forms" or "shapes" of the distribution by means of the chi-square test [2].

4 Compare changes in single valued variables, e.g. percentage incomplete-trunk-code, by use the chi-square test.

TABLE 1/E.427

(Supplement to Table 1/E.422)

Observations of international outgoing telephone calls for quality of service
Additional details regarding subscriber dialled calls

Outgoing international exchange: ....

Group of circuits: ....

Period from .... to ....

Category	Number		Percentage	
	Subtotal	Total	Subtotal	Total
Details of dialled calls [a)b)c)]				
1. Calls with errors in the dialled number [d)]				
1.1(6.1) Wrong number dialled		...		100
1.1.1 Wrong country code	...		...	
1.1.2 National trunk prefix (e.g. "0") wrongly included	...		...	
1.1.3 Wrong trunk code	...		...	
1.1.4 Wrong subscriber number	...		...	
1.2(6.2) Incomplete number dialled		...		100
1.2.1 National (significant) number not dialled or incomplete	...		...	
1.2.2 Trunk code not dialled or incomplete	...		...	
1.2.3 Subscriber number not dialled or incomplete	...		...	
2. (5.3) (6.3) Calls abandoned prematurely before receipt of a tone or announcement		...		100
Interval from end of dialling to disconnect [e)]:				
0- 5 s	...		...	
5-10 s	...		...	
10-20 s	...		...	
20-30 s	...		...	
30-50 s	...		...	
> 50 s	...		...	
3. Post dialling delay on all calls that are maintained beyond the start of a tone or announcement		...		100
Interval from end of dialling to tone or announcement:				
0- 5 s	...		...	
5-10 s	...		...	
10-20 s	...		...	
20-30 s	...		...	
30-60 s	...		...	
60-90 s	...		...	
> 90 s	...		...	
Average excluded portion [f)] ...				
4. Calls that encounter ringing tones [g)]				
4.1(1) Completed calls		...		100
Interval from beginning of tone to answer:				
0-10 s	...		...	
10-20 s	...		...	
20-30 s	...		...	
30-50 s	...		...	
> 50 s	...		...	
4.2(2.6.4) Incomplete calls		...		100
Interval from beginning of tone to disconnect:				
0-10 s	...		...	
10-20 s	...		...	
20-30 s	...		...	
30-50 s	...		...	
> 50 s	...		...	

Category	Number		Percentage	
	Subtotal	Total	Subtotal	Total
5. (3.2) Calls that encounter busy/congestion tones [g]		...		100
Interval from beginning of tone to disconnect:				
0- 2 s	...		...	
2- 5 s	...		...	
5-20 s	...		...	
20-30 s	...		...	
> 30 s	...		...	
6. (4.2) Calls that encounter tones that the observer cannot identify		...		100
Interval from beginning of tone to disconnect:				
0- 2 s	...		...	
2- 5 s	...		...	
5-10 s	...		...	
10-30 s	...		...	
> 30 s	...		...	
7. (3.3, 4.3) Calls encountering recorded announcements		...		100
Interval from beginning of announcement to disconnect:				
0-10 s	...		...	
10-20 s	...		...	
20-30 s	...		...	
> 30 s	...		...	
8. List types of errors in dialling and tone interpretation which could not be categorized				
....				
....				
9. List restrictions on subscriber sample [h]				
....				
....				
....				

[a] The term "calls" throughout this table refers to circuit seizures by outgoing traffic.

[b] The data for each called country should be collected separately and not combined with other countries.

[c] The interpretation of these results cannot be made adequately except by comparing them with similar results on national calls.

[d] The practicability of putting the observation in category 1 will depend upon the observation access point and knowledge of the national numbering plan of the outgoing country and of the destination country.

[e] 0- 5 s implies $0 \leqslant t \leqslant 5$.
5-10 s implies $5 < t \leqslant 10$.

[f] The "post-dialling delay" measurements may not represent the actual delay from the time the subscriber finishes dialling to the receipt of tone. To the extent that this measurement as observed on the trunk excludes the time from completion of dialling to seizure of trunk, the average duration of this excluded time should be reported.

[g] Identification of tone categories should be made by service observers who are trained to identify the tone categories reliably.

[h] If access to the trunks being observed is restricted to some specified population of subscribers, e.g., heavy users, non-coin users, or residents of large urban centres, such restrictions should be noted and reported with the service observations.

References

[1] MARASCUILO (L. A.), McSWEENEY (M.): Non-Parametric and Distribution-Free Methods for the Social Sciences, *Wadsworth Publishing Co.*, California, 1977.

[2] SIEGEL (S.): Non-Parametric Statistics for the Behavioural Sciences, *McGraw Hill*, New York, 1956.

CONNECTION RETENTION

1 Introduction

Connection retention is one of the parameters influencing the quality of service (QOS) after the call has been set up. As the assessment of this parameter is difficult and costly, it is recommended to investigate only after information from sources such as operator trouble reports, subscriber complaints, interviews and/or service observations indicate that there is a problem.

2 Cut-off call ratio

The cut-off call ratio is the percentage of the established calls that are released for a reason other than intentional by any of the parties involved in the call. The cut-off call ratio can only be measured by placing test calls (see Recommendation G.181).

3 Investigations required

Before measuring the cut-off call ratio the source information leading to the investigation should be utilized to the extent possible (see Annex B to Recommendation E.420) and the outcome should be compared with other relevant sources. For example, subscriber complaints can be followed up by investigating operator trouble reports.

These investigations should lead to a suspect part of the network.

4 Exchange of information

In case the suspect part of the network is outside the territory of an Administration, the Administration which has responsibility for the suspect part of the network should be contacted and informed of the results of the investigations. In the information given, the type of cut-offs should be classified by causes such as "absence of answer signal", "artificial clearback signal", etc.

5 Further investigation required

Within the suspect part of the network, the Administration should look for obvious causes of cut-offs such as exchange or facility failures. If obvious causes cannot be identified the Administration should consider test calls.

6 Test calls

One or both Administrations may decide to set up a test call programme.

International standardized test call types are contained in Recommendation E.424. As stated in that Recommendation, before applying type 3 test calls or subscriber-to-subscriber type test calls, it should be verified that there are no evident faults in the national network. This verification can easily be undertaken by applying non-standardized test calls, for example from the international centre to a subscriber number in the national network of the same country (see Recommendation E.424, 4) of § 1).

The usefulness of standardization in these kinds of test calls is under study in several Administrations.

For making test calls on connection retention, the responding equipment used should send an answer signal after 10 seconds followed by a continuous tone in conformity with Recommendation O.61. The threshold level of the detector in the directing equipment should also be in conformity with Recommendation O.61.

7 Objectives

The long-term objectives as well as the allocations are given in Recommendation G.181. In the intermediate time the following requirements should apply: the cut-off call ratio for subscriber-to-subscriber tests should, measured over 24-hour periods, be below 0.5% for 5 minute calls. In any time-consistent hour the call cut-off rate may not exceed 3%.

It should be noted that, in general, cut-off problems have a more severe impact on customers' perception of service on routes with a low answer seizure ratio than on routes with a low completion rate.

PART II

Recommendations E.500 to E.600

TRAFFIC ENGINEERING

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

SECTION 1

MEASUREMENT AND RECORDING OF TRAFFIC

Recommendation E.500

TRAFFIC INTENSITY MEASUREMENT PRINCIPLES

1 Introduction

1.1 Traffic measurements provide the data base from which the planning, operation, management and, in some cases, accounting for transit considerations of the telephone network are carried out. Different applications may exist for the same traffic measurement.

1.2 This Recommendation gives the principles for measuring carried traffic and bids on circuit groups and exchanges. The number of bids and preferably also carried traffic intensity should also be determined by individual relations (destinations). Data so obtained are applied both for operation and planning. Recommendation E.501 gives methods for estimating offered traffic from carried traffic measurements. Recommendation E.502 describes exchange requirements for traffic measurements both in national and international exchanges. Recommendation E.525 describes the traffic data analysis. Recommendation E.506 gives methods for forecasting future traffic requirements. The remainder of the E.500 Series of Recommendations describes how to utilize this data base in the operation and planning of telephone networks.

The measurements required for network management as described in the E.410 Series are generally similar to those described in this Recommendation. They will usually require a variable and shorter reporting interval.

2 Definitions

A **measurement of the amount of traffic carried** is the average Erlang value during a certain period of time (e.g. 15 min., 1 hour).

A **measurement of the number of bids** is a count of this entity during a certain period of time.

Measurements are taken continuously during the day or with exclusion of known low traffic periods. The set of days at which measurement has been taken is called the *measurement days.*

In the **yearly continuous measurement** the measurement days are post-selected from a base period with a length of the whole year. The post-selected days include the peak intensity values measured during the base period.

In the **yearly non-continuous measurement** the measurement days are scheduled (pre-selected) from a base period of a few months. The pre-selected days include the high load days of expectation or of earlier observations.

A traffic profile is defined to be *stable* when the individual daily traffic profiles differ only little in shape and traffic volume between each other.

A traffic profile is defined to be *unstable* when the individual daily traffic profiles differ in shape or traffic volume between each other.

3 Overview

Circuit group dimensioning is based on a congestion objective, on the traffic intensity values at high load time and on the forecast value of intensity until the next augmentation of circuits. Intensity is measured during a daily busy hour and averaged over a number of days, to avoid exceptional values.

If traffic measurements are taken every day of the year (yearly continuous measurements), the required averages can be calculated directly as described in § 4. If traffic measurements are taken only during a limited number of days in the year (yearly non-continuous measurements), the equivalent traffic loads may be estimated using the procedures given in § 5.

The busy hour concept is an important aspect of teletraffic engineering and may be applied in a number of ways. In the E.500 Series of Recommendations the busy hour traffic used is an average of several days with, in some cases, an allowance for day to day variations (Recommendation E.521).

Within the busy hour, traffic is considered to be stationary and thus the recorded intensity is the mean value during the busy hour.

The recommended standard method of calculating the daily average requires *continuously* measuring all quarter hours for *all* days concerned and selecting the busiest hour in the average profile for all days. This method is called the Time-Consistent Busy Hour (TCBH) and is described in detail in § 6. This method is most valuable in situations of stable traffic profiles. The daily continuous measurements provide the data necessary for confirming profile stability.

Another method of arriving at the representative average busy hour also involves *continuously* measuring all quarter hours, but only the busiest hour of *each day* is retained for averaging. This method is called the Average Daily Peak Hour (ADPH) and is described in detail in § 6 together with the relation of ADPH results to TCBH results.

The advantages of ADPH are that it requires less data storage and manipulation than TCBH and that it gives a more representative value in the situation of unstable traffic profiles.

In some situations Administrations do not measure traffic *continuously* over the day, but only for the hour or few hours expected to be busiest. This method is called the Fixed Daily Measurement Period (FDMP) or Fixed Daily Measurement Hour (FDMH) and is described in detail in § 7 together with the relation of FDMP results to TCBH results.

The advantage of FDMP is that it requires less measurement resources than TCBH or ADPH. The disadvantage is that in individual situations the difference between FDMP and TCBH results may vary widely.

In some network situations significant savings can be made by multihour dimensioning (e.g. cluster engineering, time zone differences). This requires daily continuous measurements.

4 Yearly continuous measurements

Traffic statistics should be measured for the significant period of each day of the whole year. The significant period may in principle be 24 hours of the day.

The measurements for computing normal traffic load should be the 30 highest days in a fixed 12-month period. Normally these will be working days, but in some cases separate weekend or tariff-related period measurements should be examined so that Administrations can agree bilaterally on appropriate measures to maintain a reasonable grade of service (GOS) for weekends and tariff-related periods. Recurring exceptional days (e.g. Christmas, Mother's Day, etc.) should be excluded for network dimensioning purposes although the data should be collected for network management purposes (Recommendation E.410). This method gives traffic information of relatively high accuracy and is suitable for circuits groups operated automatically or semiautomatically.

4.1 *Normal and high load levels*

Teletraffic performance objectives and dimensioning practices generally set objectives for two sets of traffic load conditions.

A normal traffic load can be considered the typical operating condition of a network for which subscribers service expectations should be met.

A high traffic load can be considered a less frequently encountered operating condition of a network for which normal subscriber expectations would not be met but for which a reduced level of performance should be achieved to prevent excessive repeat calling and spread of network congestion.

In order to estimate normal and high load levels, offered traffic intensity values should, where necessary, be estimated from daily carried traffic measurements. Estimation procedures are presented in Recommendation E.501.

Normal and high loads are defined in Table 1/E.500.

TABLE 1/E.500

Circuit groups		
Parameter	Normal load	High load
Carried traffic intensity	Mean of the 30 highest working days during a 12-month period.	Mean of the five highest days in the same period as normal load.
Number of bids	Mean of the same 30 days on which the offered traffic intensities are highest.	Mean of same five days on which the offered traffic intensities are the highest.

Exchanges		
Parameter	Normal load	High load
Carried traffic intensity	Mean of the ten highest days during a 12-month period.	Mean of the five highest days in the same period as normal load.
Number of bids	Mean of the same ten highest days (not necessarily the same as the highest offered traffic days) during a 12-month period.	Mean of the five highest days (not necessarily the same as the highest offered traffic days) in the same period as normal load.

5 Yearly non-continuous measurements

5.1 *Introduction*

This method consists in taking measurements on a limited sample of days in each year. Limited sample measurements will normally be taken on working days, but Administrations may agree bilaterally to measure weekend or reduced tariff periods separately.

Any Administration proposing to use a yearly non-continuous measurement procedure is advised to confer with other end Administrations to ensure that the maximum information is available to assist in the choice of measurement days. For example, if the other end Administration has continuous measurement capability it may be possible to identify busy seasons or consistent low-traffic days.

Table 2/E.500 shows the results of a study carried out on circuit groups within a large metropolitan network [1]. The errors shown are the under-estimates resulting if average busy hour carried traffic intensity is measured over a pre-defined two-week period of the year, rather than the actual busiest two-week period. (The pre-defined period was, in fact, the peak period of the preceding year.)

The error averages 7.6% more or less, depending on circuit group size. Had an Administration wished to estimate the true peak two-week intensity with 90% confidence, starting with the pre-defined two-week measurements, the latter would have had to be increased by amounts ranging from about 14% for large circuit groups, up to about 31% for small ones. (The magnitude of these corrections indicates how inadequate a two-week sample can be as a basis for network planning.)

TABLE 2/E.500

Weighted mean error and the upper limit of the intensity error class for a cumulative proportion of circuit groups, categorized according to traffic intensity

	Total	Low < 10 Erl	Medium 10-100 Erl	High > 100 Erl
Circuit groups	2728	1056	1564	110
Weighted mean error of the intensity value	7.6%	13.7%	7.8%	5.2%
Cumulative proportion of circuit groups				
50%	7.9%	12.9%	6.9%	3.9%
80%	16.9%	22.9%	17.9%	7.9%
90%	23.9%	30.9%	23.9%	13.9%
95%	31.9%	37.9%	34.9%	17.9%
98%	41.9%	47.9%	40.9%	26.9%

5.2 *Estimation method*

An approximate statistical method for estimating normal and high load levels from limited sample measurements is provided below.

5.2.1 *Principle of estimation method*

Measurements are taken on a limited sample of days, and the mean (M) and standard deviation (S) of the daily busy hour traffic loads are calculated. Normal and high load level estimates (L) are given by:

$$L = M + k \cdot S$$

different values of the factor k being used for normal and high load levels.

$$S = \left[\frac{1}{n-1} \sum_{i=1}^{n} (X_i - M)^2\right]^{1/2}$$

where

X_i is the time-consistent busy hour traffic measured on the ith day,

$M = \frac{1}{n} \sum_{i=1}^{n} X_i$ is the sample mean, and

n is the number of measurement days.

If the measurement period is less than 30 days then the estimate will not be very reliable. In this case Administrations should, if possible, carry out special measurement studies to determine typical values of the standard deviation (e.g. as a function of the sample mean).

5.2.2 *Base period for measurements*

It is important to determine the "base period" since the length of this period influences the values assigned to the multiplication factors k.

The base period is the set of valid days in each year from which measurement days are preselected. This period should include all days which are potential candidates for being among the 30 highest days (but excluding recurring exceptional days – see § 4).

The base period may be restricted to a busy season (which need not necessarily comprise a set of consecutive weeks) provided that the traffic is known to be consistently higher during this period than during the remainder of the year.

The base period may be the whole year, but Administrations may also decide to exclude known low-traffic days.

5.2.3 *Selection of measurement days*

Measurement days should be distributed reasonably evenly throughout the base period. If the base period extends over the whole year then the measurement sample should include some days from the busiest part of the year, if these are known. The limited sample should comprise at least 30 days to ensure reliable estimates. If this is not possible, then a minimum of 10 measurement days may be used. In this case the reliability of the estimate is poor.

5.2.4 *Multiplication factors*

Multiplication factors k for 5-day, 10-day, and 30-day load levels are given by the curves in Figure 1/E.500, as a function of the number of days in the base period. These factors are derived from tables of order statistics from the normal distribution [2].

When the base period extends over the whole year these factors may not always be reliable because of the effects of differing seasonal patterns. Individual Administrations may then prefer to use different values for the factors, if they have obtained more precise information from special measurement studies.

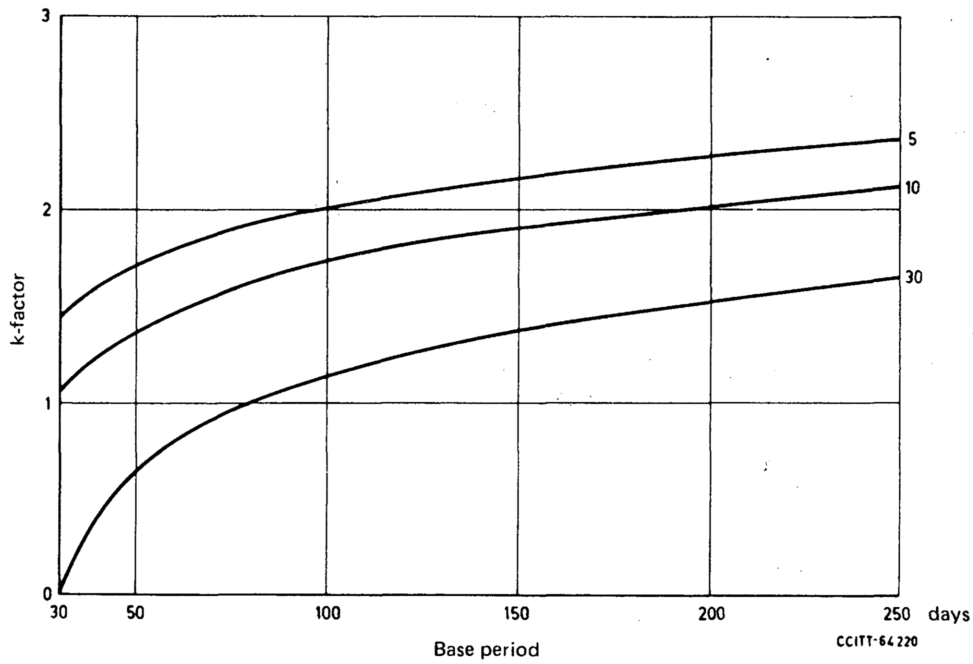


FIGURE 1/E.500

Multiplication factors for estimating mean of 5, 10 or 30 highest days from noncontinuous measurements

5.2.5 *Example*

The following data illustrate the application of this procedure to the estimation of normal and high load levels from non-continuous measurements on a circuit group over a 1-year period.

After excluding holidays and other known low traffic periods the base period which is available for measurement purposes is determined to be 220 days. The k-factors to be used are therefore (from Figure 1/E.500):

Normal (30-day) load level: $k = 1.6$

High (5-day) load level: $k = 2.3$

Measurements are taken on 50 days within the base period. The daily measured busy-hour traffic values, in Erlangs, are as follows:

21.5	20.5	18.7	15.0	18.4	21.6	18.1	24.2	26.7	22.1
21.8	17.8	17.2	19.8	15.2	20.4	16.7	20.6	23.1	23.5
19.6	18.1	21.3	15.9	15.9	17.8	17.4	20.9	25.9	20.6
20.9	19.2	17.6	12.9	14.2	18.1	16.9	24.2	22.2	26.8
22.5	22.8	19.3	19.1	18.7	19.8	18.0	26.0	22.5	27.5

The sample mean and standard deviation are:

$M = 20.11$

$S = 3.37$

The normal and high load level estimates are then calculated from $L = M + k \cdot S$ to give:

Normal load = 25.5 Erlangs

High load = 27.9 Erlangs

5.2.6 *High to normal traffic ratios*

In some circumstances, actual values of high day loads are not available. In such cases, various Administrations use standard ratios of high to normal load for forecasting for design or planning purposes.

For example, as a general order of magnitude, the following ratios of high to normal load may be used as a guide for a healthy network:

Parameter	*Circuit groups*	*Exchanges*
Offered traffic intensity	1.2	1.1
Number of call attempts	1.4	1.2

6 Daily continuous measurements

6.1 *Measurement*

It is recommended that Administrations take traffic measurements continuously over the day throughout the measurement period.

Depending on the application, a busy hour value for dimensioning should be calculated as the peak value of the mean day profile or the average of daily peak values.

6.2 *Time-consistent busy hour (TCBH)-intensity (post-selected)*

For a number of days, carried traffic values for each quarter hour for each day are recorded. The values for the same quarter hour each day are averaged.

The four consecutive quarter-hours in this average day which together give the largest sum of observed values form the TCBH with its TCBH-intensity. This is sometimes referred to as post-selected TCBH.

In the case where a stable traffic profile exists, the TCBH-intensity is used as a base method for dimensioning; if measurement methods yielding systematically lower or higher intensity values than the TCBH-method are used, adjustments to the calculations are needed.

6.3 *Average of the daily peak hours traffic, defined on quarter hour or on full hour basis*

To find the average of daily peak quarterly defined hour (ADPQH) intensity, the traffic intensity is measured continuously over a day in quarter-hour periods. The intensity values are processed daily to find out the four consecutive quarter hours with the highest intensity value sum. Only this daily peak hour traffic intensity value is registered. The average is taken over a number of working days peak intensities. The timing of peak intensity normally varies from day to day.

To find the average of daily peak full hour (ADPFH) intensity, the traffic intensity is measured continuously over a day in full-hour periods. Only the highest of these intensity values is registered. The average is taken over a number of days peak intensities.

The comparative measurements have shown that the traffic intensity values measured by the ADPFH-method, are very consistent with the values measured by the TCBH-method, whereas the ADPQH-method yields slightly (a few percent) higher values. (See Annex A.) ADPH has an advantage over TCBH when traffic profiles are unstable.

6.4 *Alternate routing networks*

When alternate routing is used, the dimensioning methods in Recommendation E.522 should be applied (multi-hour dimensioning technique). In general this requires the continuous measurement of a 24-hour profile for each traffic quantity in the alternative routing cluster.

In Annex A the differences in results between busy hours defined for individual circuit groups and for clusters indicate the advantage of continuous measurements and multi-hour dimensioning for alternative routing networks.

In circumstances where the traffic profiles are stable and similar in the whole cluster, the multi-hour dimensioning may be applied on a few selected hours of significance to the entire cluster. The stability of traffic profiles must be confirmed.

7 Daily non-continuous measurements

7.1 *Measurement*

Some Administrations may find it necessary or economically attractive to restrict measurements to a few hours or only one hour per day. Such measurements will always be less accurate than continuous measurements. The resulting busy hour values will always be less than or equal to TCBH.

The time of fixed daily measurements should be confirmed several times a year by measurement of the full daily traffic profile for every circuit group. The measurement can cover several periods daily, as well.

7.2 *Fixed daily measurement period (FDMP)*

With this method measurements are taken within a fixed period (e.g. of 3 hours) each day. This period should correspond to the highest part of the traffic profile, which is expected to include the TCBH. Measurement values are accumulated separately for each quarter-hour, and the busiest hour is determined at the end of the measurement period, as for the TCBH. This method will normally give results which are about 95% of the TCBH traffic level, when the time of fixed daily measurement is defined for every single circuit group, although major changes in the traffic profile could lead to larger errors.

In alternate routing networks with traffic profiles that are similar and stable in the whole cluster, FDMP may be used to produce measurements for multi-hour dimensioning applied on a few selected hours of significance. The stability of traffic profiles should be confirmed several times a year.

7.3 *Fixed daily measurement hour (FDMH)*

If the fixed daily measurement period is reduced to 1 hour, then it is only necessary to accumulate a single measured value from each day. This is the simplest measurement method, and it will normally give results which are about 90% of the TCBH traffic value, when the time of the fixed daily measurement is defined for single circuit groups individually. However, the variations around the average are large.

8 Flow chart for the application of the different calculation methods

The decisions represented in Figure 2/E.500 compare measurement and analysis costs to variations in the results for a single circuit group or cluster. The costs are particular to each Administration.

The preceding sections of this Recommendation indicate the amount of measurement variance that can occur in typical situations which can result in overprovisioning or a risk of poor grade of service.

In cluster engineering for alternative routing networks, measurements outside the busy hour are normally needed if the traffic profile is unstable. In situations of stable traffic load the significant traffic hours can be predicted accurately, allowing use of a FDMP method.

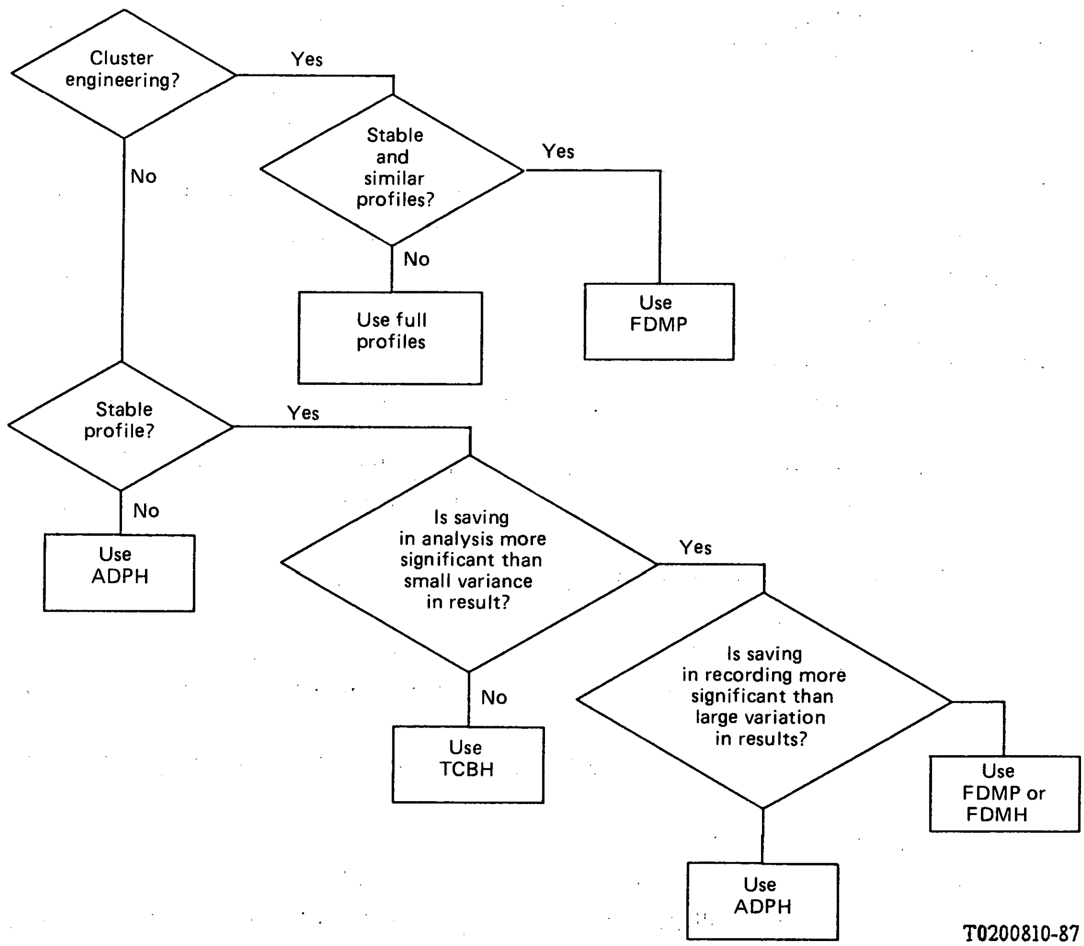


FIGURE 2/E.500

Flow chart for choosing the measurement method

ANNEX A

(to Recommendation E.500)

Example of influence of different busy hour definitions on measured traffic intensity

A.1 *Introduction*

The influence of different busy hour definitions on measured traffic intensity has been investigated by means of measurements on real traffic outgoing from an international exchange.

Three clusters with a total of 15 circuit groups have been studied. One of the clusters (Cluster 1) carries traffic between different time zones.

Traffic per quarter of an hour was measured during the whole day in 5 two-week periods (10 consecutive working days). The total elapsed time covered 9 months.

From the results of the first two-week period of daily continuous measurements the times of FDMH and FDMP have been determined:

- for each circuit group individually (ind),
- per cluster (clu), and
- for all three clusters commonly (com).

The time of FDMH is equal to the time of TCBH in the first two-week period. FDMP includes FDMH and the hour before and the hour after.

A.2 *Results of measurements*

The results of the measurements undertaken are summarized in Figures A-1/E.500 to A-5/E.500.

Figure A-1/E.500 shows how the starting time of TCBH varies between the five measurement periods:

- for each cluster, and
- for individual circuit groups in each cluster.

The following observations on the starting time of TCBH can be made:

- the starting time of TCBH is the same in not more than 2 periods. This refers to both circuit groups and clusters;
- 5 circuit groups and 1 cluster have different TCBH in all periods;
- 8 circuit groups and 2 clusters have TCBH within the same part of the day (morning or evening) in all periods;
- TCBH common to all clusters is in the evening in all periods. Only 2 periods have the same common TCBH.

In Figures A-2/E.500 to A-5/E.500 traffic intensities according to different busy hour definitions have been compared. Traffic intensity according to the TCBH definition has been used as reference value (corresponding to 100% in the figures).

Figure A-2/E.500 shows the results of comparisons on a cluster level, and Figures A-3/E.500 to A-5/E.500 on a circuit group level.

Means and variations of traffic intensities are given as:

- an average of all five periods (ADPQH and ADPFH), and
- an average of measurement periods 2, 3, 4 and 5 compared with period 1 (FDMH and FDMP).

A.3 *Results on cluster level (Figure A-2/E.500)*

ADPQH intensities over 100%, mean = 102%.

ADPFH intensities around 100%, mean = 100%.

$FDMP_{clu}$ intensities from 95 to 100%, mean = 99%.

$FDMH_{clu}$ intensities from 90 to 98%, mean = 94%.

$FDMP_{com}$ intensities from 42 to 100%, mean = 89%.

$FDMH_{com}$ intensities from 35 to 93%, mean = 83%.

A.4 *Results on circuit group level (Figures A-3/E.500 to A-5/E.500)*

ADPQH intensities over 100%, mean = 104%.

ADPFH intensities around 100%, mean = 100%.

$FDMP_{ind}$ intensities from 88 to 100%, mean = 99%.

$FDMH_{ind}$ intensities from 80 to 100%, mean = 93%.

$FDMP_{clu}$ intensities from 51 to 100%, mean = 98%.

$FDMH_{clu}$ intensities from 45 to 99%, mean = 91%.

$FDMP_{com}$ intensities from 24 to 100%, mean = 89%.

$FDMH_{com}$ intensities from 14 to 99%, mean = 81%.

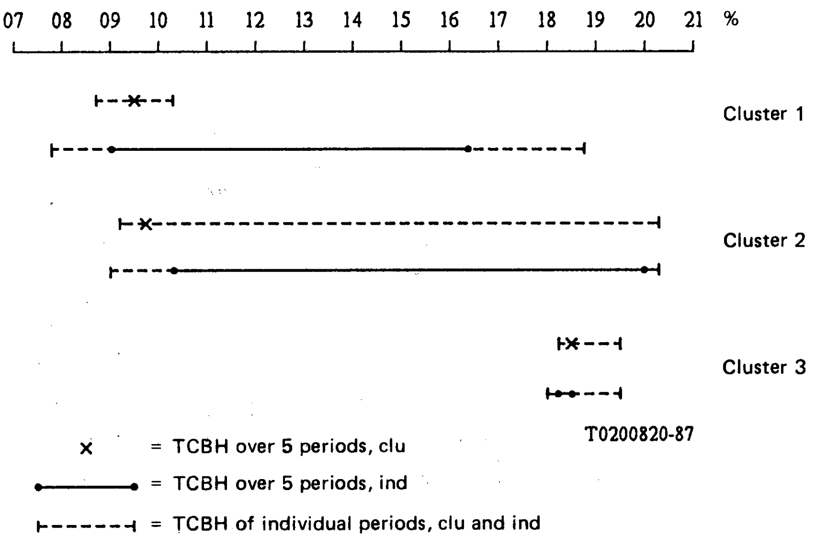


FIGURE A-1/E.500

Variations in time of TCBH

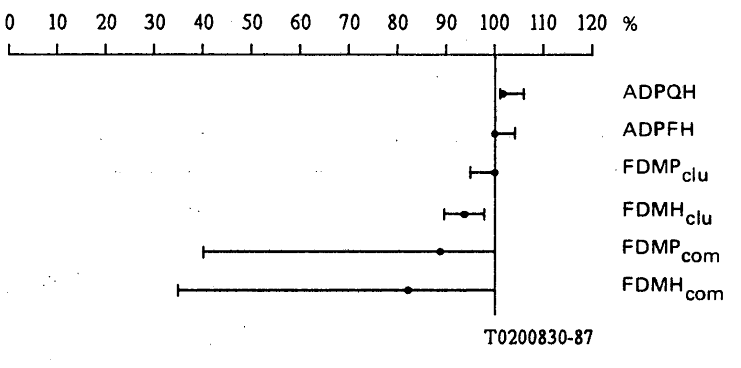


FIGURE A-2/E.500

Comparisons on cluster level

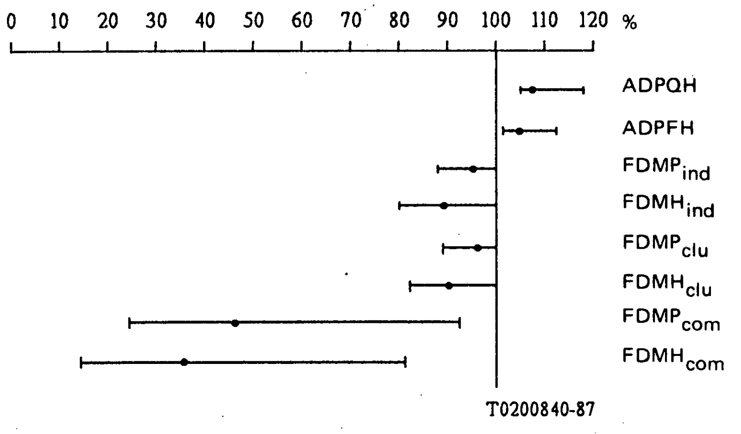


FIGURE A-3/E.500

Comparison on circuit group level (Cluster 1)

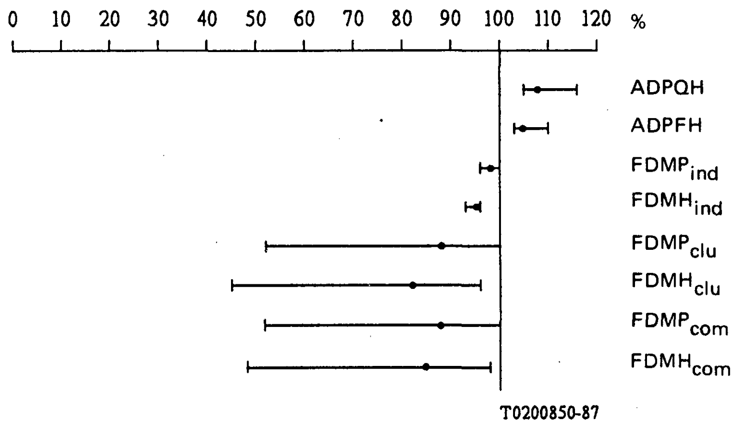


FIGURE A-4/E.500

Comparison on circuit group level (Cluster 2)

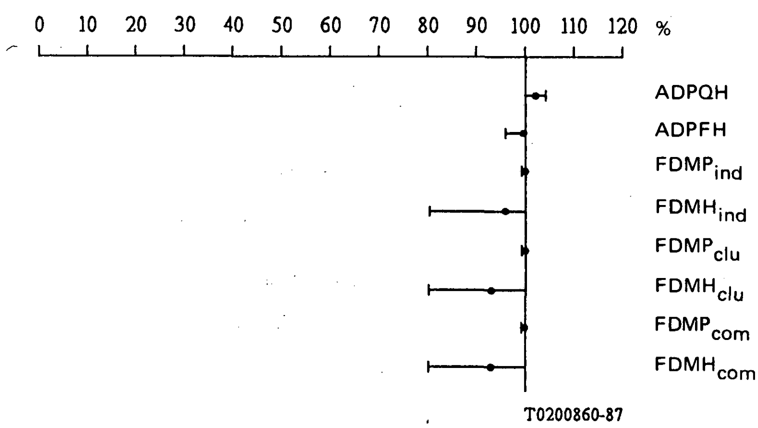


FIGURE A-5/E.500

Comparison on circuit group level (Cluster 3)

References

[1] PARVIALA (A.): The stability of telephone traffic intensity profiles and its influence on measurement schedules and dimensioning (with Appendix). 11th International Teletraffic Congress, Kyoto 1985.

[2] Biometrika Tables for Statisticians, Table 9, Vol. 2. *Cambridge University Press*, 1972.

ESTIMATION OF TRAFFIC OFFERED IN THE INTERNATIONAL NETWORK

1 Introduction

For planning the growth of the international network the following quantities must be estimated from measurements:

- traffic offered to international circuit groups,
- traffic offered to destinations, on a point-to-point basis,
- traffic offered to international exchanges,
- call attempts offered to international exchanges,
- traffic offered to signalling links.

(The term "traffic offered" as used here is different from the "equivalent traffic offered" used in the pure lost call model, which is defined in Annex B.)

These quantities are normally estimated from measurements of busy-hour carried traffic and call attempts, but there are a number of factors which may need to be taken into account within the measurement and estimation procedures:

a) Measurements may need to be subdivided, e.g. on a destination basis, or by call type (for example, calls using different signalling systems).

b) It may not be possible to obtain a complete record of traffic carried. For example, in a network with high usage and final groups it may not be possible to measure the traffic overflowing from each high usage group.

c) Measurements may be affected by congestion. This will generally result in a decrease in traffic carried, but the decrease may be affected by customer repeat attempts and by the actions (for example, automatic repeat attempts) of other network components.

d) When high levels of congestion persist for a lengthy period (many days), some customers may avoid making calls during the congested period of each day. This apparent missing component of offered traffic is known as suppressed traffic. It should be taken into account in planning since the offered traffic will increase when the equipment is augmented. At present, suitable algorithms for estimating suppressed traffic have not been defined.

Three situations should be distinguished:

i) congestion upstream of the measurement point. This is not directly observable;

ii) congestion due to the measured equipment. Congestion measurements should be used to detect this;

iii) congestion downstream of the measurement point. This can often be detected from measurements of ineffective traffic or completion ratio. Note that where groups are bothway, congestion elsewhere in the network may be both upstream and downstream of the measurement point for different parcels of traffic.

When congestion is due to the measured equipment this must be properly accounted for in the estimation of traffic offered, which is used for planning the growth of the measured equipment.

When congestion arises elsewhere in the network the planner needs to consider whether the congestion will remain throughout the considered planning period. This may be difficult if he does not have control of the congested equipment.

This Recommendation presents estimation procedures for two of the situations described above. § 2 deals with the estimation of traffic offered to a fully-operative only-route circuit group which may be in significant congestion. § 3 deals with a high-usage and final group arrangement with no significant congestion. These estimation procedures should be applied to individual busy-hour measurements. The resulting estimates of traffic offered in each hour should then be accumulated according to the procedures described in Recommendation E.500.

2 Only-route circuit group

2.1 *No significant congestion*

Traffic offered will equal traffic carried measured according to Recommendation E.500. No estimation is required.

2.2 *Significant congestion*

Let A_c be the *traffic carried* on the circuit group. Then on the assumption that augmentation of the circuit group would have no effect on the mean holding time of calls carried, or on the completion ratio of calls carried, the *traffic offered* to the circuit group may be expressed as

$$A = A_c \frac{(1 - WB)}{(1 - B)}$$

where B is the present average loss probability for all call attempts to the considered circuit group, and W is a parameter representing the effect of call repetitions. Models for W are presented in Annex A.

To facilitate the quick determination of offered traffic according to the approximate procedure in Annex A, Table A-1/E.501 including numerical values of the factor $(1 - WB)/(1 - B)$ was prepared for a wide range of B, H and r' (for the definition of H and r', see Annex A). For the use of Table A-1/E.501, see Note 2 in Annex A.

Note 1 – Annex A gives a derivation of this relationship, and also describes a more complex model which may be of use when measurements of completion ratios are available.

Note 2 – When measurements of completion ratios are not available a W value may be selected from the range 0.6-0.9. It should be noted that a lower value of W corresponds to a higher estimate of traffic offered. Administrations are encouraged to exchange the values of W that they propose to use.

Note 3 – Administrations should maintain records of data collected before and after augmentations of circuit groups. This data will enable a check on the validity of the above formula, and on the validity of the value of W used.

Note 4 – In order to apply this formula it is normally assumed that the circuit group is in a fully operative condition, or that any faulty circuits have been taken out of service. If faulty circuits, or faulty transmission or signalling equipment associated with these circuits remain in service, then the formula may give incorrect results.

3 High-usage/final network arrangement

3.1 *High-usage group with no significant congestion on the final group*

3.1.1 Where a relation is served by a high-usage and final group arrangement, it is necessary to take simultaneous measurements on both circuit groups.

Let A_H be the traffic carried on the high-usage group, and A_F the traffic overflowing from this high-usage group and carried on the final group. With no significant congestion on the final group, the traffic offered to the high-usage group is:

$$A = A_H + A_F$$

3.1.2 Two distinct types of procedure are recommended, each with several possible approaches. The method given in § 3.1.2.1 a) is preferred because it is the most accurate, although it may be the most difficult to apply. The methods of § 3.1.2.2 may be used as additional estimates.

3.1.2.1 Simultaneous measurements are taken of A_H and the total traffic carried on the final group. Three methods are given for estimating A_F, in decreasing order of preference:

a) A_F is measured directly. In most circumstances this may be achieved by measuring traffic carried on the final group on a destination basis.

b) The total traffic carried on the final group is broken down by destination in proportion to the number of effective calls to each destination.

c) The traffic carried on the final group is broken down according to ratios between the bids from the high-usage groups and the total number of bids to the final group.

3.1.2.2 Two alternative methods are given for estimating the traffic offered to the high-usage group, which in this circumstance equals the equivalent traffic offered:

a) A is estimated from the relationship

$$A_H = A[1 - E_N(A)]$$

Here $E_N(A)$ is the Erlang loss formula, N is the number of working circuits on the high-usage group. The estimation may be made by an iterative computer program, or manually by the use of tables or graphs.

The accuracy of this method may be adversely affected by the non-randomness of the offered traffic, intensity variation during the measurement period, or use of an incorrect value for N.

b) A is estimated from

$$A = A_H/(1 - B)$$

where B is the measured overflow probability. The accuracy of this method may be adversely affected by the presence of repeat bids generated by the exchange if they are included in the circuit group bid register.

It is recommended to apply both methods a) and b); any significant discrepancy would then require further investigation. It should be noted however that both of these methods may become unreliable for high-usage groups with high overflow probability: in this situation a longer measurement period may be required for reliable results.

3.2 *High-usage group with significant congestion on the final group*

In this case, estimation of the traffic offered requires a combination of the methods of §§ 2.2 and 3.1. A proper understanding of the different parameters, through further study, is required before a detailed procedure can be recommended.

ANNEX A

(to Recommendation E.501)

A simplified model for the formula presented in § 2.2

The call attempts arriving at the considered circuit group may be classified as shown in Figure A-1/E.501.

The total call attempt rate at the circuit group is

$$N = N_0 + N_{NR} + N_{LR}.$$

We must consider $N_0 + N_{NR}$ which would be the call attempt rate if there were no congestion on the circuit group.

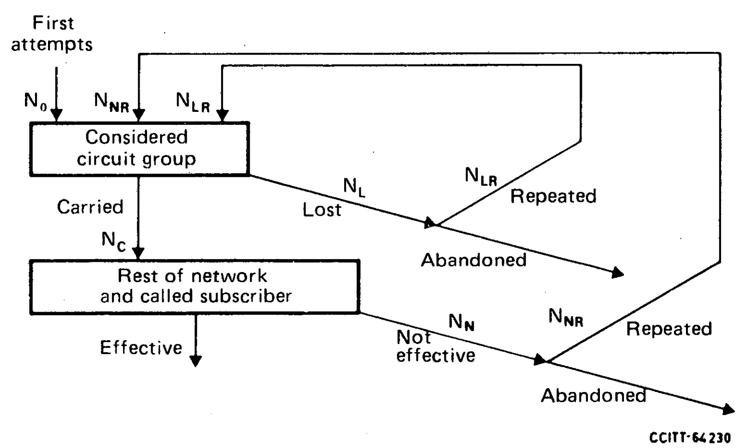
Let

$$B = \frac{N_L}{N} = \text{measured blocking probability on the circuit group.}$$

$$W = \frac{N_{LR}}{N_L} = \text{proportion of blocked call attempts that re-attempt.}$$

We have

$$N_0 + N_{NR} = N - N_{LR} = (N - N_{LR}) \frac{N_c}{N_c} = N_c \frac{(N - N_{LR})}{(N - N_L)} = N_c \frac{(1 - BW)}{(1 - B)}.$$



N_0 First attempt calls
N_C Carried calls
N_L Lost calls
N_{LR} Lost calls repeated
N_N Non-effective calls
N_{NR} Non-effective calls repeated

FIGURE A-1/E.501

Multiplying by the mean holding time of calls carried on the circuit group, h, gives

$$A = A_c \frac{(1 - BW)}{(1 - B)},$$

where

A_c the traffic carried on the circuit group.

The above model is actually a simplification since the rate N_{NR} would be changed by augmentation of the circuit group.

An alternative procedure is to estimate an equivalent persistence W from the following formulae:

$$W = \frac{r' H}{1 - H(1 - r')}$$

$$H = \frac{\beta - 1}{\beta(1 - r)}$$

$$\beta = \frac{\text{All call attempts}}{\text{First call attempts}}$$

where r' is the completion ratio for seizures on the considered circuit group and r is the completion ratio for call attempts to the considered circuit group.

These relationships may be derived by considering the situation after augmentation (see Figure A-2/E.501).

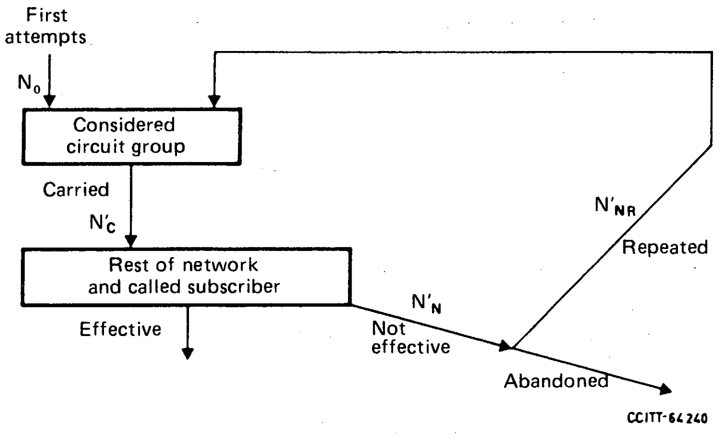


FIGURE A-2/E.501

It is required to estimate N'_c, the calls to be carried when there is no congestion on the circuit group. This may be done by establishing relationships between N_c and N_0 (before augmentation) and between N'_c, and N_0 (after augmentation), since the first attempt rate N_0 is assumed to be unchanged. We introduce the following parameters:

H = overall subscriber persistence,

r' = completion ratio for seizures on the circuit group.

Before augmentation:

$$H = \frac{N_{NR} + N_{LR}}{N_N + N_L}$$

$$r' = \frac{N_c - N_N}{N_c}$$

After augmentation:

$$H = \frac{N'_{NR}}{N'_N}$$

$$r' = \frac{N'_c - N'_N}{N'_c}$$

It is assumed for simplicity that H and r' are unchanged by the augmentation. The following two relationships may be readily derived:

$$N_0 = \frac{N_c\,[1 - H(1 - r') - r'\,BH]}{1 - B}$$

$$N_0 = N'_c\,[1 - H\ (1 - r')].$$

Hence

$$N'_c = \frac{N_c\left[1 - \left(\frac{r'H}{1 - H(1 - r')}\right)B\right]}{1 - B}$$

On multiplying by the mean call holding time, h, this provides our estimate of traffic offered in terms of traffic carried.

The relationship $$H = \frac{\beta - 1}{\beta(1 - r)}$$

is valid both before and after augmentation, as may easily be derived from the above diagrams.

Note 1 – Other Administrations may be able to provide information on the call completion ratio to the considered destination country.

Note 2 – The procedure of estimating the factor W above is based on the assumptions that H, r' and h remain unchanged after augmentation. The elimination of congestion in the group considered leads to a change in H and in practical cases this causes an underestimation of the factor W and consequently an overestimation of offered traffic in the formula of § 2.2. A relevant study in the period 1985-88 has shown that the overestimation is practically negligible if $B \leqslant 0.2$ and $r' \geqslant 0.6$. For larger B and smaller r' values, the overestimation may be significant unless other factors, not having been taken into account by the study, do not counteract. Therefore caution is required in using Table A-1/E.501 in the indicated range. In the case of dynamically developing networks the overestimation of offered traffic and relevant overprovisioning may be tolerated, but this may not be the case for stable networks.

TABLE A-1/E.501

Values of $\frac{1 - WB}{1 - B}$

$H =$	0.70	0.75	0.80	0.85	0.90	0.95
$B = 0.1$						
$r' = 0.3$	1.0653	1.0584	1.0505	1.0411	1.0300	1.0165
$r' = 0.4$	1.0574	1.0505	1.0427	1.0340	1.0241	1.0129
$r' = 0.5$	1.0512	1.0444	1.0370	1.0289	1.0202	1.0105
$r' = 0.6$	1.0462	1.0396	1.0326	1.0252	1.0173	1.0089
$r' = 0.7$	1.0421	1.0358	1.0292	1.0223	1.0152	1.0077
$r' = 0.8$	1.0387	1.0326	1.0264	1.0200	1.0135	1.0068
$B = 0.2$						
$r' = 0.3$	1.1470	1.1315	1.1136	1.0925	1.0675	1.0373
$r' = 0.4$	1.1293	1.1136	1.0961	1.0765	1.0543	1.0290
$r' = 0.5$	1.1153	1.1	1.0833	1.0652	1.0454	1.0238
$r' = 0.6$	1.1041	1.0892	1.0735	1.0568	1.0390	1.0201
$r' = 0.7$	1.0949	1.0806	1.0657	1.0503	1.0342	1.0174
$r' = 0.8$	1.0872	1.0735	1.0595	1.0451	1.0304	1.0154
$B = 0.3$						
$r' = 0.3$	1.2521	1.2255	1.1948	1.1587	1.1158	1.0639
$r' = 0.4$	1.2216	1.1948	1.1648	1.1311	1.0931	1.0498
$r' = 0.5$	1.1978	1.1714	1.1428	1.1118	1.0779	1.0408
$r' = 0.6$	1.1785	1.1530	1.1260	1.0974	1.0669	1.0345
$r' = 0.7$	1.1627	1.1382	1.1127	1.0862	1.0587	1.0299
$r' = 0.8$	1.1495	1.1260	1.1020	1.0774	1.0522	1.0264
$B = 0.4$						
$r' = 0.3$	1.3921	1.3508	1.3030	1.2469	1.1801	1.0995
$r' = 0.4$	1.3448	1.3030	1.2564	1.2040	1.1449	1.0775
$r' = 0.5$	1.3076	1.2666	1.2222	1.1739	1.1212	1.0634
$r' = 0.6$	1.2777	1.2380	1.1960	1.1515	1.1041	1.0537
$r' = 0.7$	1.2531	1.2150	1.1754	1.1342	1.0913	1.0466
$r' = 0.8$	1.2325	1.1960	1.1587	1.1204	1.0813	1.0411
$B = 0.5$						
$r' = 0.3$	1.5882	1.5263	1.4545	1.3703	1.2702	1.1492
$r' = 0.4$	1.5172	1.4545	1.3846	1.3061	1.2173	1.1162
$r' = 0.5$	1.4615	1.4	1.3333	1.2608	1.1818	1.0952
$r' = 0.6$	1.4166	1.3571	1.2941	1.2272	1.1562	1.0806
$r' = 0.7$	1.3797	1.3225	1.2631	1.2013	1.1369	1.0699
$r' = 0.8$	1.3488	1.2941	1.2380	1.1807	1.1219	1.0617

ANNEX B

(to Recommendation E.501)

Equivalent traffic offered

In the lost call model the equivalent traffic offered corresponds to the traffic which produces the observed carried traffic in accordance with the relation

$$y = A(1 - B)$$

where

y is the carried traffic,

A is the equivalent traffic offered,

B is the call congestion through the part of the network considered.

Note 1 – This is a purely mathematical concept. Physically it is only possible to detect bids whose effect on occupancies tells whether these attempts give rise to very brief seizures or to calls.

Note 2 – The equivalent traffic offered, which is greater than the traffic carried and therefore greater than the effective traffic, is greater than the traffic offered when the subscriber is very persistent.

Note 3 – B is evaluated on a purely mathematical basis so that it is possible to establish a direct relationship between the traffic carried and call congestion B and to dispense with the role of the equivalent traffic offered A.

Recommendation E.502

TRAFFIC MEASUREMENT REQUIREMENTS FOR SPC (ESPECIALLY DIGITAL) TELECOMMUNICATION EXCHANGES

1 Introduction

This Recommendation applies to all SPC (especially digital) telecommunications exchanges operating in a switched telephone network and providing basic telephony service. This Recommendation will be the basis for measurements in an Integrated Services Digital Network (ISDN).

Traffic measurements on exchanges and surrounding telephone network provide the data base from which the dimensioning, planning, operation and management of the telephone network are carried out.

Information gathered from these measurements can be used for:

– identifying traffic patterns and distributions on a route and destination basis;
– determining the amount of traffic in the exchange and the network;
– monitoring the continuity of service and the grade of service.

The above data and information are gathered with the purpose of supporting the following fundamental activities:

a) dimensioning, planning and administration of the exchange and surrounding network;
b) performance monitoring of the exchange and surrounding network;
c) network management;
d) operation and maintenance of the exchange and surrounding network;
e) tariff and marketing studies;
f) forecasting;
g) dimensioning, planning and administration of the common channel signalling network;
h) performance monitoring of the common channel signalling network.

The information generated by the exchange (see Recommendation Q.544) can be provided to the end user in either real-time or non real-time (post processed). The activities being performed by the end user will dictate the speed of this response: for example, operation and maintenance will require real-time information while the forecasting and planning information can be provided after the event in non real-time.

For these activities, the following major processing steps can be identified:

- generation, collection and storage of data;
- analysis and processing of data;
- presentation and use of the analysis results.

The generation, collection and output of raw data is achieved by continuous as well as periodic and non-periodic measurements carried out in the exchange.

The data analysis may be performed by the SPC exchange or by another system depending on the following:

- total amount of data;
- need for analysis of data from multiple exchange;
- processor load constraints.

For further information see Recommendation E.503.

2 Traffic measurement

2.1 *Traffic measurement model*

This section establishes the basic structure for a traffic measurement model that can be applied to measurements of traffic generated by the basic telephony service.

Measurements of traffic generated by ISDN services and common channel signalling systems is for further study.

A measurement is identified by three basic elements: time, entities, objects. Time includes all the necessary information to define the start, the duration and periodicity of a certain measurement. Entities describe the quantities for which data collection must be performed with a certain measurement. Objects are individual items on which the measurements are performed. Some examples of entities and objects are given below:

Entities:

- traffic volume;
- number of call attempts;
- number of seizures;
- number of successful call attempts;
- number of call attempts for which a delay exceeds a predetermined threshold value.

Objects:

- subscriber line groups;
- circuit groups;
- common control units;
- auxiliary devices;
- destinations;
- common channel signalling links;
- signal transfer points (STP).

The measurements are classified into different measurement types on the basis of a measurement matrix in which each row represents an entity and each column represents an object (Figure 1/E.502).

A measurement type is a particular combination of entities and objects corresponding to certain entries in the measurement matrix. Part of these measurement types may be standardized while the rest of them seem to be system- and/or Administration-dependent. It should be noted that all the entries in the measurement matrix cannot be used because some of them will be impossible and some others may be meaningless. In all measurement types, the entities are fixed although some entities may not be measured for some applications. Selected objects form an object list. In some measurement types, the object list is fixed. In other types one can choose for the actual measurement some or all of the allowed objects. A measurement set is a collection of measurement types.

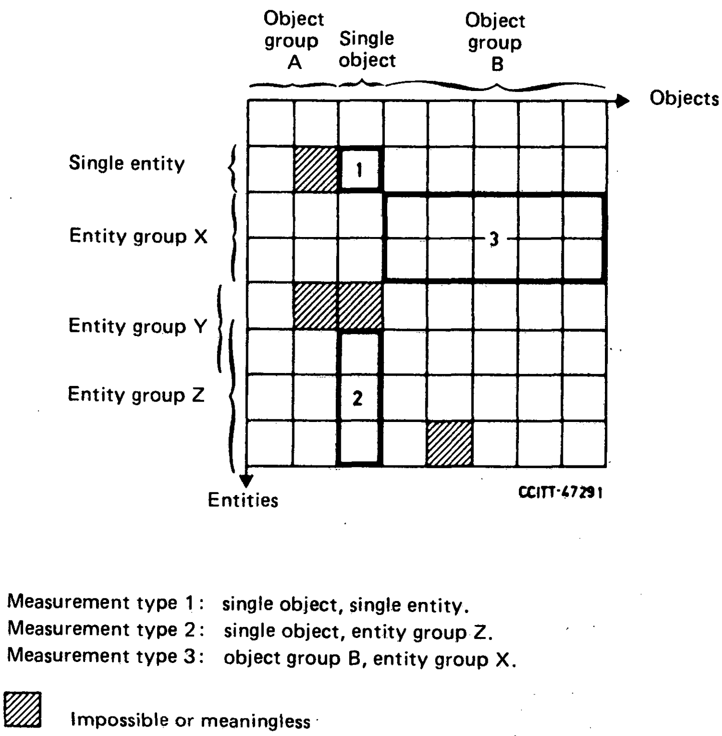


FIGURE 1/E.502

Measurement matrix

2.2 *Traffic measurement structure*

A traffic measurement consists of:

- measurement set information;
- time information;
- output routing and scheduling information (output parameters).

Measurement set information, time information and output routing and scheduling information may be predefined as well as object lists. It should be noted that predefinition characteristic are system-dependent. Time data routing and the schedule may be fixed.

2.2.1 *Measurement set information*

Measurement set information consists of one or several selected measurement types with defined object (object lists) and measurement-type-dependent parameters (e.g. sampling interval, number of events in a certain category, destination codes, etc.).

2.2.2 *Time information*

Measurements may have an undetermined duration (stop date is not prespecified), or a predetermined duration, or be taken all the time. In addition, measurements may be performed continuously or on a non-continuous basis.

For measurements of undetermined duration and performed non-continuously, the recording days must be determined on a periodic basis (periodicity pattern within a calendar week). For measurements of predetermined duration, the recording days may be determined on a periodic basis or by defining the dates of the recording days (see Figure 2/E.502).

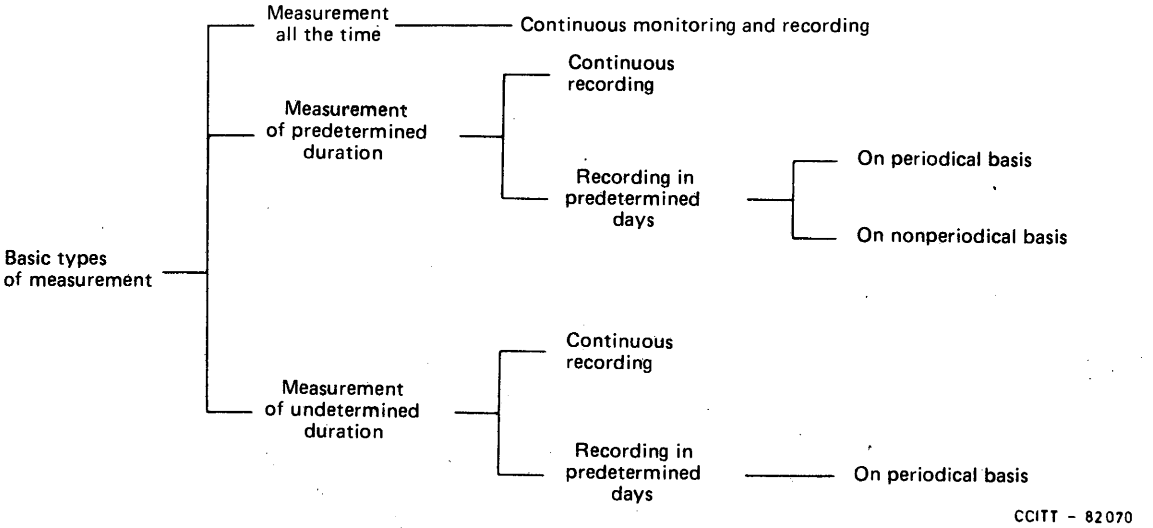


FIGURE 2/E.502

As shown in Figure 3/E.502, time data are measurement level, recording day level and recording period level.

Measurement level: Contains information about dates of recording days for non-periodic measurements or periodicity pattern for periodic measurements.

Recording day level: Contains information about the start and stop time for recording periods within a recording day.

Recording period level: Contains information about the periodicity of the data collection, controlled by the result accumulation period. The result accumulation period can be shorter than the recording period; in that case, more than one set of data is collected for each of the recording periods, to be routed toward the output media according to the results output schedule.

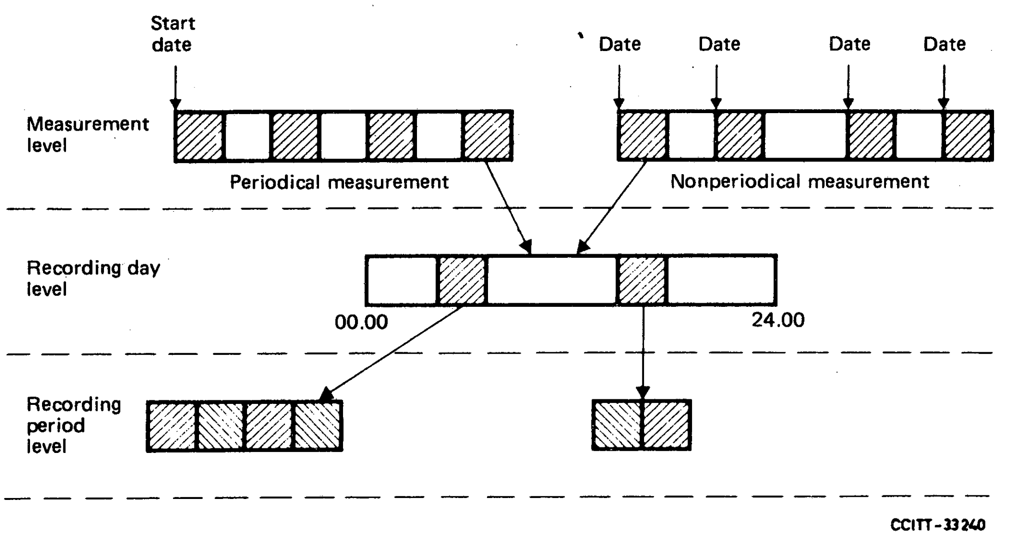


FIGURE 3/E.502

2.2.3 *Output routing and scheduling information*

Output routing information defines to what destination the produced measurement results should be routed for the recording; the output routing may be toward either a physical medium (e.g. printer) or a logical medium (e.g. file).

Output scheduling information defines when (days and time) the output of the results is to be made. The output of results may be related to the end of the result accumulation period.

3 Traffic flows

Each type of traffic flow occurring in/through the exchange can be distinguished by association with an inlet [1] or outlet [2] of the exchange, or both. The different types of traffic flow for a generalized exchange, viz. one that combines both local and transit functions and that provides operator (telephonist) service, are illustrated as shown in Figure 4/E.502:

From Figure 4/E.502 the following relations apply:

$$A = E + F + G + H + Z_1$$

$$B = I + J + K + L + Z_2$$

$$C = O + P$$

$$D = M + N + Z_3$$

where Z_1, Z_2 and Z_3 account for traffic flows corresponding to calls with incomplete or invalid dialling information, and

$$Q = M + F + K + O - d_1$$

$$R = N + G + L + P - d_2$$

$$S = H + J - d_3$$

$$T = E + I - d_4$$

where d_1, d_2, d_3 and d_4 account for traffic flows corresponding to calls that fail within the exchange owing to any of the following reasons:

a) all suitable outlets are busy or unavailable;

b) internal congestion;

c) incomplete dialling;

d) invalid destination code;

e) service barring/blocking (as a result of network management controls, for instance, or the operation of some supplementary service (e.g. absentee service), or because the calling/called party is disallowed such service).

The types of calls, viz. *system-originating* call and *system-terminating* calls, result from the operation of some of the supplementary or value-added services that SPC exchanges offer in addition to conventional telephone service. In the traffic flow diagram of Figure 4/E.502, system-originating and system-terminating calls are identified by the aggregate traffic flows C and S respectively.

4 Basic measurement types

4.1 *General*

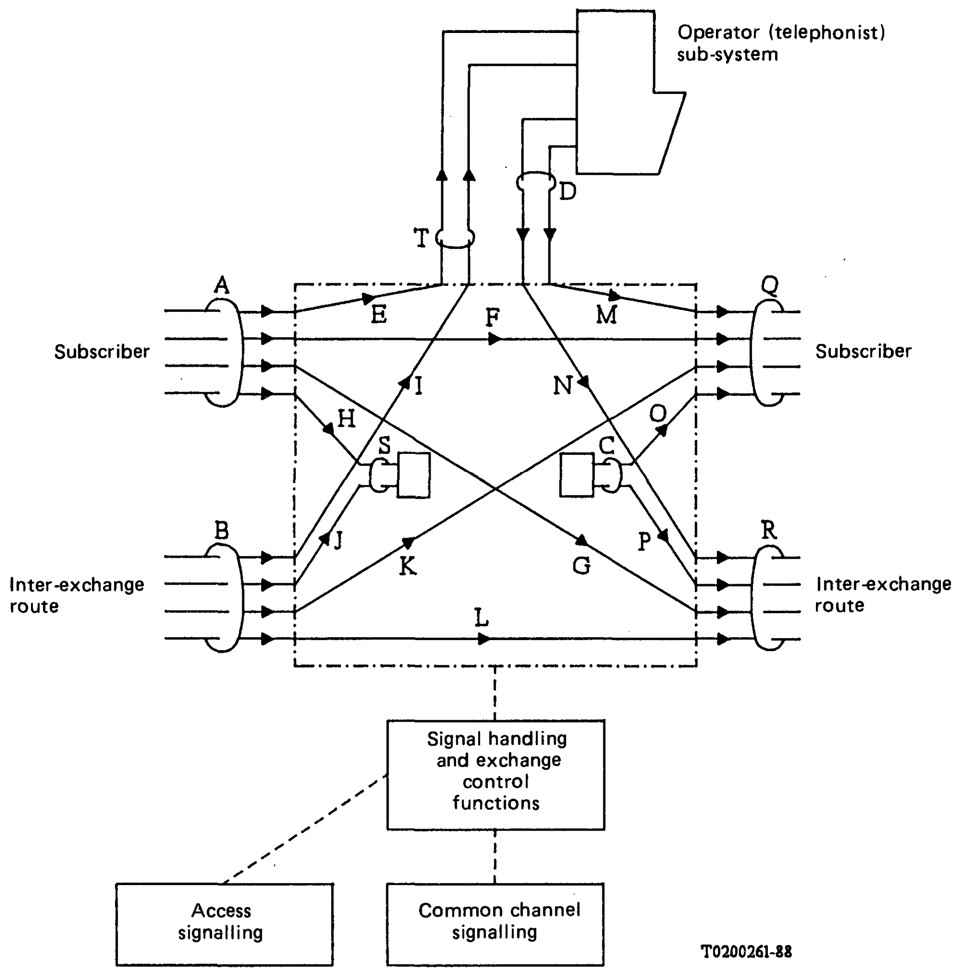
4.1.1 Depending on the activities listed in § 1, a different degree of detail may be needed.

In order to provide bulk data for each of the above-mentioned traffic categories, overall measurements can be performed on the totality of subscriber lines and/or circuits.

---

[1] Inlet is the point on or within the boundary of the exchange system where a call attempt arrives or arises.

[2] Outlet is the point on or within the boundary of the exchange system to which a call attempt bearing adequate and valid dialling information would tend to be routed.



A – Originating traffic
B – Incoming traffic
Q – Terminating traffic
R – Outgoing traffic
F – Internal traffic
G – Originating outgoing traffic
H – Originating system terminating traffic
J – Incoming system terminating traffic
K – Incoming terminating traffic
L – Transit traffic
O – System originating terminating traffic
P – System originating outgoing traffic
S – System terminating traffic
C – System originating traffic
T – Operator terminating traffic
D – Operator originating traffic
E – Originating operator terminating traffic
I – Incoming operator terminating traffic
M – Operator originating terminating traffic
N – Operator originating outgoing traffic

FIGURE 4/E.502

Main traffic flow diagram

Such overall measurements have been taken into account in this Recommendation only for the traffic items from *A* to *P* in Figure 4/E.502, while they have not been considered for items *Q*, *R*, *S* and *T* since, with the assumptions made above, it is possible to achieve the relevant information by taking into account the relationship between these items and the measured ones. It is recognized that the overall measurement results might be partitioned to cover various Administrations' needs. As an example, in an international transit exchange, the traffic data measured on the totality of incoming circuits should be split into data measured on national incoming circuits and international incoming circuits, and these in turn could be differentiated according to the relevant country.

More detailed information on traffic data relevant to the exchange and surrounding network performance can be provided by means of measurements on selected sets of circuit groups, subscriber line groups, common channel signalling links, STPs, auxiliary and control units.

Very detailed traffic data can be obtained by the analysis of call records.

These call records should be produced by the exchange, containing all data (e.g. time of occurrence of signalling event, dialled digits, etc.) characterizing each individual call attempt.

The basic measurement types are given in § 4.2 below.

Their applicability will depend on the function of the exchange (local, transit, international, etc.)

Manufacturers and Administrations are to note that the list of basic measurement types is derived from the traffic model given in Figure 4/E.502. It is not intended that every exchange system should contain all the different measurement types. The measurement types are exchange- and system-dependent, and are intended as a guide to the type of measurements required to fulfil various functions. Measurement types may be combined into a few sets to enable requirements to be met for specific exchange types, e.g. local. In particular Administrations may consider that by the use of a few measurement types it is possible to satisfy the majority of their requirements.

No single measurement type can be assumed to be exclusive to a single user or to satisfy a single requirement. More than one user may require the same information presented in different ways at the same time. As an example, measurement type 22 is required for both network management and traffic engineering purposes.

4.1.2 *Network management considerations*

4.1.2.1 Information on network management is contained in the E.410 Series of Recommendations. Network management requires "real-time" monitoring and measurement of network status and performance and the ability to take prompt action to control the flow of traffic when necessary.

4.1.2.2 *Performance reports*

Performance reports can be provided by the exchange and/or its network management operations system (OS) in the following ways, as required by the Administration:

i) automatic data – this data is provided automatically as specified in the exchange or OS program;

ii) scheduled data – this data is provided according to a schedule established by the network manager;

iii) demand data – this data is provided only in response to a specific request by the network manager. In addition to performance data, demand data includes reference data, such as the number of circuits provided or available for service, routing information, assigned threshold values, numbers of installed switching system components, etc.;

iv) exception data – this data is provided when a data count for calculation exceeds a threshold established by the network manager.

Data reports can be provided for example on a 5-minute, 15-minute or 30-minute basis. The specific interval for any data report will be determined by the network manager. Historic data relating to the previous two or three periods (5, 15 or 30-minute) must also be available.

4.1.2.3 In order to obtain information and apply controls which may be instrumental in reducing exchange congestion, Administration should ensure that network management terminals and functions should have the highest possible priority, so that network management operations can continue uninterrupted.

4.1.2.4 Information as to which network management controls, detailed in Recommendation E.412, are currently activated and whether the controls were activated by manual or automatic means should be available to all necessary points (for example, the network management centre, exchange staff).

4.1.3 *Traffic engineering*

Information on measurements for planning purposes is given in Recommendation E.500. For further details about requirements on measurement lengths over the year and the day, data reporting intervals, etc., reference should be made to that Recommendation.

4.2 *Measurements*

4.2.1 *Overall measurements*

Type 1: Overall measurements on originating traffic (A).

Object: Totality of subscriber lines.

Entities:

a) Number of originating seizures;

b) Number of call attempts not routed due to:
   - i) no dialling (including permanent signal),
   - ii) incomplete dialling[3],
   - iii) invalid address;

c) Number of call attempts lost due to internal congestion[4].

Type 2: Overall measurements on internal traffic ($E + F + H$)[5].

Object: Totality of subscriber lines.

Entities:

a) Number of internal seizures;

b) Number of call attempts lost due to internal congestion;

c) Number of call attempts:
   - i) with called-party busy,
   - ii) with called-party free/no answer[6],
   - iii) answered;
   - iv) line out of order,
   - v) vacant national number,
   - vi) transferred subscriber;

d) Number of unsuccessful call attempts due to incomplete dialling[5].

Type 3: Overall measurements on originating outgoing traffic (G).

Object: Totality of subscriber lines.

Entities:

a) Number of outgoing seizures;

b) Number of call attempts lost due to internal congestion;

c) Number of call attempts in overflow on the last choice route;

d) Number of successful call attempts getting:
   - i) no answer[7],
   - ii) answer or metering pulse(s);

e) Number of unsuccessful call attempts due to incomplete dialling[3].

---

[3] Not enough digits to discriminate if internal or outgoing call.

[4] When possible, broken down by reason of congestion, e.g. c-1 blocking through the switching network, c-2 unavailability of common resources, c-3 system faults.

[5] Entities may be broken down according to relevant traffic flows.

[6] Expiring of time-outs calling-party's abandon.

[7] Due to time-out expiring or calling-party's abandon or called-party busy.

Type 4: Overall measurements on incoming traffic (B).

Object: Totality of incoming circuits and both-way circuits.

Entities:

a) Number of incoming seizures;

b) Number of call attempts not routed due to:
   - i) incomplete dialling [8],
   - ii) invalid address;

c) Number of call attempts lost due to internal congestion.

Type 5: Overall measurements on incoming terminating traffic ($I + J + K$) [9].

Object: Totality of incoming circuits and both-way circuits.

Entities:

a) Number of incoming terminating seizures;

b) Number of call attempts lost due to internal congestion;

c) Number of successful call attempts:
   - i) with called-party busy,
   - ii) with called-party free/not answered,
   - iii) answered or metering pulse(s);

d) Number of unsuccessful call attempts due to incomplete dialling.

Type 6: Overall measurements on transit traffic (L).

Object: Totality of incoming circuits and both-way circuits.

Entities:

a) Number of incoming transit seizures;

b) Number of call attempts lost due to internal congestion;

c) Number of call attempts in overflow on the last-choice route;

d) Number of successful call attempts obtaining:
   - i) no answer [10].
   - ii) no answer or metering pulse(s);

e) Number of unsuccessful call attempts due to incomplete dialling [10].

Type 7: Overall measurements on system originating traffic ($O + P$) [9].

Object: Exchange system.

Entities:

a) Number of system originating seizures;

b) Number of call attempts lost due to internal congestion;

c) Number of successful call attempts:
   - i) with called party busy or no free outlet,
   - ii) with called party free/not answered (for O),
   - iii) answered.

---

[8] Not enough digits to discriminate if internal or outgoing call.

[9] Entities may be broken down according to relevant traffic flows.

[10] Expiring of time-out or receiving a release forward.

Type 8: Overall measurements on operator-originating traffic $(M + N)$[11].

Object: Totality of operator board trunks.

Entities:

a) Number of operator originating seizures;

b) Number of unsuccessful call attempts due to:
   - i) incomplete dialling,
   - ii) invalid address,
   - iii) internal congestion;

c) Number of successful call attempts:
   - i) with called party busy or no free outlet,
   - ii) with called party free/not answered (for M),
   - iii) answered.

4.2.2 *Measurement on selectable objects*

Type 9: Incoming traffic measurements.

Object: Each incoming circuit group and both-way circuit group.

Entities:

a) Number of incoming seizures;
b) Traffic volume;
c) Number of call attempts lost due to internal congestion[12];
d) Number of circuits in service;
e) Number of circuits out of service.

Type 10: Outgoing traffic measurements.

Object: Each outgoing circuit group and both-way circuit group.

Entities:

a) Number of outgoing seizures;
b) Traffic volume;
c) Number of call attempts in overflow;
d) Number of seizures obtaining answer;
e) Number of circuits in service;
f) Number of circuits out of service;
g) Number of dual seizures (both-way circuits only).

Type 11: Route destination traffic measurements.

Object: For destinations on each outgoing circuit group and both-way circuit group.

Entities:

a) Number of outgoing seizures;
b) Number of effective call attempts;
c) Traffic volume;
d) Number of call attempts, lost due to congestion on the circuit group;
e) Source (identity of incoming circuit group) – if available.

---

[11] Entities may be broken down according to relevant traffic flows.

[12] When possible, broken down by reason of congestion, e.g. c-1 blocking through the switching network, c-2 unavailability of common resources, c-3 system faults.

Type 12: Measurements on subscriber line groups.

Object: Set of lines composing a functional unit.

Entities:

a) Originating traffic volume;
b) Terminating traffic volume;
c) Number of originating seizures;
d) Number of terminating seizures;
e) Number of terminating call attempts.

Type 13: Measurements on auxiliary units[13]).

Object: Selected groups of auxiliary units.

Entities:

a) Number of seizures;
b) Traffic volume;
c) Numbers of non-serviced call attempts;
d) Number of units in service;
e) Number of units out of service.

4.2.3 *Measurements on control unit(s)*

Type 14: Measurements on control unit(s).

Object: Control unit(s).

These measurements are highly system-dependent and therefore no specific recommendations on relevant entities can be made. However, it is essential that systems have provisions for determining the utilization of control units as required for dimensioning, planning, and grade of service monitoring of the exchange.

4.2.4 *Measurements on call records*[14])

Type 15: Traffic dispersion and duration.

Object: Originating (by subscriber, exchange system, operator) and/or incoming seizures ($A + B + C + D$).

Entities:

a) Source of inlet (local subscriber, exchange system or incoming/both-way circuit group);
b) Time of seizure of inlet;
c) Dialled digits;
d) Service characteristic of call attempt[15]) for successful call attempt;
e) Identity of exchange outlet;
f) Time of seizure of outlet;
g) Time of occurrence of call attempt at exchange outlet;
h) Time of address-complete signal (if available);
i) Time of answer signal;
j) Time of release of outlet;
k) Time of release of inlet.

---

13) By auxiliary units it is meant multifrequency code (MFC) receivers, tone circuits, etc.

14) The collection of the totality of call attempts could cause an excessive load for the SPC system resources, therefore such measurements might be performed on a sampling basis.

15) Whether the call attempt uses or seeks to use any of the supplementary facilities of the exchange; if so, the supplementary facility concerned shall be specifically indicated.

Type 16: Quality-of-service assessment.

Object: Originating (by subscriber, exchange system, operator) and/or incoming seizures ($A + B + C + D$).

Entities:

a) Source or inlet (local subscriber, exchange system or incoming/both-way inter-office circuit group);

b) Time of seizures of inlet;

c) Dialled digits.

For unsuccessful call attempt, specify causes of failure:

d) No dialling;

e) Incomplete dialling;

f) Invalid address;

g) No free outlet;

h) Internal congestion;

i) Due to network management action.

For successful call attempt:

j) Order of routing choice (first, second, . . ., last) (when considering the automatic repeated attempts and/or rerouting);

k) Time of address-complete signal (undifferentiated subscriber free, subscriber busy, backward congestion) (if available);

l) Result of call attempt (answer, release due to abandon, release due to congestion).

4.2.5 *Delay grade-of-service (GOS) monitoring*

Measuring delays on a per call basis could produce severe cost penalties to the exchange. Since the accuracy requirements from the statistical viewpoint are not very high, call sampling procedures or test calls are normally sufficient for GOS monitoring purposes. For this reason these measurement types are listed separately even if types 16 and 17 should belong to § 4.1 and measurement type 18 to § 4.2.

4.2.5.1 *On a per exchange basis*

Type 17: Overall delay grade-of-service parameters monitoring.

Object: Totality of subscriber lines.

Entities:

a) Total number of originating seizures;

b) Total number of originating seizures for which the required information for setting up a through connection is available for processing in the exchange;

c) Total number of originating seizures for which sufficient address information has been received, which are addressed to a certain outgoing circuit group and for which the seizing signal or the corresponding address information is sent to the subsequent exchange;

d) Total number of originating seizures for which the dial tone delay exceeds a predetermined threshold value;

e) Seizures already counted in b) for which the through-connection delay exceeds a predetermined threshold value;

f) Seizures already counted in c) for which the call set-up delay exceeds a predetermined threshold value.

Type 18: Overall delay grade-of-service parameters monitoring.

Object: Totality of incoming or both-way circuit groups.

Entities:

a) Total number of incoming seizures;

b) Total number of incoming seizures for which the required information for setting up a through connection is available for processing in the exchange for a certain circuit group;

c) Total number of incoming seizures for which sufficient address information has been received, which are addressed to a certain outgoing circuit group and for which the seizing signal or the corresponding address information is sent to the subsequent exchange;

d) Total number of incoming seizures for which the incoming response delay exceeds a predetermined threshold value;

e) Seizures already counted in b) for which the through-connection delay exceeds a predetermined threshold value;

f) Seizures already counted in c) for which the call set-up delay exceeds a predetermined threshold value.

4.2.5.2 *On per circuit group basis*

Type 19: Delay grade-of-service parameters monitoring.

Object: Each incoming or both-way circuit group.

Entities:

a) Total number of incoming seizures;

b) Total number of incoming seizures for which the required information for setting up a through connection is available for processing in the exchange for a certain circuit group;

c) Total number of incoming seizures for which sufficient address information has been received, which are addressed to a certain outgoing circuit group and for which the seizing signal or the corresponding address information is sent to the subsequent exchange;

d) Total number of incoming seizures for which the incoming response delay exceeds a predetermined threshold value;

e) Seizures already counted in b) for which the through-connection delay exceeds a predetermined threshold value;

f) Seizures already counted in c) for which the call set-up delay exceeds a predetermined threshold value.

4.2.6 *Network performance monitoring*

Type 20: Network management.

Object: Total exchange and its major components, e.g. processor.

Entities:

a) Bids;

b) Incoming call queue length and overflows;

c) Number and percentage of bids encountering switching delays;

d) Percentage of processor capacity available or in use;

e) Cross exchange delay measurements;

f) Switching loss;

g) Counts of calls blocked by automatic load shedding actions.

Type 21: Network management.[16)]

Object: Common channel signalling system and links.

Entities:

a) Counts of signalling units and percent occupancy of signal links;

b) Counts of outgoing Initial Address Messages (IAMs) and incoming answer signals (ANC and ANN);

16) Although measurement type 21 is identified as being for network management, it is also required for traffic engineering purposes.

c) Counts of incoming Initial Address Messages (IAMs) and outgoing answer signals (ANC and ANN);

d) Counts of changeovers;

e) Counts of occurrences and duration of terminal buffer overflow conditions;

f) Counts of circuit group congestion (CGC), National Network Congestion (NNC), and/or Switching Equipment Congestion (SEC) indications sent and received on the signalling link;

g) Counts of calls overflowed or lost due to terminal buffer overflow;

h) Counts of Transfer-Prohibited (TFP) signals sent and received on the link.

Type 22: Network management.

Object: Each circuit group.

Entities:

a) Bids;

b) Seizures – outgoing and incoming;

c) Answer signals received;

d) Overflows;

e) Traffic carried;

f) Number of circuits made busy to traffic;

g) Transit bids;

h) Incoming transit seizures;

i) Counts of calls affected by network management control, by type of control.

Type 23: Network management.

Object: Destinations.

Entities:

a) Bids;

b) Seizures;

c) Answer signals received;

d) Overflows;

e) Counts of calls affected by network management controls, by type of control (*Note* – This includes code block and call gap controls).

4.2.7 *Measurement of the performance of common channel signalling systems*

Measurement Type 21 (see § 4.2.6) is required. Other measurement types are for further study.

4.2.8 *Measurement of the integrated services digital network and its services*

For further study.

5 Related Recommendations

The use of the analysed results will be dependent on the procedures in each Administration. The list of Recommendations below are those currently existing and covering many operational aspects. They are offered only as a guide rather than a comprehensive and complete set.

– Recommendation E.500 defines the traffic intensity measurement principles;

– Recommendation E.175 defines the network model for planning purposes;

– E.410 Series of Recommendations provide information for network management;

– E.420 Series of Recommendations describe checking the quality of the international telephone service;
– Recommendation E.506 defines the forecasting methods for international traffic;
– Recommendation E.543 defines the grade of service in digital international telephone exchanges;
– Recommendation E.503 defines the traffic measurement data analysis;
– Recommendation E.504 defines the traffic measurement administration;
– the O Series of Recomendations outline specifications of measuring equipment;
– the M Series of Recommendations detail many maintenance aspects of international carrier and circuits;
– The Q Series of Recommendations cover all aspects relating to common channel signalling.
– Recommendation Q.544 deals with exchange measurements.

ANNEX A

(to Recommendation E.502)

The purpose of this Annex is to identify the measurements to be made at exchanges and the criteria needed to satisfy the basic measurement requirements, and is produced to assist the exchange designers to ensure that these measurements can be made.

Considering that an SPC digital exchange is mainly composed of software with few physical entities which can be identified as specific measurement points, it is not possible to identify exactly where measurements should be taken. However, the basic measurement types given in § 4.2 require that it be possible to differentiate between events occurring:

i) from a customer/previous exchange node, arriving at an exchange.

ii) from an exchange to another exchange node/customer.

iii) within an exchange.

In the three segments indicated above it is necessary to have the ability to record the entities independently in each segment, as well as being able to associate entities between segments.

The entities recorded are:

– bids;
– seizures;
– effective calls;
– congested bids;
– traffic volume.

An exchange should categorize failed call attempts according to the reason for the failure. However, the information available to the exchange for this purpose may depend on the signalling system used and the function and position of the exchange in the network relative to the failed call attempts.

It should be noted that measurement type 15 is a call record which has to be generated wholly within an exchange system. Also, measurement types 20, 21, 22 and 23 are specific to network management requiring slightly different criteria.

It shall be possible for any of the basic measurement types to be amalgamated to form a unique measurement program to meet an Administration's requirements. It shall also be possible to output measurement information to more than one user. As an example, measurements may be in progress continuously for traffic engineering purposes and, at a particular time (say for one hour), measurements of the same type may be required for maintenance purposes. The output or recording of these two measurements must not interfere with each other or with any other measurements being made at the same time, e.g. for network management.

TRAFFIC MEASUREMENT DATA ANALYSIS

1 Introduction

The aim of traffic measurements is to provide data that can be used by an Administration for planning, engineering and managing its network. The resulting measured data can be used to support various activities as stated in Recommendation E.502. In order to reduce the amount of data transfer and off-line processing, the exchange or operations system can be used to make preliminary analyses for purposes of:

- eliminating unnecessary data values;
- replacing missing or wrong values in an approriate way;
- performing simple calculations on the values of the basic measurement entities to derive characteristic parameter values of the traffic;
- storing some measured or calculated values, in particular, traffic data records;
- producing appropriate user friendly report printouts.

For each measurement object, there is a data record in which a certain number of traffic values are stored. Also, some calculated values, e.g. moving average, can be stored and updated in this data record area.

The internal functions of the analysis are not specified here. They depend on the requirements for the output results which are specified by the Administration. An acceptable method may be to collect and store the data in real time, either in a temporary data base file or directly in the traffic data record, and later perform the calculations and report printout during periods of low exchange processing activity. Alternatively, the records can be transferred to an off-line system for processing, to reduce the load on the exchange.

2 Potential applications

In order to provide bulk data for traffic and operational analysis, overall measurements can be performed on the totality of subscriber lines and/or circuits.

More specific information on traffic data relevant to the exchange and surrounding network performance can be provided by means of measurements on selected sets of circuit groups, subscriber line groups, common channel signalling links, auxiliary and control units.

Very detailed traffic data can be obtained by analysis of call records. These call records should be produced by the exchange, containing all the data (e.g. time of occurrence of signalling event, dialled digits, etc.) characterizing each individual call attempt.

The relationships between the above measurements and the potential applications are shown in Table 1/E.503. The basic measurement types are given in Recommendation E.502. Their applicability will depend on the function of the exchange (local, transit, international, etc.).

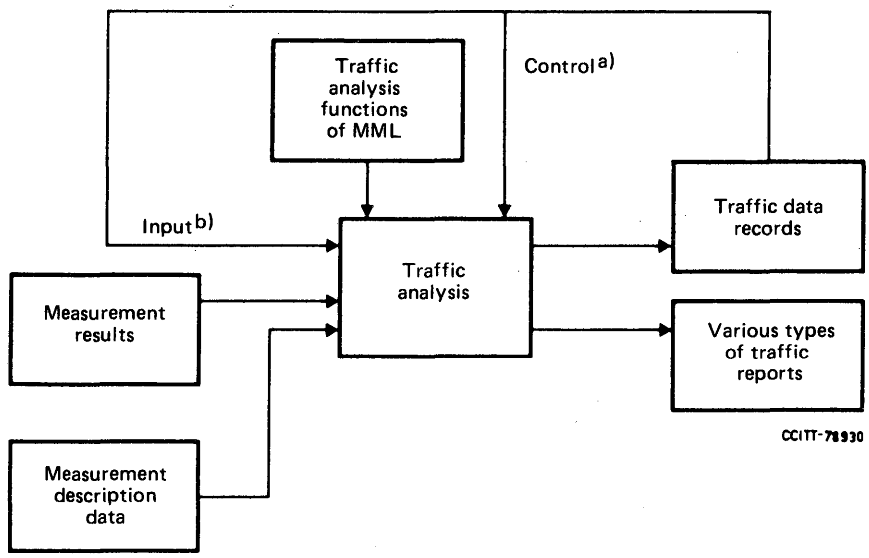
3 Traffic analysis model

Corresponding to a variety of measurements, there are a variety of analyses, some of which are typically running continuously from day to day. From the viewpoint of a particular measurement, there are one or more analyses for which the measured data are written in particular files which are included in the output device list of a measurement as logical devices. These files are input files from the viewpoint of a traffic analysis and the process can be regarded as a transformation of the measurement entities into desired output information to the traffic analyst to aid in making various decisions.

For example, various criteria for dimensioning and verification of the grade of service could be produced by one or more analyses. A schematic picture of the flow of information is presented in Figure 1/E.503 as an activity diagram.

TABLE 1/E.503

Measurements basis \ Potential applications	Exchange dimensioning, planning and administration	Network dimensioning, planning and administration	Exchange performance monitoring	Network performance monitoring	Support to maintenance	Network management	Tariff and marketing studies
Overall traffic	X	X	X	X	X	X	
Circuit groups	X	X	X	X	X	X	
Subscriber line groups	X		X				
Auxiliary units	X		X		X		
Control units	X		X		X	X	
Common channel signalling	X	X	X	X	X	X	
Call records	X	X	X	X	X	X	X



a) The traffic values in the data record may have an effect on the internal functional steps.

b) There is a traffic data record for each individual measurement object which is included in the analysis. The past traffic values, and also calculated values, are used as input when updating the contents of the record at the time of a new traffic value.

FIGURE 1/E.503

Activity diagram of the information flows associated with traffic analysis

The following information is associated to each traffic analysis:

- identities of the related measurements;
- parameter values which are user-selectable to define the desired option or mode of the analysis;
- report dates of such report types for which the user must define the printout schedule;
- output devices for all report types.

4 Traffic analysis administration

4.1 In order to administer traffic analysis, the operator should perform a series of related activities and the system should support such activities by suitable system functions. Details are given below.

4.2 *List of tasks*

The following list of tasks is not intended to be complete; it aims to cover the operator's main activities in the area of traffic analysis administration:

a) to define parameter values in the parameter list of the analysis and to modify old values;

b) to define report dates for each type of report in a report date list as required and to modify it;

c) to define output routing for each type of report by an output routing list, as required, and to modify the dates;

d) to activate and/or deactivate the performance of the analysis;

e) to retrieve different kinds of information related to the existing traffic analysis;

f) to administer traffic data records of the measurement object which are included in the analysis.

4.3 *List of system functions*

The system should offer the following functions to support the jobs of the operator and the analysis itself:

a) transfer of the measured data to the analysis;

b) scheduling of various functions within the analysis, e.g. end-of-day calculation, report printout on report dates, etc.

c) management of traffic data records;

d) management of analysis description data;

e) transfer of the identification and capacity information of the measurement object to the analysis, e.g. title of a circuit group and the number of circuits assigned to it[1];

f) management of the printout of reports;

g) supervision control on the time delay of the various operations associated with the analysis.

4.4 *List of man-machine language (MML) functions*

Only a preliminary list of MML functions is presented below, and the complete specifications of such functions will appear in the Z-Series Recommendations:

- define analysis parameters;
- define a report date list;
- define an output routing list;
- administer traffic data records;
- activate a traffic analysis;
- deactivate a traffic analysis;
- interrogate a traffic analysis;
- interrogate a traffic analysis versus measurements;
- interrogate an output routing list;
- interrogate analysis parameters;
- interrogate a report date list.

Recommendation E.504

TRAFFIC MEASUREMENT ADMINISTRATION

1 Introduction

Traffic measurement administration includes the scheduling and control of traffic data collection, and production of reports for analysis. The data collected by means of traffic measurements performed by the exchange is output in a form suitable for on-line or deferred analysis.

It may be useful to consider the concept of a generic Traffic Measurement System (TMS) for purposes of administering traffic measurements. Such a system may comprise elements of an exchange working in conjunction with some combinations of remote data processors and associated devices for output of measurement reports.

In order to administer traffic measurements, a series of related man-machine activities (referred to as "tasks") will need to be performed through one or more man-machine interfaces, and supported by appropriate system functions. Details are given below.

The traffic measurement output should contain the measured data together with reference information about network conditions at the time of the measurement which would assist in the data analysis, for example the number of blocked devices on a route or temporary alternative routing in effect.

[1] All this information may or may not be available in the collection of the measured data.

2 List of tasks

The following list of tasks is not intended to be complete; however, it aims to cover the essential activities in the area of the traffic measurements administration. The TMS will provide functions to support these tasks:

a) to create new measurements or measurement components and to modify old ones, by selecting the measurement types, schedules, object identities and parameters of the measurements (WHAT, WHEN and HOW to measure);

b) to delete measurements or measurement components which are no longer useful;

c) to define output routing and scheduling of measurement results (WHEN and WHERE the result will be output);

d) to activate and/or to deactivate the scheduling of the measurements that have been previously defined;

e) to retrieve the required categories of data related to the existing measurements.

3 List of system functions

To support the man-machine tasks, the TMS should offer the following functions:

a) a menu of traffic measurements;

b) scheduling of traffic measurement execution and results output;

c) management of measurement's description data;

d) retrieving of measurement's description data.

4 Man-machine functions

A preliminary list of man-machine functions needed to control the TMS functions previously given is listed below; the complete specification of such functions appears in the Z-Series Recommendations:

- create a measurement;
- create a measurement set;
- create an object list;
- create a time data list;
- create an output routing list;
- create a results output schedule;
- modify a measurement;
- modify a measurement set;
- modify an object list;
- modify a time data list;
- modify an output routing list;
- modify a results output schedule;
- delete a measurement;
- delete a measurement set;
- delete an object list;
- delete a time data list;
- delete an output routing list;
- delete a results output schedule;
- activate a measurement;
- deactivate a measurement;
- interrogate a measurement;
- interrogate a measurement set;
- interrogate a measurement type;
- interrogate an object list;
- interrogate a time data list;
- interrogate an output routing list;
- interrogate a results output schedule.

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

SECTION 2

FORECASTING OF TRAFFIC

Recommendation E.506

FORECASTING INTERNATIONAL TRAFFIC[1)]

1 Introduction

This Recommendation is the first in a series of three Recommendations that cover international telecommunications forecasting.

In the operation and administration of the international telephone network, proper and successful development depends to a large degree upon estimates for the future. Accordingly, for the planning of equipment and circuit provision and of telephone plant investments, it is necessary that Administrations forecast the traffic which the network will carry. In view of the heavy capital investments in the international network, the economic importance of the most reliable forecast is evident.

The purpose of this Recommendation is to give guidance on some of the prerequisites for forecasting international telecommunications traffic. Base data, not only traffic and call data but also economic, social and demographic data, are of vital importance for forecasting. These data series may be incomplete; strategies are recommended for dealing with missing data. Different forecasting approaches are presented including direct and composite methods, matrix forecasting, and top down and bottom up procedures.

Recommendation E.507 provides guidelines for building forecasting models and contains an overview of various forecasting techniques. Recommendation E.508 covers the forecasting of new international telecommunications services.

2 Base data for forecasting

An output of the international traffic forecasting process is the estimated number of circuits required for each period in the forecast horizon. To obtain these values, traffic engineering techniques are applied to forecast Erlangs, a measure of traffic. Figure 1/E.506 outlines two different approaches for determining forecasted Erlangs.

The two different strategies for forecasting are the direct strategy and the composite strategy. The first step in either process is to collect raw data. These raw data, perhaps adjusted, will be the base data used to generate the traffic forecasts. Base data may be hourly, daily, monthly, quarterly, or annual. Most Administrations use monthly accounting data for forecasting purposes.

With the direct strategy, the traffic carried in Erlangs, or measured usage, for each relation would be regarded as the base data in forecasting traffic growth. These data may be adjusted to account for such occurrences as regeneration (see Recommendation E.500).

1) The old Recommendation E.506 which appeared in the *Red Book* was split into two Recommendations, revised E.506 and new E.507 and considerable new material was added to both.

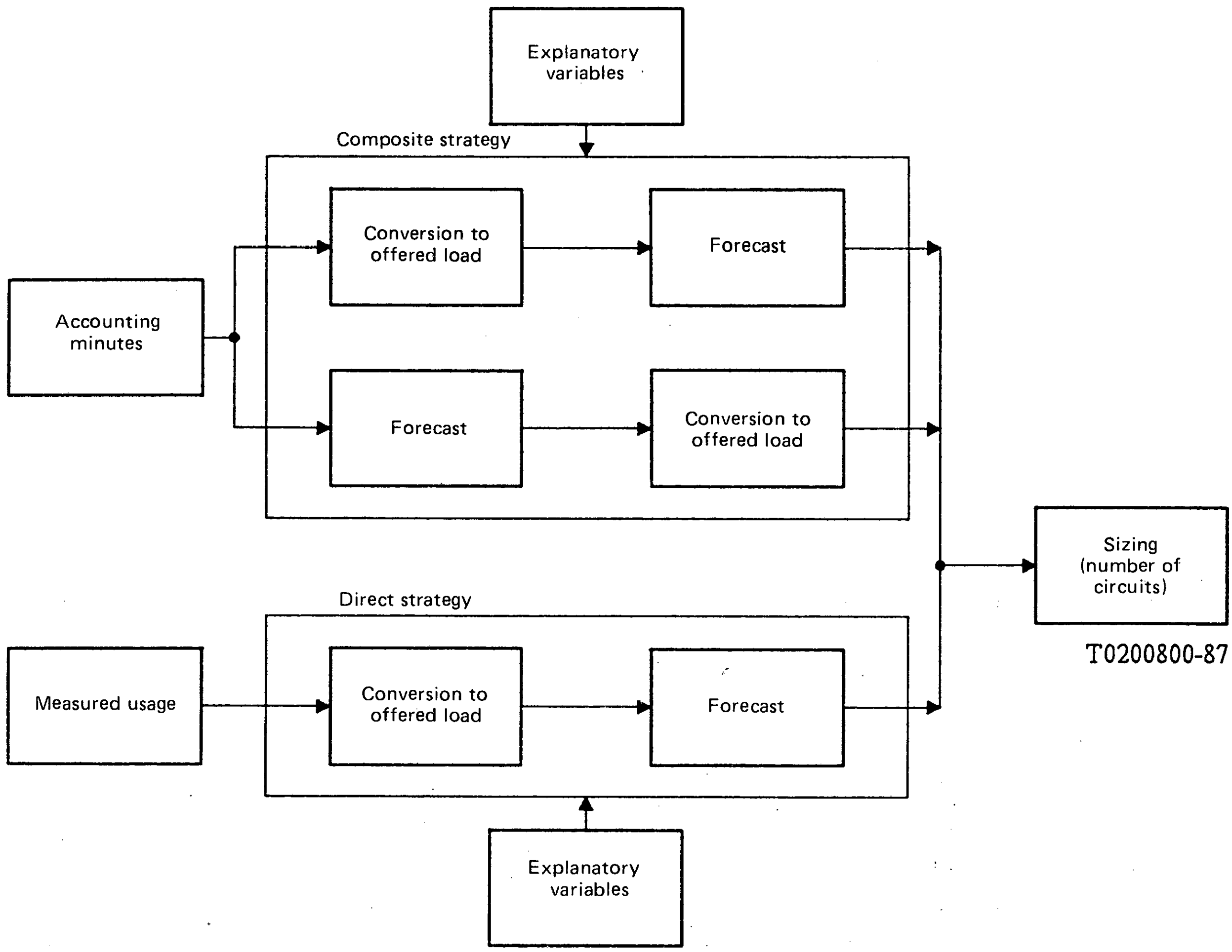


FIGURE 1/E.506

Direct and composite strategy

In both strategies (direct and composite) it is necessary to convert the carried traffic into offered traffic Erlangs. The conversion formula can be found in Recommendation E.501 for the direct strategy and in this Recommendation for the composite strategy.

Composite forecasting uses historical international accounting data of monthly paid minute traffic as the base data. The data may be adjusted by a number of factors, either before or after the forecasting process, that are used for converting paid minutes on the basis of the accounting data into busy-hour Erlang forecasts.

As seen in Figure 1/E.506, the forecasting process is common to both the direct and composite strategy. However, the actual methods or models used in the process vary. Forecasts can be generated, for example, using traffic matrix methods (see § 4), econometric models or autoregressive models (see § 3, Recommendation E.507). There are various other data that are input to the forecasting process. Examples of these are explanatory variables, market segmentation information and price elasticities.

Wherever possible, both the direct and composite forecasting strategies should be used and compared. This comparison may reveal irregularities not evident from the use of only one method. Where these are significant, in particular in the case of the busy hour, the causes for the differences should be identified before the resulting forecast is adopted.

In econometric modelling especially, explanatory variables are used to forecast international traffic. Some of the most important of these variables are the following:

- exports,
- imports,
- degree of automation,
- quality of service,
- time differences between countries,
- tariffs,
- consumer price index, and
- gross national product.

Other explanatory variables, such as foreign business travellers and nationals living in other countries, may also be important to consider. It is recommended that data bases for explanatory variables should be as comprehensive as possible to provide more information to the forecasting process.

Forecasts may be based on market segmentation. Base data may be segmented, for example, along regional lines, by business, non-business, or by type of service. Price elasticities should also be examined, if possible, to quantify the impact of tariffs on the forecasting data.

3 Composite strategy – Conversion method

The monthly paid-minutes traffic is converted to busy-hour Erlangs for dimensioning purposes by the application of a number of traffic related conversion factors for each service category. The conversion is carried out in accordance with the formula:

$$A = Mdh/60e \tag{3-1}$$

where

A is the estimated mean traffic in the busy hour,

M is the monthly paid-minutes,

d is day-to-month ratio,

h is the busy hour-to-day ratio, and

e is the efficiency factor.

The formula is described in detail in Annex A.

4 Procedures for traffic matrix forecasting

4.1 *Introduction*

To use traffic matrix or point-to-point forecasts the following procedures may be used:

- Direct point-to-point forecasts,
- Kruithof's method,
- Extension of Kruithof's method,
- Weighted least squares method.

It is also possible to develop a Kalman Filter model for point-to-point traffic taking into account the aggregated forecasts. Tu and Pack describe such a model in [16].

The forecasting procedures can be used to produce forecasts of internal traffic within groups of countries, for example, the Nordic countries. Another application is to produce forecasts for national traffic on various levels.

4.2 *Direct point-to-point forecasts*

It is possible to produce better forecasts for accumulated traffic than forecast of traffic on a lower level.

Hence, forecasts of outgoing traffic (row sum) or incoming traffic (column sum) between one country and a group of countries will give a relatively higher precision than the separate forecasts between countries.

In this situation it is possible to adjust the individual forecasts by taking into account the aggregated forecasts.

On the other hand, if the forecasts of the different elements in the traffic matrix turn out to be as good as the accumulated forecasts, then it is not necessary to adjust the forecasts.

Evaluation of the relative precision of forecasts may be carried out by comparing the ratios $\hat{\sigma}(X)/X$ where X is the forecast and $\hat{\sigma}(X)$ the estimated forecast error.

4.3 *Kruithof's method*

Kruithof's method [11] is well known. The method uses the last known traffic matrix and forecasts of the row and column sum to make forecasts of the traffic matrix. This is carried out by an efficient iteration procedure.

Kruithof's method does not take into account the change over time in the point-to-point traffic. Because Kruithof's method only uses the last known traffic matrix, information on the previous traffic matrices does not contribute to the forecasts. This would be disadvantageous. Especially when the growth of the distinct point-to-point traffic varies. Also when the traffic matrices reflect seasonal data, Kruithof's method may give poor forecasts.

4.4 *Extension of Kruithof's method*

The traditional Kruithof's method is a projection of the traffic based on the last known traffic matrix and forecasts of the row and column sums.

It is possible to extend Kruithof's method by taking into account not only forecasts of the row and column but also forecasts of point-to-point traffic. Kruithof's method is then used to adjust the point-to-point traffic forecasts to obtain consistency with the forecasts of row and column sums.

The extended Kruithof's method is superior to the traditional Kruithof's method and is therefore recommended.

4.5 *Weighted least squares method*

Weighted least squares method is again an extension of the last method. Let $\{C_{ij}\}$, $\{C_{i\cdot}\}$ and $\{C_{\cdot j}\}$ be forecasts of point-to-point traffic, row sums and column sums respectively.

The extended Kruithof's method assumes that the row and column sums are "true" and adjust $\{C_{ij}\}$ to obtain consistency.

The weighted least squares method [2] is based on the assumption that both the point-to-point forecasts and the row and column sum forecasts are uncertain. A reasonable way to solve the problem is to give the various forecasts different weights.

Let the weighted least squares forecasts be $\{D_{ij}\}$. The square sum Q is defined by:

$$Q = \sum_{ij} a_{ij} (C_{ij} - D_{ij})^2 + \sum_{i} b_i (C_{i\cdot} - D_{i\cdot})^2 + \sum_{j} c_j (C_{\cdot j} - D_{\cdot j})^2 \qquad (4\text{-}1)$$

where $\{a_{ij}\}$, $\{b_i\}$, $\{c_j\}$ are chosen constants or weights.

The weighted least squares forecast is found by:

$$\min_{D_{ij}} Q(D_{ij})$$

subject to

$$D_{i\cdot} = \sum_{j} D_{ij} \qquad i = 1, 2, \ldots \qquad (4\text{-}2)$$

and

$$D_{\cdot j} = \sum_{i} D_{ij} \qquad j = 1, 2, \ldots$$

A natural choice of weights is the inverse of the variance of the forecasts. One way to find an estimate of the standard deviation of the forecasts is to perform ex-post forecasting and then calculate the root mean square error.

The properties of this method are analyzed in [14].

5 Top down and bottom up methods

5.1 *Choice of model*

The object is to produce forecasts for the traffic between countries. For this to be a sensible procedure, it is necessary that the traffic between the countries should not be too small, so that the forecasts may be accurate. A method of this type is usually denoted as "bottom up".

Alternatively, when there is a small amount of traffic between the countries in question, it is better to start out with forecasting the traffic for a larger group of countries. These forecasts are often used as a basis for forecasts for the traffic to each country. This is done by a correction procedure to be described in more detail below. Methods of this type are called "top down". The following comments concern the preference of one method over another.

Let σ_T^2 be the variance of the aggregated forecast, and σ_i^2 be the variance of the local forecast No. i and γ_{ij} be the covariance of the local forecast No. i and j. If the following inequality is true:

$$\sigma_T^2 < \sum_i \sigma_i^2 + \sum_{i \neq j} \sum \gamma_{ij} \tag{5-1}$$

then, in general, it is not recommended to use the bottom up method, but to use the top down method.

In many situations it is possible to use a more advanced forecasting model on the aggregated level. Also, the data on an aggregated level may be more consistent and less influenced by stochastic changes compared to data on a lower level. Hence, in most cases the inequality stated above will be satisfied for small countries.

5.2 *Bottom up method*

As outlined in § 5.1 the bottom up method is defined as a procedure for making separate forecasts of the traffic between different countries directly. If the inequality given in § 5.1 is not satisfied, which may be the case for large countries, it is sufficient to use the bottom up method. Hence, one of the forecasting models mentioned in Recommendation E.507 can be used to produce traffic forecasts for different countries.

5.3 *Top down procedure*

In most cases the top down procedure is recommended for producing forecasts of international traffic for a small country. In Annex D a detailed example of such a forecasting procedure is given.

The first step in the procedure is to find a forecasting model on the aggregated level, which may be a rather sophisticated model. Let X_T be the traffic forecasts on the aggregated level and σ_T the estimated standard deviation of the forecasts.

The next step is to develop separate forecasting models of traffic to different countries. Let X_i be the traffic forecast to the i^{th} country and $\hat{\sigma}_i$ the standard deviation. Now, the separate forecasts $[X_i]$ have to be corrected by taking into account the aggregated forecasts X_T. We know that in general

$$X_T \neq \sum_i X_i \tag{5-2}$$

Let the corrections of $[X_i]$ be $[X'_i]$, and the corrected aggregated forecast then be $X'_T = \Sigma X'_i$.

The procedure for finding $[X'_i]$ is described in Annex C.

6 Forecasting methods when observations are missing

6.1 *Introduction*

Most forecasting models are based on equally spaced time series. If one observation or a set of observations are missing, it is necessary either to use an estimate of missing observations and then use the forecasting model or to modify the forecasting model.

All smoothing models are applied on equally spaced observations. Also autoregressive integrated moving average (ARIMA)-models operate on equally spaced time series, while regression models work on irregularly spaced observations without modifications.

In the literature it is shown that most forecasting methods can be formulated as dynamic linear models (DLM). The Kalman Filter is a linear method to estimate states in a time series which is modelled as a dynamic linear model. The Kalman Filter introduces a recursive procedure to calculate the forecasts in a DLM which is optimal in the sense of minimizing the mean squared one step ahead forecast error. The Kalman Filter also gives an optimal solution in the case of missing data.

6.2 *Adjustment procedure based on comparable observations*

In situations when some observations are missing, it may be possible to use related data for estimating the missing observations. For instance, if measurements are carried out on a set of trunk groups in the same area, then the traffic measurements on various trunk groups are correlated, which means that traffic measurements on a given trunk group to a certain degree explain traffic measurements on other trunk groups.

When there is high correlation between two time series of traffic measurements, the relative change in level and trend will be of the same size.

Suppose that a time series x_t of equidistant observations from 1 to n has an inside gap · x_t is, for instance, the yearly increase. The gap consists of k missing observations between r and $r + k + 1$.

A procedure for estimating the missing observations is given by the following steps:

i) Examine similar time series to the series with missing observations and calculate the cross correlation.

ii) Identify time series with high cross correlation at lag zero.

iii) Calculate the growth factor Δ_{r+i} between r and $r + k$ of the similar time series y_t:

$$\Delta_{r+i} = \frac{y_{r+i} - y_r}{y_{r+k+1} - y_r} \qquad i = 1, 2, \ldots k \tag{6-1}$$

iv) Estimates of the missing observations are then given by:

$$\hat{x}_{r+i} = x_r + \Delta_{r+i} (x_{r+k+1} - x_r) \qquad i = 1, 2, \ldots k \tag{6-2}$$

Example

Suppose we want to forecast the time series x_t. The series is observed from 1 to 10, but the observations at time 6, 7 and 8 are missing. However a related time series y_t is measured. The measurements are given in Table 1/506.

TABLE 1/E.506

Measurements of two related time series; one with missing observations

t	1	2	3	4	5	6	7	8	9	10
x_t	100	112	125	140	152	–	–	–	206	221
y_t	300	338	380	422	460	496	532	574	622	670

The last known observation of x_t before the gap at time 5 is 152, while the first known observation after the gap at time 9 is 206.

Hence $r = 5$ and $k = 3$. The calculation gives:

$$\Delta_6 = \frac{496 - 460}{622 - 460} = \frac{36}{162}$$

$$\Delta_7 = \frac{532 - 460}{622 - 460} = \frac{72}{162}$$

$$\Delta_8 = \frac{574 - 460}{622 - 460} = \frac{114}{162}$$

$$\hat{x}_6 = 152 + \frac{36}{162} (206 - 152) = \underline{164}$$

$$\hat{x}_7 = 152 + \frac{72}{162} (206 - 152) = \underline{176}$$

$$\hat{x}_8 = 152 + \frac{114}{162} (206 - 152) = \underline{190}$$

6.3 *Modification of forecasting models*

The other possibility for handling missing observations is to extend the forecasting models with specific procedures. When observations are missing, a modified procedure, instead of the ordinary forecasting model, is used to estimate the traffic.

To illustrate such a procedure we look at simple exponential smoothing. The simple exponential smoothing model is expressed by:

$$\hat{\mu}_t = (1 - a)\, y_t + a\hat{\mu}_{t-1} \tag{6-3}$$

where

y_t is the measured traffic at time t

$\hat{\mu}_t$ is the estimated level at time t

a is the discount factor [and $(1 - a)$ is the smoothing parameter].

Equation (6-3) is a recursive formula. The recursion starts at time 1 and ends at n if no observation is missing. Then a one step ahead forecast is given by:

$$\hat{y}_t(1) = \hat{\mu}_t \tag{6-4}$$

If some observations lying in between 1 and n are missing, then it is necessary to modify the recursion procedure. Suppose now that $y_1, y_2, \ldots, y_r, y_{r+k+1}, y_{r+k+2}, \ldots, y_n$ are known and $y_{r+1}, y_{r+2}, \ldots, y_{r+k}$ are unknown. Then the time series has a gap consisting of k missing observations.

The following modified forecasting model for simple exponential smoothing is proposed in Aldrin [2].

$$\hat{\mu}_t = \begin{cases} (1 - a)\, y_t + a\, \hat{\mu}_{t-1} & t = 1, 2, \ldots, r \\ (1 - a_k)\, y_t + a_k \hat{\mu}_t & t = r+k+1 \\ (1 - a)\, y_t + a\, \hat{\mu}_{t-1} & t = r+k+2, \ldots, n \end{cases} \tag{6-5}$$

where

$$a_k = \frac{a}{1 + k(1-a)^2} \tag{6-6}$$

By using the (6-5) and (6-6) it is possible to skip the recursive procedure in the gap between r and $r + k + 1$.

In Aldrin [2] similar procedures are proposed for the following forecasting models:

- Holt's method,
- Double exponential smoothing,
- Discounted least squares method with level and trend,
- Holt-Winters seasonal methods.

Wright [17] and [18] also suggests specific procedures to modify the smoothing models when observations are missing.

As mentioned in the first paragraph, regression models are invariant of missing observations. When using the least squares method, all observations are given the same weight. Hence, missing observations do not affect the estimation procedure and forecast are made in the usual way.

On the other hand it is necessary to modify ARIMA models when observations are missing. In the literature several procedures are suggested in the presence of missing data. The basic idea is to formulate the ARIMA model as a dynamic linear model. Then the likelihood function is easy to obtain and the parameters in the model can be estimated recursively. References to work on this field are Jones [9] and [10], Harvey and Pierse [8], Ansley and Kohn [3] and Aldrin [2].

State space models or dynamic linear models and the Kalman Filter are a large class of models. Smoothing models, ARIMA models and regression models may be formulated as dynamic linear models. This is shown, for instance, in Abraham and Ledolter [1]. Using dynamic linear models and the Kalman Filter the parameters in the model are estimated in a recursive way. The description is given, for instance, in Harrison and Stevens [7], Pack and Whitaker [13], Moreland [12], Szelag [15] and Chemouil and Garnier [6].

In Jones [9] and [10], Barham and Dunstan [4], Harvey and Pierse [8], Aldrin [2] and Bølviken [5] it is shown how the dynamic linear models and the Kalman Filter handle missing observations.

ANNEX A

(to Recommendation E.506)

Composite strategy

A.1 *Introduction*

This annex describes a method for estimating international traffic based on monthly paid-minutes and a number of conversion factors. It demonstrates the method by examining the factors and showing their utility.

The method is seen to have two main features:

1) Monthly paid-minutes exchanged continuously between Administrations for accounting purposes provide a large and continuous volume of data.
2) Traffic conversion factors are relatively stable, when compared with traffic growth and change slowly since they are governed by customers' habits and network performance. By separately considering the paid minutes and the traffic conversion factors, we gain an insight into the nature of traffic growth which cannot be obtained by circuit occupancy measurements alone. Because of the stability of the conversion factors, these may be measured using relatively small samples, thus contributing to the economy of the procedure.

A.2 *Basic procedure*

A.2.1 *General*

The composite strategy is carried out for each stream, for each direction and generally for each service category.

The estimated mean offered busy-hour traffic (in Erlangs) is derived from the monthly paid-minutes using the formula:

$$A = Mdh/60e \tag{A-1}$$

where

A is the estimated mean traffic in Erlangs offered in the busy hour,

M is the total monthly paid-minutes,

d is the day/month ratio, i.e. the ratio of average weekday paid-time to monthly paid-time,

h is the busy-hour/day ratio, i.e. the ratio of the busy-hour paid-time to the average daily paid-time,

e is the efficiency factor, i.e. the ratio of busy-hour paid-time to busy-hour occupied-time.

A.2.2 *Monthly paid-minutes (M)*[2)]

The starting point for the composite strategy is paid minutes. Sudden changes in subscriber demand, for example, resulting from improvements in transmission quality, have a time constant of the order of several months, and on this basis paid minutes accumulated over monthly intervals appear to be optimum in terms of monitoring traffic growth. A longer period (e.g. annually) tends to mask significant changes, whereas a shorter period (e.g. daily) not only increases the amount of data, but also increases the magnitude of fluctuations from one period to the next. A further advantage of the one-month period is that monthly paid-minute figures are exchanged between Administrations for accounting purposes and consequently historical records covering many years are normally readily available.

It should be recognized, however, that accounting information exchanges between Administrations often take place after the event, and it may take some time to reach full adjustments (e.g. collect call traffic).

A.2.3 *Day/month ratio (d)*

This ratio is related to the amount of traffic carried on a typical weekday compared with the total amount of traffic carried in a month.

As the number of weekdays and non-weekdays (weekends and holidays) varies month by month, it is not convenient to refer to a typical month, but it should be possible to compute the ratio for the month for which the busy hour traffic is relevant.

Hence if:

X denotes the number of weekdays in the related month

Y denotes the number of non-weekdays (weekend days and holidays) in the selected month, then

$$\frac{1}{d} = X + Y \cdot r \tag{A-2}$$

where

$$r = \frac{\text{Average non-weekday traffic}}{\text{Average weekday traffic}}$$

The relative amount of non-weekday traffic is very sensitive to the relative amount of social contact between origin and destination. (Social calls, are, in general, made more frequently on weekends.) Since changes in such social contact would be very slow, r or d are expected to be the most stable conversion factors, which in general vary only within relatively narrow limits. However, tariff policies such as reduced weekend rates can have a significant effect on r and d.

2) In a situation where only yearly paid-minutes are available, this may be converted to M by a suitable factor.

When r is in the region of 1, the Sunday traffic may exceed the typical weekday level. If this is the case, consideration should be given to dimensioning the route to cater for the additional weekend (Sunday) traffic or adopting a suitable overflow routing arrangement.

A.2.4 *Busy-hour/day ratio (h)*

The relative amount of average weekday traffic in the busy hour primarily depends on the difference between the local time at origin and destination. Moderately successful attempts have been made to predict the diurnal distribution of traffic based on this information together with supposed "degree of convenience" at origin and destination. However, sufficient discrepancies exist to warrant measuring the diurnal distribution, from which the busy-hour/day ratio may be calculated.

Where measurement data is not available, a good starting point is Recommendation E.523. From the theoretical distributions found in Recommendation E.523, one finds variations in the busy-hour/day ratio from 10% for 0 to 2 hours time difference and up to 13.5% for 7 hours time difference.

As described above, the composite strategy is implemented as an accounting-based procedure. However, it may be more practical for some Administrations to measure d and h based on occupied time, derived from available call recording equipment.

A.2.5 *Efficiency factor (e)*

The efficiency factor (ratio of busy-hour paid time to busy-hour occupied time, e) converts the paid time into a measure of total circuit occupancy. It is therefore necessary to include all occupied circuit time in the measurement of this ratio, and not merely circuit time taken up in establishing paid calls. For example, the measurement of total circuit occupied time should include the occupied time for paid calls (time from circuit seizure to circuit clearance) and, in addition, the occupied time for directory inquiry calls, test calls, service calls, ineffective attempts and other classes of unpaid traffic handled during the busy hour.

There is a tendency for the efficiency to change with time. In this regard, efficiency is mainly a function of operating method (manual, semi-automatic, international subscriber dialling), the B-subscriber's availability, and the quality of the distant network.

Forecasts of the efficiency can be made on the basis of extrapolation of past trends together with adjustments for planned improvements.

The detailed consideration of efficiency is also an advantage from an operational viewpoint in that it may be possible to identify improvements that may be made, and quantify the benefits deriving from such improvements.

It should be noted that the practical limit for e is generally about 0.8 to 0.9 for automatic working.

A.2.6 *Mean offered busy hour traffic (A)*

It should be noted that A is the mean offered busy-hour traffic expressed in Erlangs.

A.2.7 *Use of composite strategy*

In the case of countries with lower traffic volumes and manual operation, the paid-time factors (d and h) would be available from analysis of call vouchers (dockets). For derivation of the efficiency e, the manual operator would have to log the busy-hour occupied time as well as the paid time during the sampling period.

In countries using stored-program controlled exchanges with associated manual assistance positions, computer analysis may aid the composite forecasting procedure.

One consequence of the procedure is that the factors d and h give a picture of subscriber behaviour, in that unpaid time (inquiry calls, test calls, service calls, etc.) are not included in the measurement of these factors. The importance of deriving the efficiency, e, during the busy hour, should also be emphasized.

ANNEX B

(to Recommendation E.506)

Example using weighted least squares method

B.1 *Telex data*

The telex traffic between the following countries has been analyzed:

- Germany (D)
- Denmark (DNK)
- USA (USA)
- Finland (FIN)
- Norway (NOR)
- Sweden (S)

The data consists of yearly observations from 1973 to 1984 [19].

B.2 *Forecasting*

Before using the weighted least squares method, separate forecasts for the traffic matrix have to be made. In this example a simple ARIMA (0,2,1) model with logarithmic transformed observations without explanatory variables is used for forecasting. It may be possible to develop better forecasting models for the telex traffic between the various countries. However the main point in this example only is to illustrate the use of the weighted least squares technique.

Forecasts for 1984 based on observations from 1973 to 1983 are given in Table B-1/E.506.

TABLE B-1/E.506

Forecasts for telex traffic between Germany (D), Denmark (DNK), USA (USA), Finland (FIN), Norway (NOR) and Sweden (S) in 1984

To \ From	D	DNK	USA	FIN	NOR	S	Sum	Forecasted sum
D	–	4869	12 630	2879	2397	5230	28 005	27 788
DNK	5196	–	1655	751	1270	1959	10 831	10 805
USA	11 103	1313	–	719	1657	2401	17 193	17 009
FIN	2655	715	741	–	489	1896	6496	6458
NOR	2415	1255	1821	541	–	1548	7580	7597
S	4828	1821	2283	1798	1333	–	12 063	12 053
Sum	26 197	9973	19 130	6688	7146	13 034		
Forecasted sum	26 097	9967	19 353	6659	7110	12.914		

It should be noticed that there is no consistency between row and column sum forecasts and forecasts of the elements in the traffic matrix. For instance, the sum of forecasted outgoing telex traffic from Germany is 28 005, while the forecasted row sum is 27 788.

To adjust the forecasts to get consistency and to utilize both row/column forecasts and forecasts of the traffic elements the weighted least squares method is used.

B.3 *Adjustment of the traffic matrix forecasts*

To be able to use the weighted least squares method, the weights and the separate forecasts are needed as input. The separate forecasts are found in Table B-2/E.506, while the weights are based on the mean squared one step ahead forecasting errors.

Let y_t be the traffic at time t. The ARIMA (0,2,1) model with logarithmic transformed data is given by:

$$z_t = (1 - B)^2 \ln y_t = (1 - \theta B)\, a_t$$

or

$$z_t = a_t - \theta a_{t-1}$$

where

$z_t = \ln y_t - 2 \ln y_{t-1} + \ln y_{t-2}$

a_t is white noise,

θ is a parameter,

B is the backwards shift operator.

The mean squared one step ahead forecasting error of z_t is:

$$MSQ = \frac{1}{n} \Sigma \left(z_t - \hat{z}_{t-1}(1)\right)^2$$

where

$\hat{z}_{t-1}(1)$ is the one step ahead forecast.

The results of using the weighted least squares method is found in Table B-3/E.506 and show that the factors in Table B-1/E.506 have been adjusted. In this example only minor changes have been performed because of the high conformity in the forecasts of row/column sums and traffic elements.

TABLE B-2/E.506

Inverse weights as mean as squared one step ahead forecasting errors of telex traffic (100^{-4}) between Germany (D), Denmark (DNK), USA (USA), Finland (FIN), Norway (NOR) and Sweden (S) in 1984

To \ From	D	DNK	USA	FIN	NOR	S	Sum
D	–	28.72	13.18	11.40	8.29	44.61	7.77
DNK	5.91	–	43.14	18.28	39.99	18.40	10.61
USA	23.76	39.19	–	42.07	50.72	51.55	21.27
FIN	23.05	12.15	99.08	–	34.41	19.96	17.46
NOR	21.47	40.16	132.57	24.64	–	17.15	20.56
S	6.38	12.95	28.60	28.08	8.76	–	6.48
Sum	6.15	3.85	14.27	9.55	12.94	8.53	

TABLE B-3/E.506

Adjusted telex forecasts using the weighted least squares method

To \ From	D	DNK	USA	FIN	NOR	S	Sum
D	–	4850	12 684	2858	2383	5090	27 865
DNK	5185	–	1674	750	1257	1959	10 825
USA	11 001	1321	–	717	1644	2407	17 090
FIN	2633	715	745	–	487	1891	6471
NOR	2402	1258	1870	540	–	1547	7617
S	4823	1817	2307	1788	1331	–	12 066
Sum	26 044	9961	19 280	6653	7102	12 894	

ANNEX C

(to Recommendation E.506)

Description of a top down procedure

Let

X_T be the traffic forecast on an aggregated level,

X_i be the traffic forecast to country i,

$\hat{\sigma}_T$ the estimated standard deviation of the aggregated forecast,

$\hat{\sigma}_i$ the estimated standard deviation of the forecast to country i.

Usually

$$X_T \neq \sum_i X_i, \qquad \text{(C-1)}$$

so that it is necessary to find a correction

$$[X'_i] \text{ of } [X_i] \text{ and } [X'_T] \text{ of } [X_T]$$

by minimizing the expression

$$Q = \alpha_0(X_T - X'_T)^2 + \sum_i \alpha_i(X_i - X'_i)^2 \qquad \text{(C-2)}$$

subject to

$$X'_T = \sum_i X'_i \qquad \text{(C-3)}$$

where α and $[\alpha_i]$ are chosen to be

$$\alpha_0 = \frac{1}{\hat{\sigma}^2_T} \text{ and } \alpha_i = \frac{1}{\hat{\sigma}^2_i} \qquad i = 1, 2, \ldots \qquad \text{(C-4)}$$

The solution of the optimization problem gives the values $[X'_i]$:

$$X'_i = X_i - \hat{\sigma}_i^2 \frac{\sum_i X_i - X_T}{\sum_i \hat{\sigma}_i^2 + \hat{\sigma}_T^2} \qquad \text{(C-5)}$$

A closer inspection of the data base may result in other expressions for the coefficients $[\alpha_i]$, $i = 0, 1, \dots$ On some occasions, it will also be reasonable to use other criteria for finding the corrected forecasting values $[X'_i]$. This is shown in the top down example in Annex D.

If, on the other hand, the variance of the top forecast X_T is fairly small, the following procedure may be chosen:

The corrections $[X_i]$ are found by minimizing the expression

$$Q' = \sum_i \alpha_i (X_i - X'_i)^2 \qquad \text{(C-6)}$$

subject to

$$X_T = \sum X'_i \qquad \text{(C-7)}$$

If α_i, $i = 1, 2, \dots$ is chosen to be the inverse of the estimated variances, the solution of the optimization problem is given by

$$X'_i = X_i - \hat{\sigma}_i^2 \frac{\sum X_i - X_T}{\sum \hat{\sigma}_i^2} \qquad \text{(C-8)}$$

ANNEX D

(to Recommendation E.506)

Example of a top down modelling method

The model for forecasting telephone traffic from Norway to the European countries is divided into two separate parts. The first step is an econometric model for the total traffic from Norway to Europe. Thereafter, we apply a model for the breakdown of the total traffic on each country.

D.1 *Econometric model of the total traffic from Norway to Europe*

With an econometric model we try to explain the development in telephone traffic, measured in charged minutes, as a function of the main explanatory variables. Because of the lack of data for some variables, such as tourism, these variables have had to be omitted in the model.

The general model may be written:

$$X_t = e^K \cdot GNP_t^a \cdot P_t^b \cdot A_t^c \cdot e^{u_t} \qquad (t = 1, 2, \dots, N) \qquad \text{(D-1)}$$

where:

X_t is the demand for telephone traffic from Norway to Europe at time t (charged minutes).

GNP_t is the gross national product in Norway at time t (real prices).

P_t is the index of charges for traffic from Norway to Europe at time t (real prices).

A_t is the percentage direct-dialled telephone traffic from Norway to Europe (to take account of the effect of automation). For statistical reasons (i.e. impossibility of taking logarithm of zero) A_t goes from 1 to 2 instead of from 0 to 1.

K is the constant.

a is the elasticity with respect to *GNP*.

b is the price elasticity.

c is the elasticity with respect to automation.

u_t is the stochastic variable, summarizing the impact of those variables that are not explicitly introduced in the model and whose effects tend to compensate each other (expectation of $u_t = 0$ and var $u_t = \sigma^2$).

By applying regression analysis (OLSQ) we have arrived at the coefficients (elasticities) in the forecasting model for telephone traffic from Norway to Europe given in Table D-1/E.506 (in our calculations we have used data for the period 1951-1980).

The t statistics should be compared with the Student's Distribution with $N - d$ degrees of freedom, where N is the number of observations and d is the number of estimated parameters. In this example, $N = 30$ and $d = 4$.

The model "explains" 99.7% of the variation in the demand for telephone traffic from Norway to Europe in the period 1951-1980.

From this logarithmic model it can be seen that:

- an increase in GNP of 1% causes an increase in the telephone traffic of 2.80%,
- an increase of 1% in the charges, measured in real prices, causes a decrease in the telephone traffic of 0.26%, and
- an increase of 1% in A_t causes an increase in the traffic of 0.29%.

We now use the expected future development in charges to Europe, in GNP, and in the future automation of traffic to Europe to forecast the development in telephone traffic from Norway to Europe from the equation:

$$X_t = e_t^{-16.095} \cdot GNP_t^{2.80} \cdot P_t u^{-0.26} \cdot A_t^{0.29} \tag{D-2}$$

TABLE D-1/E.506

Coefficients	Estimated values	t statistics
K	−16.095	−4.2
a	2.799	8.2
b	− 0.264	−1.0
c	0.290	2.1

D.2 *Model for breakdown of the total traffic from Norway to Europe*

The method of breakdown is first to apply the trend to forecast the traffic to each country. However, we let the trend become less important the further into the period of forecast we are, i.e. we let the trend for each country converge to the increase in the total traffic to Europe. Secondly, the traffic to each country is adjusted up or down, by a percentage that is equal to all countries, so that the sum of the traffic to each country equals the forecasted total traffic to Europe from equation (D-2).

Mathematically, the breakdown model can be expressed as follows:

Calculation of the trend for country i:

$$R_{it} = b_i + a_i \cdot t, \qquad i = 1, \ldots, 34 \qquad t = 1, \ldots, N \qquad \text{(D-3)}$$

where

$R_{it} = \frac{X_{it}}{X_t}$, i.e country i's share of the total traffic to Europe.

X_{it} is the traffic to country i at time t

X_t is the traffic to Europe at time t

t is the trend variable

a_i and b_i are two coefficients specific to country i; i.e. a_i is country i's trend. The coefficients are estimated by using regression analysis, and we have based calculations on observed traffic for the period 1966-1980.

The forecasted shares for country i is then calculated by

$$R_{it} = R_{iN} + a_i \cdot (t - N) \cdot e^{-\frac{t-5}{40}} \qquad \text{(D-4)}$$

where N is the last year of observation, and e is the exponential function.

The factor $e^{-\frac{t-5}{40}}$ is a correcting factor which ensures that the growth in the telephone traffic to each country will converge towards the growth of total traffic to Europe after the adjustment made in Equation (D-6).

To have the sum of the countries' shares equal one, it is necessary that

$$\sum_i R_{it} = 1 \qquad \text{(D-5)}$$

This we obtain by setting the adjusted share, $\tilde{R}_{it}$, equal to

$$\tilde{R}_{it} = R_{it} \frac{1}{\sum_i R_{it}} \qquad \text{(D-6)}$$

Each country's forecast traffic is then calculated by multiplying the total traffic to Europe, X_t, by each country's share of the total traffic:

$$X_{it} = \tilde{R}_{it} \times X_t \qquad \text{(D-7)}$$

D.3 *Econometric model for telephone traffic from Norway to Central and South America, Africa, Asia, and Oceania.*

For telephone traffic from Norway to these continents we have used the same explanatory variables and estimated coefficients. Instead of gross national product, our analysis has shown that for the traffic to these continents the number of telephone stations within each continent are a better and more significant explanatory variable.

After using cross-section/time-series simultaneous estimation we have arrived at the coefficients in Table D-2/E.506 for the forecasting model for telephone traffic from Norway to these continents (for each continent we have based our calculations on data for the period 1961-1980):

TABLE D-2/E.506

Coefficients	Estimated values	t statistics
Charges	– 1.930	– 5.5
Telephone stations	2.009	4.2
Automation	0.5	–

We then have $R^2 = 0.96$. The model may be written:

$$X_t^k = e^K \cdot (TS_t^k)^{2.009} \cdot (P_t^k)^{1.930} \cdot (A_t^k)^{0.5} \tag{D-8}$$

where

X_t^k is the telephone traffic to continent k (k = Central America, . . ., Oceania) at time t,

e^K is the constant specific to each continent. For telephone traffic from Norway to:

Central America: $K^1 = -11.025$
South America: $K^2 = -12.62$
Africa: $K^3 = -11.395$
Asia: $K^4 = -15.02$
Oceania: $K^5 = -13.194$

TS_t^k is the number of telephone stations within continent k at time t,

P_t^k is the index of charges, measured in real prices, to continent k at time t, and

A_t^k is the percentage direct-dialled telephone traffic to continent k.

Equation (D-8) is now used – together with the expected future development in charges to each continent, future development in telephone stations on each continent and future development in automation of telephone traffic from Norway to the continent – to forecast the future development in telephone traffic from Norway to the continent.

References

[1] ABRAHAM (A.) and LEDOLTER (J.): Statistical methods for forecasting. *J. Wiley*, New York, 1983.

[2] ALDRIN (M.): Forecasting time series with missing observations. *Stat 15/86 Norwegian Computing Center*, 1986.

[3] ANSLEY (C. F.) and KOHN (R.): Estimation, filtering and smoothing in state space models with incomplete specified initial conditions. *The Annals of Statistics*, 13, pp. 1286-1316, 1985.

[4] BARHAM (S. Y.) and DUNSTAN (F. D. J.): Missing values in time series. *Time Series Analysis: Theory and Practice 2*: Anderson, O. D., ed., pp. 25-41, North Holland, Amsterdam, 1982.

[5] BØLVIKEN (E.): Forecasting telephone traffic using Kalman Filtering: Theoretical considerations. *Stat 5/86 Norwegian Computing Center*, 1986.

[6] CHEMOUIL (P.) and GARNIER (B.): An adaptive short-term traffic forecasting procedure using Kalman Filtering. *XI International Teletraffic Congress*, Kyoto, 1985.

[7] HARRISON (P. J.) and STEVENS (C. F.): Bayesian forecasting. *Journal of Royal Statistical Society*. Ser B 37, pp. 205-228, 1976.

[8] HARVEY (A. C.) and PIERSE (R. G.): Estimating missing observations in econometric time series. *Journal of American Statistical As.*, 79, pp. 125-131, 1984.

[9] JONES (R. H.): Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics*, 22, No. 3, pp. 389-396, 1980.

[10] JONES (R. H.): Time series with unequally spaced data. *Handbook of Statistics 5*. ed. Hannah, E. J., et al., pp. 157-177, North Holland, Amsterdam, 1985.

[11] KRUITHOF (J.): Telefoonverkeersrekening. *De Ingenieur*, 52, No. 8, 1937.

[12] MORELAND (J. P.): A robust sequential projection algorithm for traffic load forecasting. *The Bell Technical Journal*, 61, pp. 15-38, 1982.

[13] PACK (C. D.) and WHITAKER (B. A.): Kalman Filter models for network forecasting. *The Bell Technical Journal*, 61, pp. 1-14, 1982.

[14] STORDAHL (K.) and HOLDEN (L.): Traffic forecasting models based on top down and bottom up models. *ITC 11*, Kyoto, 1985.

[15] SZELAG (C. R.): A short-term forecasting algorithm for trunk demand servicing. *The Bell Technical Journal*, 61, pp. 67-96, 1982.

[16] TU (M.) and PACK (D.): Improved forecasts for local telecommunications network. 6th International Forecasting Symposium, Paris, 1986.

[17] WRIGHT (D. H.): Forecasting irregularly spaced data: An extension of double exponential smoothing. *Computer and Engineering*, 10, pp. 135-147, 1986.

[18] WRIGHT (D. H.): Forecasting data published at irregular time intervals using an extension of Holt's method. *Management science*, 32, pp. 499-510, 1986.

[19] *Table of international telex relations and traffic*, ITU, Geneva, 1973-1984.

Recommendation E.507 [1)]

MODELS FOR FORECASTING INTERNATIONAL TRAFFIC

1 Introduction

Econometric and time series model development and forecasting requires familiarity with methods and techniques to deal with a range of different situations. Thus, the purpose of this Recommendation is to present some of the basic ideas and leave the explanation of the details to the publications cited in the reference list. As such, this Recommendation is not intended to be a complete guide to econometric and time series modelling and forecasting.

The Recommendation also gives guidelines for building various forecasting models: identification of the model, inclusion of explanatory variables, adjustment for irregularities, estimation of parameters, diagnostic checks, etc.

In addition the Recommendation describes various methods for evaluation of forecasting models and choice of model.

2 Building the forecasting model

This procedure can conveniently be described as four consecutive steps. The first step consists in finding a useful class of models to describe the actual situation. Examples of such classes are simple models, smoothing models, autoregressive models, autoregressive integrated moving average (ARIMA) models or econometric models. Before choosing the class of models, the influence of external variables should be analyzed. If special external variables have significant impact on the traffic demand, one ought to include them in the forecasting models, provided enough historical data are available.

The next step is to identify one tentative model in the class of models which have been chosen. If the class is too extensive to be conveniently fitted directly to data, rough methods for identifying subclasses can be used. Such methods of model identification employ data and knowledge of the system to suggest an appropriate parsimonious subclass of models. The identification procedure may also, in some occasions, be used to yield rough preliminary estimates of the parameters in the model. Then the tentative model is fitted to data by estimating the parameters. Usually, maximum likelihood estimators or least square estimators are used.

The next step is to check the model. This procedure is often called diagnostic checking. The object is to find out how well the model fits the data and, in case the discrepancy is judged to be too severe, to indicate possible remedies. The outcome of this step may thus be acceptance of the model if the fit is acceptable. If on the other hand it is inadequate, it is an indication that new tentative models may in turn be estimated and subjected to diagnostic checking.

---

1) The old Recommendation E.506 which appeared in the *Red Book* was split into two Recommendations, revised E.506 and new E.507, and considerable new material was added to both.

In Figure 1/E.507 the steps in the model building procedure are illustrated.

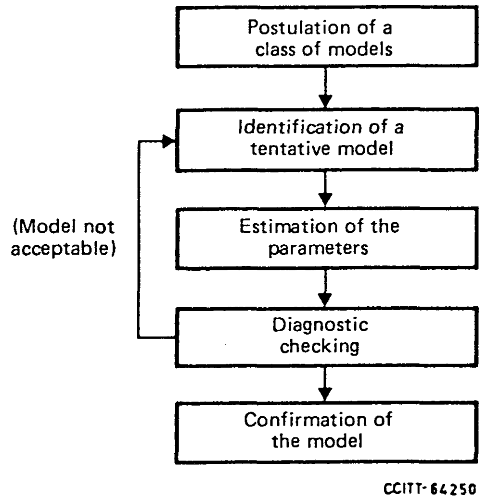


FIGURE 1/E.507

Steps in the model building procedure

3 Various forecasting models

The objective of § 3 is to give a brief overview of the most important forecasting models. In the GAS 10 Manual on planning data and forecasting methods [5], a more detailed description of the models is given.

3.1 *Curve fitting models*

In curve fitting models the traffic trend is extrapolated by calculating the values of the parameters of some function that is expected to characterize the growth of international traffic over time. The numerical calculations of some curve fitting models can be performed by using the least squares method.

The following are examples of common curve fitting models used for forecasting international traffic:

Linear: $Y_t = a + bt$ (3-1)

Parabolic: $Y_t = a + bt + ct^2$ (3-2)

Exponential: $Y_t = ae^{bt}$ (3-3)

Logistic: $Y_t = \frac{M}{1 + ae^{bt}}$ (3-4)

Gompertz: $Y_t = M(a)^{b^t}$ (3-5)

where

Y_t is the traffic at time t,

a, b, c are parameters,

M is a parameter describing the saturation level.

The various trend curves are shown in Figures 2/E.507 and 3/E.507.

The logistic and Gompertz curves differ from the linear, parabolic and exponential curves by having saturation or ceiling level. For further study see [10].

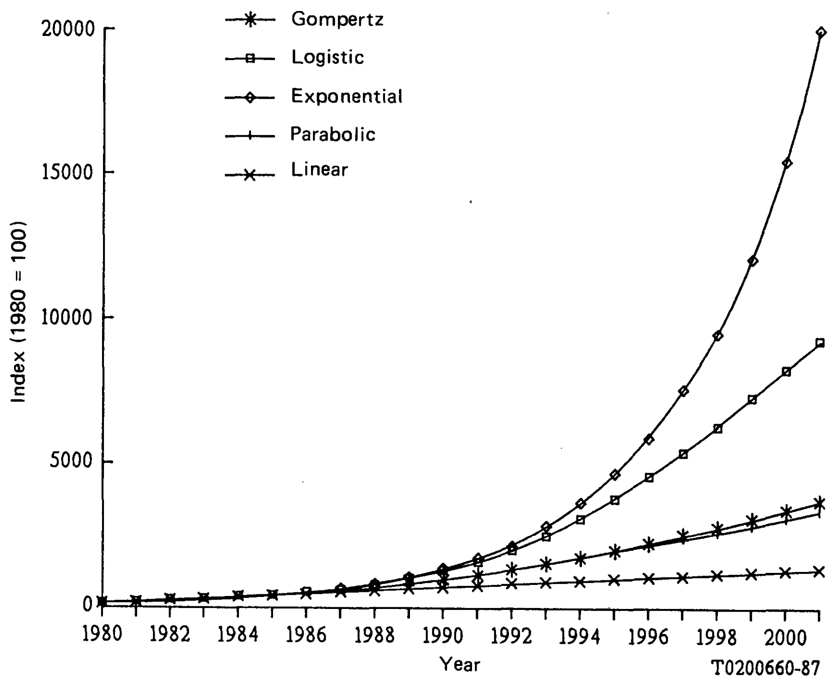


FIGURE 2/E.507

Example of fitting international telephone traffic using different models

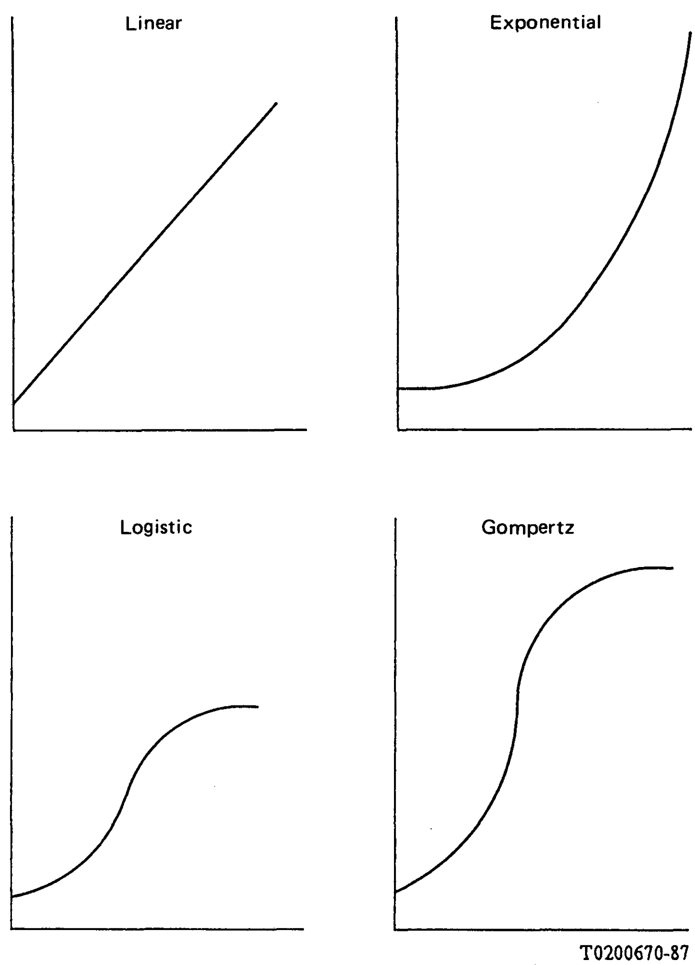


FIGURE 3/E.507

Curve fitting examples

3.2 *Smoothing models*

By using a smooth process in curve fitting, it is possible to calculate the parameters of the models to fit current data very well but not necessarily the data obtained from the distant past.

The best known smoothing process is that of the moving average. The degree of smoothing is controlled by the number of most recent observations included in the average. All observations included in the average have the same weight.

In addition to moving average models, there exists another group of smoothing models based on weighting the observations. The most common models are:

- simple exponential smoothing,
- double exponential smoothing,
- discounted regression,
- Holt's method, and
- Holt-Winters' seasonal models.

For example, in the method of exponential smoothing the weight given to previous observations decreases geometrically with age according to the following equation:

$$\hat{\mu}_t = (1 - a)Y_t + a\hat{\mu}_{t-1} \tag{3-6}$$

where:

Y_t is the measured traffic at time t,

μ_t is the estimated level at time t, and

a is the discount factor [and $(1 - a)$ is the smoothing parameter].

The impact of past observations on the forecasts is controlled by the magnitude of the discount factor.

Use of smoothing models is especially appropriate for short-term forecasts. For further studies see [1], [5] and [9].

3.3 *Autoregressive models*

If the traffic demand, X_t, at time t can be expressed as a linear combination of earlier equidistant observations of the past traffic demand, the process is an autoregressive process. Then the model is defined by the expression:

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + a_t \tag{3-7}$$

where

a_t is white noise at time t;

Φ_k, $k = 1, \dots p$ are the autoregressive parameters.

The model is denoted by $AR(p)$ since the order of the model is p.

By use of regression analysis the estimates of the parameters can be found. Because of common trends the exogenous variables $(X_{t-1}, X_{t-2}, \dots X_{t-p})$ are usually strongly correlated. Hence the parameter estimates will be correlated. Furthermore, significance tests of the estimates are somewhat difficult to perform.

Another possibility is to compute the empirical autocorrelation coefficients and then use the Yule-Walker equations to estimate the parameters $[\Phi_k]$. This procedure can be performed when the time series $[X_t]$ are stationary. If, on the other hand, the time series are non stationary, the series can often be transformed to stationarity e.g., by differencing the series. The estimation procedure is given in Annex A, § A.1.

3.4 *Autoregressive integrated moving average (ARIMA) models*

An extention of the class of autoregressive models which include the moving average models is called autoregressive moving average models (ARMA models). A moving average model of order q is given by:

$$X_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \dots - \theta_q a_{t-q} \qquad (3\text{-}8)$$

where

a_t is white noise at time t;

$[\theta_k]$ are the moving average parameters.

Assuming that the white noise term in the autoregressive models in § 3.3 is described by a moving average model, one obtains the so-called ARMA (p, q) model:

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \dots - \theta_q a_{t-q} \qquad (3\text{-}9)$$

The ARMA model describes a stationary time series. If the time series is non-stationary, it is necessary to difference the series. This is done as follow:

Let Y_t be the time series and B the backwards shift operator, then

$$X_t = (1 - B)^d Y_t \qquad (3\text{-}10)$$

where

d is the number of differences to have stationarity.

The new model ARIMA (p, d, q) is found by inserting equation (3-10) into equation (3-9).

The method for analyzing such time series was developed by G. E. P. Box and G. M. Jenkins [3]. To analyze and forecast such time series it is usually necessary to use a time series program package.

As indicated in Figure 1/E.507 a tentative model is identified. This is carried out by determination of necessary transformations and number of autoregressive and moving average parameters. The identification is based on the structure of the autocorrelations and partial autocorrelations.

The next step as indicated in Figure 1/E.507 is the estimation procedure. The maximum likelihood estimates are used. Unfortunately, it is difficult to find these estimates because of the necessity to solve a nonlinear system of equations. For practical purposes, a computer program is necessary for these calculations. The forecasting model is based on equation (3-9) and the process of making forecasts l time units ahead is shown in § A.2.

The forecasting models described so far are univariate forecasting models. It is also possible to introduce explanatory variables. In this case the system will be described by a transfer function model. The methods for analyzing the time series in a transfer function model are rather similar to the methods described above.

Detailed descriptions of ARIMA models are given in [1], [2], [3], [5], [11], [15] and [17].

3.5 *State space models with Kalman Filtering*

State space models are a way to represent discrete-time process by means of difference equations. The state space modelling approach allows the conversion of any general linear model into a form suitable for recursive estimation and forecasting. A more detailed description of ARIMA state space models can be found in [1].

For a stochastic process such a representation may be of the following form:

$$X_{t+1} = \Phi X_t + Z_t + \omega_t \tag{3-11}$$

and

$$Y_t = HX_t + \nu_t \tag{3-12}$$

where

X_t is an s-vector of state variables in period t,

Z_t is an s-vector of deterministic events,

Φ is an $s \times s$ transition matrix that may, in general, depend on t,

ω_t is an s-vector of random modelling errors,

Y_t is a d-vector of measurements in period t,

H is a $d \times s$ matrix called the observation matrix, and

ν_t is a d-vector of measurement errors.

Both ω_t in equation (3-11) and ν_t in equation (3-12) are additive random sequences with known statistics. The expected value of each sequence is the zero vector and ω_t and ν_t satisfy the conditions:

$$\begin{aligned} E\left[\omega_t \omega_j^T\right] &= Q_t\,\delta_{tj} \text{ for all } t, j, \\ E\left[\nu_t \nu_j^T\right] &= R_t\,\delta_{tj} \text{ for all } t, j, \end{aligned} \tag{3-13}$$

where

Q_t and R_t are nonnegative definite matrices,[2)]

and

δ_{tj} is the Kronecker delta.

Q_t is the covariance matrix of the modelling errors and R_t is the covariance matrix of the measurement errors; the ω_t and the ν_t are assumed to be uncorrelated and are referred to as white noise. In other words:

$$E\left[\nu_t\,\omega_j^T\right] = 0 \text{ for all } t, j, \tag{3-14}$$

and

$$E\left[\nu_t\,X_0^T\right] = 0 \text{ for all } t. \tag{3-15}$$

Under the assumptions formulated above, determine $X_{t,t}$ such that:

$$E\left[(X_{t,t} - X_t)^T(X_{t,t} - X_t)\right] = \text{minimum}, \tag{3-16}$$

where

$X_{t,t}$ is an estimate of the state vector at time t, and

X_t is the vector of true state variables.

---

2) A matrix A is nonnegative definite, if and only if, for all vectors z, $z^TAz \geqslant 0$.

The Kalman Filtering technique allows the estimation of state variables recursively for on-line applications. This is done in the following manner. Assuming that there is no explanatory variable Z_t, once a new data point becomes available it is used to update the model:

$$X_{t,t} = X_{t,t-1} + K_t(Y_t - HX_{t,t-1}) \tag{3-17}$$

where

K_t is the Kalman Gain matrix that can be computed recursively [18].

Intuitively, the gain matrix determines how much relative weight will be given to the last observed forecast error to correct it. To create a k-step ahead projection the following formula is used:

$$X_{t+k,t} = \Phi^k X_{t,t} \tag{3-18}$$

where

$X_{t+k,t}$ is an estimate of X_{t+k} given observations $Y_1, Y_2, \ldots, Y_t$.

Equations (3-17) and (3-18) show that the Kalman Filtering technique leads to a convenient forecasting procedure that is recursive in nature and provides an unbiased, minimum variance estimate of the discrete time process of interest.

For further studies see [4], [5], [16], [18], [19] and [22].

The Kalman Filtering works well when the data under examination are seasonal. The seasonal traffic load data can be represented by a periodic time series. In this way, a seasonal Kalman Filter can be obtained by superimposing a linear growth model with a seasonal model. For further discussion of seasonal Kalman Filter techniques see [6] and [20].

3.6 *Regression models*

The equations (3-1) and (3-2) are typical regression models. In the equations the traffic, Y_t, is the dependent (or explanatory) variable, while time t is the independent variable.

A regression model describes a linear relation between the dependent and the independent variables. Given certain assumptions ordinary least squares (OLS) can be used to estimate the parameters.

A model with several independent variables is called a multiple regression model. The model is given by:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + u_t \tag{3-19}$$

where

Y_t is the traffic at time t,

β_i, $i = 0, 1, \ldots, k$ are the parameters,

X_{it}, $i_e = 1, 2, \ldots, k$ is the value of the independent variables at time t,

u_t is the error term at time t.

Independent or explanatory variables which can be used in the regression model are, for instance, tariffs, exports, imports, degree of automation. Other explanatory variables are given in § 2 "Base data for forecasting" in Recommendation E.506.

Detailed descriptions of regression models are given in [1], [5], [7], [15] and [23].

3.7 *Econometric models*

Econometric models involve equations which relate a variable which we wish to forecast (the dependent or endogenous variable) to a number of socio-economic variables (called independent or explanatory variables). The form of the equations should reflect an expected casual relationship between the variables. Given an assumed model form, historical or cross sectional data are used to estimate coefficients in the equation. Assuming the model remains valid over time, estimates of future values of the independent variables can be used to give forecasts of the variables of interest. An example of a typical econometric model is given in Annex C.

There is a wide spectrum of possible models and a number of methods of estimating the coefficients (e.g., least squares, varying parameter methods, nonlinear regression, etc.). In many respects the family of econometric models available is far more flexible than other models. For example, lagged effects can be incorporated, observations weighted, ARIMA residual models subsumed, information from separate sections pooled and parameters allowed to vary in econometric models, to mention a few.

One of the major benefits of building an econometric model to be used in forecasting is that the structure or the process that generates the data must be properly identified and appropriate causal paths must be determined. Explicit structure identification makes the source of errors in the forecast easier to identify in econometric models than in other types of models.

Changes in structures can be detected through the use of econometric models and outliers in the historical data are easily eliminated or their influence properly weighted. Also, changes in the factors affecting the variables in question can easily be incorporated in the forecast generated from an econometric model.

Often, fairly reliable econometric models may be constructed with less observations than that required for time series models. In the case of pooled regression models, just a few observations for several cross-sections are sufficient to support a model used for predictions.

However, care must be taken in estimating the model to satisfy the underlying assumptions of the techniques which are described in many of the reference works listed at the end of this Recommendation. For example the number of independent variables which can be used is limited by the amount of data available to estimate the model. Also, independent variables which are correlated to one another should be avoided. Sometimes correlation between the variables can be avoided by using differenced or detrended data or by transformation of the variables. For further studies see [8], [12], [13], [14] and [21].

4 Discontinuities in traffic growth

4.1 *Examples of discontinuities*

It may be difficult to assess in advance the magnitude of a discontinuity. Often the influence of the factors which cause discontinuities is spread over a transitional period, and the discontinuity is not so obvious. Furthermore, discontinuities arising, for example, from the introduction of international subscriber dialling are difficult to identify accurately, because changes in the method of working are usually associated with other changes (e.g. tariff reductions).

An illustration of the bearing of discontinuities on traffic growth can be observed in the graph of Figure 4/E.507.

Discontinuities representing the doubling – and even more – of traffic flow are known. It may also be noted that changes could occur in the growth trend after discontinuities.

In short-term forecasts it may be desirable to use the trend of the traffic between discontinuities, but for long-term forecasts it may be desirable to use a trend estimate which is based on long-term observations, including previous discontinuities.

In addition to random fluctuations due to unpredictable traffic surges, faults, etc., traffic measurements are also subject to systematic fluctuations, due to daily or weekly traffic flow cycles, influence of time differences, etc.

4.2 *Introduction of explanatory variables*

Identification of explanatory variables for an econometric model is probably the most difficult aspect of econometric model building. The explanatory variables used in an econometric model identify the main sources of influence on the variable one is concerned with. A list of explanatory variables is given in Recommendation E.506, § 2.

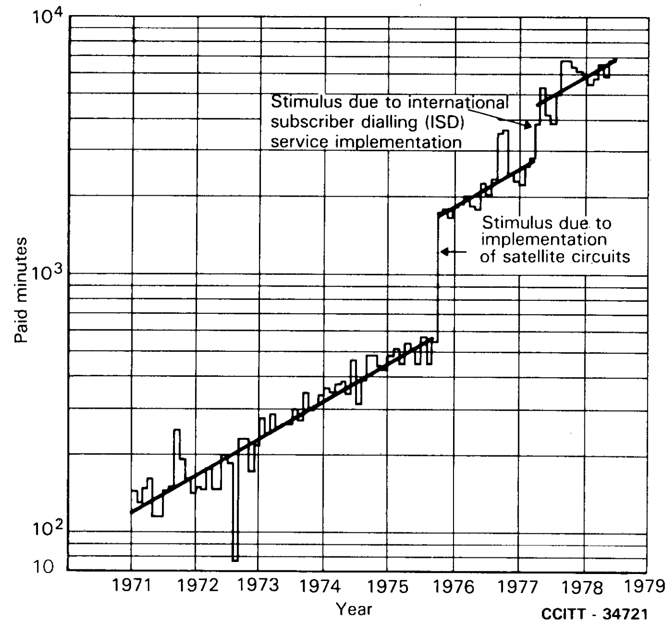


FIGURE 4/E.507

Outgoing telephone paid-minutes from Australia to Sri Lanka

Economic theory is the starting point for variable selection. More specifically, demand theory provides the basic framework for building the general model. However, the description of the structure or the process generating the data often dictate what variables enter the set of explanatory variables. For instance, technological relationships may need to be incorporated in the model in order to appropriately define the structure.

Although there are some criteria used in selecting explanatory variables [e.g., $\overline{R}^2$, Durbin-Watson (D-W) statistic, root mean square error (RMSE), ex-post forecast performance, explained in the references], statistical problems and/or availability of data (either historical or forecasted) limit the set of potential explanatory variables and one often has to revert to proxy variables. Unlike pure statistical models, econometric models admit explanatory variables, not on the basis of statistical criteria alone but, also, on the premise that causality is, indeed, present.

A completely specified econometric model will capture turning points. Discontinuities in the dependent variable will not be present unless the parameters of the model change drastically in a very short time period. Discontinuities in the growth of telephone traffic are indications that the underlying market or technological structure have undergone large changes.

Sustained changes in the growth of telephone demand can either be captured through varying parameter regression or through the introduction of a variable that appears to explain the discontinuity (e.g., the introduction of an advertising variable if advertising is judged to be the cause of the structural change). Once-and-for-all, or step-wise discontinuities, cannot be handled by the introduction of explanatory changes: dummy variables can resolve this problem.

4.3 *Introduction of dummy variables*

In econometric models, qualitative variables are often relevant; to measure the impact of qualitative variables, dummy variables are used. The dummy variable technique uses the value 1 for the presence of the qualitative attribute that has an impact on the dependent variable and 0 for the absence of the given attribute.

Thus, dummy variables are appropriate to use in the case where a discontinuity in the dependent variable has taken place. A dummy variable, for example, would take the value of zero during the historical period when calls were operator handled and one for the period for which direct dial service is available.

Dummy variables are often used to capture seasonal effects in the dependent variable or when one needs to eliminate the effect of an outlier on the parameters of a model, such as a large jump in telephone demand due to a postal strike or a sharp decline due to facility outages associated with severe weather conditions.

Indiscriminate use of dummy variables should be discouraged for two reasons:

1) dummy variables tend to absorb all the explanatory power during discontinuities, and

2) they result in a reduction in the degrees of freedom.

5 Assessing model specification

5.1 *General*

In this section methods for testing the significance of the parameters and also methods for calculating confidence intervals are presented for some of the forecasting models given in § 3. In particular the methods relating to regression analysis and time series analysis will be discussed.

All econometric forecasting models presented here are described as regression models. Also the curve fitting models given in § 3.1 can be described as regression models.

An exponential model given by

$$Z_t = ae^{bt} \cdot u_t \tag{5-1}$$

may be transformed to a linear form

$$\ln Z_t = \ln a + bt + \ln u_t \tag{5-2}$$

or

$$Y_t = \beta_0 + \beta_1 X_t + a_t \tag{5-3}$$

where

Y_t = $\ln Z_t$

β_0 = $\ln a$

β_1 = b

X_t = t

a_t = $\ln u_t$ (white noise).

5.2 *Autocorrelation*

A good forecasting model should lead to small autocorrelated residuals. If the residuals are significantly correlated, the estimated parameters and also the forecasts may be poor. To check whether the errors are correlated, the autocorrelation function r_k, $k = 1, 2, \ldots$ is calculated. r_k is the estimated autocorrelation of residuals at lag k. A way to detect autocorrelation among the residuals is to plot the autocorrelation function and to perform a Durbin-Watson test. The Durbin-Watson statistic is:

$$\text{D-W} = \frac{\sum_{t=2}^{N} (e_t - e_{t-1})^2}{\sum_{t=1}^{N} e_t^2} \tag{5-4}$$

where

e_t is the estimated residual at time t,

N is the number of observations.

5.3 *Test of significance of the parameters*

One way to evaluate the forecasting model is to analyse the impact of different exogenous variables. After estimating the parameters in the regression model, the significance of the parameters has to be tested.

In the example of an econometric model in Annex C, the estimated values of the parameters are given. Below these values the estimated standard deviation is given in parentheses. As a rule of thumb, the parameters are considered as significant if the absolute value of the estimates exceeds twice the estimated standard deviation. A more accurate way of testing the significance of the parameters is to take into account the distributions of their estimators.

The multiple correlation coefficient (or coefficient of determination) may be used as a criterion for the fitting of the equation.

The multiple correlation coefficient, R^2, is given by:

$$R^2 = \frac{\sum_{i=1}^{N} (\hat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{N} (Y_i - \overline{Y})^2} \tag{5-5}$$

If the multiple correlation coefficient is close to 1 the fitting is satisfactory. However, a high R^2 does not imply an accurate forecast.

In time series analysis, the discussion of the model is carried out in another way. As pointed out in § 3.4, the number of autoregressive and moving average parameters in an ARIMA model is determined by an identification procedure based on the structure of the autocorrelation and partial autocorrelation function.

The estimation of the parameters and their standard deviations is performed by an iterative nonlinear estimation procedure. Hence, by using a time series analysis computer program, the estimates of the parameters can be evaluated by studying the estimated standard deviations in the same way as in regression analysis.

An overall test of the fitting is based on the statistic

$$Q_{N-d} = \sum_{i=1}^{N} r_i^2 \tag{5-6}$$

where r_i is the estimated autocorrelation at lag i and d is the number of parameters in the model. When the model is adequate, Q_{N-d} is approximately chi-square distributed with $N - d$ degrees of freedom. To test the fitting, the value Q_{N-d} can be compared with fractiles of the chi-square distribution.

5.4 *Validity of exogenous variables*

Econometric forecasting models are based on a set of exogenous variables which explain the development of the endogenous variable (the traffic demand). To make forecasts of the traffic demand, it is necessary to make forecasts of each of the exogenous variables. It is very important to point out that an exogenous variable should not be included in the forecasting model if the prediction of the variable is less confident than the prediction of the traffic demand.

Suppose that the exact development of the exogenous variable is known which, for example, is the case for the simple models where time is the explanatory variables. If the model fitting is good and the white noise is normally distributed with expectation equal to zero, it is possible to calculate confidence limits for the forecasts. This is easily done by a computer program.

On the other hand, the values of most of the explanatory variables cannot be predicted exactly. The confidence of the prediction will then decrease with the number of periods. Hence, the explanatory variables will cause the confidence interval of the forecasts to increase with the number of the forecast periods. In these situations it is difficult to calculate a confidence interval around the forecasted values.

If the traffic demand can be described by an autoregressive moving average model, no explanatory variables are included in the model. Hence, if there are no explanatory variable in the model, the confidence limits of the forecasting values can be calculated. This is done by a time series analysis program package.

5.5 *Confidence intervals*

Confidence intervals, in the context of forecasts, refer to statistical constructs of forecast bounds or limits of prediction. Because statistical models have errors associated with them, parameter estimates have some variability associated with their values. In other words, even if one has identified the correct forecasting model, the influence of endogenous factors will cause errors in the parameter estimates and the forecast. Confidence intervals take into account the uncertainty associated with the parameter estimates.

In causal models, another source of uncertainty in the forecast of the series under study are the predictions of the explanatory variables. This type of uncertainty cannot be handled by confidence intervals and is usually ignored, even though it may be more significant than the uncertainty associated with coefficient estimates. Also, uncertainty due to possible outside shocks is not reflected in the confidence intervals.

For a linear, static regression model, the confidence interval of the forecast depends on the reliability of the regression coefficients, the size of the residual variance, and the values of the explanatory variables. The 95% confidence interval for a forecasted value Y_{N+1} is given by:

$$\hat{Y}_N(1) - 2\hat{\sigma} \leqslant Y_{N+1} \leqslant \hat{Y}_N(1) + 2\hat{\sigma} \qquad (5\text{-}7)$$

where $\hat{Y}_N(1)$ is the forecast one step ahead and $\hat{\sigma}$ is the standard error of the forecast.

This says that we expect, with a 95% probability, that the actual value of the series at time $N + 1$ will fall within the limits given by the confidence interval, assuming that there are no errors associated with the forecast of the explanatory variables.

6 Comparison of alternative forecasting models

6.1 *Diagnostic check — Model evaluation*

Tests and diagnostic checks are important elements in the model building procedure. The quality of the model is characterized by the residuals. Good forecasting models should lead to small autocorrelated residuals, the variance of the residuals should not decrease or increase and the expectation of the residuals should be zero or close to zero. The precision of the forecast is affected by the size of the residuals which should be small.

In addition the confidence limits of the parameter estimates and the forecasts should be relatively small. And in the same way, the mean square error should be small compared with results from other models.

6.2 *Forecasts of levels versus forecasts of changes*

Many econometric models are estimated using levels of the dependent and independent variables. Since economic variables move together over time, high coefficients of determination are obtained. The collinearity among the levels of the explanatory variables does not present a problem when a model is used for forecasting purposes alone, given that the collinearity pattern in the past continues to exist in the future. However, when one attempts to measure structural coefficients (e.g., price and income elasticities) the collinearity of the explanatory variables (known as multicollinearity) renders the results of the estimated coefficients unreliable.

To avoid the multicollinearity problem and generate benchmark coefficient estimates and forecasts, one may use changes of the variables (first difference or first log difference which is equivalent to a percent change) to estimate a model and forecast from that model. Using changes of variables to estimate a model tends to remove the effect of multicollinearity and produce more reliable coefficient estimates by removing the common effect of economic influences on the explanatory variables.

By generating forecasts through levels of and changes in the explanatory variables, one may be able to produce a better forecast through a reconciliation process. That is, the models are adjusted so that the two sets of forecasts give equivalent results.

6.3 *Ex-post forecasting*

Ex-post forecasting is the generation of a forecast from a model estimated over a sub-sample of the data beginning with the first observation and ending several periods prior to the last observation. In ex-post forecasting, actual values of the explanatory variables are used to generate the forecast. Also, if forecasted values of the explanatory variables are used to produce an ex-post forecast, one can then measure the error associated with incorrectly forecasted explanatory variables.

The purpose of ex-post forecasting is to evaluate the forecasting performance of the model by comparing the forecasted values with the actuals of the period after the end of the sub-sample to the last observation. With ex-post forecasting, one is able to assess forecast accuracy in terms of:

1) percent deviations of forecasted values from actual values,
2) turning point performance,
3) systematic behaviour of deviations.

Deviations of forecasted values from actual values give a general idea of the accuracy of the model. Systematic drifts in deviations may provide information for either re-specifying the model or adjusting the forecast to account for the drift in deviations. Of equal importance in evaluating forecast accuracy is turning point performance, that is, how well the model is able to forecast changes in the movement of the dependent variable. More criteria for evaluating forecast accuracy are discussed below.

6.4 *Forecast performance criteria*

A model might fit the historical data very well. However, when the forecasts are compared with future data that are not used for estimation of parameters, the fit might not be so good. Hence comparison of forecasts with actual observations may give additional information about the quality of the model. Suppose we have the time series, $Y_1, Y_2, \ldots, Y_N, Y_{N+1}, \ldots, Y_{N+M}$.

The M last observations are removed from the time series and the model building procedure. The one-step-ahead forecasting error is given by:

$$e_{N+t} = Y_{N+t} - \hat{Y}_{N+t-1}(1) \qquad t = 1, 2, \ldots, M \tag{6-1}$$

where

$\hat{Y}_{N+t-1}(1)$ is the one-step-ahead forecast.

Mean error

The mean error, ME, is defined by

$$\mathrm{ME} = \frac{1}{M} \sum_{t=1}^{M} e_{N+t} \tag{6-2}$$

ME is a criterium for forecast bias. Since the expectation of the residuals should be zero, a large deviation from zero indicates bias in the forecasts.

Mean percent error

The mean percent error, MPE, is defined by

$$\mathrm{MPE} = \frac{100}{M} \sum_{t=1}^{M} \frac{e_{n+t}}{Y_{N+t}} \tag{6-3}$$

This statistic also indicates possible bias in the forecasts. The criterium measures percentage deviation in the bias. It is not recommended to use MPE when the observations are small.

Root mean square error

The root mean square error, RMSE, of the forecast is defined as

$$\text{RMSE} = \left[\frac{1}{M} \sum_{t=1}^{M} e^2_{N+t} \right]^{1/2} \tag{6-4}$$

RMSE is the most commonly used measure for forecasting precision.

Mean absolute error

The mean absolute error, MAE, is given by

$$\text{MAE} = \frac{1}{M} \sum_{t=1}^{M} \left| e_{N+t} \right| \tag{6-5}$$

Theil's inequality coefficient

Theil's inequality coefficient is defined as follows:

$$U = \left[\sum_{t=1}^{M} \frac{e^2_{N+t}}{Y^2_{N+t}} \right]^{1/2} \tag{6-6}$$

Theil's U is preferred as a measure of forecast accuracy because the error between forecasted and actual values can be broken down to errors due to:

1) central tendency,
2) unequal variation between predicted and realized changes, and
3) incomplete covariation of predicted and actual changes.

This decomposition of prediction errors can be used to adjust the model so that the accuracy of the model can be improved.

Another quality that a forecasting model must possess is ability to capture turning points. That is, a forecast must be able to change direction in the same time period that the actual series under study changes direction. If a model is estimated over a long period of time which contains several turning points, ex-post forecast analysis can generally detect a model's inability to trace closely actuals that display turning points.

7 Choice of forecasting model

7.1 *Forecasting performance*

Although the choice of a forecasting model is usually guided by its forecasting performance, other considerations must receive attention. Thus, the length of the forecast period, the functional form, and the forecast accuracy of the explanatory variables of an econometric model must be considered.

The length of the forecast period affects the decision to use one type of a model versus another, along with historical data limitations and the purpose of the forecasting model. For instance, ARIMA models may be appropriate forecasting models for short-term forecasts when stability is not an issue, when sufficient historical data are available, and when causality is not of interest. Also, when the structure that generates the data is difficult to identify, one has no choice but to use a forecasting model which is based on historical data of the variable of interest.

The functional form of the model must also be considered in a forecasting model. While it is true that a more complex model may reduce the model specification error, it is also true that it will, in general, considerably increase the effect of data errors. The model form should be chosen to recognize the trade-off between these sources of error.

Availability of forecasts for explanatory variables and their reliability record is another issue affecting the choice of a forecasting model. A superior model using explanatory variables which may not be forecasted accurately can be inferior to an average model whose explanatory variables are forecasted accurately.

When market stability is an issue, econometric models which can handle structural changes should be used to forecast. When causality matters, simple models or ARIMA models cannot be used as forecasting tools. Nor can they be used when insufficient historical data exist. Finally, when the purpose of the model is to forecast the effects associated with changes in the factors that influence the variable in question, time series models may not be appropriate (with the exception, of course, of transfer function and multiple time series models).

7.2 *Length of forecast period*

For normal extensions of switching equipment and additions of circuits, a forecast period of about six years is necessary. However, a longer forecast period may be necessary for the planning of new cables or other transmission media or for major plant installations. Estimates in the long term would necessarily be less accurate than short-term forecasts but that would be acceptable.

In forecasting with a statistical model, the length of the forecast period is entirely determined by:

a) the historical data available,

b) the purpose or use of the forecast,

c) the market structure that generates the data,

d) the forecasting model used,

e) the frequency of the data.

The historical data available depends upon the period over which it has been collected and the frequency of collection (or the length of the period over which data is aggregated). A small historical data base can only support a short prediction interval. For example, with 10 or 20 observations a model can be used to forecast 4-5 periods past the sample (i.e. into the future). On the other hand, with 150-200 observations, potentially reliable forecasts can be obtained for 30 to 50 periods past the sample – other things being equal.

Certainly, the purpose of the forecast affects the number of predicted periods. Long range facility planning requires forecasts extending 15-20 or more years into the future. Rate change evaluations may only require forecasts for 2-3 years. Alteration of routing arrangements could only require forecasts extending a few months past the sample.

Stability of a market, or lack thereof, also affect the length of the forecast period. With a stable market structure one could conceivably extend the forecast period to equal the historical period. However, a volatile market does not afford the same luxury to the forecaster; the forecast period can only consist of a few periods into the future.

The forecasting models used to generate forecasts do, by their nature, influence the decision on how far into the future one can reasonably forecast. Structural models tend to perform better than other models in the long run, while for short-run predictions all models seem to perform equally well.

It should be noted that while the purpose of the forecast and the forecasting model affect the length of the forecast, the number of periods to be forecasted play a crucial role in the choice of the forecasting model and the use to which a forecast is put.

ANNEX A

(to Recommendation E.507)

Description of forecasting procedures

A.1 *Estimation of autoregressive parameters*

The empirical autocorrelation at lag k is given by:

$$r_k = \frac{v_k}{v_0} \qquad \text{(A-1)}$$

where

$$v_k = \frac{1}{N-1} \sum_{t=1}^{N-k} (X_t - \overline{X})(X_{t+k} - \overline{X}) \qquad \text{(A-2)}$$

and

$$\overline{X} = \frac{1}{N} \sum_{t=1}^{N} X_t \qquad \text{(A-3)}$$

N being the total number of observations.

The relation between $[r_k]$ and the estimates $[\hat{\Phi}_k]$ of $[\Phi_k]$ is given by the Yule-Walker equations:

$$\begin{aligned} r_1 &= \hat{\Phi}_1 + \hat{\Phi}_2 r_1 + \ldots + \hat{\Phi}_p r_{p-1} \\ r_2 &= \hat{\Phi}_1 r_1 + \hat{\Phi}_2 r_2 + \ldots + \hat{\Phi}_p r_{p-2} \\ &\cdot \\ &\cdot \\ &\cdot \\ r_p &= \hat{\Phi}_1 r_{p-1} + \hat{\Phi}_2 r_{p-2} + \ldots + \hat{\Phi}_p \end{aligned} \qquad \text{(A-4)}$$

Hence the estimators $[\hat{\Phi}_k]$ can be found by solving this system of equations.

For computations, an alternative to directly solving the equations is the following recursive procedure. Let $[\hat{\Phi}_{k,j}]_j$ be estimators of the parameters at lag $j = 1, 2, \ldots, k$ given that the total number of parameters are k. The estimators $[\hat{\Phi}_{k+1,j}]_j$ are then found by

$$\hat{\Phi}_{k+1,\, k+1} = \frac{r_{k+1} \sum_{j=1}^{k} \hat{\Phi}_{k,j}\, r_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\Phi}_{k,j}\, r_j} \qquad \text{(A-5)}$$

$$\hat{\Phi}_{k+1,\, j} = \hat{\Phi}_{k,j} - \hat{\Phi}_{k+1,\, k+1}\, \hat{\Phi}_{k,k-j+1} \qquad j = 1, 2, \ldots, k \qquad \text{(A-6)}$$

Defining $\hat{\Phi}_{p,j} = \hat{\Phi}_j$, $j = 1, 2, \ldots, p$, the forecast of the traffic demand at time $t+1$ is expressed by:

$$X_{t+1} = \hat{\Phi}_1 X_t + \hat{\Phi}_2 X_{t-1} + \ldots + \hat{\Phi}_p X_{t-p} \qquad \text{(A-7)}$$

A.2 *Forecasting with ARIMA models*

The forecast l time units ahead is given by:

$$\begin{aligned}\hat{X}_t(l) = \hat{\Phi}_1 [X_{t+l-1}] + \hat{\Phi}_2 [X_{t+l-2}] \\ + \ldots + \hat{\Phi}_p [X_{t+l-p}] \\ + [a_{t+l}] - \hat{\theta}_1 [a_{t+l-1}] \\ - \hat{\theta}_2 [a_{t+l-2}] - \ldots - \hat{\theta}_q [a_{t+l-q}],\end{aligned} \tag{A-8}$$

where $$[\hat{X}_j] = \begin{cases} \hat{X}_t (j - t) & \text{if } j > t \\ X_j & \text{if } j \leq t \end{cases} \tag{A-9}$$

$$[a_j] = \begin{cases} 0 & \text{if } j > t \\ X_j - \hat{X}_j & \text{if } j \leq t, \end{cases} \tag{A-10}$$

which means that $[X_j]$ is defined as a forecast when $j > t$ and otherwise as an actual observation and that $[a_j]$ is defined as 0 when $j > t$ since white noise has expectation 0. If the observations are known ($j \leq t$), then $[a_j]$ is equal to the residual.

ANNEX B

(to Recommendation E.507)

Kalman Filter for a linear trend model

To model telephone traffic, it is assumed that there are no deterministic changes in the demand pattern. This situation can be modelled by setting the deterministic component Z_t to zero. Then the general state space model is:

$$\begin{aligned} X_{t+1} &= \varphi X_t + \omega_t \\ Y_t &= H X_t + \nu_t \end{aligned} \tag{B-1}$$

where

X_t is an s-vector of state variables in period t,

Y_t is a vector of measurements in year t,

φ is an $s \times s$ transition matrix that may, in general, depend on t,

and

ω_t is an s-vector of random modelling errors,

ν_t is the measurement error in year t.

For modelling telephone traffic demand, adapt a simple two-state, one-data variable model defined by:

$$X_{t+1} = \begin{bmatrix} x_{t+1} \\ \dot{x}_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ \dot{x}_t \end{bmatrix} + \begin{bmatrix} \omega_t \\ \dot{\omega}_t \end{bmatrix} \tag{B-2}$$

and

$$y_t = x_t + \nu_t \tag{B-3}$$

where

x_t is the true load in year t,

$\dot{x}_t$ is the true incremental growth in year t,

y_t is the measured load in year t,

ν_t is the measurement error in year t.

Thus, in our model

$$\varphi = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad \text{and } H = 1. \tag{B-4}$$

The one-step-ahead projection is written as follows:

$$X_{t+1,t} = \begin{bmatrix} x_{t+1,t} \\ \dot{x}_{t+1,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{t,t} \\ \dot{x}_{t,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{t,t-1} + \alpha_t(y_t - x_{t,t-1}) \\ \dot{x}_{t,t-1} + \beta_t(y_t - x_{t,t-1}) \end{bmatrix} \tag{B-5}$$

where

$X_{t+1,t}$ is the projection of the state variable in period $t + 1$ given observations through year t.

The α_t and β_t coefficients are the Kalman gain matrices in year t. Rewriting the above equation yields:

$$x_{t,t} = (1-\alpha_t)x_{t,t-1} + \alpha_t y_t \tag{B-6}$$

and

$$\dot{x}_{t,t} = (1-\beta_t)\dot{x}_{t,t-1} + \beta_t(y_t - x_{t-1,t-1}) \tag{B-7}$$

The Kalman Filter creates a linear trend for each time series being forecast based on the current observation or measurement of traffic demand and the previous year's forecast of that demand. The observation and forecasted traffic load are combined to produce a smoothed load that corresponds to the level of the process, and a smoothed growth increment. The Kalman gain values α_t and β_t can be either fixed or adaptive. In [16] Moreland presents a method for selecting fixed, robust parameters that provide adequate performance independent of system noise, measurement error, and initial conditions. For further details on the proper selection of these parameters see [6], [20] and [22].

ANNEX C

(to Recommendation E.507)

Example of an econometric model

To illustrate the workings of an econometric model, we have chosen the model of United States billed minutes to Brazil. This model was selected among alternative models for three reasons:

a) to demonstrate the introduction of explanatory variables,

b) to point out difficulties associated with models used for both the estimation of the structure and forecasting purposes, and

c) to show how transformations may affect the results.

The demand of United States billed minutes to Brazil (*MIN*) is estimated by a log-linear equation which includes United States billed messages to Brazil (*MSG*), a real telephone price index (*RPI*), United States personal income in 1972 prices (*YP*72), and real bilateral trade between the United States and Brazil (*RTR*) as explanatory variables. This model is represented as:

$$\ln(MIN)_t = \beta_0 + \beta_1 \ln(MSG)_t + \beta_2 \ln(RPI)_t + \beta_3 \ln(YP72)_t + \beta_4 \ln(RTR)_t + u_t \tag{C-1}$$

where u_t is the error term of the regression and where, $\beta_1 > 0$, $\beta_2 < 0$, $\beta_3 > 0$ and $\beta_4 > 0$ are expected values.

Using ridge regression to deal with severe multicollinearity problems, we estimate the equation over the 1971 : 1 (i.e. first quarter of 1971) to 1979 : 4 interval and obtain the following results:

$$\ln(MIN)_t = -3.489 + \underset{(0.035)}{0.619} \ln(MSG)_t - \underset{(0.095)}{0.447} \ln(RPI)_t + \underset{(0.269)}{1.166} \ln(YP72)_t + \underset{(0.084)}{0.281} \ln(RTR)_t \quad \text{(C-2)}$$

$$\bar{R}^2 = 0.985, \; SER = 0.083, \; D\text{-}W = 0.922, \; k = 0.10 \quad \text{(C-3)}$$

where $\bar{R}^2$ is the adjusted coefficient of determination, *SER* is the standard error of the regression, D-W is the Durbin-Watson statistic, and k is the ridge regression constant. The values in parentheses under the equation are the estimated standard deviation of the estimated parameters $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, $\hat{\beta}_4$.

The introduction of messages as an explanatory variable in this model was necessitated by the fact that since the mid-seventies transmission quality has improved and completion rates have risen while, at the same time, the strong growth in this market has begun to dissipate. Also, the growth rates for some periods could not have been explained by rate activity on either side or real United States personal income. The behaviour of the message variable in the minute equation was able to account for all these factors.

Because the model serves a dual purpose – namely, structure estimation and forecasting – at least one more variable is introduced than if the model were to be used for forecasting purposes alone. The introduction of additional explanatory variables results in severe multicollinearity and necessitates employing ridge regression which lowers $\bar{R}^2$ and the Durbin-Watson statistic. Consequently, the predictive power of the model is reduced somewhat.

The effect of transforming the variables of a model are shown in the ex-post forecast analysis performed on the model of United States billed minutes to Brazil. The deviations using levels of the variables are larger than those of the logarithms of the variables which were used to obtain a better fit (the estimated RMSE for the log-linear regression model is 0.119 827). The forecast results in level and logarithmic form are shown in Table C-1/E.507.

TABLE C-1/E.507

	Logarithms			Levels		
	Forecast	Actual	% deviation	Forecast	Actual	% deviation
1980: 1	14.858	14.938	–0.540	2 836 269	3 073 697	– 7.725
2	14.842	14.972	–0.872	2 791 250	3 180 334	–12.234
3	14.916	15.111	–1.296	3 005 637	3 654 092	–17.746
4	14.959	15.077	–0.778	3 137 698	3 529 016	–11.089
1981: 1	15.022	15.102	–0.535	3 341 733	3 621 735	– 7.731
2	14.971	15.141	–1.123	3 175 577	3 762 592	–15.601
3	15.395	15.261	0.879	4 852 478	4 244 178	14.333
4	15.405	15.302	0.674	4 901 246	4 421 755	10.844
1982: 1	15.365	15.348	0.110	4 709 065	4 630 238	1.702
2	15.326	15.386	–0.387	4 528 947	4 807 901	– 5.802

References

[1] ABRAHAM (A.) and LEDOLTER (J.): Statistical methods for forecasting *J. Wiley*, New York, 1983.

[2] ANDERSON (O. D.): Time series analysis and forecasting. The Box-Jenkins approach. *Butterworth*, London, 1976.

[3] BOX (G. E. P.) and JENKINS (G. M.): Time Series Analysis: Forecasting and Control, *Holden-Day*, San Francisco, 1976.

[4] BROWN (R. G.): Introduction to random signal analysis and Kalman Filtering. *John Wiley & Sons*, New York, 1983.

[5] CCITT: Manual planning data and forecasting methods, Vol. I and II, *ITU*, Geneva, 1988.

[6] CHEMOUIL (P.) and GARNIER (B.): An Adaptive Short-Term Traffic Forecasting Procedure Using Kalman Filtering. *ITC 11*, Tokyo, 1985.

[7] DRAPER (N.) and SMITH (H.): Applied Regression Analysis, Second Edition, *John Wiley & Sons*, New York, 1981.

[8] DUTTA (M.): Econometric Methods, *South-Western Publishing Co.*, Cincinnati, 1975.

[9] GARDNER (E. S. Jr.): Exponential smoothing the state of art. *Journal of forecasting*, 4, pp. 1-28, 1985.

[10] GILCHRIST W.: Statistical forecasting. *John Wiley & Sons*, New York, 1976.

[11] GRANGER (C. W. J.) and NEWBOLD (P.): Forecasting Economic Time Series, *Academic Press*, New York, 1977.

[12] JOHNSTON (J.): Econometric Methods, Second Edition, *McGraw-Hill*, New York, 1972.

[13] JUDGE (G. G.) *et al.*: The Theory and Practice of Econometrics, *John Wiley & Sons*, New York, 1980.

[14] KMENTA (J.): Elements of Econometrics, *Macmillan Publishing Company*, New York, 1971.

[15] MAKRIDAKIS (S.), WHEELWRIGHT (S. C.), McGEE (V. .E.): Forecasting methods and applications Second Edition. *John Wiley & Sons*, New York, 1983.

[16] MORELAND (J. P.): A robust sequential projection algorithm for traffic load forecasting. *The Bell Technical Journal*, Vol. 61, No. 1, 1982.

[17] NELSON (C. R.): Applied Time Series Analysis for Managerial Forecasting, *Holden-Day*, San Francisco, 1973.

[18] PACK (C. D.) and WHITAKER (B. A.): Kalman Filter models for network forecasting. *The Bell Technical Journal*, Vol. 61, No. 1, pp. 1-9, 1982.

[19] SORENSON (H. W.): Kalman filtering techniques. Advances in control systems theory and applications. *Academic Press*, Vol. 3, pp. 219-292, 1966.

[20] SZELAG (C. R.): A short-term forecasting algorithm for trunk demand servicing. *The Bell Technical Journal*, Vol. 61, No. 1, pp. 67-96, 1982.

[21] THEIL (H.): Principles of Econometrics, *John Wiley & Sons*, New York, 1971.

[22] TOME (F. M.) and CUNHA (J. A.): Traffic forecasting with a state space model. *ITC 11*, Tokyo, 1985.

[23] WONNACOTT (T. H.) and WONNACOTT (R. J.): Regression. *John Wiley & Sons*, New York, 1981.

Bibliography

PINDYCK (R. S.) and RUBINFELD (D. F.): Econometric Models and Econometric Forecasts, *McGraw-Hill*, New York, 1981.

SASTRI, (T.): A state space modelling approach for time series forecasting. *Management Science*, Vol. 31, No. 11, pp. 1451-1470, 1985.

FORECASTING NEW INTERNATIONAL SERVICES

1 Introduction

The operation and administration of an international telecommunications network should include the consideration of subscriber demands for new services which may have different characteristics than the traditional traffic (i.e. peak busy hours, bandwidth requirements, and average call durations may be different). By addressing these new demands, Administrations can be more responsive to customer requirements for innovative telecommunications services. Based on the type of service and estimated demand for a service, network facilities and capacity may have to be augmented. An augmentation of the international network could require large capital investments and additional administrative functions and responsibilities. Therefore, it is appropriate that Administrations forecast new international services within their planning process.

This Recommendation presents methods for forecasting new services. The definitions of some of the characteristics of these services, together with their requirements, are covered in § 2, followed by base data requirements in § 3. § 4 discusses research to identify the potential market. Presentation of forecasting methods are contained in § 5. § 6 concludes with forecast tests and adjustments.

2 New service definitions

2.1 A distinction exists between those services which are enhancements of existing services carried on the existing network and those services which are novel.

Many of the services in this latter category will be carried on the Integrated Services Digital Network (ISDN). It is not the purpose of this section to provide an exhaustive list of services but rather to establish a framework for their classification. This framework is required because different base data and forecasting strategies may be necessary in each case.

2.2 enhanced services offered over the existing network

These are services which are offered over the existing network, and which offer an enhancement of the original use for which the network was intended. Services such as the international freephone service, credit card calling and closed user groups are examples of enhancements of voice services; while facsimile, telefax and videotex are examples of non-voice services. These services may be carried over the existing network and, therefore, data will concern usage or offered load specific to the enhancement. Arrangements can be established for the measurement of this traffic, such as the use of special network access codes for non-voice applications or by sampling outgoing circuits for the proportion of non-voice to voice traffic.

2.3 novel services

Novel services are defined as totally new service offerings many of which may be carried over the ISDN. In the case of ISDN, Recommendation I.210 divides telecommunications services into two broad categories: bearer services and teleservices. Recommendation I.210 further defines supplementary services which modify or supplement a basic telecommunications service. The definition of bearer services supported by the ISDN is contained in Recommendations I.210 and I.211, while that for teleservices is found in Recommendations I.210 and I.212. Bearer services may include circuit switched services from 64 kbit/s to 2 Mbit/s and packet services. Circuit switched services above 2 Mbit/s are for further study.

Teleservices may include Group 4 facsimile, mixed mode text and facsimile, 64 kbit/s Teletex and Videotex, videophone, videoconferencing, electronic funds transfer and point of sale transaction services. These lists are not exhaustive but indicate the nature and scope of bearer services and teleservices. Examples of new services are diagrammatically presented in Table 1/E.508.

TABLE 1/E.508

Examples of enhanced and novel services

Enhancement of existing services	"Novel" services	
	Bearer services	Teleservices
Teletex	Packet	Group 4 facsimile
Facsimile		Mixed mode
Videotex		Videophone
Message handling systems	Circuit switched services	Videoconferencing
International freephone	– 64 kbit/s	Electronic funds transfer
Credit cards	– 2 Mbit/s	Point of sale transactions
Closed user groups		Teletex (64 kbit/s) Videotex (64 kbit/s)

3 Base data for forecasting

3.1 *Measurement of enhanced services*

Measurements for existing services are available in terms of calls, minutes, Erlangs, etc. These procedures are covered in Recommendation E.506, § 2. In order to measure/identify enhanced service data from other traffic data on the same network it may be necessary to establish sampling or other procedures to aid in the estimation of this traffic, as described in § 4 and § 5.

3.2 *Novel services*

Novel services, as defined in § 2, may be carried on the ISDN. In the case of the ISDN, circuit switched bearer services and their associated teleservices will be measured in 64 kbit/s increments. Packet switched bearer services and associated teleservices will be measured by a unit of throughput, for example, kilocharacters or kilopackets per second. Other characteristics needed will reflect service quality measurements such as: noise, echo, post-dialing delay, clipping, bit-error rate, holding time, set-up time, error-free seconds, etc.

4 Market research

Market research is conducted to test consumer response and behaviour. This research employs the methods of questionnaires, market analysis, focus groups and interviews. Its purpose is to determine consumers' intentions to purchase a service, attitudes towards new and existing services, price sensitivity and cross service elasticities. Market research helps make decisions concerning which new services should be developed. A combination of the qualitative and quantitative phases of market research can be used in the initial stages of forecasting the demand for a new service.

The design of market research considers a sampling frame, customer/market stratification, the selection of a statistically random sample and the correction of results for non-response bias. The sample can be drawn from the entire market or from subsegments of the market. In sampling different market segments, factors which characterize the segments must be alike with respect to consumer behaviour (small intragroup variance) and should differ as much as possible from other segments (large intergroup variance); each segment is homogeneous while different segments are heterogeneous.

The market research may be useful in forecasting existing services or the penetration of new services. The research may be used in forecasting novel services or any service which has no historical series of demand data. It is important that potential consumers be given a complete description of the new service, including the terms and conditions which would accompany its provisioning. It is also important to ask the surveyees whether they would purchase the new service under a variety of illustrative tariff structures and levels. This aspect of market research will aid in redimensioning the demand upon final determination of the tariff structure and determining the customers' initial price sensitivity.

5 Forecasting procedures

5.1 *General*

The absence of historical data is the fundamental difference between forecasting new services and forecasting existing services. The forecast methodology is dependent on the base data. For example, for a service that is planned but has not been introduced, market research survey data can be used. If the service is already in existence in some countries, forecasting procedures for its introduction to a new country will involve historical data on other countries, its application to the new country and comparison of characteristics between countries.

5.2 *Sampling and questionnaire design*

The forecasting procedure for novel services based on market research is made up of five consecutive steps. The first of these consists in defining the scope of the study.

The second step involves the definition and selection of a sample from the population, where the population includes all potential customers which can be identified by qualitative market research developed through interviews at focus groups. The research can use stratified samples which involves grouping the population into homogeneous segments (or strata) and then sampling within each strata. Stratification prevents the disproportionate representation of some parts of the population that can result by chance with simple random sampling. The sample can be structured to include specified numbers of respondents having characteristics that are known, or believed, to affect the subject of the research. Examples of customer characteristics would be socio-economic background and type of business.

The third step is the questionnaire design. A trade-off exists between obtaining as much information as practical and limiting the questionnaire to a reasonable length, as determined by the surveyor. Most questionnaires have three basic sections:

1) qualifying questions to determine if a knowledgeable person has been contacted;
2) basic questions including all questions which constitute the body of the questionnaire;
3) classification questions collecting background on demographic information.

The fourth step involves the implementation of the research – the actual surveying portion. Professional interviewers, or firms specializing in market research should be employed for interviewing.

The fifth and final step is the tabulation and analysis of the survey data. § 5.3-5.7 describe this process in detail.

5.3 *Conversion ratios for the sample*

Conversion ratios are used in estimating the proportion of respondents expressing an interest in the service who will eventually subscribe.

The analysis of the market research data based on a sample survey, where a stratified sample is drawn across market segments, for a service that is newly introduced or is planned, is discussed below:

Let

X_{1i} = the proportion of firms in market segment i that are very interested in the service.

X_{2i} = the proportion of firms in market segment i that are interested in the service.

X_{3i} = the proportion of firms in market segment i that are not interested in the service.

X_{4i} = the proportion of firms in market segment i that cannot decide whether they are interested or not.

The above example has 4 categories of responses. Greater of fewer categories may be used depending on the design of the questionnaire.

Notice that

$$\sum_{j} X_{ji} = 1,$$

where j = the index of categories of responses.

Market research firms sometimes determine conversion ratios for selected product/service types. Conversion ratios depend on the nature of the service, the type of respondents, and the questionnaire and its implementation. Conversion ratios applied to the sample will estimate the expected proportion of firms *in the survey* that will eventually subscribe, over the planning period. For studies related to the estimation of conversion ratios, refer to [1], [3] and [5].

Then,

$c_1 X_{1i}$ = the proportion of firms in market segment i that expressed a strong interest and are expected to subscribe.

$c_2 X_{2i}$ = the proportion of firms in market segment i that expressed an interest and are expected to subscribe.

$c_3 X_{3i}$ = the proportion of firms in market segment i that expressed no interest but are expected to subscribe.

$c_4 X_{4i}$ = the proportion of undecided firms in market segment i that are expected to subscribe.

where c_j = conversion ratio for response j.

The proportion of firms in market segment i, P_i, that are expected to subscribe to the service, equals

$$P_i = \sum_{j=1}^{4} c_j X_{ji} \tag{5-1}$$

The conversion ratio is based on the assumption that there is a 100% market awareness. That is, all surveyees are fully informed of the service availability, use, tariffs, technical parameters, etc. P_i, therefore, can be interpreted as the long-run proportion of firms in market segment i that are expected to subscribe to the service at some future time period, T.

Two issues arise in the estimation of the proportion of customers that subscribe to the service:

1) while P_i refers to the sample surveyed, the results need to be extrapolated to represent the population.

2) P_i is the long-run (maximum) proportion of firms expected to subscribe. We are interested in predicting no just the eventual number of subscribers but, also, those at intermediate time periods before the service reaches a saturation point.

5.4 *Extrapolation from sample to population*

To extrapolate the data from the sample to represent the population, let

N_i = size of market segment i (measured for example, by the number of firms in market segment i)

Then S_i, the expected number of subscribers in the planning horizon, equals:

$$S_i = P_i N_i \tag{5-2}$$

5.5 *Market penetration over time*

To determine the expected number of subscribers at various points in time before the service reaches maturity, let

p_{it} = the proportion of firms in market segment i that are expected to subscribe at time t.

Clearly,

$p_{it} < P_i$

and $p_{it} \rightarrow P_i$ as $t \rightarrow T$

The relation between p_{it} and P_i can be explicitly defined as:

$$p_{it} = a_{it} \cdot P_i \tag{5-3}$$

a_{it} is a penetration function, reflecting changing market awareness and acceptance of the service over time, in market segment i. An appropriate functional form for a_{it} should be bounded in the interval (0,1).

As an example, let a_{it} be a logistic function:

$$a_{it} = \frac{1}{1 + e^{b_i t}} \tag{5-4}$$

$b_i \leqslant 0$ is the speed with which p_{it} approaches P_i in market segment i, as illustrated in Figure 1/E.508.

For other examples of non-linear penetration functions, refer to the Annex A.

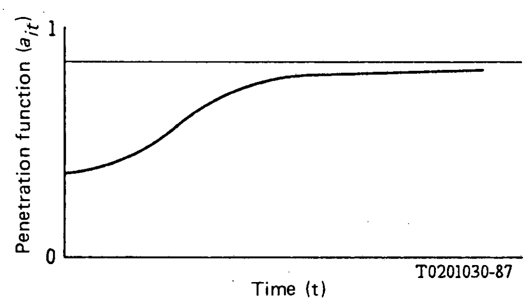


FIGURE 1/E.508

Rate of market penetration

The introduction of a new service will usually differ according to the market segment. The rate of penetration may be expressed as a function of time, and the speed of adjustment (b_i) may vary across segments. Lower absolute values of b_i, for the logistic function will imply faster rates of penetration.

While the form of the penetration function relating the rate of penetration to time is the same for all segments, the parameter b_i varies across segments, being greater in segments with a later introduction of the new service.

Let t_{0i} = time period of introduction of service in market segment i.

Then, $t - t_{0i}$ = time period elapsed since service was introduced in market segment i.

In the diagramatic illustration, of Figure 2/E.508, the service has achieved the same level of market penetration a_0, in t_C periods after its introduction in market C as it did in t_A periods after its introduction in market segment A. Later introductions may not necessarily lead to faster rates of penetration across segments. However, within the same market segment, across countries with similar characteristics, such an expectation is reasonable.

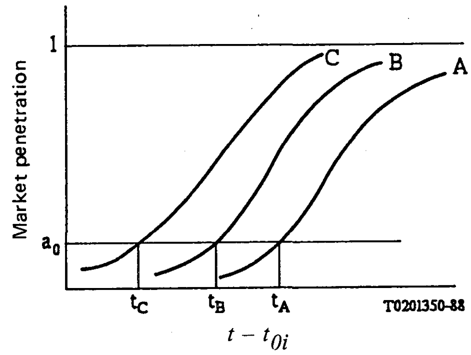


FIGURE 2/E.508

Market awareness as a function of time of introduction of service

5.6 *Growth of market segment over time*

The above discussion has accounted for gradual market penetration of the new service, by allowing p_{it} to adjust to P_i over time. The same argument can be extended to the size of market segment i over time.

Let n_{it} = size of market segment i at time t.

Then, the expected number of subscribers at time t in market segment i, equals:

$$s_{it} = a_{it} \cdot p_{it} \cdot n_{it} \tag{5-5}$$

and

$S_t = \sum_i s_{it}$ = expected number of subscribers across all market segments at time t.

5.7 *Quantities forecasted*

The above procedure forecasts the expected number of customers for a new service within a country. Other quantities of interest may include lines, minutes, messages, revenue, packets, kilobits, etc. The most straightforward forecasting method for some of these quantities is to assume constant relationships such as:

expected access lines = (average access lines) × expected number of subscribers

expected minutes = (average use per line) × expected access lines

expected messages = expected minutes/(average length of conversation)

expected revenue = (average rate per minute) × expected minutes

The constants, appearing in parentheses, above, can be determined through 1) the process of market research, or 2) past trends in similar services.

5.8 *Forecasting with historical data: application analysis*

After a new service has been introduced, historical data can be analyzed to forecast demand for expanded availability to other countries. Development of a new service will follow trends based on applications, such as data transmission, travel reservations, intracompany communications, and supplier contact. Applications of a service vary widely and no single variable may be an adequate indicator of total demand.

The following procedure links demand to country characteristics for forecasting expanded availability of a new service to other countries.

Let $D = (D_i, D_2, \ldots\ldots, D_n)'$

represent a vector of country-specific annual demand for the service across n countries, where the service currently exists. Let C = matrix of m characteristics relating to each of the n countries that are reasonable explanatory variables of demand. The components of m would vary depending on the nature of the service and its application.

Some essential components of m would be the price of the service (or an index representing its price) and some proxy for market awareness. As discussed in earlier sections, market awareness is one of the key determinants of the rate of market penetration of the service. Reasonable proxies would be advertising expenditures and time (measured as $t^* = t - t_0$) where t^* would measure time elapsed since the service was first introduced at time t_0. Market awareness can be characterized as some non-linear function of t^*, as presented in § 5.5. Other components of m may include socio-economic characteristics of the customers, market size and location of customers.

The model that is estimated is:

$$D = C\beta + u \tag{5-6}$$

where

C is a $(n \times m)$ matrix of country characteristics

D is a $(n \times 1)$ vector of demand

β is a $(m \times 1)$ vector of coefficients corresponding to each of the m characteristics

$u = (n \times 1)$ vector of error terms

The estimated regression is:

$$\hat{D} = C\hat{\beta} \tag{5-7}$$

Traditional methods of estimating regressions will be applied. Equation (5-7) can be used for predicting demand for any country where the service is being newly introduced, as long as elements of the matrix C are available.

5.9 *Forecasting with limited information*

In the extreme case where no market research data is available (or is uneconomical given resource constraints), or country characteristics that affect demand are not easily available or quantifiable, other methods of forecasting need to be devised.

For example, to forecast the demand for a new international private line service using digital technology, the following elements should be taken into account in the development of reasonable estimates of the expected number of lines:

a) discussions with foreign telephone companies,

b) discussions with very large potential customers regarding their future needs,

c) service inquiries from customers,

d) customer letters of intent, and

e) any other similar qualitative information.

6 Forecast tests and adjustments

6.1 *General*

Forecast tests and adjustments are dependent on the methodology applied. For example, in the case of a market research based forecast, it is important to track the forecast of market size, awareness and rate of penetration over time and to adjust forecasts accordingly. However, for an application-based methodology, traditional tests and adjustments applicable to regression methods will be employed, as discussed below.

6.2 *Market research based analysis*

This section discusses adjustments to forecasts based on the methodology described in §§ 5.2 to 5.8. The methodology was based on quantification of responses from a sample survey.

The forecast was done in two parts:

a) extrapolating the sample to the population, using market size, N_i;

b) allowing for gradual market penetration (awareness), a_{it} of the new service over time.

The values attributed to n_{it} (which represents the size of market segment i at time t) and a_{it} can be tracked over time and forecast adjustments made in the following manner:

a) As an example for n_{ii}, the segments could be categorized as travel or financial services. The size of the segment would be the number of tourists, and the number of large banks. Historical data, where available, on these units of measurement can be used to forecast their sizes at any point of time in the future. Where history is not available, reasonable growth factors can be developed through subject matter experts and past experiences. The forecast of n_{it} should be tracked against actual measured values and adjusted for large deviations.

b) For a_{it}, testing with only a few observations since the introduction of the service is more difficult.

Given that,

$$a_{it} = \frac{p_{it}}{P_i} \tag{6-1}$$

and P_i is assumed fixed (in the long run), testing a_{it} is equivalent to testing p_{it}. p_{it} can be tracked by observing the proportion of respondents that actually subscribe to the service at time t. This assumes the need to track the same individuals who were originally in the survey, as is customary in a panel survey. Panel data is collected through sample surveys of cross-sections of the same individuals, over time. This method is commonly used for household socio-economic surveys. Having observed p_{it} for a new period, values of a_{it} can be plotted against time to study the nature of the penetration function, a_{it}, and the most appropriate functional form that fits the data should be chosen. At very early stages of service introduction, traditional functional forms for market penetration, such as a logistic function (as illustrated in the example in § 5.5), will be a reasonable form to assume. Other variations of the functional form depicting market penetration would be the Gompertz or Gauss growth curves. The restriction is that the penetration function should be bounded in the interval (0,1). See Annex A for an algebraic depiction of functional forms.

There are various statistical forms that may be chosen as representations for the penetration function. The appropriate functional form should be based on some theoretical based information such as the expected nature of penetration of the specific service over time.

Continuous tracking of n_{it}, p_{it} and a_{it} over time will enable adjustments to these values whenever necessary and enable greater confidence in the forecasts.

6.3 *Application based analysis*

The application based analysis is a regression based approach and traditional forecast tests for a regression model will apply. For instance, hypothesis tests on each of the explanatory variables included in the model will be necessary. Corrections may be needed for hetero-elasticity, serial correlation and multicollinearity, when suspect. The methodology for performing such tests are described in most econometrics text books. In particular, references [2] and [4] can be used as guidelines. Recommendation E.507 also discusses these corrections.

Adjustments need to be made for variables that should be included in the regression model but are not easily quantifiable. For example, market awareness that results from advertising and promotional campaigns plays an important role in the growth of a new service, but data on such expenditures or the associated awareness may not be readily available. Some international services are targeted towards international travelers, and fluctuations in exchange rates will be a determining factor. Such variables, while not impossible to measure, may be expensive to acquire. However, expectations of future trends in such variables can enable the forecaster to arrive at some reasonable estimates of their impact on demand. Unexpected occurrences such as political turmoil and natural disasters in particular countries will also necessitate post forecast adjustments based upon managerial judgement.

Another important adjustment that may be necessary is the expected competition from other carriers offering similar or substitutable services. Competitor prices, if available, may be used as explanatory variables within the model and allow the measurement of a cross-price impact. In most situations, it is difficult to obtain competitor prices. In such cases, other methods of calculating competitor market shares need to be developed.

Regardless of forecasting methodology, the final forecasts will have to be reviewed by management responsible for planning the service as well as by network engineers in order to assess the feasibility both from a planning implementation and from a technical point of view.

ANNEX A

(to Recommendation E.508)

Penetration functions (growth curves)

Some examples of non-linear penetration functions are illustrated below:

A.1 *Logistic curve*

$$a_{it} = \alpha \ / \ \{1 + e^{-bt}\} \tag{A-1}$$

For $\alpha = 1$, the curve is bounded in the interval (0,1). Changing b will alter the steepness of the curve. The higher the value of b, the faster the rate of penetration. This curve is S-shaped and is symmetrical about its point of inflection, the latter being where;

$$\frac{d^2 a_{it}}{dt^2} = 0 \tag{A-2}$$

A.2 *Gompertz curve*

$$a_{it} = \alpha \exp \left\{ - b e^{-kt} \right\} \tag{A-3}$$

As $t \to \infty$ $a_{it} \to \alpha$, the limiting growth.

Holding $k = 1$ and $\alpha = 1$, higher values of b will imply slower rates of penetration. This curve is also S-shaped like the logistic curve, but is not symmetrical about its inflection point.

When $t = 0$, then $a_{it} = \alpha e^{-b}$, which is the initial rate of penetration.

A.3 *Gauss curve*

$$a_{it} = \alpha \left(1 - e^{-bt^2}\right) \tag{A-4}$$

As $t \to \infty$, then $a_{it} \to \alpha$

As $t \to 0$, then $a_{it} \to 0$.

Choosing $\alpha = 1$, the curve is bounded in the interval (0,1).

References

[1] AXELROD (J. N.): Attitude measures that predict purchase, *Journal of Advertising Research*, Vol. 8, No. 1, pp. 3-17, New York, March 1968.

[2] JOHNSTON (J.): Econometric methods, Second Edition, *McGraw-Hill*, New York, 1972.

[3] KALWANI (M. U.), SILK, (A. J.): On the reliability and predictive validity of purchase intention measures, *Marketing Science*, Vol. 1, No. 3, pp. 243-286, Providence, RI, Summer 1982.

[4] KMENTA (J.): Elements of econometrics, *Macmillan Publishing Co.*, New York, 1971.

[5] MORRISON (D. G.): Purchase intentions and purchase behavior, *Journal of Marketing*, Vol. 43, pp. 65-74, Chicago, Ill., Spring 1979.

Bibliography

BEN-AKIVA (M.) and LERMAN (S. R.): Discrete choice analysis.

DRAPER (N.) and SMITH (H.): Applied regression analysis, Second Edition, *John Wiley & Sons*, New York, 1981.

SECTION 3

DETERMINATION OF THE NUMBER OF CIRCUITS IN MANUAL OPERATION

Recommendation E.510 [1)]

DETERMINATION OF THE NUMBER OF CIRCUITS IN MANUAL OPERATION

1 The quality of an international manual demand service should be defined as the percentage of call requests which, during the average busy hour (as defined later under § 3) cannot be satisfied immediately because no circuit is free in the relation considered.

By *call requests satisfied immediately* are meant those for which the call is established by the same operator who received the call, and within a period of two minutes from receipt of that call, whether the operator (when she does not immediately find a free circuit) continues observation of the group of circuits, or whether she makes several attempts in the course of this period.

Ultimately, it will be desirable to evolve a corresponding definition based on the *average speed* of establishing calls in the busy hour, i.e. the average time which elapses between the moment when the operator has completed the recording of the call request and the moment when the called subscriber is on the line, or the caller receives the advice *subscriber engaged, no reply*, etc. But for the moment, in the absence of information about the operating time in the European international service, such a definition cannot be established.

2 The number of circuits it is necessary to allocate to an international relation, in order to obtain a given grade of service, should be determined as a function of the *total holding time* of the group in the busy hour.

The total holding time is the product of the number of calls in the busy hour and a factor which is the sum of the average call duration and the average operating time.

These durations will be obtained by means of a large number of observations made during the busy hours, by agreement between the Administrations concerned. If necessary, the particulars entered on the tickets could also serve to determine the average duration of the calls.

The average call duration will be obtained by dividing the total number of minutes of conversation recorded by the recorded number of effective calls.

The average operating time will be obtained by dividing the total number of minutes given to operating (including ineffective calls) by the number of effective calls recorded.

3 The number of calls in the busy hour will be determined from the average of returns taken during the busy hours on a certain number of busy days in the year.

Exceptionally busy days, such as those which occur around certain holidays, etc., will be eliminated from these returns. The Administrations concerned should plan, whenever possible, to put additional circuits into service for these days.

In principle, these returns will be taken during the working days of two consecutive weeks, or during ten consecutive working days. If the monthly traffic curve shows only small variations, they will be repeated twice a year only. They will be taken three or four times a year or more if there are material seasonal variations, so that the average established is in accordance with all the characteristic periods of traffic flow.

---

1) This Recommendation dates from the XIIIth Plenary Assembly of the CCIF (London, 1946) and has not been fundamentally revised since. It was studied under Question 13/II in the Study Period 1968-1972 and was found to be still valid.

4 The total occupied time thus determined should be increased by a certain amount determined by agreement between the Administrations concerned according to the statistics of traffic growth during earlier years, to take account of the probable growth in traffic and the fact that putting new circuits into service takes place some time after they are first found to be necessary.

5 The total holding time of the circuits thus obtained, in conjunction with a suitable table (see Table 1/E.510), will enable the required number of circuits to be ascertained.

6 In the international manual telephone service, the following Tables A and B should be used as a basis of minimum allocation:

Table A corresponds to about 30% of calls failing at the first attempt because of all circuits being engaged and to about 20% of the calls being deferred.

Table B, corresponding to about 7% of calls deferred, will be used whenever possible.

These tables do not take account of the fact that the possibility of using secondary routes permits, particularly for small groups, an increase in the permissible occupation time.

TABLE 1/E.510

Capacity of circuit groups
(See Supplement No. 2 at the end of this fascicle)

Number of circuits	Table A		Table B	
	Percentage of circuit usage	Minutes of circuit usage possible in the busy hour	Percentage of circuit usage	Minutes of circuit usage possible in the busy hour
1	65.0	39	–	–
2	76.7	92	46.6	56
3	83.3	150	56.7	102
4	86.7	208	63.3	152
5	88.6	266	68.3	205
6	90.0	324	72.0	259
7	91.0	382	74.5	313
8	91.7	440	76.5	367
9	92.2	498	78.0	421
10	92.6	556	79.2	475
11	93.0	614	80.1	529
12	93.4	672	81.0	583
13	93.6	730	81.7	637
14	93.9	788	82.3	691
15	94.1	846	82.8	745
16	94.2	904	83.2	799
17	94.3	962	83.6	853
18	94.4	1020	83.9	907
19	94.5	1078	84.2	961
20	94.6	1136	84.6	1015

Note – Tables A and B can be extended for groups comprising more than 20 circuits by using the values given for 20 circuits.

SECTION 4

DETERMINATION OF THE NUMBER OF CIRCUITS IN AUTOMATIC AND SEMIAUTOMATIC OPERATION

Recommendation E.520

NUMBER OF CIRCUITS TO BE PROVIDED IN AUTOMATIC AND/OR SEMIAUTOMATIC OPERATION, WITHOUT OVERFLOW FACILITIES

This Recommendation refers to groups of circuits used:

- in automatic operation;
- in semiautomatic operation;
- in both automatic and semiautomatic operations on the same group of circuits.

1 General method

1.1 The CCITT recommends that the number of circuits needed for a group should be read from tables or curves based on the classical Erlang B formula (see Supplements Nos. 1 and 2 at the end of this fascicle which refers to full availability groups). Recommended methods for traffic determination are indicated in Recommendation E.500.

For *semi-automatic operation* the loss probability p should be based on 3% during the mean busy hour.

For *automatic operation* the loss probability p should be based on 1% during the mean busy hour.

Semiautomatic traffic using the same circuits as automatic traffic is to be added to the automatic traffic and the same parameter value of $p = 1\%$ should be used for the total traffic.

The values of 3% and 1% quoted above refer to the Erlang B formula and derived tables and curves. The 3% value should not be considered as determining a grade of service because with semiautomatic operation there will be some smoothing of the traffic peaks; it is quoted here only to determine the value of the parameter p (loss probability) to use in the Erlang B tables and curves.

1.2 In order to provide a satisfactory grade of service both for the mean busy-hour traffic and for the traffic on exceptionally busy days, it is recommended that the proposed number of circuits should, if necessary, be increased to ensure that the loss probability shall not exceed 7% during the mean busy hour for the average traffic estimated for *the five busiest days* as specified in Recommendation E.500.

1.3 For *small groups of long intercontinental circuits* with automatic operation some relaxation could be made in respect to loss probability. It is envisaged that such circuits would be operated on a both-way basis and that a reasonable minimum for automatic service would be a group of six circuits. A table providing relaxation in Annex A is based on a loss probability of 3% for six circuits, with a smooth progression to 1% for 20 circuits. The general provision for exceptional days remains unchanged.

For exceptional circumstances in which very small groups (less than six intercontinental circuits) are used for automatic operation, dimensioning of the group should be based on the loss probability of 3%.

2 Time differences

Time differences at the two terminations of intercontinental circuits are likely to be much more pronounced than those on continental circuits. In order to allow for differences on groups containing both-way circuits it will be desirable to acquire information in respect to traffic flow both during the mean busy hour for both directions and during the mean busy hour for each direction.

It is possible that in some cases overflow traffic can be accepted without any necessity to increase the number of circuits, in spite of the fact that this overflow traffic is of a peaky nature. Such circumstances may arise if there is no traffic overflowing from high-usage groups during the mean busy hour of the final group.

3 Both-way circuits

3.1 With the use of both-way circuits there is a danger of simultaneous seizure at both ends; this is particularly the case on circuits with a long propagation time. It is advisable to arrange the sequence of selection at the two ends so that such double seizure can only occur when a single circuit remains free.

When all the circuits of a group are operated on a both-way basis, time differences in the directional mean busy hours may result in a total mean busy-hour traffic flow for the group which is not the sum of the mean busy-hour traffic loads in each direction. Furthermore, such differences in directional mean busy hour may vary with seasons of the year. However, the available methods of traffic measurement can determine the traffic flow during mean busy hour for this total traffic.

3.2 Some intercontinental groups may include one-way as well as both-way operated circuits. It is recommended that in all cases the one-way circuits should be used, when free, in preference to the both-way circuits. The number of circuits to be provided will depend upon the one-way and total traffic.

The total traffic will need to be determined for:

a) each direction of traffic;

b) both-way traffic.

This determination is to be made for the busy hour or the busy hours corresponding to the two cases a) and b) above.

In the cases where the number of one-way circuits is approximately equal for each direction, no special procedure is necessary, and the calculation can be treated as for a simple two-group grading [1].

If the number of one-way circuits is quite different for the two directions, some correction may be needed for the difference in randomness of the flow of calls from the two one-way circuit groups to the both-way circuit group. The general techniques for handling cases of this type are quoted in Recommendation E.521.

ANNEX A

(to Recommendation E.520)

Table A-1/E.520 may be applied to small groups of long intercontinental circuits. The values in column 2 are suitable for a random offered traffic with full availability access.

The table is based on 1% loss probability for 20 circuits and increases progressively to a loss probability of 2% at 9 circuits and 3% at 6 circuits (loss probabilities for these three values being based on the Erlang loss formula: see Supplement No. 1). The traffic flow values obtained from a smoothing curve coincide very nearly with those determined by equal marginal utility theory, i.e. an improvement factor of 0.05 Erlang for an additional circuit.

For groups requiring more than 20 circuits the table for loss probability of 1%, mentioned in Supplement No. 1, should be used.

TABLE A-1/E.520

Number of circuits	Traffic flow (in erlangs)		
	Offered	Carried	Encountering congestion
(1)	(2)	(3)	(4)
6	2.54	2.47	0.08
7	3.13	3.05	0.09
8	3.73	3.65	0.09
9	4.35	4.26	0.09
10	4.99	4.90	0.09
11	5.64	5.55	0.10
12	6.31	6.21	0.10
13	6.99	6.88	0.10
14	7.67	7.57	0.10
15	8.37	8.27	0.11
16	9.08	8.96	0.11
17	9.81	9.69	0.11
18	10.54	10.42	0.11
19	11.28	11.16	0.12
20	12.03	11.91	0.12

Reference

[1] TÅNGE (I.): Optimal use of both-way circuits in cases of unlimited availability, *TELE*, English Edition, No. 1, 1956.

Recommendation E.521

CALCULATION OF THE NUMBER OF CIRCUITS IN A GROUP CARRYING OVERFLOW TRAFFIC

A calculation of the number of circuits in a group carrying overflow traffic should be based on this Recommendation and on Recommendation E.522 dealing with high-usage groups.

The objective grade of service used is that the average blocking during the busy-hour of the 30 busiest days of the year will not exceed 1%.

To determine the number of circuits in a group carrying overflow traffic, three traffic parameters are required: the average traffic offered to the group, the weighted peakedness factor, and the level of day-to-day traffic variations.

The level of day-to-day traffic variations indicates the degree to which the daily busy-hour traffic deviates from the overall mean traffic, and is determined by the sample variance of the 30 busy-hour traffic.

The peakedness factor indicates the degree to which the variability of the traffic deviates from pure chance traffic within a single hour, and in statistical terms is the variance-to-mean ratio of the distribution of simultaneous overflow traffic.

1 Determination of the level of day-to-day traffic variations

Let $M_1, M_2, \ldots, M_{30}$ denote the 30 busy-hour loads of the traffic offered to the final group. Determine the mean traffic M of the daily traffic by

$$M = \frac{1}{30} \sum_{j=1}^{30} M_j$$

Determine the sample variance V_d of the daily traffic by

$$V_d = \frac{1}{29} \sum_{j=1}^{30} (M_j - M)^2$$

Determine the point (M, V_d) on Figure 1/E.521; M on the horizontal axis, and V_d on the vertical axis.

i) If the point (M, V_d) is below the bottom curve, the level of variation is *Null.*

ii) If the point is between the lower two curves, the level of variation is *Low.*

iii) If the point is between the upper two curves, the level of variation is *Medium.*

iv) If the point is above the highest curve, the level of variation is *High.*

Default procedures: if the data are not available to compute the variance V_d use the following guidelines:

a) If no more than 25 per cent of the traffic offered to the final group is overflow from other groups, assume the level of day-to-day variation is Low.

b) Otherwise, assume a Medium level of variation.

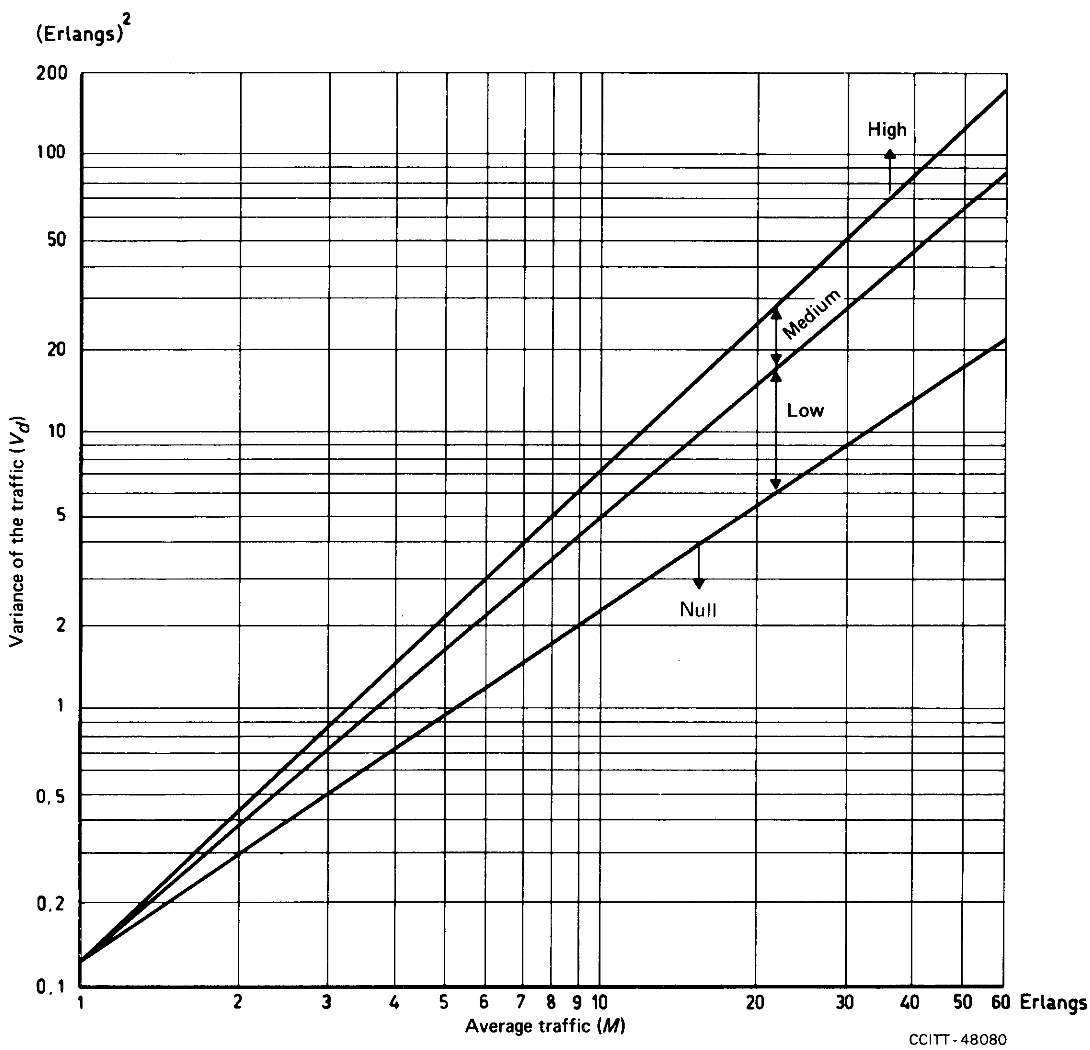


FIGURE 1/E.521

Determination of the level of day-to-day traffic variation

2 Determination of peakedness factor z

Peakedness factors depend principally upon the number of high-usage circuits over which random traffic has access. When the number of such high-usage circuits does not exceed 30, the actual peakedness of the traffic overflowing from a high-usage group will be only slightly below the maximum peakedness values[1), 2)]. The maximum peakedness values are given in Table 1/E.521.

TABLE 1/E.521

Maximum peakedness factor z_i

Number of high-usage circuits (n_i)	Peakedness factor (z_i)	Number of high-usage circuits (n_i)	Peakedness factor (z_i)
1	1.17	16	2.44
2	1.31	17	2.49
3	1.43	18	2.55
4	1.54	19	2.61
5	1.64	20	2.66
6	1.73	21	2.71
7	1.82	22	2.76
8	1.90	23	2.81
9	1.98	24	2.86
10	2.05	25	2.91
11	2.12	26	2.96
12	2.19	27	3.00
13	2.26	28	3.05
14	2.32	29	3.09
15	2.38	30	3.14

For more than 30 circuits, the peakedness of the traffic overflowing from a high-usage group i of n_i circuits is given by

$$z_i = 1 - \beta_i + \frac{A_i}{n_i + 1 + \beta_i - A_i}$$

where

A_i is the mean (random) traffic offered to the n_i circuits and

β_i is the traffic overflowing. The overflow traffic β_i is found by employing the standard Erlang loss formula $E_{1,\, n_i}\ (A_i)$:

$$\beta_i = A_i E_{1,\, n_i}(A_i).$$

The weighted mean peakedness factor z, is then calculated from:

$$z = \frac{\sum_{i=1}^{h} \beta_i z_i}{\sum_{i=1}^{h} \beta_i}$$

for the h parcels of traffic being offered to the final group.

Note that for the traffic directly offered to the final group, the peakedness factor is $z_i = 1$.

---

1) Tables giving:
- the exact mean of the overflow traffic, and
- the difference between variance and mean of the overflow

have been computed and are set out in [1].

2) Curves giving the exact mean and variance of overflow traffic are given in [2]. See also a more detailed description of the method in [3] and [4].

3 Determination of the mean traffic offered to the final group and the number of circuits required

3.1 For planning future network requirements, the traffic overflowing to a final group should be determined theoretically from forecasts of traffics offered to the high-usage groups.

The mean traffic overflowing to the final group from a high-usage group is determined in two steps:

i) the "single-hour" overflow traffic β_i overflowing from n_i circuits is given as above by

$$\beta_i = A_i E_{i,\, n_i}(A_i),$$

when A_i is the forecast of traffic offered to the i^{th} high-usage group;

ii) the average overflow traffic $\bar{\beta}_i$ overflowing from the n_i circuits is then determined by adjusting the single-hour traffic β_i for the effect of day-to-day traffic variations.

$$\bar{\beta}_i = r_i \, \beta_i$$

The adjustment factor r_i is given in Table 2/E.521; it is a function of:

- the offered traffic A_i,
- the traffic $A_i E_{i,\, n_i - 1}(A_i) - \beta_i$ carried by the last trunk i, and
- the level of day-to-day variations of the traffic offered to the high-usage group.

This level can be determined using the method described in § 1 above, but applying it to measurements of traffic offered to the high-usage group. If such measurements are not available a *medium* level can be used.

The mean traffic offered to the final group is then the sum of all $\bar{\beta}_i$ over the h parcels of traffic:

$$M = \sum_{i=1}^{h} \bar{\beta}_i$$

It can be assumed that the level of day-to-day traffic variations on the final group remains constant over the forecast time period.

Using the level of day-to-day traffic variation as determined in § 1 above on the final group and the peakedness factor of § 2 above, the appropriate table of Tables 3/E.521 to 6/E.521 is used to derive the number of circuits required.

Note 1 – This method of calculation of the mean traffic offered to the final group is valid only if the overflow traffic due to blocking encountered in the exchange in the attempts to connect to a high-usage, is negligible.

Note 2 – Table 3/E.521 differs slightly from the previous tables published by CCITT, although in Table 3.1/E.521 there is no allowance for day-to-day variations. The new table takes into account a systematic bias in the measurement procedure that is based on a finite period of time (1 hour), instead of an infinite period as was assumed in the previous table [5].

Note 3 – Tables 4/E.521, 5/E.521 and 6/E.521 are based on the calculation of the average blocking from the formula:

$$\bar{\beta} = \int B(m)\, f(m) dm,$$

where

$B(m)$ is the single-hour expected blocking and

$f(m)$ is the density distribution of day-to-day traffic (m), assuming a Pearson Type III distribution:

$$\left[f(m) = \frac{(M/V)^{(M^2/V_d)}}{\gamma\, ^{(M^2/V^d)}} m^{[(M^2/V_d)\, -\, 1]}\, e^{-M_m/V_d} \right]$$

M and V_d are the mean and day-to-day variance of the traffic as calculated [5] in § 1 above.

TABLE 2/E.521

Overflow adjustment for high-usage trunk groups
Factor r_i

Offered traffic A_i	Last trunk traffic														
	Low daily variation					Medium daily variation					High daily variation				
	0.25	0.3	0.4	0.5	0.6	0.25	0.3	0.4	0.5	0.6	0.25	0.3	0.4	0.5	0.6
3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
5	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.0	1.0	1.2	1.2	1.1	1.1	1.0
7	1.0	1.0	1.0	1.0	1.0	1.2	1.2	1.1	1.1	1.0	1.4	1.3	1.2	1.1	1.1
10	1.1	1.1	1.1	1.0	1.0	1.3	1.2	1.2	1.1	1.1	1.5	1.4	1.3	1.2	1.1
15	1.2	1.1	1.1	1.1	1.0	1.5	1.4	1.2	1.2	1.1	1.8	1.6	1.4	1.3	1.1
20	1.2	1.2	1.1	1.1	1.0	1.6	1.5	1.3	1.2	1.1	2.0	1.8	1.5	1.3	1.2
25	1.3	1.2	1.2	1.1	1.1	1.8	1.6	1.4	1.3	1.1	2.3	2.0	1.7	1.4	1.2
30	1.3	1.3	1.2	1.1	1.1	1.8	1.7	1.4	1.3	1.2	2.4	2.1	1.7	1.5	1.3

TABLE 3/E.521

Single-hour capacity, in Erlangs, as a function of the number of trunks and of the peakedness factor

Parameters: – Blockage 0.01;
– *No* allowance for day-to-day variation;
– Weighted mean peakedness factor.

Number of trunks required	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0	3.4	3.8	4.0
1	0.06	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.53	0.33	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.94	0.69	0.50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	1.42	1.14	0.89	0.67	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	1.97	1.64	1.36	1.08	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	2.56	2.19	1.86	1.58	1.31	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	3.19	2.81	2.44	2.11	1.81	1.53	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	3.83	3.42	3.03	2.67	2.36	2.03	1.75	1.50	0.0	0.0	0.0	0.0	0.0	0.0
10	4.53	4.08	3.67	3.28	2.92	2.58	2.28	2.00	1.75	0.0	0.0	0.0	0.0	0.0
11	5.22	4.75	4.31	3.89	3.53	3.17	2.83	2.53	2.25	1.97	0.0	0.0	0.0	0.0
12	5.94	5.44	4.97	4.56	4.14	3.78	3.42	3.08	2.78	2.47	2.22	0.0	0.0	0.0
13	6.67	6.14	5.64	5.19	4.81	4.39	4.03	3.67	3.33	3.03	2.72	0.0	0.0	0.0
14	7.42	6.86	6.36	5.89	5.44	5.03	4.67	4.28	3.94	3.61	3.28	2.69	0.0	0.0
15	8.17	7.58	7.06	6.58	6.11	5.69	5.31	4.92	4.56	4.19	3.86	3.22	0.0	0.0
16	8.94	8.33	7.78	7.28	6.81	6.36	5.94	5.56	5.17	4.81	4.44	3.81	3.19	0.0
17	9.72	9.08	8.50	8.00	7.50	7.06	6.61	6.19	5.81	5.42	5.06	4.39	3.75	3.44
18	10.50	9.83	9.25	8.72	8.22	7.75	7.31	6.86	6.44	6.06	5.69	4.97	4.31	4.00
19	11.31	10.61	10.00	9.44	8.92	8.44	7.97	7.53	7.11	6.72	6.33	5.58	4.89	4.58
20	12.08	11.39	10.78	10.19	9.67	9.14	8.67	8.22	7.81	7.39	6.97	6.22	5.50	5.17
21	12.89	12.19	11.53	10.94	10.39	9.86	9.39	8.92	8.47	8.06	7.64	6.86	6.11	5.78
22	13.72	13.00	12.31	11.69	11.14	10.61	10.08	9.61	9.17	8.72	8.31	7.50	6.75	6.39
23	14.53	13.78	13.08	12.47	11.89	11.36	10.81	10.33	9.86	9.42	8.97	8.17	7.39	7.00
24	15.36	14.58	13.89	13.22	12.64	12.08	11.56	11.03	10.56	10.11	9.67	8.83	8.03	7.64
25	16.19	15.39	14.67	14.00	13.39	12.83	12.28	11.78	11.28	10.81	10.36	9.50	8.69	8.31
26	17.03	16.22	15.47	14.81	14.17	13.58	13.03	12.50	12.00	11.53	11.06	10.19	9.36	8.94
27	17.86	17.03	16.28	15.58	14.94	14.33	13.78	13.22	12.72	12.22	11.75	10.86	10.03	9.61
28	18.69	17.86	17.08	16.36	15.72	15.11	14.53	13.97	13.44	12.94	12.47	11.56	10.69	10.28
29	19.56	18.69	17.89	17.17	16.50	15.86	15.28	14.72	14.19	13.67	13.19	12.28	11.39	10.94
30	20.39	19.53	18.72	17.97	17.28	16.64	16.06	15.47	14.92	14.42	13.92	12.97	12.08	11.64
31	21.25	20.36	19.53	18.78	18.08	17.42	16.81	16.22	15.67	15.14	14.64	13.69	12.78	12.33
32	22.11	21.19	20.36	19.58	18.89	18.22	17.58	17.00	16.42	15.89	15.36	14.39	13.47	13.03
33	22.97	22.06	21.19	20.39	19.67	19.00	18.36	17.75	17.19	16.64	16.11	15.11	14.17	13.72
34	23.83	22.89	22.00	21.22	20.47	19.81	19.14	18.53	17.94	17.39	16.86	15.86	14.89	14.42
35	24.69	23.75	22.83	22.03	21.28	20.58	19.92	19.31	18.69	18.14	17.61	16.58	15.61	15.14
36	25.58	24.58	23.69	22.86	22.11	21.39	20.72	20.08	19.47	18.89	18.36	17.31	16.31	15.83
37	26.44	25.44	24.53	23.69	22.92	22.19	21.50	20.86	20.25	19.67	19.11	18.06	17.06	16.56
38	27.31	26.31	25.36	24.53	23.72	23.00	22.31	21.64	21.03	20.44	19.86	18.81	17.78	17.28
39	28.19	27.17	26.22	25.36	24.56	23.81	23.11	22.44	21.81	21.19	20.64	19.53	18.50	18.00
40	29.08	28.03	27.06	26.19	25.39	24.61	23.89	23.22	22.58	21.97	21.39	20.28	19.25	18.72
41	29.94	28.89	27.92	27.03	26.19	25.44	24.69	24.03	23.36	22.75	22.17	21.06	19.97	19.47
42	30.83	29.75	28.78	27.86	27.03	26.25	25.53	24.81	24.17	23.53	22.94	21.81	20.72	20.19
43	31.72	30.64	29.61	28.72	27.86	27.08	26.33	25.61	24.94	24.31	23.69	22.56	21.47	20.94
44	32.61	31.50	30.47	29.56	28.69	27.89	27.14	26.42	25.75	25.11	24.50	23.33	22.22	21.69
45	33.50	32.39	31.33	30.42	29.53	28.72	27.94	27.22	26.56	25.89	25.28	24.08	22.97	22.42
46	34.39	33.25	32.19	31.25	30.39	29.56	28.78	28.03	27.33	26.69	26.06	24.86	23.72	23.17
47	35.28	34.14	33.08	32.11	31.22	30.39	29.58	28.86	28.14	27.47	26.83	25.64	24.47	23.92
48	36.17	35.00	33.94	32.97	32.06	31.22	30.42	29.67	28.94	28.28	27.64	26.42	25.25	24.69
49	37.06	35.89	34.81	33.81	32.92	32.06	31.25	30.47	29.75	29.08	28.42	27.19	26.00	25.44
50	37.97	36.78	35.67	34.67	33.75	32.89	32.08	31.31	30.58	29.89	29.22	27.97	26.78	26.19

TABLE 4/E.521

Single-hour capacity, in Erlangs, as a function of the number of trunks and of the peakedness factor

Parameters: – Blockage 0.01;
– *Low* day-to-day variation allowance;
– Weighted mean peakedness factor.

Number of trunks required	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0	3.4	3.8	4.0
1	0.06	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.53	0.33	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.94	0.69	0.50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	1.39	1.14	0.89	0.67	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	1.89	1.64	1.36	1.08	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	2.44	2.14	1.86	1.58	1.31	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	3.03	2.69	2.42	2.11	1.81	1.53	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	3.64	3.28	2.97	2.67	2.36	2.03	1.75	1.50	0.0	0.0	0.0	0.0	0.0	0.0
10	4.25	3.89	3.56	3.22	2.92	2.58	2.28	2.00	1.75	0.0	0.0	0.0	0.0	0.0
11	4.92	4.53	4.17	3.83	3.50	3.17	2.83	2.53	2.25	1.97	0.0	0.0	0.0	0.0
12	5.58	5.17	4.78	4.44	4.08	3.78	3.42	3.08	2.78	2.47	2.22	0.0	0.0	0.0
13	6.25	5.81	5.42	5.06	4.69	4.36	4.03	3.67	3.33	3.03	2.72	0.0	0.0	0.0
14	6.94	6.50	6.08	5.69	5.33	4.97	4.64	4.28	3.94	3.61	3.28	2.69	0.0	0.0
15	7.64	7.17	6.75	6.33	5.97	5.61	5.25	4.92	4.56	4.19	3.86	3.22	0.0	0.0
16	8.33	7.86	7.42	7.00	6.61	6.25	5.89	5.53	5.17	4.81	4.44	3.81	3.19	0.0
17	9.06	8.56	8.11	7.67	7.28	6.89	6.53	6.17	5.81	5.42	5.06	4.39	3.75	3.44
18	9.81	9.28	8.81	8.36	7.94	7.56	7.17	6.81	6.44	6.06	5.69	4.97	4.31	4.00
19	10.53	10.00	9.50	9.06	8.61	8.22	7.83	7.44	7.08	6.72	6.33	5.58	4.89	4.58
20	11.28	10.72	10.22	9.75	9.31	8.89	8.50	8.11	7.72	7.36	6.97	6.22	5.50	5.17
21	12.03	11.44	10.94	10.44	10.00	9.56	9.17	8.78	8.39	8.03	7.64	6.86	6.11	5.78
22	12.78	12.19	11.67	11.17	10.69	10.25	9.83	9.44	9.06	8.67	8.31	7.56	6.75	6.39
23	13.53	12.94	12.39	11.89	11.42	10.94	10.53	10.11	9.72	9.33	8.94	8.19	7.39	7.00
24	14.31	13.69	13.14	12.61	12.11	11.67	11.22	10.81	10.39	10.00	9.61	8.86	8.03	7.64
25	15.08	14.44	13.86	13.33	12.83	12.36	11.92	11.50	11.08	10.67	10.28	9.50	8.67	8.31
26	15.86	15.22	14.61	14.08	13.56	13.08	12.61	12.19	11.75	11.36	10.94	10.17	9.33	8.94
27	16.64	15.97	15.36	14.81	14.28	13.81	13.33	12.89	12.44	12.03	11.64	10.83	10.00	9.61
28	17.42	16.75	16.14	15.56	15.03	14.53	14.06	13.58	13.14	12.72	12.31	11.50	10.67	10.28
29	18.22	17.53	16.89	16.31	15.78	15.25	14.78	14.31	13.86	13.42	13.00	12.19	11.36	10.94
30	19.00	18.31	17.67	17.06	16.50	16.00	15.50	15.03	14.56	14.11	13.69	12.86	12.06	11.64
31	19.81	19.08	18.44	17.83	17.25	16.72	16.22	15.72	15.28	14.83	14.39	13.56	12.75	12.33
32	20.61	19.89	19.19	18.58	18.00	17.47	16.94	16.47	16.00	15.53	15.11	14.25	13.44	13.03
33	21.39	20.67	19.97	19.36	18.78	18.22	17.69	17.19	16.72	16.25	15.81	14.94	14.14	13.72
34	22.22	21.47	20.75	20.11	19.53	18.97	18.42	17.92	17.44	16.97	16.53	15.67	14.83	14.42
35	23.03	22.25	21.56	20.89	20.28	19.72	19.17	18.67	18.17	17.69	17.22	16.36	15.56	15.11
36	23.83	23.06	22.33	21.67	21.06	20.47	19.92	19.39	18.89	18.42	17.94	17.08	16.25	15.81
37	24.64	23.86	23.14	22.44	21.83	21.25	20.67	20.14	19.64	19.14	18.67	17.78	16.94	16.50
38	25.47	24.67	23.92	23.25	22.61	22.00	21.44	20.89	20.36	19.89	19.42	18.50	17.64	17.19
39	26.28	25.47	24.72	24.03	23.39	22.78	22.19	21.64	21.11	20.61	20.14	19.22	18.33	17.89
40	27.11	26.28	25.53	24.81	24.17	23.53	22.94	22.39	21.86	21.36	20.86	19.94	19.06	18.61
41	27.92	27.08	26.31	25.61	24.94	24.31	23.72	23.14	22.61	22.11	21.61	20.67	19.78	19.31
42	28.75	27.92	27.11	26.39	25.72	25.08	24.47	23.92	23.36	22.83	22.33	21.39	20.47	20.03
43	29.58	28.72	27.92	27.19	26.50	25.86	25.25	24.67	24.11	23.58	23.08	22.11	21.19	20.75
44	30.42	29.56	28.75	28.00	27.31	26.64	26.03	25.44	24.89	24.33	23.83	22.86	21.92	21.44
45	31.25	30.36	29.56	28.81	28.08	27.44	26.81	26.22	25.64	25.11	24.58	23.58	22.64	22.17
46	32.08	31.19	30.36	29.61	28.89	28.22	27.58	26.97	26.42	25.86	25.33	24.33	23.36	22.89
47	32.92	32.03	31.17	30.42	29.69	29.00	28.36	27.75	27.17	26.61	26.08	25.06	24.11	23.64
48	33.75	32.83	32.00	31.22	30.47	29.81	29.14	28.53	27.94	27.39	26.83	25.81	24.83	24.36
49	34.58	33.67	32.81	32.03	31.28	30.58	29.94	29.31	28.72	28.14	27.58	26.56	25.56	25.08
50	35.44	34.50	33.64	32.83	32.08	31.39	30.72	30.08	29.50	28.92	28.36	27.31	26.31	25.83

TABLE 5/E.521

Single-hour capacity, in Erlangs, as a function of the number of trunks and of the peakedness factor

Parameters: – Blockage 0.01;
– *Medium* day-to-day variation allowance;
– Weighted mean peakedness factor.

Number of trunks required	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0	3.4	3.8	4.0
1	0.06	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.53	0.33	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.94	0.69	0.50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	1.39	1.14	0.89	0.67	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	1.86	1.61	1.36	1.08	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	2.39	2.11	1.83	1.58	1.31	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	2.94	2.64	2.36	2.08	1.81	1.53	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	3.53	3.19	2.89	2.61	2.33	2.03	1.75	1.50	0.0	0.0	0.0	0.0	0.0	0.0
10	4.11	3.78	3.47	3.17	2.86	2.58	2.28	2.00	1.75	0.0	0.0	0.0	0.0	0.0
11	4.72	4.39	4.03	3.72	3.42	3.14	2.83	2.53	2.25	1.97	0.0	0.0	0.0	0.0
12	5.36	4.97	4.64	4.31	4.00	3.69	3.39	3.08	2.78	2.47	2.22	0.0	0.0	0.0
13	6.00	5.61	5.25	4.89	4.56	4.25	3.94	3.67	3.33	3.03	2.72	0.0	0.0	0.0
14	6.64	6.22	5.86	5.50	5.17	4.83	4.53	4.22	3.92	3.61	3.28	2.69	0.0	0.0
15	7.31	6.89	6.47	6.11	5.78	5.42	5.11	4.78	4.47	4.19	3.86	3.22	0.0	0.0
16	7.97	7.53	7.11	6.75	6.39	6.03	5.69	5.39	5.06	4.75	4.44	3.81	3.19	0.0
17	8.64	8.19	7.78	7.36	7.00	6.64	6.31	5.97	5.64	5.33	5.03	4.39	3.75	3.44
18	9.33	8.86	8.42	8.03	7.64	7.28	6.92	6.58	6.25	5.92	5.61	4.97	4.31	4.00
19	10.03	9.53	9.08	8.67	8.28	7.89	7.53	7.19	6.86	6.53	6.19	5.58	4.89	4.58
20	10.69	10.19	9.75	9.33	8.92	8.53	8.17	7.81	7.47	7.14	6.81	6.17	5.50	5.17
21	11.42	10.89	10.42	9.97	9.56	9.17	8.81	8.44	8.08	7.75	7.42	6.75	6.11	5.78
22	12.11	11.58	11.11	10.64	10.22	9.83	9.44	9.06	8.69	8.36	8.03	7.36	6.72	6.39
23	12.83	12.28	11.78	11.33	10.89	10.47	10.08	9.69	9.33	8.97	8.64	7.97	7.33	7.00
24	13.53	13.00	12.47	12.00	11.56	11.14	10.72	10.36	9.97	9.61	9.25	8.58	7.94	7.61
25	14.25	13.69	13.17	12.69	12.25	11.81	11.39	11.00	10.61	10.25	9.89	9.19	8.56	9.19
26	14.97	14.42	13.86	13.39	12.92	12.47	12.06	11.64	11.28	10.89	10.53	9.83	9.17	8.81
27	15.69	15.11	14.58	14.08	13.61	13.14	12.72	12.31	11.92	11.53	11.17	10.44	9.78	9.42
28	16.44	15.83	15.28	14.78	14.28	13.83	13.39	12.97	12.58	12.19	11.81	11.08	10.39	10.06
29	17.17	16.56	16.00	15.47	14.97	14.53	14.08	13.64	13.25	12.83	12.47	11.72	11.03	10.67
30	17.92	17.28	16.72	16.17	15.67	15.19	14.75	14.31	13.92	13.50	13.11	12.36	11.64	11.31
31	18.64	18.03	17.42	16.89	16.39	15.89	15.44	15.00	14.58	14.17	13.78	13.03	12.28	11.94
32	19.39	18.75	18.14	17.58	17.08	16.58	16.11	15.67	15.25	14.83	14.44	13.67	12.92	12.56
33	20.14	19.47	18.86	18.31	17.78	17.28	16.81	16.36	15.92	15.50	15.11	14.33	13.58	13.19
34	20.89	20.22	19.61	19.03	18.50	18.00	17.50	17.06	16.61	16.17	15.78	14.97	14.22	13.86
35	21.64	20.97	20.33	19.75	19.22	18.69	18.19	17.75	17.28	16.86	16.44	15.64	14.86	14.50
36	22.39	21.69	21.06	20.47	19.92	19.42	18.92	18.44	17.97	17.53	17.11	16.31	15.53	15.14
37	23.14	22.44	21.81	21.19	20.64	20.11	19.61	19.14	18.67	18.22	17.81	16.97	16.19	15.81
38	23.89	23.19	22.53	21.94	21.36	20.83	20.31	19.83	19.36	18.92	18.47	17.64	16.86	16.47
39	24.64	23.94	23.28	22.67	22.08	21.56	21.03	20.53	20.06	19.61	19.17	18.33	17.53	17.11
40	25.42	24.69	24.03	23.39	22.81	22.25	21.75	21.25	20.75	20.31	19.86	19.00	18.19	17.78
41	26.17	25.44	24.78	24.14	23.56	22.97	22.44	21.94	21.47	21.00	20.56	19.69	18.86	18.44
42	26.94	26.19	25.50	24.86	24.28	23.72	23.17	22.67	22.17	21.69	21.25	20.36	19.53	19.11
43	27.72	26.97	26.25	25.61	25.00	24.44	23.89	23.36	22.86	22.39	21.94	21.06	20.19	19.81
44	28.47	27.72	27.00	26.36	25.75	25.17	24.61	24.08	23.58	23.08	22.64	21.75	20.89	20.47
45	29.25	28.47	27.78	27.11	26.47	25.89	25.33	24.81	24.31	23.81	23.33	22.44	21.56	21.14
46	30.03	29.25	28.53	27.86	27.22	26.64	26.06	25.53	25.00	24.50	24.03	23.14	22.25	21.83
47	30.81	30.00	29.28	28.61	27.97	27.36	26.78	26.25	25.72	25.22	24.75	23.83	22.94	22.50
48	31.58	30.78	30.03	29.36	28.72	28.11	27.53	26.97	26.44	25.94	25.44	24.53	23.64	23.19
49	32.36	31.56	30.81	30.11	29.44	28.83	28.25	27.69	27.17	26.64	26.17	25.22	24.33	23.89
50	33.14	32.31	31.56	30.86	30.19	29.58	29.00	28.42	27.89	27.36	26.86	25.92	25.03	24.58

TABLE 6/E.521

Single-hour capacity, in Erlangs, as a function of the number of trunks and of the peakedness factor

Parameters:
– Blockage 0.01;
– *High* day-to-day variation allowance;
– Weighted mean peakedness factor.

Number of trunks required	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0	3.4	3.8	4.0
1	0.06	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.53	0.33	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.94	0.69	0.50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	1.36	1.14	0.89	0.67	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	1.86	1.61	1.36	1.08	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	2.36	2.08	1.83	1.58	1.31	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	2.89	2.61	2.33	2.06	1.81	1.53	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	3.44	3.14	2.86	2.58	2.31	2.03	1.75	1.50	0.0	0.0	0.0	0.0	0.0	0.0
10	4.03	3.69	3.39	3.11	2.83	2.56	2.28	2.00	1.75	0.0	0.0	0.0	0.0	0.0
11	4.61	4.25	3.94	3.64	3.36	3.08	2.81	2.53	2.25	1.97	0.0	0.0	0.0	0.0
12	5.19	4.83	4.50	4.19	3.89	3.61	3.33	3.06	2.78	2.47	2.22	0.0	0.0	0.0
13	5.81	5.42	5.08	4.78	4.44	4.17	3.86	3.58	3.31	3.03	2.72	0.0	0.0	0.0
14	6.42	6.03	5.67	5.33	5.03	4.72	4.42	4.14	3.83	3.58	3.28	2.69	0.0	0.0
15	7.03	6.64	6.28	5.92	5.61	5.28	4.97	4.69	4.39	4.11	3.83	3.22	0.0	0.0
16	7.67	7.25	6.86	6.53	6.19	5.86	5.56	5.25	4.94	4.67	4.36	3.81	3.19	0.0
17	8.31	7.86	7.47	7.11	6.78	6.44	6.11	5.81	5.50	5.22	4.92	4.36	3.75	3.44
18	8.94	8.50	8.11	7.72	7.36	7.03	6.69	6.39	6.08	5.78	5.47	4.89	4.31	4.00
19	9.58	9.14	8.72	8.33	7.97	7.64	7.31	6.97	6.64	6.33	6.03	5.44	4.89	4.58
20	10.22	9.78	9.36	8.94	8.58	8.22	7.89	7.56	7.22	6.92	6.61	6.00	5.44	5.14
21	10.89	10.42	9.97	9.58	9.19	8.83	8.50	8.14	7.83	7.50	7.19	6.58	6.00	5.69
22	11.53	11.06	10.61	10.22	9.83	9.44	9.08	8.75	8.42	8.08	7.78	7.17	6.56	6.25
23	12.19	11.72	11.28	10.83	10.44	10.06	9.69	9.36	9.00	8.67	8.36	7.72	7.14	6.83
24	12.86	12.36	11.92	11.47	11.08	10.69	10.31	9.94	9.61	9.28	8.94	8.31	7.69	7.39
25	13.53	13.03	12.56	12.11	11.69	11.31	10.94	10.56	10.22	9.89	9.56	8.92	8.28	7.97
26	14.19	13.69	13.22	12.75	12.33	11.94	11.56	11.19	10.83	10.47	10.14	9.50	8.86	8.56
27	14.89	14.36	13.86	13.42	12.97	12.58	12.19	11.81	11.44	11.08	10.75	10.08	9.44	9.14
28	15.56	15.03	14.53	14.06	13.64	13.22	12.81	12.42	12.06	11.69	11.36	10.69	10.03	9.72
29	16.25	15.69	15.19	14.72	14.28	13.86	13.44	13.06	12.69	12.33	11.97	11.31	10.64	10.31
30	16.92	16.36	15.86	15.36	14.92	14.50	14.08	13.69	13.31	12.94	12.58	11.89	11.22	10.92
31	17.61	17.06	16.53	16.03	15.58	15.14	14.72	14.33	13.94	13.56	13.19	12.50	11.83	11.50
32	18.31	17.72	17.19	16.69	16.22	15.78	15.36	14.94	14.56	14.19	13.83	13.11	12.44	12.11
33	18.97	18.42	17.86	17.36	16.89	16.44	16.00	15.58	15.19	14.81	14.44	13.72	13.06	12.69
34	19.67	19.08	18.53	18.03	17.56	17.08	16.67	16.25	15.83	15.44	15.08	14.36	13.67	13.31
35	20.36	19.78	19.22	18.69	18.22	17.75	17.31	16.89	16.47	16.08	15.69	14.97	14.28	13.92
36	21.06	20.47	19.89	19.36	18.89	18.42	17.97	17.53	17.11	16.72	16.33	15.61	14.89	14.53
37	21.75	21.14	20.58	20.06	19.56	19.08	18.61	18.19	17.78	17.36	16.97	16.22	15.50	15.14
38	22.44	21.83	21.25	20.72	20.22	19.72	19.28	18.83	18.42	18.00	17.61	16.86	16.14	15.78
39	23.17	22.53	21.94	21.39	20.89	20.39	19.94	19.50	19.06	18.64	18.25	17.50	16.75	16.39
40	23.86	23.22	22.64	22.08	21.56	21.06	20.58	20.14	19.72	19.31	18.89	18.11	17.39	17.00
41	24.56	23.92	23.33	22.75	22.22	21.75	21.25	20.81	20.36	19.94	19.53	18.75	18.00	17.64
42	25.28	24.61	24.00	23.44	22.92	22.42	21.92	21.47	21.03	20.58	20.19	19.39	18.64	18.29
43	25.97	25.31	24.69	24.14	23.58	23.08	22.58	22.14	21.67	21.25	20.83	20.03	19.28	18.89
44	26.67	26.03	25.39	24.81	24.28	23.75	23.25	22.78	22.33	21.92	21.47	20.67	19.89	19.53
45	27.39	26.72	26.08	25.50	24.94	24.44	23.94	23.44	23.00	22.56	22.14	21.33	20.53	20.17
46	28.08	27.42	26.78	26.19	25.64	25.11	24.61	24.14	23.67	23.22	22.78	21.97	21.17	20.81
47	28.81	28.14	27.47	26.89	26.33	25.81	25.28	24.81	24.33	23.89	23.44	22.61	21.81	21.44
48	29.53	28.83	28.19	27.58	27.00	26.47	25.97	25.47	25.00	24.56	24.11	23.28	22.47	22.08
49	30.22	29.53	28.89	28.28	27.69	27.17	26.64	26.14	25.67	25.19	24.75	23.92	23.11	22.72
50	30.94	30.25	29.58	28.97	28.39	27.83	27.31	26.81	26.33	25.86	25.42	24.58	23.75	23.36

3.2 *Computer implementation*

When computer facilities are available, it is possible to automate the use of Tables 3/E.521 to 6/E.521. For that purpose, numerical algorithms have been developed and are described in [5].

4 Example

4.1 *Level of day-to-day traffic variations*

If the traffics offered to a final group over the 30 busiest days are given (M_1 to M_{30}) and if the mean load and variance are calculated to be 10 and 20 respectively, then applying Figure 1/E.521, a ***high*** level of day-to-day traffic variations should be used.

4.2 *Future traffic offered to the final group and peakedness factor*

If the forecast of future traffics indicates that three parcels of traffic will be offered to the final group:

- the overflow from 6 circuits offered 7.8 Erlangs,
- the overflow from 12 circuits offered 10 Erlangs,
- 7 Erlangs offered directly,

then Table 7/E.521 can be developed.

TABLE 7/E.521

Number of parcels of traffic i	Traffic offered to high-usage groups A_i	Number of high-usage circuits n_i	Single-hour overflow β_i	Last trunk traffic	Peakedness factor z_i	$\beta_i z_i$	Adjustment factor r_i	Average overflow $\bar{\beta}_i = r_i \beta_i$
1	7.8	6	2.95	0.69	1.73	5.1	1.0	2.95
2	10.0	12	1.20	0.44	2.19	2.6	1.2	1.44
3	7.0	0	7.0	–	1.0	7.0	1.0	7.00
						14.7		
			$\sum_{i=1}^{h} \beta_i = 11.15$			$z = \dfrac{\sum_{i=1}^{h} \beta_i z_i}{\sum_{i=1}^{h} \beta_i} = \dfrac{14.7}{11.15} = 1.3$		$M = \sum_{i=1}^{h} \bar{\beta}_i = 11.39$

Note that the values of r_i are derived from Table 2/E.521 for *medium* level of day-to-day traffic variations; if the 30 busiest day traffics for each of the high-usage groups were available, a more appropriate level could be used for each group.

Now all the information required is available: using the capacity Table 6/E.521 for ***high*** level of day-to-day traffic variations, the average traffic offered to the final group $M = 11.39$ and a peakedness factor $z = 1.3$ (from interpolating between $z = 1.2$ and $z = 1.4$), it is calculated that 23 circuits are required.

Note that if the measurements used in § 4.1 above were not available, then to determine the level of day-to-day traffic variations it would have been necessary to use the default procedure of § 1 above.

Overflow traffic offered to the final group = 4.15 Erlangs.

Total traffic offered to the final group = 11.15 Erlangs.

The ratio 4.15/11.15 = 0.37 is higher than 0.25 and hence a *medium* level of day-to-day traffic variations would have been used.

References

[1] *Tabellen für die Planung von Fernsprecheinrichtungen*, Siemens u. Halske, München, 1961.

[2] WILKINSON (R. I.): Theories for toll traffic engineering in the USA (Figures 12 and 13), *Bell System Technical Journal*, Vol. 35, March 1956.

[3] WILKINSON (R. I.): Simplified engineering of single stage alternate routing systems, *Fourth International Teletraffic Congress*, London, 1964.

[4] WILKINSON (R. I.): Non-random traffic curves and tables, *Bell Telephone Laboratories*, 1970.

[5] HILL (D. W.) and NEAL (S. R.): The traffic capacity of a probability-engineered trunk group, *Bell System Technical Journal*, September 1976.

Recommendation E.522

NUMBER OF CIRCUITS IN A HIGH-USAGE GROUP

1 Introduction

For the economic planning of an alternate routing network the number of circuits in a high-usage group should be determined so that the annual charges for the whole network arrangement are at a minimum. This is done under the constraint that given requirements for the grade of service are fulfilled. In the optimum arrangement, the cost per erlang of carrying a marginal amount of traffic over the high-usage route or over the alternative route is the same.

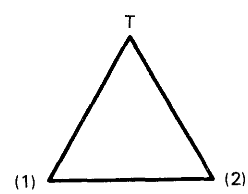


FIGURE 1/E.522

The optimum number of high-usage circuits, n, from one exchange (1) to another exchange (2) is therefore obtained from the following expression when the overflow traffic is routed over a transit exchange T (route 1-T-2, see Figure 1/E.522).

$$F_n(A) = A\,\{E_{1,\,n}(A) - E_{1,\,(n+1)}(A)\} = M \times \frac{\text{annual charge (1-2)}}{\text{annual charge (1-T-2)}}$$

A is the traffic flow offered, for the relation "1-2", in the Erlang loss formula for a full availability group. The expression $F_n(A)$ gives the marginal occupancy[1)] (improvement function) for the high-usage group, if one more circuit were added.

M is the *marginal utilization factor*[2)] *for the final route* "1-T-2" (which has nothing to do with cost ratio), if one additional circuit were provided. The annual charges are marginal charges for adding one additional circuit to route "1-2" and likewise to route "1-T-2".

Planning of an alternate routing network is described in the technical literature (see [1] to [10]).

Annual charge as used in this Recommendation refers to investment costs.

[1)] Marginal occupancy is often called LTC (last trunk capacity).

[2)] Marginal utilization factor is often called ATC (additional trunk capacity).

2 Recommended practical method

2.1 *Field of application*

It must be recognized that the conditions applying to alternative routing will vary widely between the continental network and the intercontinental network. Significant differences between the two cases apply to the length and cost of circuits, the traffic flow and the different times at which the busy hours occur. The method described attempts to take account of these factors in so far as it is practicable to do so in any simplified procedure.

2.2 *Traffic statistics*

The importance of reliable traffic estimates should be emphasized. Traffic estimates are required for each of the relations in question, for both the busy hour of the relation and for the busy hour of each link of the routes to which the traffic overflows. Since this may be affected by the high-usage arrangements finally adopted, it will be necessary to have traffic estimates for each relation covering most of the significant hours of the day. This applies particularly to the intercontinental network where the final routes carry traffic components with widely differing busy hours.

2.3 *Basis of the recommended method*

The method is based on a simplification of the economic dimensioning equations described under 1. Introduction. The simplifying assumptions are:

i) the ratios of the alternative high-usage annual charges are grouped in classes and a single ratio selected as representative for each class. This is acceptable because total network costs are known to be relatively insensitive to changes in the annual charges ratio;

ii) the marginal utilization factor M applicable to the overflow routes is regarded as constant within a range of circuit group sizes;

Size of group (number of circuits)	Value of M
For less than 10	0.6
For 10 or more	0.8

iii) each high-usage group will be dimensioned against the cheapest alternative route to which traffic overflows. (That is, the effect of parallel alternative routes is ignored.)

Where greater precision is required in either network or individual route dimensioning, more sophisticated methods may be employed (see [5] and [7]).

2.4 *Determination of cost ratio*

In continental and intercontinental working, the number of circuits to be provided in high-usage circuit groups depends upon the ratio of the annual charges estimated by the Administrations involved. The annual charge ratio (see Table 1/E.522) is defined as:

$$R = \frac{\text{annual charge of one additional circuit on the alternative route}}{\text{annual charge of one additional circuit on the high-usage route}}$$

The "annual charge of one additional circuit on the alternative route" is calculated by summing:

– the annual charge per circuit of each link comprising the alternative route, and

– the annual charge of switching one circuit at each intermediate switching centre.

When a third Administration is involved, it may be necessary to calculate the annual charge for switching at the intermediate centre from the transit switching charge per holding minute[3]. This may be done as follows:

Annual charges for switching = $M \times 60 \times F \times 26 \times 12 \times$ transit switching charge per holding minute.

In the calculation of the conversion factor F from busy hour to day, its dependence on the traffic offered to the high usage route, the overflow probability and the time difference should be taken into account. As a guideline, Table 1/E.522, which is calculated using the standard traffic profiles of Table 1/E.523, may be used.

TABLE 1/E.522

Offered traffic (erlangs)	Overflow probability (%)	Time difference												
		0	1	2	3	4	5	6	7	8	9	10	11	12
	1	2.6	3.2	3.7	3.8	2.7	2.3	2.3	1.7	3.2	2.4	2.2	2.0	2.7
	10	3.7	4.5	4.8	4.7	3.5	3.1	3.0	2.5	4.1	3.2	2.9	2.8	3.6
	20	4.5	5.2	5.4	5.3	4.0	3.7	3.5	3.1	4.7	3.8	3.4	3.4	4.2
5	30	5.1	5.8	6.0	5.8	4.6	4.2	4.0	3.7	5.1	4.3	3.9	4.0	4.8
	40	5.7	6.4	6.5	6.3	5.1	4.7	4.5	4.2	5.6	4.8	4.4	4.6	5.3
	50	6.3	6.9	7.0	6.8	5.6	5.2	5.0	4.7	6.0	5.3	5.0	5.1	5.8
	1	2.1	2.6	3.3	3.5	2.5	2.1	2.1	1.4	2.8	2.0	2.0	1.8	2.4
	10	3.2	4.0	4.4	4.3	3.1	2.7	2.6	2.1	3.8	2.8	2.6	2.4	3.2
	20	4.0	4.8	5.1	4.9	3.6	3.3	3.1	2.7	4.3	3.4	3.0	3.0	3.8
10	30	4.7	5.4	5.6	5.4	4.2	3.8	3.6	3.3	4.8	3.9	3.4	3.6	4.4
	40	5.3	6.0	6.1	5.9	4.7	4.4	4.2	3.8	5.3	4.4	4.0	4.2	4.9
	50	5.9	6.6	6.7	6.4	5.3	4.9	4.7	4.4	5.7	5.0	4.6	4.8	5.5
	1	1.6	2.0	2.8	3.1	2.2	1.8	2.0	1.2	2.4	1.7	1.8	1.6	2.1
	10	2.7	3.3	3.9	3.9	2.7	2.4	2.3	1.7	3.3	2.4	2.3	2.0	2.7
	20	3.5	4.2	4.6	4.4	3.2	2.8	2.7	2.2	3.9	3.0	2.6	2.5	3.3
25	30	4.2	5.0	5.2	5.0	3.7	3.4	3.2	2.8	4.4	3.5	3.0	3.1	3.9
	40	4.8	5.6	5.8	5.5	4.3	3.9	3.8	3.4	4.9	4.0	3.5	3.7	4.5
	50	5.5	6.2	6.3	6.1	4.9	4.5	4.3	4.0	5.4	4.6	4.1	4.4	5.1
	1	1.3	1.7	2.4	2.9	2.1	1.6	2.0	1.1	2.1	1.5	1.6	1.4	2.0
	10	2.3	2.8	3.5	3.6	2.5	2.2	2.1	1.4	3.1	2.2	2.2	1.8	2.4
	20	3.1	3.9	4.3	4.2	3.0	2.6	2.4	1.9	3.7	2.7	2.5	2.2	3.0
50	30	3.9	4.7	5.0	4.8	3.4	3.1	2.9	2.5	4.2	3.3	2.8	2.8	3.6
	40	4.6	5.4	5.6	5.3	4.0	3.7	3.5	3.2	4.7	3.8	3.2	3.5	4.3
	50	5.3	6.0	6.1	5.9	4.7	4.3	4.2	3.8	5.2	4.3	3.8	4.2	4.9

Note – Linear interpolation may be used to obtain intermediate results.

[3] It may be necessary to calculate transit switching charge per holding minute from charge per conversation minute (efficiency factor is described in Recommendation E.506).

The value determined for R should then be employed to select in Table 2/E.522 the precise (or next higher) value of annual charges ratio for use in traffic tables. The value of annual charges ratios may be grouped in the following general sets:

a) Within a single continent or other smaller closely connected land mass involving distances up to 1000 miles, high traffic and frequently one-way operation:

Annual charges ratio: R = 1.5; $\underline{\underline{2.0}}$; 3.0; 4.0; 5.0; 6.0 and 7.0 [4)]

b) Intercontinental working involving long distances, small traffic and usually two-way operation:

Annual charges ratio: R = 1.1; $\underline{\underline{1.3}}$; 1.5; 2.0; 3.0; 4.0 and 5.0. [4)]

2.5 *Use of method*

High-usage circuit groups carrying random traffic can be dimensioned from Table 2/E.522.

Step 1 – Estimate the annual charges ratio R as described under 2.4 above. (There is little difference between adjacent ratios.) If this ratio is difficult to estimate, the values underlined in a) and b) of § 2.4 above, should be used.

Step 2 – Consult Table 2/E.522 to determine the number of high-usage circuits N.

Note – When two values of N are given the right-hand figure applies to alternative routes of more than 10 circuits, the left-hand figure applies to smaller groups. The left-hand figure is omitted when it is no longer possible for the alternative route to be small.

3 **24-hour traffic profiles**

The traffic value used in the method in § 2 should be the value of traffic offered to the high-usage route during the busy hour of the final route. In the case that some of the busy hours of the circuit groups or links forming an alternative route do not coincide with the busy hour of the relation, the ensuing method should be followed to take 24-hour traffic profiles into account (see [6], [8] and [9]).

The method consists of the following three basic steps:

i) prepare hourly traffic demands for which dimensioning is to be done;

ii) size all circuit groups, high usage and final, for one hourly traffic demand;

iii) iterate the process in step ii) for each additional hourly matrix.

3.1 *Preparation of hourly traffic demands*

Each Administration gathers historical traffic data on an hourly basis in accordance with Recommendations E.500 and E.523. Using historical data and information contained in Recommendation E.506, hourly traffic demand forecasts are made, resulting in a series of hourly demands for each exchange to every other exchange.

3.2 *Sizing circuit groups for one-hourly traffic demand*

Using the methods in § 2 and Recommendation E.521, trunk group sizes are prepared for the first hourly traffic demand disregarding other hourly traffic demands.

---

4) These values are tentative. Ranges and representative values of annual charges ratio require further study.

TABLE 2/E.522

Number of high-usage circuits for different values of offered traffic, annual charges ratios and sizes of overflow groups

Traffic offered during network busy hour (erlangs)	Annual charges ratios									Number of circuits if there is no overflow route, for $p = 0.01$
	1.1	1.3	1.5	2.0	3.0	4.0	5.0	6.0	7.0	
	Minimum circuit occupancies for high-usage traffic									
	0.545/0.727	0.46/0.615	0.4/0.53	0.3/0.4	0.2/0.26	0.15/0.2	0.12/0.16	0.1/0.13	0.085/0.114	
	N, number of high usage circuits *A/B*, where *A* is for less than 10 circuits in the overflow group ($M = 0.6$) *B* is for 10 or more circuits in the overflow group ($M = 0.8$)									
1.5	1/0	1/0	2/1	2/2	3/2	3/3	4/3	4/3	4/4	6
1.75	1/0	2/1	2/1	3/2	3/3	4/3	4/4	4/4	4/4	6
2.0	1/0	2/1	2/2	3/2	4/3	4/4	4/4	5/4	5/5	7
2.25	2/0	2/1	3/2	3/3	4/4	5/4	5/4	5/5	5/5	7
2.5	2/0	3/1	3/2	4/3	5/4	5/5	5/5	6/5	6/5	7
2.75	2/1	3/2	3/2	4/3	5/4	5/5	6/5	6/6	6/6	8
3.0	3/1	3/2	4/3	4/4	5/5	6/5	6/6	6/6	7/6	8
3.5	3/1	4/2	4/3	5/4	6/5	7/6	7/6	7/7	7/7	9
4.0	4/2	4/3	5/4	6/5	7/6	7/7	8/7	8/7	8/8	10
4.5	4/2	5/3	6/4	6/6	7/7	8/7	8/8	9/8	9/8	10
5.0	5/3	6/4	6/5	7/6	8/7	9/8	9/9	9/9	10/9	11
5.5	5/3	6/5	7/5	8/7	9/8	9/9	10/9	10/10	10/10	12
6.0	6/3	7/5	7/6	8/7	9/9	10/9	11/10	11/10	11/11	13
7.0	7/4	8/6	8/7	10/8	11/10	11/11	12/11	12/12	13/12	14
8.0	8/5	9/7	10/8	11/10	12/11	13/12	13/13	14/13	14/13	15
9.0	/6	/8	/9	/11	/12	/13	/14	/14	/15	17
10.0	/7	/9	/10	/12	/14	/15	/15	/16	/16	18
12.0	/9	/11	/12	/14	/16	/17	/18	/18	/19	20
15.0	/12	/14	/16	/18	/20	/21	/21	/22	/22	24
20.0	/16	/19	/21	/23	/25	/27	/28	/28	/29	30
25.0	/21	/24	/26	/29	/31	/33	/33	/34	/35	36
30.0	/26	/29	/31	/34	/37	/38	/39	/40	/41	42
40.0	/36	/39	/42	/45	/48	/50	/51	/52	/52	53
50.0	/45	/49	/52	/55	/59	/61	/62	/63		64
60.0	/55	/60	/62	/66	/70	/72	/73			75
80.0	/74	/80	/83	/87	/92	/94	/95			96
100.0	/94	/100	/103	/108	/113	/116				117
120.0	/113	/120	/124	/129	/134	/137	Direct final circuit groups are economical within this area.			138
150.0	/143	/150	/154	/160	/166	/169				170
200.0	/192	/200	/205	/212	/219					221
250.0	/241	/250	/256	/263	/271					273
300.0	/290	/300	/306	/315	/323					324

3.3 *Iterating for each additional hourly traffic matrix*

In sizing the circuit groups for the second hourly traffic demand, the method is provided with the circuit quantities resulting from the previous step, and is constrained solely to increasing circuit group sizes; i.e., if the circuit group sizes for the first hourly traffic demand were greater than for the second hourly demand, then the circuit group sizes for the first hourly traffic demand would be retained.

All additional hourly traffic demands are processed in the same iterative manner. The resulting circuit group sizes then satisfy the traffic demands for all hours being considered (see Annex A for a computational example).

3.4 *Processing sequence*

Processing may start with the first hour of traffic demand, however, experiments have indicated that efficiencies of the network can be improved if processing starts with the hour with the smallest total traffic demand. It should be noted that this method gives us suboptimal networks, which may be improved by manual refinements.

4 Minimum outlay alternate routing networks

The method below allows Administrations to adjust alternate routing networks to take into account existing revenue accounting divisions.

The method consists of the following steps:

i) Obtain 24-hour traffic profiles in accordance with Recommendations E.500 and E.523;

ii) Compute circuit quantities and costs for a no-overflow network in accordance with Recommendation E.520;

iii) Compute monthly overflow minutes (holding time) at varying percentages of busy-hour overflow. This is done by applying three conversion factors to the busy hour overflow erlangs:

- Ratio of holding minutes to erlangs: a fixed value of 60.
- Daily overflow to busy-hour overflow ratio: a value that depends on the 24-hour traffic profile and the degree of overflow.
- Monthly overflow to daily overflow ratio (Recommendation E.506): a value that depends on the day-to-day pattern within a month and the degree of overflow.

iv) Starting with the network calculated in step ii):

- reduce the high usage circuits by one circuit,
- calculate overflow to final circuit groups,
- dimension final circuit groups in accordance with Recommendation E.521,
- calculate circuit costs and transit charges;

v) Iterate step iv) until the minimum outlay (circuit costs plus transit charges) for terminal administrations is reached (see Annex B for computational example).

5 Service considerations

On intercontinental circuits, where both-way operation is employed, a minimum of two circuits may be economical. Service considerations may also favour an increase in the number of direct circuits provided, particularly where the annual charges ratio approaches unity.

Although the dimensioning of high-usage groups is normally determined by traffic flows and annual charges ratios, it is recognized that such groups form part of a network having service requirements relative to the subscriber. The ability to handle the offered traffic with acceptable traffic efficiency should be tempered by the overall network considerations on quality of service.

The quality of service feature, which is of primary importance in a system of high-usage and final circuit groups, is the advantage derived from direct circuits versus multi-link connections. A liberal use of direct high-usage circuit groups, taking into account the economic factors, favours a high quality of service to the subscriber. It is recommended that new high-usage groups should be provided whenever the traffic flow and cost ratios are not conclusive. This practice may result in direct high-usage groups of two circuits or more.

The introduction of high-usage groups improves the overall grade of service and provides better opportunities of handling traffic during surges and breakdown conditions. When high-usage links bypass the main final routes the introduction of high-usage routes can assist in avoiding expenses which might otherwise be incurred in keeping below the maximum number of long-distance links in series. In the future, more measurements of traffic flows may be necessary for international accounting purposes and high-usage circuits should make this easier.

ANNEX A

(to Recommendation E.522)

Example of network dimensioning taking into account 24-hour traffic profiles

A.1 *Assumptions* (see also Figure A-1/E.522)

Calculations are performed under the following conditions:

1) Time difference:

   A is 9 hours west of B
   C is 5 hours west of A
   B is 10 hours west of C

2) Traffic profiles:

   24-hour traffic profiles as per Table 1/E.523 are used.

3) Busy hour traffic:

   A-B 50 erlangs
   A-C 100 erlangs
   C-B 70 erlangs

4) Cost ratio:

   $R = 1.3$

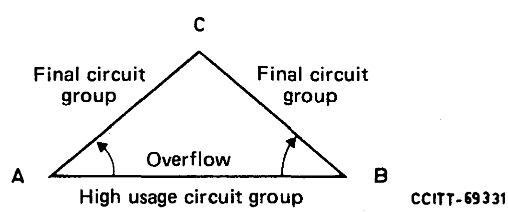


FIGURE A-1/E.522

Triangular network for numerical examples (Example 1)

A.2 *Numerical results*

24 hourly traffic demands are processed. The order of processing are from the hour with the smallest total traffic demand to the hour with the largest total traffic demand. Computational results are given in Table A-1/E.522.

TABLE A-1/E.522

Numerical results

Hour	Hourly traffic demand			Number of circuits obtained by single hour dimensioning (disregarding lower bounds imposed by the previous iterative stage)			Number of circuits obtained considering lower bounds imposed by the previous iterative stage			Number of circuits required to meet multiple hourly traffic demands		
	A-B	A-C	C-B	A-B	A-C	C-B	A-B	A-C	C-B	A-B	A-C	C-B
6	17.50	5.00	3.50	17	19	17	17	19	17	17	19	17
7	20.00	5.00	3.50	19	20	18	19	20	18	19	20	18
5	2.50	5.00	28.00	1	14	41	19	11	39	19	20	39
4	2.50	5.00	35.00	1	14	49	19	11	47	19	20	47
8	37.50	5.00	3.50	37	23	22	19	38	37	19	38	47
9	40.00	5.00	3.50	39	24	23	19	41	40	19	41	47
3	2.50	5.00	45.50	1	14	61	19	11	59	19	41	59
18	2.50	50.00	3.50	1	66	12	19	64	9	19	64	59
10	50.00	5.00	3.50	49	26	25	9	61	59	19	64	59
19	2.50	60.00	3.50	1	77	12	19	75	9	19	75	59
20	2.50	60.00	3.50	1	77	12	19	75	9	19	75	59
22	12.50	30.00	24.50	12	45	39	12	45	39	19	75	59
2	2.50	5.00	63.00	1	14	80	19	11	78	19	75	78
17	2.50	70.00	3.50	1	87	12	19	85	9	19	85	78
1	2.50	5.00	70.00	1	14	87	19	11	85	19	85	85
23	20.00	20.00	42.00	19	36	60	19	36	60	19	85	85
11	47.50	25.00	17.50	47	46	38	3	85	77	19	85	85
21	12.50	55.00	24.50	12	73	39	12	73	39	19	85	85
12	42.50	30.00	21.00	42	50	41	3	85	76	19	85	85
16	2.50	90.00	3.50	1	109	12	19	107	9	19	107	85
0	20.00	20.00	66.50	19	36	87	19	36	87	19	107	87
13	30.00	65.00	35.00	29	86	54	5	107	76	19	107	87
15	17.50	100.00	28.00	17	121	44	19	120	43	19	120	87
14	27.50	95.00	38.50	27	117	57	19	124	64	19	124	87

This example relates to an intercontinental network where busy hours of the three traffic relations are widely different among each other. The busy hour of the relation A-C, i.e. hour 15, is a low traffic period for the relations A-B and C-B. The busy hour of the relation C-B, i.e. hour 1, is a low traffic period for the relations A-B and A-C. Similarly, the busy hour of the relation A-B, i.e. hour 10, is a low traffic period for the relations A-C and C-B.

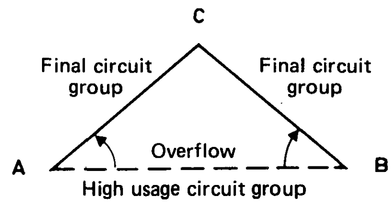
In this case, the single hour dimensioning method, where traffic data during the busy hour of the final circuit group are used for dimensioning, cannot be applied. If the single hour dimensioning method is applied, this results in considerable under-dimensioning.

If all the circuit groups are dimensioned as final, the required number of circuits are 64, 117 and 85 for the circuit groups A-B, A-C and C-B, respectively. About 14% of the total number of circuits is saved by the use of alternate routing.

ANNEX B

(to Recommendation E.522)

Example of minimum outlay network dimensioning



CCITT-69321

FIGURE B-1/E.522

Triangular network for numerical example (Example 2)

B.1 *Assumptions* (see also Figure B-1/E.522)

Calculations are performed under the following conditions:

1) Time difference:

A is 3 hours west of B
A is 3 hours west of C
No time difference between B and C

2) Traffic profiles:

24-hour traffic profiles as per Table 1/E.523 are used.

3) Busy hour traffic:

A-B 16 erlangs
A-C 33 erlangs
C-B 33 erlangs

4) Each Administration monthly cost per circuit:

A-B 1000 units
A-C 1000 units
C-B 800 units

5) Transit charge per holding minute to each terminal Administration:

1/2 unit

6) Conversion factors:

i) Holding minutes/erlangs: 60

ii) Daily overflow/busy hour overflow

This conversion factor (F) is calculated according to the guideline given in § 2.4.

iii) Monthly overflow/daily overflow: 26

where medium social contact per Recommendation E.502 is assumed.

7) Grade-of-service (GOS) on final circuit groups: 0.01

B.2 *Numerical results*

Numerical results are shown in Table B-1/E.522. The number of circuits C-B does not increase because of the 24-hour traffic profiles matching. The number of high usage circuits A-B in the minimum outlay network is larger than that in the minimum cost network. The impact of considering transit charges in dimensionings is always in the direction of less overflow.

TABLE B-1/E.522

Numerical results

Network results				Economic results (× 1000 units/month)								
Busy-hour overflow probability	Number of circuits			Circuit costs			Transit charges			Total outlay		
	A-B	A-C	C-B	A	B	C	A	B	C	A	B	C
0.0000	25	45	45	70	61	81	–	–	–	70.0	61.0	81.0
0.0090	25	45	45	70	61	81	0.3	0.3	(0.7)	70.3	61.3	80.3
0.0151	24	45	45	69	60	81	0.6	0.6	(1.3)	69.6	60.6	79.7
0.0221	23	45	45	68	59	81	0.9	0.9	(1.9)	68.9	59.9	79.1
0.0331	22	46	45	68	58	82	1.4	1.4	(2.9)	69.4	59.4	79.1
0.0471	21	46	45	67	57	82	2.1	2.1	(4.2)	69.1	59.1	77.8
0.0641	20	46	45	66	56	82	3.0	3.0	(6.0)	69.0	59.0	76.0
										Minimum outlay for A and B		
0.0861	19	47	45	66	55	83	4.2	4.2	(8.4)	70.2	59.2	74.5
0.1121	18	47	45	65	54	83	5.7	5.7	(11.5)	70.7	59.7	71.5
										Minimum cost network		
0.142	17	48	45	65	53	84	7.6	7.6	(15.1)	72.6	60.6	68.9
0.175	16	49	45	65	52	85	9.7	9.7	(19.4)	74.7	61.7	65.6

References

[1] WILKINSON (R. I.): Theories for toll traffic engineering in the USA, *Bell Syst. Tech. J.*, 1956, No. 35, pp. 421-514.

[2] WILKINSON (R. I.): Simplified engineering of single stage alternate routing systems, *Fourth International Teletraffic Congress*, London, 1964.

[3] RAPP (Y.): Planning of junction network in a multi-exchange area. 1. General Principles, *Ericsson Tech*; 1964, No. 20, 1, pp. 77-130.

[4] LEVINE (S. W.) and WERNANDER (M. A.): Modular engineering of trunk groups for traffic requirements, *Fifth International Teletraffic Congress*, New York, 1967.

[5] PRATT (C. W.): The concept of marginal overflow in alternate routing, *Fifth International Teletraffic Congress*, New York, 1967.

[6] EISENBERG (M.): Engineering traffic networks for more than one busy hour, *Bell System Tech. J.*, 1977, Vol. 56, pp. 1-20.

[7] AKIMARU (H.) *et al.*: Derivatives of Wilkinson formula and their application to optimum design of alternative routing systems, *Ninth International Teletraffic Congress*, Torremolinos, 1979.

[8] HORN (R. W.): A simple approach to dimensioning a telecommunication network for many hours of traffic demand, *International Conference on Communications*, Denver, 1981.

[9] BESHAI (M. E.): Traffic data reduction for multiple-hour network dimensioning, *Second International Network Planning Symposium*, Brighton, 1983.

[10] LINDBERGER (K.): Simple approximations of overflow system quantities for additional demands in the optimization, *Tenth International Teletraffic Congress*, Montreal, 1983.

STANDARD TRAFFIC PROFILES FOR INTERNATIONAL TRAFFIC STREAMS

The worldwide nature of the international telephone network, spanning as it does all possible time zones, has stimulated studies of the traffic streams between countries in different relative time locations. These studies have led to the development of standardized 24-hour traffic profiles which, theoretically based and verified by measurements, would be useful for engineering purposes. In fact, these concepts can be applied to a variety of network situations:

i) variable access satellite working where a large number of traffic streams with possibly differing traffic profiles share the pool of satellite circuits;

ii) combining of traffic streams on groups of terrestrial circuits which may be either high-usage or final choice routes;

iii) detour routing of traffic between origin and destination countries to take advantage of prevailing low load conditions on the detour path.

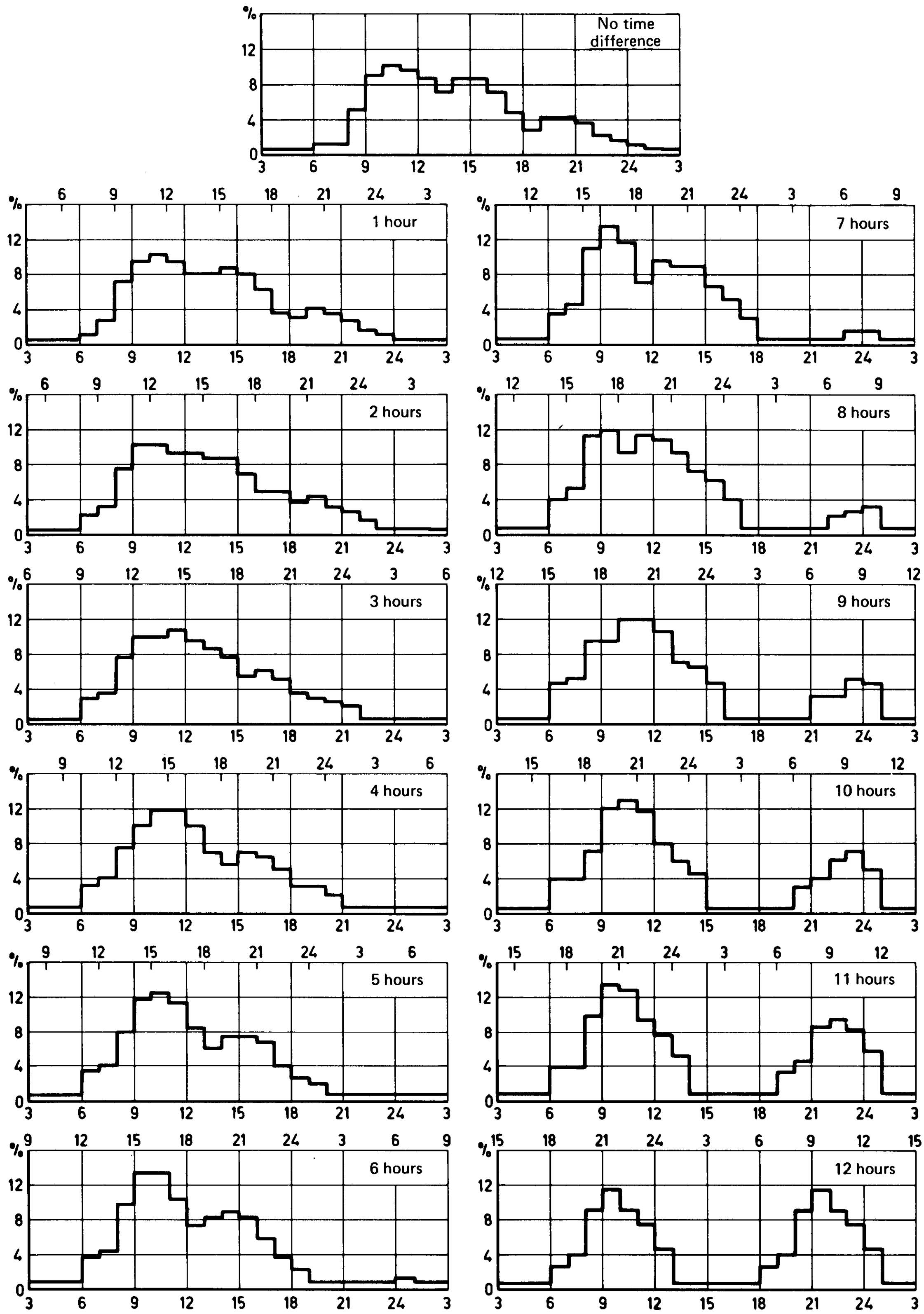
In developing any such applications, account must be taken of the International Routing Plan (Recommendation E.171 [1]) and of accepted accounting principles (Recommendation D.150 [2]).

It must be recognized that the preferred basis for dimensioning consists of traffic profiles based on real traffic. Nevertheless, many countries have found the standard profiles presented in this Recommendation very useful where streams are too small to obtain reliable measurements or where no measurements are available.

For both-way profiles, two equivalent methods of presentation are given in chart and tabular form. In Figure 1/E.523 hour-by-hour traffic volumes are shown in diagrammatically as percentages of the total daily traffic volume; such percentages are particularly convenient for tariff studies. In Table 1/E.523, hourly traffics are expressed as percentages of the busy hour traffic, and this is convenient for engineering purposes. Time zone differences are given in whole hours only. Directional profiles are given in Tables 2/E.523 and 3/E.523.

Although tables are given for both-way and directional traffic streams, it must be emphasized that at this stage only the both-way profiles can be regarded as soundly supported by measurement. The directional profiles are theoretically based and supported by some measurements, but should be used with caution until adequate verification has been achieved.

The theoretical basis for the profiles presented here is contained in Annex A. It depends on a convenience function $f(t)$ which represents the profile of *local* daily traffic, where of course no time zone difference exists. The function $f(t)$ used for computation of the standard profile was derived by mathematical manipulation of measurements of the Tokyo-Oakland and Tokyo-Vancouver streams. Although these results have been supported by other measurements, it leaves open the possibility that the convenience function may vary from one country to another and that, strictly, these should be derived independently and then used to obtain a calculated profile for the international relation. It also seems that the convenience function for the country of destination should be given greater weight than that for the country of origin. These remarks suggest possible refinements, but are not quantified in this Recommendation.



CCITT - 48101

Note – The vertical scale gives the hourly traffic volume as a percentage of the daily traffic volume. The horizontal scales show the local times.

FIGURE 1/E.523

Standard hourly both-way traffic distribution patterns

TABLE 1/E.523

Standard hourly bothway traffic patterns

Time difference (in hours) between two countries	Local time in the more westerly country: 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	BH %
0	5	5	5	5	5	5	10	10	50	90	100	95	85	70	85	85	70	45	25	40	40	35	20	15	10.0
1	5	5	5	5	5	5	10	25	70	95	100	90	80	80	85	80	60	35	30	40	35	25	15	10	10.0
2	5	5	5	5	5	5	20	30	75	100	100	90	90	85	85	65	45	45	35	40	30	25	15	5	10.0
3	5	5	5	5	5	5	25	35	75	100	95	100	95	80	70	50	60	45	35	30	25	15	5	5	10.4
4	5	5	5	5	5	5	25	35	65	85	100	100	85	60	50	60	55	40	25	25	20	5	5	5	11.5
5	5	5	5	5	5	5	25	30	65	95	100	90	70	50	60	60	55	30	20	20	5	5	5	5	12.4
6	10	5	5	5	5	5	25	30	75	100	100	75	55	60	65	60	40	25	15	5	5	5	5	5	13.1
7	10	5	5	5	5	5	25	35	80	100	85	55	70	65	65	50	40	20	5	5	5	5	5	10	13.5
8	25	5	5	5	5	5	35	45	95	100	80	95	90	75	60	50	35	5	5	5	5	5	20	20	11.7
9	40	5	5	5	5	5	35	40	75	80	100	95	85	60	55	35	5	5	5	5	5	25	25	40	12.1
10	40	5	5	5	5	5	35	35	60	95	100	90	65	50	40	5	5	5	5	5	25	30	50	55	12.5
11	40	5	5	5	5	5	30	25	75	100	95	70	55	35	5	5	5	5	5	25	30	65	70	60	12.3
12	40	5	5	5	5	5	20	35	80	100	80	65	40	5	5	5	5	5	20	35	60	100	80	65	11.3

Note 1 – The 24-hour profile of both-way traffic between any two countries is read from left to right from the appropriate row of the table; all time differences can be expressed in the range 0-12 hours. Each entry is expressed as a percentage of the busy hour traffic.

Note 2 – The *more westerly* country of a traffic relation is the one from which we can proceed eastwards to the other through time zones not exceeding 12 hours.

Note 3 – For network planning studies, UTC (Universal Coordinated Time) would normally be used so that all traffic streams are processed time consistently. Clearly if the more westerly country is W hours ahead of UTC (ignoring the international dateline), then the traffic at 0000-0100 UTC is obtained from the row corresponding to the time difference between the two countries at the column headed W. Alternatively, the first entry in the appropriate row gives the relative traffic intensity for the hour (24- W) to (25- W).

Example: For the traffic stream between the U.K. (UTC + 1 hour) and the central zone of USA (UTC + 18 hours), the time difference is 7 hours and the USA is regarded as the more westerly country, hence $W = 18$. Thus from the table, the traffic during 0000-0100 UTC is 5 % of the busy hour traffic, and the busy hour is 1500-1600 UTC.

Note 4 – The column headed "BH %" gives the busy hour traffic volume as a percentage of the daily traffic volume.

TABLE 2/E.523

Diurnal distributions of eastbound international telephone traffic

Local time in the more westerly country

Time difference (in hours) between two countries	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0	10	5	5	5	5	5	10	10	50	90	100	95	85	70	85	85	70	45	25	40	40	35	20	15
1	5	5	5	5	5	5	10	30	80	95	100	90	80	80	85	80	60	35	30	40	35	25	15	10
2	5	5	5	5	5	5	25	40	85	100	100	90	90	85	85	60	40	45	35	40	25	20	15	5
3	5	5	5	5	5	5	40	50	90	100	95	100	95	80	65	40	55	45	35	25	20	10	5	5
4	5	5	5	5	5	5	35	50	70	85	100	100	85	60	40	50	50	40	25	20	15	5	5	5
5	5	5	5	5	5	5	30	40	70	95	100	90	65	45	50	50	50	25	20	15	5	5	5	5
6	10	5	5	5	5	5	40	45	85	100	100	65	45	55	55	50	30	20	15	5	5	5	5	5
7	10	5	5	5	5	5	40	50	90	100	75	40	60	55	55	40	30	10	5	5	5	5	5	10
8	25	5	5	5	5	5	55	65	100	100	70	90	85	70	45	35	25	5	5	5	5	5	20	20
9	50	5	5	5	5	5	40	45	70	75	100	100	85	55	50	35	5	5	5	5	5	25	35	60
10	65	5	5	5	5	5	45	45	60	95	100	90	60	45	35	5	5	5	5	5	25	30	75	100
11	65	5	5	5	5	5	40	40	75	90	80	55	40	25	5	5	5	5	5	20	25	80	100	95
12	55	5	5	5	5	5	20	40	65	70	50	40	20	5	5	5	5	5	20	25	70	100	90	80

Note – This table is based on $p = 1.4$, $q = 0.6$, i.e. greater weight is given to the convenience function of the called party (see Annex A).

TABLE 3/E.523

Diurnal distributions of westbound international telephone traffic

Local time in the more westerly country

Time difference (in hours) between two countries	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0	10	5	5	5	5	5	10	10	50	90	100	95	85	70	85	85	70	45	25	40	40	35	20	15
1	5	5	5	5	5	5	10	20	60	95	100	90	80	80	85	80	60	35	30	40	35	25	15	10
2	5	5	5	5	5	5	15	20	65	100	100	90	90	85	85	70	50	45	35	40	35	30	15	5
3	5	5	5	5	5	5	10	20	60	100	95	100	95	80	75	60	65	45	35	35	30	15	5	5
4	5	5	5	5	5	5	15	20	60	85	100	100	85	60	60	70	60	40	25	30	25	5	5	5
5	5	5	5	5	5	5	20	20	60	95	100	90	75	55	70	70	60	35	20	25	5	5	5	5
6	10	5	5	5	5	5	10	15	65	100	100	85	65	65	75	70	50	30	15	5	5	5	5	5
7	10	5	5	5	5	5	10	20	70	100	95	70	80	75	75	60	50	30	5	5	5	5	5	10
8	20	5	5	5	5	5	15	25	90	100	90	95	95	80	75	65	45	5	5	5	5	5	20	20
9	25	5	5	5	5	5	30	35	80	85	100	95	85	65	60	35	5	5	5	5	5	20	20	25
10	10	5	5	5	5	5	25	25	60	95	100	90	70	55	45	5	5	5	5	5	25	30	25	10
11	15	5	5	5	5	5	10	10	65	95	100	80	65	45	5	5	5	5	5	25	35	40	35	25
12	20	5	5	5	5	5	20	25	70	100	90	80	55	5	5	5	5	5	20	40	65	70	50	40

Note – This table is based on $p = 1.4$, $q = 0.6$, i.e. greater weight is given to the convenience function of the called party (see Annex A).

ANNEX A

(to Recommendation E.523)

Mathematical expression for the influence of time differences on the traffic flow

A telephone call is initiated when a person wishes to call someone else, but both parties have to be on the line before the call is established. It is considered that a telephone call is made at a time which tends to be convenient for both the calling and called parties. The *degree of convenience* for making a telephone call is considered to be a periodical function of time t, whose period is 24 hours. When the time difference between both parties is zero, the degree of convenience is denoted by $f(t)$, where t is local standard time. The graphic shape of the basic function $f(t)$ will be determined by the daily pattern of human activities, and will resemble, or fairly closely coincide with, the hour by hour traffic distribution in the national (or local) telephone network.

It is assumed that the hourly traffic distribution $F_\tau(t)$, when a time difference of τ hours exists between the originating and called locations, is expressed as the geometric mean of convenience functions of two locations τ hours apart:

$$F_{\tau,}(t) = k \left\{ f(t) \cdot f(t + \tau) \right\}^{1/2}$$

where

$$k = 1/ \int_{24\text{ hours}} \left\{ f(t) \cdot f(t + \tau) \right\}^{1/2} dt \qquad \text{(A-1)}$$

The sign of τ is positive when the time at the destination is ahead of the reference time, and negative when the time of destination is behind the reference time.

The distribution of equation (A-1) represents the sum of the outgoing and incoming traffics. Expressions for the one-way hourly traffic distributions can also be obtained by extending the concept of convenience function as follows.

Define convenience functions both for the caller $f_0(t)$ and for the called party $f_i(t)$.Then the one-way traffic distributions of east-bound and west-bound telephone calls, for the case of τ hour time-difference, are similarly expressed as follows:

$$F_{\tau,\,\text{east}}(t) = k \left\{ f_0(t) \cdot f_i(t + \tau) \right\}^{1/2}$$

$$k = 1/ \int_{24\text{ hours}} \left\{ f_0(t) \cdot f_i(t + \tau) \right\}^{1/2} dt \qquad \text{(A-2)}$$

$$F_{\tau,\,\text{west}}(t) = k \left\{ f_i(t) \cdot f_0(t + \tau) \right\}^{1/2}$$

$$k = 1/ \int_{24\text{ hours}} \left\{ f_i(t) \cdot f_0(t + \tau) \right\}^{1/2} dt \qquad \text{(A-3)}$$

where

t is the local standard time of the west station and

τ is positive.

It is natural that a caller makes a call considering the convenience of the called person, and therefore the convenience function of the called person f_i contributes more than the convenience of the caller f_0 to the directional distribution F. They can be written as follows:

$$f_i(t) = k_1 \left\{ f(t) \right\}^p, \qquad f_0(t) = k_2 \left\{ f(t) \right\}^q, \qquad \text{(A-4)}$$

where

$$p > q \quad \text{and} \quad p + q = 2,$$

and where k_1 and k_2 are normalizing coefficients to ensure that:

$$\int_{24 \text{ hours}} f_i(t)\, dt = 1, \qquad \int_{24 \text{ hours}} f_0(t)\, dt = 1.$$

As to the values of p and q in equation (A-4), it has been found empirically that the convenience of the called side p is considerably larger than that of originating side q, and appropriate values are roughly $p = 1.4$ and consequently $q = 0.6$.

References

[1] CCITT Recommendation *International telephone routing plan*, Rec. E.171.

[2] CCITT Recommendation *New system for accounting in international telephony*, Rec. D.150.

Bibliography

CASEY (J. Jr.) and SHIMASAKI (N.): Optimal dimensioning of a satellite network using alternate routing concepts, *Sixth International Teletraffic Convention*, Munich, 1970.

RAPP (Y.): Planning of a junction network with non-coincident busy hours, *Ericsson Technics*, No. 1, 1971.

CABALLERO (P. A.) and DÍAZ (F.): Optimization of networks of hierarchical structure with non-coincident busy hours, *Seventh International Teletraffic Convention*, Stockholm, 1973.

OTHA (T.): Network efficiency and network planning considering telecommunication traffic influenced by time difference, *Seventh International Teletraffic Convention*, Stockholm, 1973.

Recommendation E.524

OVERFLOW APPROXIMATIONS FOR NON-RANDOM INPUTS

1 Introduction

This Recommendation introduces approximate methods for the calculation of blocking probabilities for individual traffic streams in a circuit group arrangement. It is based on contributions submitted in the Study Period 1984-1988 and is expected to be amended and expanded in the future (by adding the latest developments of methods).

The considered methods are necessary complements to those included in the existing Recommendation E.521 when it is required to take into account concepts such as cluster engineering with service equalization, service protection and end-to-end grade of service. Recommendation E.521 is then insufficient as it is concerned with the grade of service for only one non-random traffic stream in a circuit group.

Design methods concerning the above-mentioned areas are subject to further study and this Recommendation will serve as a reference when, in the future, Recommendation E.521 is complemented or replaced.

In this Recommendation the proposed methods are evaluated in terms of accuracy, processing time, memory requirements and programming effort. Other criteria may be relevant and added in the future.

The proposed methods are described briefly in § 2. Section 3 defines a set of examples of circuit group arrangements with exactly calculated (exact resolution of equations of state) individual blocking probabilities, to which the result of the methods can be compared. This leads to a table in § 4, where for each method the important criteria are listed. The publications cited in the reference section at the end contain detailed information about the mathematical background of each of the methods.

2 Proposed methods

The following methods are considered:

a) Interrupted Poisson Process (IPP) method,

b) Equivalent Capacity (EC) method,

c) Approximative Wilkinson Wallström (AWW) method.

2.1 *IPP method*

IPP (Interrupted Poisson Process) is a Poisson process interrupted by a random switch. The on-/off-duration of the random switch has a negative exponential distribution. Overflow traffic from a circuit group can be accurately approximated by an IPP, since IPP can represent bulk characteristics of overflow traffic. IPP has three parameters, namely, on-period intensity and mean on-/off-period durations. To approximate overflow traffic by an IPP, those three parameters are determined so that some moments of overflow traffic will coincide with those of IPP.

The following two kinds of moment match methods are considered in this Recommendation:

– three-moment match method [1] – where IPP parameters are determined so that the first three moments of IPP will coincide with those of overflow traffic;

– four-moment ratio match method [2] where IPP parameters are determined so that the first moment and the ratios of the 2nd/3rd and 7th/8th binomial moments of IPP will coincide with those of overflow traffic.

To analyze a circuit group where multiple Poisson and overflow traffic streams are simultaneously offered, each overflow stream is approximated by an IPP. The IPP method is well suited to computer calculation. State transition equations of the circuit group with IPP inputs can be solved directly and no introduction of equivalent models is necessary. Characteristics of overflow traffic can be obtained from the solution of state transition equations. The main feature of the IPP method is that the individual means and variances of the overflow traffic can be solved.

2.2 *EC method*

The EC (Equivalent Capacity) method [3] does not use the traffic-moments but the transitional behaviour of the primary traffic, by introducing a certain function $\rho(n)$ versus the equivalent capacity (n) of the partial overflow traffic, as defined by the recurrent process:

$$\left[\begin{array}{l} \rho(o) = Em(\alpha) \qquad \text{[Erlang loss formula]} \\ \dfrac{n}{\rho(n)} = (m + n - a) + \alpha \cdot \rho(n - 1) \end{array}\right. \tag{2-1}$$

if n is a positive integer and approximated by linear interpolation, if not.

A practical approximation, considering the predominant overflow congestion states only, leads to the equations:

$$\frac{n_i}{n} = \frac{a_i\, \rho_i(n_i)/D_i\,(n_i + 1)}{\sum\limits_{k=1}^{x} a_k\, \rho_k(n_k)/D_k\,(n_k + 1)} \tag{2-2}$$

with:

$$D_i(n) = 1 + a_i \left[\rho_i(n) - \rho_i(n - 1)\right] \tag{2-3}$$

defining the equivalent capacity (n_i) of the partial overflow traffic labelled i, and influenced by the mutual dependency between the partial overflow traffic streams.

The mean value of the partial second overflow is:

$$O_i = a_i\, \pi\, \rho_i(n_i) \tag{2-4}$$

where π is the time congestion of the overflow group.

The partial GOS (grade of service) equalization is fulfilled if:

$$\rho_i\,(n_i) = C \tag{2-5}$$

C being a constant to be chosen.

2.3 *AWW method*

The AWW (Approximative Wilkinson Wallström) method uses an approximate ERT (Equivalent Random Traffic) model based on an improvement of Rapp's approximation. The total overflow in traffic is split up in the individual parts by a simple expression, see Equations (2-7) and (2-9). To calculate the total overflow traffic, any method can be used. An approximate Erlang formula calculation for which the speed is independent of the size of the calculated circuit group is given in [4].

The following notations are used:

M mean of total offered traffic;

V variance of total offered traffic;

Z V/M;

B mean blocking of the studied group;

m_i, v_i, z_i, b_i corresponding quantities for an individual traffic stream;

~ is used for overflow quantities.

2.3.1 *Blocking of overflow traffic*

For overflow calculations, an approximate ERT-model is used. By numerical investigations, a considerable improvement has been found to Rapp's classical approximation for the fictitious traffic. The error added by the approximation is small compared to the error of the ERT-model. It is known that ERT underestimates low blockings when mixing traffic of diverse peakedness [2]. The formula, which was given in [4] (although with one printing error), is for $Z > 1$:

$$A^* \approx V + Z(Z - 1)(2 + \gamma^{\beta})$$

where

$$\gamma = (2.36\ Z - 2.17) \log \{1 + (z - 1)/[M(Z + 1.5)]\}$$

and

$$\beta = Z/(1.5M + 2Z - 1.3) \qquad (2\text{-}6)$$

2.3.2 *Wallström formula for individual blocking*

There has been much interest in finding a simple and accurate formula for the individual blocked traffic $\tilde{m}_i$. Already in 1967, Katz [5] proposed a formula of the type

$$\tilde{m}_i = m_i B(1 - w + wz_i/Z) \qquad (2\text{-}7)$$

with w being a suitable expression. Wallström proposed a very simple one but with reasonable results [6], [2]:

$$w = 1 - B \qquad (2\text{-}8)$$

One practical problem is, however, that a small peaked substream could have a blocking $b_i > 1$ with this formula. To avoid such unreasonable results a modification is used in this case. Let z_{max} be the largest individual z_i.

Then the value used is

$$w = \begin{cases} 1 - B & \text{if } z_{max} < Z(1+B)/B \\ Z(1-B)/(B(z_{max} - Z)) & \text{otherwise} \end{cases} \qquad (2\text{-}9)$$

2.3.3 *Handling of overflow variances*

For the calculation of a large network it would be very cumbersome to keep track of all covariances. The normal case is that the overflow traffic from one trunk group is either lost or is offered to a secondary group without splitting up. Therefore it is practical to include covariances in the individual overflow parameters $\tilde{v}_i$ so that they sum up to the total variance. The quantities v_i are obtained from the total overflow variance $\tilde{V}$ by a simple splitting formula:

$$\tilde{v}_i = \tilde{V}_i v_i / V \qquad (2\text{-}10)$$

One can prove that Wallström's splitting formula (2-8) and formula (2-10) together with the ERT-model satisfies a certain consistency requirement. One will obtain the same values for the individual blocked traffic when calculating a circuit group of $N_1 + N_2$ circuits as when calculating first the N_1 circuits and then offering the overflow to the N_2 circuits.

Since the individual variances are treated in this manner, they are not comparable with the results reported in Table 2/E.524.

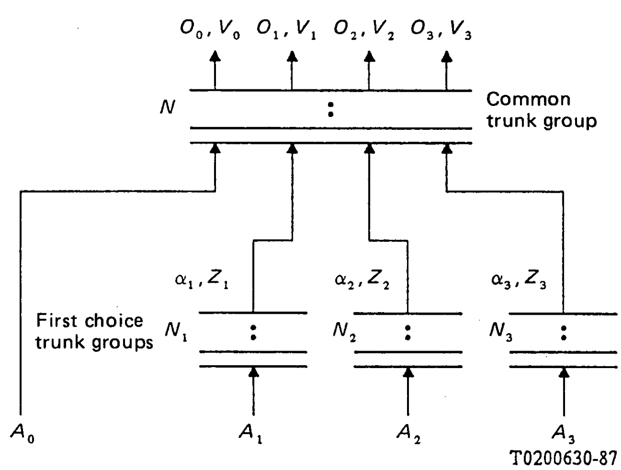
3 Examples and criteria for comparison

The defined methods are tested by calculating the examples given in Table 1/E.524.

The calculation model is given in Figure 1/E.524.

For comparison, the following criteria are established:

- accuracy of the overflow traffic mean and variance (mean and standard deviation),
- computational criteria (processor time, memory requirements, programming effort).



A_i: Offered Poisson traffic volume
N_i: Number of first choice trunk group
α_i: Mean of overflow traffic from first choice trunk group
Z_i: Peakedness of overflow traffic from first choice trunk group
N: Number of common trunk group
O_i: Mean of overflow traffic from common trunk group
V_i: Variance of overflow traffic from common trunk group

FIGURE 1/E.524

Calculation model

TABLE 1a/E.524

Exactly calculated mean and variance of individual overflow traffic – Three first choice circuit groups

Case	A_1	A_2	A_3	α_1	α_2	α_3	A_0	N	O_0	O_1	O_2	O_3
	N_1	N_2	N_3	Z_1	Z_2	Z_3			V_0	V_1	V_2	V_3
1	7.036	26.688	64.169	3.003	3.001	3.000	–	11	–	0.4337	0.7490	1.091
	5	28	70	1.573	3.022	4.527			–	0.7656	2.110	4.441
2	7.036	26.688	64.169	3.003	3.001	3.000	–	16	–	0.1149	0.2758	0.4944
	5	28	70	1.573	3.022	4.527			–	0.2436	0.7328	1.911
3	7.036	26.688	64.169	3.003	3.001	3.000	–	25	–	0.01369	0.02846	0.06627
	5	28	70	1.573	3.022	4.527			–	0.02041	0.06461	0.2205
4	7.036	10.176	13.250	3.003	5.003	7.002	–	14	–	0.7459	1.262	1.785
	5	6	7	1.573	1.567	1.559			–	1.193	2.292	3.624
5	7.036	10.176	13.250	3.003	5.003	7.002	–	19	–	0.2884	0.4857	0.6832
	5	6	7	1.573	1.567	1.559			–	0.4636	0.9089	1.460
6	7.036	10.176	13.250	3.003	5.003	7.002	–	26	–	0.03570	0.05915	0.08237
	5	6	7	1.573	1.567	1.559			–	0.05358	0.1026	0.1621
7	7.036	32.395	77.617	3.003	5.002	7.001	–	16	–	0.4516	1.176	2.344
	5	31	77	1.573	3.029	4.511			–	0.7434	3.466	10.39

TABLE 1a/E.524 *(cont.)*

Case	A_1	A_2	A_3	α_1	α_2	α_3	A_0	N	O_0	O_1	O_2	O_3
	N_1	N_2	N_3	Z_1	Z_2	Z_3			V_0	V_1	V_2	V_3
8	7.036	32.395	77.617	3.003	5.002	7.001	–	23	–	0.1538	0.4294	0.9739
	5	31	77	1.573	3.029	4.511			–	0.2427	1.200	4.219
9	7.036	32.395	77.617	3.003	5.002	7.001	–	35	–	0.01303	0.03984	0.1006
	5	31	77	1.573	3.029	4.511			–	0.1841	0.09378	0.3690
10	64.169	32.395	13.250	3.000	5.002	7.002	–	15	–	1.157	1.456	1.320
	70	31	7	4.527	3.029	1.559			–	4.442	4.256	2.850
11	64.169	32.395	13.250	3.000	5.002	7.002	–	21	–	0.5564	0.5849	0.4749
	70	31	7	4.527	3.029	1.559			–	2.026	1.675	1.023
12	64.169	32.395	13.250	3.000	5.002	7.002	–	32	–	0.06907	0.05265	0.03848
	70	31	7	4.527	3.029	1.559			–	0.2167	0.1295	0.07165
13	7.036	26.688	64.169	3.003	3.001	3.000	3.000	13	0.4064	0.5038	0.8274	1.160
	5	28	70	1.573	3.022	4.527			0.5578	0.8566	2.243	4.574
14	7.036	26.688	64.169	3.003	3.001	3.000	3.000	18	0.1460	0.1840	0.3384	0.5729
	5	28	70	1.573	3.022	4.527			0.1992	0.3043	0.8779	2.163

TABLE 1a/E.524 *(cont.)*

Case	A_1	A_2	A_3	α_1	α_2	α_3	A_0	N	O_0	O_1	O_2	O_3
	N_1	N_2	N_3	Z_1	Z_2	Z_3			V_0	V_1	V_2	V_3
15	7.036	26.688	64.169	3.003	3.001	3.000	3.000	28	0.01170	0.01506	0.03086	0.07035
	5	28	70	1.573	3.022	4.527			0.01472	0.02218	0.06861	0.2287
16	7.036	32.395	77.617	3.003	5.002	7.001	1.000	17	0.1253	0.4451	1.156	2.304
	5	31	77	1.573	3.029	4.511			0.1392	0.7266	3.366	10.10
17	7.036	32.395	77.617	3.003	5.002	7.001	1.000	24	0.04250	0.1536	0.4275	0.9674
	5	31	77	1.573	3.029	4.511			0.04696	0.2409	1.183	4.148
18	7.036	32.395	77.617	3.003	5.002	7.001	1.000	35	0.004542	0.01687	0.05106	0.1282
	5	31	77	1.573	3.029	4.511			0.004891	0.02398	0.1214	0.4751
19	64.169	32.395	13.250	3.000	5.002	7.002	9.000	21	1.761	1.251	1.654	1.630
	70	31	7	4.527	3.029	1.559			3.052	4.517	4.406	3.103
20	64.169	32.395	13.250	3.000	5.002	7.002	9.000	28	0.6761	0.6501	0.7389	0.6427
	70	31	7	4.527	3.029	1.559			1.253	2.225	1.956	1.279
21	64.169	32.395	13.250	3.000	5.002	7.002	9.000	40	0.06219	0.09577	0.07978	0.06069
	70	31	7	4.527	3.029	1.559			0.1054	0.2884	0.1887	0.1099

TABLE 1b/E.524

Exactly calculated mean and variance of individual overflow traffic – Two first choice circuit groups

A_1	N_1	A_2	N_2	N	O_1	V_1	O_2	V_2
8.2	5	30.0	30	10	0.6155	1.1791	1.1393	3.4723
				5	1.8068	3.2634	2.4656	7.4312
				21	0.0188	0.0304	0.0485	0.1240
				14	0.2108	0.3898	0.4624	1.3701
		14.3	7	22	0.0470	0.0771	0.0929	0.1983
				16	0.3743	0.6602	0.7546	1.7626
				12	0.9282	1.6137	1.8320	4.2120
				7	2.0023	3.2718	4.0953	7.8064
		42.0	37	27	0.0230	0.0354	0.0978	0.2984
				19	0.2136	0.3683	0.8356	2.9450
				8	1.4984	2.6161	4.4363	14.6018
				13	0.6940	1.2375	2.4148	8.4923
30.0	30	14.3	7	25	0.0653	0.1613	0.0541	0.1112
				18	0.4664	1.2990	0.4662	1.0879
				12	1.3746	3.9321	1.7390	4.0015
				7	2.4255	6.9941	3.8063	7.6277
8.2	5	67.9	65	30	0.0160	0.0242	0.0979	0.3548
				20	0.1839	0.3141	0.9739	4.1953
				14	0.5385	0.9676	2.4438	10.7208
				8	1.3598	1.4401	4.7035	19.7109
51.5	54	14.3	7	27	0.0735	0.2239	0.0399	0.0802
				19	0.6404	1.2499	0.4699	1.1030
				13	1.4033	5.0795	1.3609	3.2229
				7	2.5873	9.6136	3.6744	7.5139

TABLE 1c/E.524

Exactly calculated mean and variance of individual overflow traffic – One first choice circuit group

A_1	N_1	A_0		N	O_1	V_1	O_0	V_0
8.2	5	4.0		16	0.0499	0.0872	0.0331	0.0479
				11	0.4859	0.9154	0.3494	0.5382
				9	1.1692	2.1202	0.9011	1.3274
				5	2.1422	3.5883	1.8018	2.3694
30.0	30			20	0.0601	0.1565	0.0167	0.023
				13	0.5804	1.7427	0.1990	0.3062
				9	1.3997	4.2546	0.5988	0.9338
				5	2.5579	5.6196	1.5661	2.1991
51.5	54			22	0.9751	0.2497	0.0144	0.0197
				15	0.5141	1.8924	0.1209	0.1819
				10	1.8820	5.3004	0.4297	0.6790
				5	2.4294	3.2974	1.1450	1.7255

4 Summary of results

The available methods and the performance measures with respect to the criteria are listed in Table 2/E.524.

TABLE 2/E.524

Comparison of different approximation methods

Functions / Method	Input	Output	Comparison						
			Overflow traffic error				Computational effort		
	Required higher moments	Highest moments of overflow traffic	Mean		Variance				
			Mean	Standard deviation	Mean	Standard deviation	Processor time	Memory requirements	Programming effort
IPP method									
a) 3 moment match	3	3	−0.0045	0.0585	−0.0210	0.0922			
b) 4 moment ratio	8	∞	0.0008	0.0255	−0.0053	0.0373			
EC method	1	1	−0.0661	0.1527					
AWW method	2	2	−0.0448	0.1647					

References

[1] MATSUMOTO (J.) and WATANABE (Y.): Analysis of individual traffic characteristics for queuing systems with multiple Poisson and overflow inputs. *Proc. 10th ITC*, paper 5.3.1, Montreal, 1983.

[2] RENEBY (L.): On individual and overall losses in overflow systems. *Proc. 10th ITC*, paper 5.3.5, Montreal, 1983.

[3] LE GALL (P.): Overflow traffic combination and cluster engineering. *Proc. 11th ITC*, paper 2.2B-1, Kyoto, 1985.

[4] LINDBERG (P.), NIVERT, (K.), SAGERHOLM, (B.): Economy and service aspects of different designs of alternate routing networks. *Proc. 11th ITC*, Kyoto, 1985.

[5] KATZ (S.): Statistical performance analysis of a switched communications network. *Proc. 5th ITC*, New York, 1967.

[6] LINDBERGER (K.): Simple approximations of overflow system quantities for additional demands in the optimization. *Proc. 10th ITC*, Montreal, 1983.

SERVICE PROTECTION METHODS

1 Introduction

The objective of service protection methods is to control the grade of service for certain streams of traffic by restricting the access to circuit groups. Several methods are available, with the common feature that they may reject certain call attempts when the considered circuit group has little spare capacity. Service protection is generally used in alternative routing networks to restrict overflow traffic, but can also be used to give priority service to one class of traffic over another.

Failure or overload conditions may require temporary changes to service protection parameters. This is considered to be network management action which is described in the E.400 Series Recommendations.

Applications of service protection methods are described in § 2, and the available methods are described in § 3.

The use of service protection generally increases the complexity of dimensioning algorithms. Appropriate dimensioning algorithms are presented in § 4.

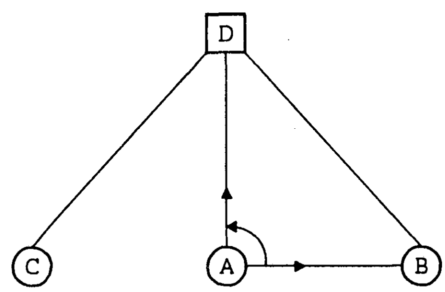
The choice between available methods will generally depend on performance characteristics and ease of implementation. These are discussed in §§ 5 and 6.

2 Applications

2.1 *Traffic routing*

2.1.1 *Overflow routing strategies – General principles*

Routing strategies that involve overflow often have direct first-choice (high-usage) routes, and indirect alternative routes. In conditions of traffic overload the proportion of alternatively-routed traffic increases rapidly, with the risk of severe degradation of network performance. Service protection methods should be used to prevent calls overflowing from a direct route to an alternative route when circuit groups on the alternative route are heavily loaded. In Figure 1/E.525, which shows a hierarchical case only for the sake of simplification, calls from A to B have a direct first-choice route and an alternative route via D. Exchange A should apply service protection on the circuit group AD. When AD is occupied over a certain limit, overflow calls (e.g. from AB) are rejected and priority is given to first-choice traffic (e.g. from A to C).



T0200640-87

FIGURE 1/E.525

In Figure 1/E.525 traffic from A to B has access to two possible routes but traffic from A to C has only one. In this situation traffic from A to B is likely to experience a much better end-to-end grade of service unless service protection is used to restrict its access to AD. This control of grade of service allows optimal (minimum cost) dimensioning for planned traffic loads in addition to giving protection against heavy overloads.

2.1.2 *Fixed hierarchical alternative routing*

An example of fixed hierarchical alternative routing is illustrated in Figure 1/E.525. Here exchange D is a tandem exchange at a higher hierarchical level than A, B and C. Direct routes at the lower level (e.g. AB) overflow via the hierarchical route (ADB). This hierarchical route is always the final alternative routing. In such networks it is highly recommended to apply service protection to restrict traffic overflowing to final choice routes.

2.1.3 *Fixed non-hierarchical alternative routing*

This term describes routing strategies which are based on fixed sequences of alternative routes (as in hierarchical alternative routing) but which do not have a hierarchical final-choice route for all overflow traffic. Figure 2/E.525 may be used to illustrate some simple but common cases. Traffic from A to B has a first-choice route AB and an alternative ACB which is final to this traffic, while traffic from A to C may use AC as a first choice and then overflow to ADC. Traffic from D to B is either first offered to the route DAB and then overflowing to DCB or vice versa. The latter routing principle is commonly known as mutual overflow.

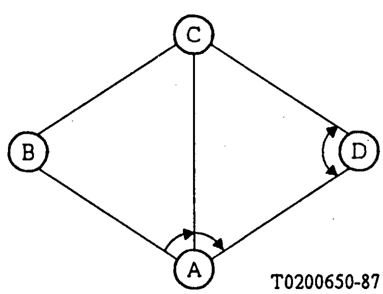


FIGURE 2/E.525

In both routing schemes a certain hierarchy is distinguishable. They are however non-hierarchical, in the sense that no hierarchical trunk group that is final to all its carried traffic can be found. The application of service protection methods may be less simple than for hierarchical routing, but the general principles presented in § 2.1.1 still apply.

2.1.4 *Dynamic routing*

Many different forms of preplanned or adaptive dynamic routing are possible, with either centralized or distributed control (see Recommendation E.170). One feature that is common to most dynamic routing schemes is the availability of a large number of potential alternative routes for any given connection. With this type of routing scheme, service protection is of crucial importance and has several special features:

- Protection should be stronger than with other overflow routing schemes (i.e. larger reservation parameters should be used).
- If possible, service protection should be applied on all circuit groups in an alternative route. This requires a certain amount of information-passing between exchanges or to a central processor.

- In connection with adaptive routing, the service protection concept can be used not only to block overflow calls but also in the selection of a good alternative route (generally this will be a route on which all circuit groups have at least a requested number of free circuits).

2.2 *Priority service*

Service protection methods can also be used to give priority service to one type of traffic, for example in a multiservice network, e.g. ISDN.

2.3 *Stability*

In order to provide stability in networks with non-hierarchical routing schemes, service protection should be used to restrict overflow traffic if that traffic overflows to an alternative route which is shared with first-choice traffic.

3 Available methods

3.1 *Split circuit group*

A straightforward technique is to divide a circuit group into two components. Priority traffic is allowed access to the whole circuit group, while non-priority (usually overflow) traffic is only allowed access to one component. Normally the priority traffic is offered first to the reserved component – this is then equivalent to a separate high usage group.

3.2 *Trunk reservation*

This method is also known as priority reservation system. Non-priority calls are accepted on the considered circuit group only when the momentary number of free circuits in that group observed at the arrival of a non-priority call exceeds a specified lower limit (irrespective of which particular circuits are free). Priority calls are always accepted if any circuits are free.

Trunk reservation may also be applied selectively, for example, to restrict call attempts to hard-to-reach destinations. This method is known as selective trunk reservation.

4 Evaluation and dimensioning

4.1 *Cluster engineering concept*

In hierarchical automatic alternative routing (AAR) a cluster consists of a final-choice circuit group together with those high usage groups from which traffic overflows to the final group. This cluster should be engineered as a whole. This implies firstly that grade-of-service (GOS) criteria should be applied to the whole cluster rather than separately to final groups. Secondly, the question of high-load dimensioning must be considered for the whole cluster. In order to meet normal and high load cluster GOS criteria in the most efficient way, the parameters of service protection methods must be determined appropriately as a part of the dimensioning process.

4.2 *Split circuit group*

With hierarchical AAR, the split final circuit group creates a separate high-usage group for first-choice traffic. This should be dimensioned so as to achieve the cluster GOS criteria. Standard evaluation methods that can be used include the Wilkinson Equivalent Random Traffic theory [1]. Interrupted Poisson Process methods can be used to give more precise evaluation [2], [3] and to evaluate network performance [4].

Split circuit groups may be useful to control GOS in non-hierarchical routing. The precise dimensioning and evaluation depends on the individual situation and it is generally more practical to use 1-moment methods of analysis [5], [6].

4.3 *Trunk reservation*

With hierarchical AAR, a trunk reservation parameter should be applied to the final group so as to achieve the cluster GOS criteria optimally for all traffic offered to the cluster. In most situations a small value of this parameter is appropriate. For evaluation of Poisson streams a recursion method is available which may be extended, using equivalent random traffic (ERT) techniques, into overflow situations [7]. Interrupted Poisson Process [3] methods can be used to give a more precise evaluation and to evaluate network performance [8].

For non-hierarchical strategies, 1-moment evaluation methods are again recommended. Simple recursion formulas are available for a circuit group using trunk reservation and offered Poisson traffics. 1-moment [7] methods can also be extended to give better accuracy by taking account of downstream blocking and traffic correlations [6] and [8].

5 Performance characteristics

5.1 *Efficiency*

Efficiency can be measured by traffic capacity at normal load subject to GOS criteria. In this respect there is little to choose between trunk reservation and split circuit group methods, provided each is correctly dimensioned.

5.2 *Overload protection*

The two service protection methods, trunk reservation and split final with a reserved high usage group, provide considerably better overload protection for first-choice-final traffic in cases of general and overflow overload than does the less usual method, split final with reserved final group.

5.3 *Robustness*

A significant advantage of trunk reservation is that it provides a robust performance profile with respect to traffic load variations (decreasing high priority traffic in combination with increasing low priority traffic) and reservation parameter settings. Independent of the trunk group size, traffic variations (which have not been forecasted) are relatively well carried.

With trunk reservation the same parameter value is likely to be optimal for a wide range of configurations at both normal load and overloads.

In contrast, the reserved section of a split circuit group should be redimensioned for different configurations and (when dimensioned according to the normal traffic load pattern) will not give optimal values at overload.

5.4 *Peakedness*

Changes in the peakedness of overflowing traffic have a slightly greater impact on the blocking probabilities within split circuit group arrangements in comparison with trunk reservation.

6 Implementation consequences

6.1 *Dimensioning methods*

Methods for the calculation of a split circuit group or a trunk reservation parameter are available [7], [9], [10].

6.2 *Traffic measurements*

Both service protection methods require the estimates of the first-choice-final traffic which is to be protected and the overflowing traffic from the high-usage trunk group(s) (i.e. measurements on a per destination basis).

With the split circuit group method, first routed traffic can be easily measured on a trunk group basis. With the trunk reservation method, measurements other than traditional are required to identify the first offered traffic.

6.3 *Operational aspects*

Since trunk reservation is a software controlled technique, protection for different traffic streams can easily be changed by changing parameters in the software. This allows temporary changes to be made under network management control. Precautions should be taken in such situations to restore design parameter values.

6.4 *Technology requirements*

Split circuit group methods can be installed in both electromechanical and processor controlled exchanges.

Trunk reservation may, in practice, only be realized in software as a conditional overflow facility and consequently only be installed in SPC exchanges.

Both methods require that the exchange have the ability to distinguish between priority and non-priority traffic.

References

[1] WILKINSON (R. I).: Theories for toll traffic engineering in the USA. *Bell System Technical Journal*, Vol. 35, March 1956.

[2] MATSUMOTO (J.), WATANABE (Y.): Analysis of individual traffic characteristics for queueing systems with multiple Poisson and overflow inputs. *Proc. 10th ITC*, paper 5.3.1, Montreal, 1983.

[3] KUCZURA (A.): The interrupted Poisson Process as an overflow processor. *Bell System Technical Journal*, Vol. 52, No. 3, 1973.

[4] MANSFIELD (D. R.), DOWNS (T.): A moment method for the analysis of telephone traffic networks by decomposition. *Proc. 9th ITC*, paper 2.4.4, Torremolinos, 1979.

[5] MANSFIELD (D. R.), DOWNS (T.): On the one-moment analysis of telephone traffic networks. *IEEE Trans. Comm. 27*, pp. 1169-1174, 1979.

[6] LE GALL (F.), BERNUSSOU (J.): An analytical formulation for grade-of-service determination in telephone networks. *IEE Trans. Comm., 31*, pp. 420-424, 1983.

[7] COOPER (R. B.): Introduction to queueing theory. *North Holland*, 1977.

[8] SONGHURST (D. J.): Protection against traffic overload in hierarchical networks employing alternative routing. *Proc. Telecommunication Network Planning Symposium*, pp. 214-220, Paris, 1980.

[9] LEBOURGES (M.), PASSERON (A.): Contribution to a network sizing procedure using probability distributions of traffic data, *Networks '86*, Tarpon Springs, 1986.

[10] LINDBERG (K.): Simple approximations of overflow system quantities for additional demands in the optimization. *Proc. 10th ITC*, Montreal, 1983.

Bibliography

LEBOURGES (M.), BECQUE (C. R.), SONGHURST, (D. J.): Analysis and dimensioning on non-hierarchical telephone networks. *Proc. 11th ITC*, paper 2.28-4, Kyoto, 1985.

SECTION 5

GRADE OF SERVICE

Recommendation E.540

OVERALL GRADE OF SERVICE OF THE INTERNATIONAL PART OF AN INTERNATIONAL CONNECTION

1 The International Routing Plan envisages that international traffic relations may be served by any of the following routing arrangements:

a) direct circuits;

b) transit operation involving one or more transit centres for all connections,

c) direct high-usage circuits with overflow via one or more transit centres.

In principle there would be merit in dimensioning international facilities to provide the same grade of service for all relations, however served. Practical considerations make it advisable to depart from one universal value.

2 Direct circuit groups are dimensioned, according to Recommendation E.520 on the basis of $p = 1\%$ loss probability during the mean busy hour. An exception is permitted for small groups of very long international circuits for which $p = 3\%$ loss probability is accepted for six or fewer circuits. As the traffic increases the grade of service improves progressively until $p = 1\%$ loss value is reached for 20 circuits.

3 For the relations served exclusively by transit operation the grade of service will deteriorate with the number of transit centres in the connection. Measurements made on congestion in such circumstances suggest that the overall grade of service for up to six links in tandem is less than twice the congestion of any of the six links in the chain. Hence, for a series of routes, each dimensioned for $p = 1\%$, the overall grade of service should seldom exceed 2%. An East-West type of connection would have the advantage of different busy hours on the various links. Corresponding advantage would not apply to North-South circuits.

In the case of relations served by high-usage circuits the overflow traffic will route over at least two links and, hence, will be subject to the same deterioration of service as in the case for transit traffic. However, a substantial part of the traffic will be connected over the high-usage circuits and the overall grade of service will approximate that of the relations served solely by direct circuits.

It is desirable that at least one high-usage circuit should always be provided between a CT3 and its homing CT1, even though the circuit may not be wholly justified on economic considerations alone. However, such a circuit should not be provided unless there is a measurable amount of traffic which exists, or can be foreseen in the busy hour. The provision of such circuits would improve the transmission as well as the grade of service; these considerations should encourage an increase both in traffic and in the revenue-earning capacity of the circuits provided.

The overall grade of service for the international part of a connection is a contributory factor to the overall grade of service from the calling party in one country to the called party in another.

OVERALL GRADE OF SERVICE FOR INTERNATIONAL CONNECTIONS (SUBSCRIBER-TO-SUBSCRIBER)

1 Introduction

1.1 The overall grade of service (subscriber-to-subscriber) on international connections – relating only to the phenomena of congestion in the entire network as a result of the traffic flow – depends on a number of different factors, such as the routing arrangements in the national and international parts of the connection, congestion allowed per switching stage, the methods used to measure traffic and compute the traffic base, and the time differences between the busy hours of the various links involved in the connection.

1.2 The most satisfactory way in which this grade of service could be described would be to give its distribution. The design average grade of service during the busy hour of the complete connection would be the most useful single parameter. However, until such time as continuous traffic measurements are carried out during the busy season in all parts of the network on a routine basis, it is not possible to compute this average grade of service. Therefore, at this stage it cannot be used as a criterion for the dimensioning of the network.

1.3 The only practical way of ensuring an acceptable overall grade of service on international calls is to specify an upper limit on the design loss probability per connecting link in the national network as is done for the links in the international network (see Recommendation E.540).

2 General considerations

2.1 Since the success of the international automatic service is highly dependent on the grade of service of all links involved in the connection from subscriber-to-subscriber, it is desirable that the originating and terminating national network involved in the connection has grade of service standards comparable with those of the international network.

2.2 It is especially important that the links in the country of destination should have a good grade of service for handling the traffic, since high congestion in the terminating national network could have serious effects on the international network. High congestion in the network of the country of destination causes added retrials with consequent increased loading on common switching devices as well as increased occupation of the routes with ineffective calls.

3 Design objectives

3.1 It is recommended that the links in the national network should be designed for a loss probability [1] not exceeding 1 per cent per link in the final choice route during its applicable busy hour. It is recognized, however, that in some countries additional congestion is permitted for the internal switching stages of the transit exchanges. It is also recognized that, where this recommended grade of service is not provided for the national service, it may not be economically feasible to provide it for international relations.

3.2 The maximum number of links in tandem used by an international connection is defined by Recommendation E.171 [1].

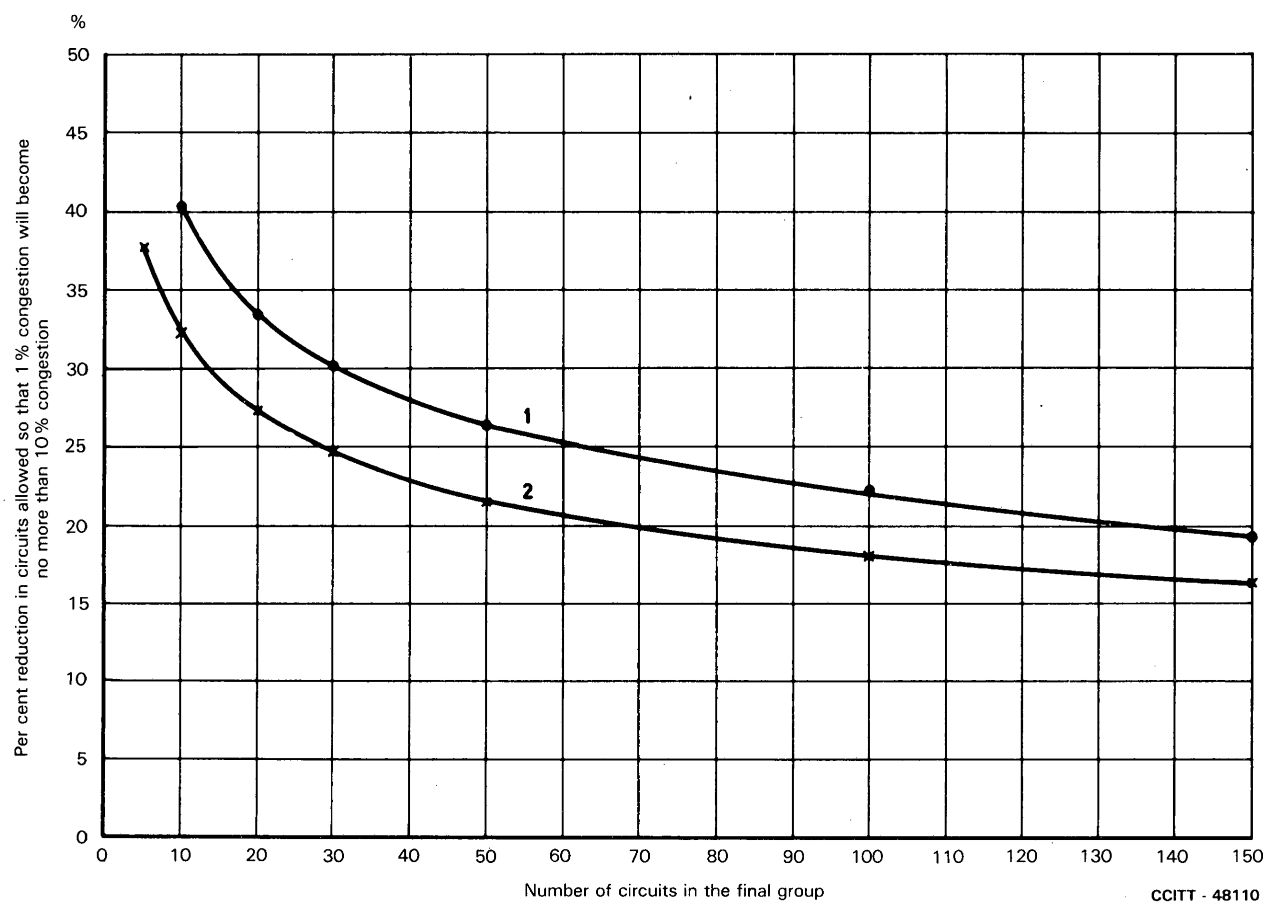
3.3 Although the worst overall grade of service would be approximated by the sum of loss probabilities for individual links connected in tandem, on most calls the overall grade of service will be significantly better.

[1] The loss probability mentioned refers to busy hour traffic values as defined in Recommendation E.500.

4 Maximum traffic loading

4.1 An acceptable automatic service on a final circuit group is difficult to maintain if the traffic loading on the group exceeds a level corresponding to a calculated Erlang grade of service of 10 per cent. Beyond this traffic loading, service on the route may rapidly deteriorate. This condition will be accentuated under the cumulative effect of repeat attempt calls if these should occur.

4.2 The curves of Figure 1/E.541 indicate the proportionate reduction in circuits that may be tolerated for a short period, 15 minutes for example, under normal busy-hour conditions, on a full-availability circuit group dimensioned for 1 per cent Erlang loss, in accordance with the above traffic overload criterion. Table 1/E.541 gives the figures used to plot the curves.



1: peakedness factor = 2.5
2: random traffic (peakedness factor = 1.0)

FIGURE 1/E.541

Proportionate reduction in the number of circuits in a final group in the event of a breakdown if the calculated Erlang grade of service is not to exceed 10 per cent

TABLE 1/E.541

Percentage reduction in the number of circuits if the calculated Erlang grade of service is not to exceed 10 %

Number of circuits	If originally operating at 1 % congestion, % reduction in circuits allowed to yield 10 % congestion	
	Random traffic (peakedness factor = 1.0)	Peakedness factor = 2.5
5	37.7	–
10	32.3	40.2
20	27.2	33.3
30	24.8	30.1
50	21.7	26.5
100	18.3	22.4
150	16.7	19.7

4.3 The curves of Figure 1/E.541 are intended merely as a guide. If the breakdown occurs during an exceptionally busy hour, the permissible proportionate reduction will be less. Conversely, if the breakdown occurs during an hour of light traffic, a higher proportionate reduction in circuits could be tolerated. A higher reduction might also be acceptable after an appropriate oral announcement has been introduced. In the general case, a knowledge of the circuit occupancy will enable an estimate to be made of the prevailing Erlang loss figure with the reduced number of circuits.

The permissible reduction in the case of large groups should not be exceeded; otherwise very serious congestion can result from repeated attempts.

5 General notes

Note 1 – Teletraffic implications for international switching and operational procedures under failure of a transmission facility are discussed in Supplement No. 5 of this fascicle.

Note 2 – Alternative routing in the national and in the international networks provides on average a grade of service that is better than that provided in the theoretical final route.

Note 3 – Non-coincidence of traffic peaks in the national and international networks will provide reduction in the overall grade of service compared with the sum of the design grade of service values per link.

Note 4 – Time differences will also improve the resulting grade of service.

Note 5 – The methods of measuring and calculating the traffic base for provisioning purposes in the national networks may be different in various countries and differ from the methods for the international network given in Recommendation E.500. This means that the national traffic values are not always comparable among themselves or with the values of the international network. Each Administration must estimate how its design traffic level compares with that recommended for the international network.

Note 6 – The design grade of service value of each link will only apply if the traffic at each switching stage is equal to the forecast. In practice, such a situation will seldom occur. Furthermore, the planning procedure normally is such that the specified grade of service should not be exceeded until the end of the planning period. In a growing network, this means that the circuit groups during almost the whole planning period give a better service than the specified critical standard.

In conclusion, the overall grade of service depends on the accuracy of forecasts made and the planning procedure used, i.e. it depends on the interval between plant additions and on the specific traffic value in future to which the grade of service is related.

Reference

[1] CCITT Recommendation *International routing plan,* Rec. E.171.

GRADES OF SERVICE IN DIGITAL INTERNATIONAL TELEPHONE EXCHANGES

1 Introduction

1.1 The grade of service (GOS) parameters and values to be used as dimensioning standards and as performance objectives for international telephone exchanges are indicated below. Procedures to monitor the actual GOS performance of the exchanges are also recommended.

1.2 The GOS standards for international telephone exchanges defined in this Recommendation assume "fully operative" conditions for the exchange and they are based on the load levels specified in Recommendation E.500.

2 Scope of the Recommendation

2.1 The GOS standards are specified for an exchange as a whole, i.e. neither the delay nor the loss parameters are associated solely with the control area or with the connecting network, so that no particular system concept is favoured.

2.2 Although the GOS parameters defined in this Recommendation apply to the digital as well as to the analogue exchanges, the numerical values recommended for these parameters are primarily intended for digital exchanges. The GOS may be too stringent for analogue exchanges and Administrations are advised to make suitable allowances when applying to the analogue exchanges.

Administrations may also consider these GOS values for dimensioning the national transit exchanges so that the end-to-end GOS performance for international connections is maintained at a high level.

3 Grade of service parameters

The loss and delay GOS standards are defined as follows:

3.1 *Loss grade of service*

internal loss probability: for any call attempt, it is the probability that an overall connection cannot be set up between a given incoming circuit and any suitable free outgoing circuit within the switching network.

The loss grade of service is to be met by every pair of incoming and outgoing trunk groups averaged over all inlets of the incoming group.

This approach takes explicit account of the fact that the Administrations will take actions such as the favourable loading of switch blocks in order to balance access to all trunk groups. These actions will minimize the impact of the worst case upon the traffic flow capacity of the switch, by confining the necessary adjustments to localized regions of the switching network.

These actions should ensure that the switching system operates as efficiently as possible within the constraints imposed by this loss standard.

3.2 *Delay grade of service in case of channel-associated signalling*

incoming response delay: The interval from the instant when an incoming seizure signal has arrived at the incoming side of the exchange to the instant when a proceed-to-send signal is returned to the preceding exchange by the receiving exchange.

The incoming response delay may affect the holding time of the preceding trunks and of the common control equipment in the preceding exchange(s). It may also be perceived by the subscriber as dial-tone delay, in case of special dial tone for international calls in outgoing international exchanges, or may contribute to the post-dialling delay experienced by the subscriber in all other cases. The contribution to post-dialling delay does not necessarily comprise the whole of the incoming response delay.

Note – The above definition of incoming response delay does not explicitly mention that it includes receiver attachment delay. However, for the purpose of this Recommendation, it is assumed that receiver attachment delay is a part of the incoming response delay.

3.3 *Delay grade of service in case of any combination of channel-associated and common channel signalling*

exchange call set-up delay: The interval from the instant when the address information required for setting up a call is received at the incoming side of the exchange to the instant when the seizing signal or the corresponding address information is sent to the subsequent exchange.

through-connection delay (end-to-end channel associated or common channel signalling): the interval from the instant when the information required for setting up a through-connection in an exchange is available for processing in the exchange to the instant when the switching network through-connection is established between the incoming and outgoing circuits.

through-connection delay (link-by-link channel associated signalling): the interval from the completion of outpulsing to the establishment of a communication path through the exchange between the incoming and the outgoing circuits.

4 Grade of service standards

The values shown in Table 1/E.543 are recommended for GOS standards of international digital telephone exchanges. The normal and high load levels are the ones defined in Recommendation E.500.

TABLE 1/E.543

	Normal load	High load
Incoming response delay [a] [b]	P(> 0.5 sec.) ⩽ 5%	P(> 1 sec.) ⩽ 5%
Exchange call set-up delay [b]	P(> 0.5 sec.) ⩽ 5%	P(> 1 sec.) ⩽ 5%
Through-connection delay [b]	P(> 0.5 sec.) ⩽ 5%	P(> 1 sec.) ⩽ 5%
Internal loss probability [c]	0.002	0.01

[a] See Note in § 3.2.

[b] The determination of the number of bids for the different devices or exchange modules at normal and high load levels should be made according to Rec. E.500. Circuit group or exchange load levels will be used according to the devices or exchange modules affected.

[c] The values of traffic offered to the circuit group and to the switching network of the exchange, to be used for loss probability evaluation, should correspond to the traffic flow levels defined for circuit groups and exchanges, respectively, in Rec. E.500.

In case of differences between exchange and circuit group busy hours it is recommended to use models which can take account of the different traffic values in the different parts of the exchange. For example, models used for dimensioning the auxiliary equipment could take advantage of the differences of busy hour of the different circuit groups using the same auxiliary equipment.

5 Measurements to monitor exchange GOS performance

In the context of traffic Administration, monitoring the GOS performance in an exchange is a means of detecting potential problems which can affect the GOS performance of that exchange. By analysing deviations from previously established GOS performance thresholds, problem areas can be detected. After having identified the problems, actions such as load balancing, fault removal, extensions, etc., can be derived from GOS performance monitoring. These actions are not taken on a real-time basis, and consequently the data collection and analysis do not have real-time constraints. The traffic measurements recommended below do not separate the causes of call attempt failure or excessive delay.

When the values of the GOS performance are consistently worse than the GOS standards specified in § 4, it will be necessary to identify the causes of such a situation through the analysis of ad-hoc measurement procedures. Considering the above framework, errors in GOS estimation are only important to the extent that they can generate over- or under-reactions to exchange situations.

For each of the GOS parameters a statistical estimator has been defined. The measurements must be made on a per circuit group and per exchange basis. Eventually, savings could be derived from delay measurements made on the basis of signalling types when several circuit groups share the same auxiliary devices. All measurements described below refer to a specific measurement period.

5.1 *Delay measurements*

5.1.1 *Incoming response delay*

The exchange GOS performance with respect to this parameter can be estimated by means of the ratio:

$$p = \frac{B}{A},$$

where

A is the number of call attempts accepted for processing from a given incoming circuit group

B is the number of call attempts out of the set *A*, for which the incoming response delay exceeded the predetermined value *X*

Note – In SPC exchanges a certain time may elapse from the moment that the incoming seizure signal appears at the incoming circuit until the moment that the processor accepts the call attempt for processing. Measuring this delay would require external equipment to the call handling processors. The above measurement only provides an indication of the incoming response delay after the call has been accepted for call processing. In the case where this delay is significant, it should be taken into account in dimensioning and should be subtracted from the total time allowed for the incoming response delay.

5.1.2 *Exchange call set-up delay*

The exchange performance with respect to this parameter can be measured by means of the following ratio:

$$q = \frac{D}{C},$$

where

C is the number of call attempts for which sufficient address information has been received at the incoming side of the exchange, which are addressed to a certain outgoing circuit group and for which the seizing signal or the corresponding address information is sent to the subsequent exchange.

D is the number of call attempts already counted in *C* for which the call set-up delay exceeds the predetermined value, *T*.

5.1.3 *Through-connection delay*

The exchange performance with respect to this parameter can be measured by means of the following ratio:

$$r = \frac{F}{E},$$

where

E (for end-to-end channel associated and common channel signalling) is the number of call attempts for which the required information for setting up a through-connection is available for processing in the exchange for a certain circuit group.

E (for link-by-link channel associated signalling) is the number of call attempts which have completed outpulsing in a certain circuit group.

F is the number of call attempts already counted in *E* for which the through-connection delay has exceeded the predetermined value *V*.

Note 1 – The loss of call attempts caused by the exchange itself, premature subscriber release or time-out expiration in an upstream exchange may modify the outcome of the above delay measurements. However, the effect will only be significant under abnormal conditions which should be investigated separately.

Note 2 – It is recommended that values for X, T, and V be either 0.5 s (normal load) or 1 s (high load).

Note 3 – Measuring delays on a per call basis could produce severe cost penalties to the exchange. Since the accuracy requirements from the statistical viewpoint are not very high, call sampling procedures or test calls can be sufficient for GOS monitoring purposes.

5.2 *Loss measurements*

One estimator of this parameter per circuit group is:

$$s = \frac{H}{G},$$

where:

G is the number of call attempts which require a connection from an inlet to the desired outgoing circuit group having at least one free circuit and for which sufficient call handling information was made available to the exchange.

H is the number of those call attempts described by G which failed to build up the required connection.

Note – The loss of call attempts caused by premature subscriber release or time-out expiration in an upstream exchange, may modify the outcome of the above measurement.

Recommendation E.550

GRADE-OF-SERVICE AND NEW PERFORMANCE CRITERIA UNDER FAILURE CONDITIONS IN INTERNATIONAL TELEPHONE EXCHANGES

1 Introduction

1.1 This Recommendation is confined to failures in a single exchange and their impact on calls within that exchange – network impacts are not covered in these Recommendations.

1.2 This Recommendation from the viewpoint of exchange Grade of Service (GOS) has been established.

1.3 In conformity with Recommendation E.543 for transit exchanges under normal operation, this Recommendation applies primarily to international digital exchanges. However, Administrations may consider these Recommendations for their national networks.

1.4 The GOS seen by a subscriber (blocking and/or delay in establishing calls) is not only affected by the variations in traffic loads but also by the partial or complete faults of network components. The concept of customer-perceived GOS is not restricted to specific fault and restoration conditions. For example, the customer is usually not aware of the fact that a network problem has occurred, and he is unable to distinguish a failure condition from a number of other conditions such as peak traffic demands or equipment shortages due to routine maintenance activity. It is therefore necessary that suitable performance criteria and GOS objectives for international telephone exchanges be formulated that take account of the impact of partial and total failures of the exchange. Further, appropriate definitions, models and measurement and calculation methods need to be developed as part of this activity.

1.5 From the subscriber's point of view, the GOS should not only be defined by the level of unsatisfactory service but also by the duration of the intervals in which the GOS is unsatisfactory and by the frequency with which it occurs. Thus, in its most general form the performance criteria should take into account such factors as: intensity of failures and duration of resulting faults, traffic demand at time of failures, number of subscribers affected by the failures and the distortions in traffic patterns caused by the failures.

However, from a practical viewpoint, it will be desirable to start with simpler criteria that could be gradually developed to account for all the factors mentioned above.

1.6 Total or partial failures within the international part of the network have a much more severe effect than similar failures in the national networks because the failed components in the national networks can be isolated and affected traffic can be rerouted.

Failures in the international part of the network may therefore lead to degraded service in terms of increased blocking delays and even complete denial of service for some time. The purpose of this Recommendation is to set some service objectives for international exchanges so that the subscribers demanding international connections are assured a certain level of service.

It should be noted however that where there are multi-gateway exchanges providing access to and from a country, with diversity of circuits and provision for restoration, the actual GOS will be better than that for the single exchange.

2 General considerations

2.1 The new performance criteria being sought involve concepts from the field of "availability" (intensity of failures and duration of faults) and "traffic congestion" (levels of blocking and/or delay). It is therefore necessary that the terminology, definitions and models considered should be consistent with the appropriate CCITT Recommendations on terminology and vocabulary.

2.2 During periods of heavy congestion, caused either by traffic peaks or due to malfunction in the exchange, a significant increase in repeated attempts is likely to occur. Further, it is expected that due to accumulated demands during a period of complete faults, the exchange will experience a heavy traffic load immediately after a failure condition has been removed and service restored. The potential effects of these phenomena on the proposed GOS under failure conditions should be taken into account (for further study).

3 Exchange performance characteristics under fault situations

3.1 The exchange is considered to be in a fault situation if any failure in the exchange (hardware, software, human errors) reduces its throughput when it is needed to handle traffic. The following four classes of exchange faults are included in this Recommendation:

a) complete exchange faults;

b) partial faults resulting in capacity reduction in all traffic flows to the same extent;

c) partial faults in which traffic flows to or from a particular point are restricted or totally isolated from their intended route;

d) intermittent fault affecting a certain proportion of calls.

3.2 To the extent practical, an exchange should be designed so that the failure of a unit (or units) within the exchange should have as little as possible adverse affect on its throughput. In addition, the exchange should be able to take measures within itself to lessen the impact of any overload resulting from failure of any of its units. Units within an exchange whose failure reduces the exchange throughput by greater amounts than other units should have proportionally higher availability (Recommendation Q.504, § 4).

3.3 When a failure reduces exchange throughput and congestion occurs, the exchange should be able to initiate congestion control indications to other exchanges and network management systems so as to help control the offered load to the exchange, (Recommendations E.410 and Q.506).

4 GOS and applicable models

4.1 In this section, the terms "accessible" and "inaccessible" are used in the sense defined in Recommendation G.106 (*Red Book*). The GOS for exchanges under failure conditions can be formulated at the following two conceptual levels from a subscriber's viewpoint:

4.1.1 *Instantaneous service accessibility (inaccessibility)*

At this level, one focuses on the probability that the service is accessible (not accessible) to the subscriber at the instant he places a demand.

4.1.2 *Mean service accessibility (inaccessibility)*

At this level, one extends the concept of "downtime" used in availability specifications for exchanges to include the effects of partial failures and traffic overloads over a long period of time.

4.2 Based on the GOS concept outlined in § 4.1, the GOS parameters for exchanges under failure conditions are defined as follows:

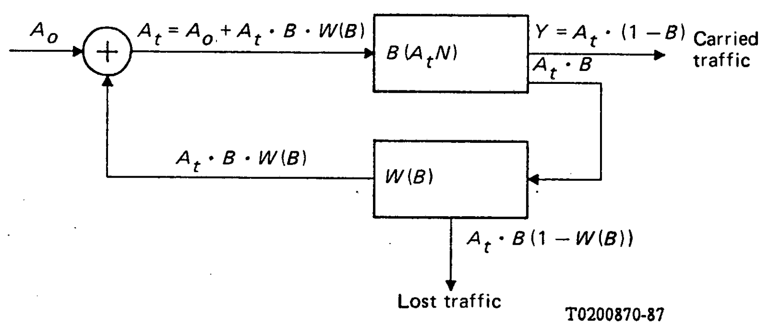
4.2.1 **instantaneous exchange inaccessibility** is the probability that the exchange in question cannot perform the required function (i.e. cannot successfully process calls) under stated conditions at the time a request for service is placed.

4.2.2 **mean exchange service inaccessibility** is the average of instantaneous exchange service inaccessibility over a prespecified observation period (e.g. one year).

4.2.3 *Note 1* – The GOS model in the case of instantaneous exchange inaccessibility parallels the concept of the call congestion in traffic theory and needs to be extended to include the call congestion caused by exchange failures classified in § 3.1. The GOS value can then be assigned on a basis similar to Recommendation E.543 for transit exchanges under normal operation.

Note 2 – A model for estimating the mean exchange inaccessibility is provided in Annex A. Though the model provides a simple and hence attractive approach, some practical issues related to measurement and monitoring and the potential effects of network management controls and scheduled maintenance on the GOS need further study.

4.3 The model in Figure 1/E.550 outlines the change in the nature of traffic offered under failure conditions.



where:

A_o = offered traffic

A_t = total traffic

B = congestion factor (call attempts not processed) which may include the effects of network management controls

N = resources

Y = carried traffic

W = proportion of blocked call attempts that reattempt

FIGURE 1/E.550

Model for traffic offered under failure conditions

In normal conditions the congestion factor B is low and there should be few repeat attempts: as a consequence the traffic A_t approximates A_o.

Under failure conditions there is a reduction in resources and the congestion factor B increases. This provokes the phenomenon of repeat attempts and hence the load A_t on the exchange becomes greater than the original A_o.

Therefore it is necessary to evaluate the congestion with the new load A_t assuming system stability exists, which may not always be the case.

Recommendation E.501 furnishes the appropriate models to detect the traffic offered from the carried traffic taking into account the repeat attempts.

4.4 The impact on the GOS for each of the exchange fault modes can be characterized by:

- load in Erlangs (A_t) and busy hour call attempts (BHCA);
- inaccessibility (instantaneous and mean), congestion and delay parameters (call set-up, through-connection, etc.);
- fault duration;
- failure intensity.

5 GOS standards and inaccessibility

5.1 Exchange fault situations can create similar effects to overload traffic conditions applied to an exchange under fault free conditions.

In general, digital exchanges operating in the network should be capable of taking action to ensure maximum throughput when they encounter an overload condition, including any that have been caused by a fault condition within the exchange.

Calls that have been accepted for processing by the exchange should continue to be processed as expeditiously as possible, consistent with the overload protection strategies recommended in § 3 of Recommendation Q.543.

5.2 One of the actions the exchange may take to preserve call processing capacity is to initiate congestion controls and/or other network management actions, to control the load offered to the exchange (Recommandations E.410, E.413 and Q.506). The most obvious impact from the caller's viewpoint may be a lowering of the probability that the network as a whole will be able to complete some portion of the call attempts that the exchange is unable to accept during the failure condition.

5.3 International exchanges occupy a prominent place in the network and it is important that their processing capacity have high availability. There are likely to be many variations in exchange architectures and sizes that will have different impacts in the categories of failure and the resulting loss of capacity.

In general, failures that cause large proportions of exchange capacity to be lost must have a low probability of occurring and a short downtime. It is important that maintenance procedures to achieve appropriate exchange availability performance be adopted.

5.4 The formal expression of the criterion of mean exchange service inaccessibility is as follows:

Let:

$y(t)$: Intensity of call attempts gaining access through the exchange assuming no failures.

$s(t)$: Intensity of call attempts actually given access through the exchange, taking into account the fault conditions which occur in the exchange.

Then the mean exchange service inaccessibility during a period of time T is given by

$$P = \frac{1}{T} \int_0^T \frac{y(t) - s(t)}{y(t)} dt$$

Annex A describes a practical implementation of this criterion.

For periods in which the exchange experiences a complete fault, i.e. $s(t) = 0$, the expression:

$$\frac{y(t) - s(t)}{y(t)} \quad \text{is equal to 1.}$$

The contribution of such periods to the total criterion P may then be expressed simply as the fraction P_{total} of the evaluation period T during which complete exchange outage due to failure occurred.

The objective for P_{total} is given as P_{total} not more than 0.4 hours per year.

For the period of partial failure, it is convenient to also express the objective as equivalent hours per year – the term equivalent is used because the duration of partial faults is weighted by the fraction:

$$\frac{y(t) - s(t)}{y(t)}$$

of call attempts denied access. The objectives for the contribution of period of partial exchange faults to the total criterion P is given by:

$P_{partial}$ not more than 1.0 equivalent hours per year.

Note that by definition $P = P_{total} + P_{partial}$

The inaccessibility criterion does not cover:

- planned outages
- faults with duration of less than 10 seconds
- accidental damage to equipment during maintenance
- external failures such as power failures, etc.

It does cover failures resulting from both hardware and software faults.

In addition, the objectives relate to the exchange under normal operating conditions and do not include failures just after cutover of an exchange or those during the end of the period it is in service, i.e. the well known "bath tub" distribution.

6 Performance monitoring

Certain failure conditions [i.e. the type mentioned in § 3.1, b)] usually will be reflected in the normal GOS performance measurements called for in Recommendation E.543.

Other failure conditions [i.e. the type mentioned in § 3.1, c)] can result in a reduced performance for a portion of traffic flows but with little or no impact on measured exchange GOS. For example if a trunk module in a digital exchange fails, the traffic normally associated with that module is completely blocked, but since the attempts are also not measured the failure does not change the monitoring of the exchange GOS.

For this second situation, the mean inaccessibility can be calculated using direct measurement of unit outages to provide m_i and t_i information and estimates of b_i together with the model of Annex A. (See Annex A for an explanation of these symbols.)

The estimates of b_i can incorporate both fixed factors based on exchange architecture and variable factors based on traffic measurements just prior to the time of failure.

ANNEX A

(to Recommendation E.550)

A model for mean exchange inaccessibility

A.1 Let P be the probability that a call attempt is not processed due to a fault in the exchange, then:

$$P = \sum_{i=1}^{N} p_i b_i \quad \text{(A-1)}$$

where:

p_i is the probability of fault mode i. Each fault mode denotes a specific combination of faulty exchange components

N is the number of the fault mode

b_i is the average proportion of traffic which cannot be processed due to the fault mode i. It is a function of the specific fault present and the offered traffic load at the time of the failure condition.

During a period of time T, the fault probability p_i may be estimated by:

$$p_i = \frac{m_i \cdot t_i}{T} \qquad i = 1, 2, \dots N \quad \text{(A-2)}$$

where:

m_i is the number of occurrences of fault mode i during the period T

t_i is the average duration of occurrences of fault mode i

As a practical matter, one may wish to exclude from the calculation faults of duration less than 15 seconds.

Note 1 – A given fault mode causes the exchange to enter the corresponding fault state, which is characterized by a given mean duration and a function b_i giving the proportion of offered traffic affected. In principle, the possible number of fault modes can be very large because of the number of combinations which can occur. In practice this number can be reduced by considering all fault modes with the same b_i and t_i as equivalent.

Note 2 – b_i should take into account the distribution of traffic during a day and the probability of fault mode i occurring in a given time period. The value assigned in the above model should be the average b_i value for all hours considered in these distributions. For example, a partial fault affecting 20% of the exchange traffic throughput in the busy hour and 2 similar hours, could be evaluated to effect a 10% reduction in 4 other moderately busy hours and to have negligible impact during all other hours. If this fault is considered to be equally probable in time, the average value of b_i can be obtained as follows:

$$b_i = \text{Sum of} \left(\frac{\text{Percentage of traffic affected} \times \text{number of relevant busy hours}}{24 \text{ hours}} \right) =$$

$$= \frac{0.2 \times 3}{24} + \frac{0.1 \times 4}{24} + \frac{0.0 \times 17}{24} = 0.025 + 0.0167 = 0.0417$$

Note 3 – The probability that a call attempt is not processed relates to the category of traffic affected by the fault. Other traffic will experience a different GOS depending on system architecture which is not taken into account in this Recommendation. For example, partial faults which remove from service blocks of trunks connected to an exchange have the effect of reducing the total traffic offered to the exchange. The traffic flows not using the failed trunks could thus have a slightly improved GOS.

A.2 *Example for calculating the inaccessibility, P*

See Table A-1/E.550.

TABLE A.1/E.550

An example of using the model for calculating the inaccessibility P

(T = 1 year = 8760 hours)

b_i	m_i	t_i	$p_i \cdot b_i$
Average proportion of traffic which cannot be processed	Number of failures of type i per year	Average duration of failure type i (hours)	Probability that a call attempt is not processed ($\times 10^{-5}$)
1.00	2	0.2	4.56
0.40	3	0.22	3.01
0.20	4	0.3	2.74
0.10	6	0.4	2.74
0.05	10	0.5	2.85

The value of P is the sum of the individual $p_i \cdot b_i$ terms in Table A-1/E.550. In this example $P = 15.90 \times 10^{-5}$ which is equivalent to 1.39 hours of inaccessibility per year ($1.39 = 15.90 \times 10^{-5} \times 8760$). P decomposes as follows:

P_{total} = 0.40 hours per year ($4.56 \times 10^{-5} \times 8760$)

$P_{partial}$ = 0.99 hours per year (the remaining part of P)

A.3 As a further example consider a circuit group where exchange failures may occur which disable one or more circuits (see Figure A-1/E.550). It is possible to expand the formula (A-1).

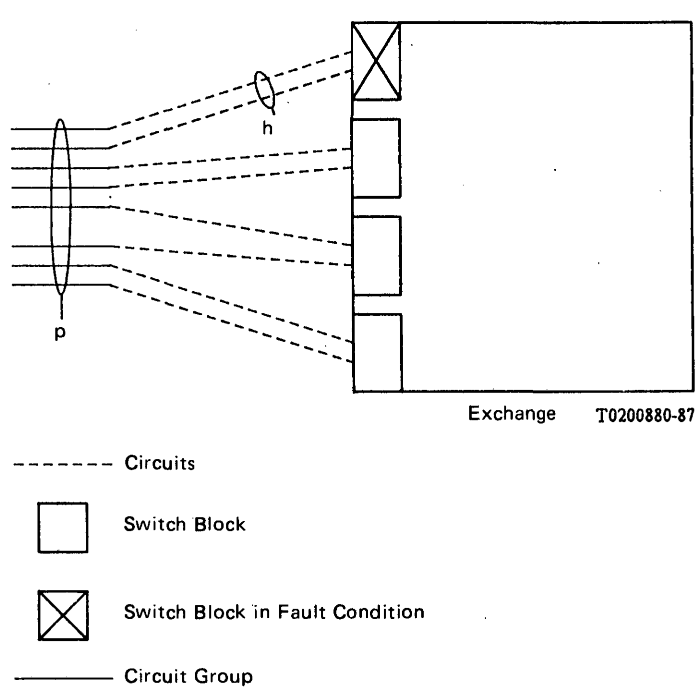


FIGURE A-1/E.550

Exchange failure disabling one or more circuits

The average proportion of traffic $b(n, k, A)$, which cannot be processed due to failures on circuits is now a function of:

- n, the size of the circuit group;
- k, number of circuits out of order because of the failure;
- A, the mean traffic offered to the circuit group, in the absence of faults.

Let the throughput of a circuit group of size n with a traffic offered A be $C_n(A)$ – then the throughput of the same circuit group is $C_{n-k}(A)$ where k circuits are out of order – hence the average proportion of traffic $b(n, k, A)$ which cannot be processed because of the failure is given by:

$$b(n, k, A) = \frac{[C_n(A) - C_{n-k}(A)]}{C_n(A)} \qquad \text{(A-3)}$$

Let

$f(k, A)$ be the probability for having k circuits in a fault condition and the mean offered traffic A. The probability, P_n, that a call attempt is not processed due to a failure on a circuit group of size n, is given by:

$$P_n = \sum_{k, A} f(k, A) \cdot b(n, k, A) \qquad k = 1, 2, \dots n \qquad \text{(A-4)}$$

If k and A are independent then

$$f(k, A) = f_1(k) \cdot f_2(A) \qquad \text{(A-5)}$$

where $f_1(k)$ may satisfy a binomial distribution and $f_2(A)$ a Poisson distribution.

Suppose the traffic follows an Erlang distribution, $C_n(A)$ is proportional to $A \cdot (1 - E_n(A))$, where $E_n(A)$ is the blocking probability expressed by the Erlang loss formula. Hence:

$$b(n, k, A) = \frac{E_{n-k}(A) - E_n(A)}{1 - E_n(A)} \qquad \text{(A-6)}$$

can be found by using the Erlang tables and then inserting the value into equation (A-4).

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

SECTION 6

DEFINITIONS

Recommendation E.600

TERMS AND DEFINITIONS OF TRAFFIC ENGINEERING

Introduction

This Recommendation provides terms and definitions for use in the field of traffic engineering. Traffic engineering includes measurements, forecasting, planning, dimensioning and performance monitoring. Traffic engineering has a goal of ensuring trafficability performance objectives for telecommunications services. Trafficability performance is one of the major factors in Quality of Service (QOS). Recommendation E.800 explains the relation of various Quality of Service factors and gives terms and definitions for Quality of Service concepts and for availability and reliability aspects.

The purpose of this vocabulary is to aid in the understanding of traffic engineering and related Recommendations. The terms defined here may also be defined differently for applications outside the area of traffic engineering.

Alternatives for the preferred terms are given following a semi-colon.

LIST OF TERMS

1 *General theory*

1.1 Communication
1.2 Connection
1.3 Resource
1.4 User
1.5 Telecommunications traffic, teletraffic
1.6 Observed traffic
1.7 Poisson traffic; pure chance traffic
1.8 Peakedness factor
1.9 Smooth traffic
1.10 Peaked traffic
1.11 Traffic of volume
1.12 Erlang
1.13 Bid
1.14 Seizure
1.15 Idle (state)
1.16 Busy (state)
1.17 Release
1.18 Holding time
1.19 Blocked mode of operation
1.20 Delay mode of operation
1.21 Call congestion
1.22 Time congestion
1.23 Waiting time; queuing time

2 *Calls*

3 *Circuits*

4 *Grade of service*

5 *Traffic engineering*

1 General theory

1.1 communication

F: communication

S: comunicación

Transfer of information according to agreed conventions. The information flow need not be bidirectional.

1.2 **connection**

F: connexion

S: conexión

An association of resources providing means for communication between two or more devices in, or attached to, a telecommunication network.

1.3 **resource**

F: ressource

S: órgano

Any set of physically or conceptually identifiable entities within a telecommunications network, the use of which can be unambiguously determined.

1.4 **user**

F: usager

S: usuario

Any entity external to the network which utilizes connections through the network for communication.

1.5 **telecommunications traffic; teletraffic**

F: trafic de télécommunications: télétrafic

S: tráfico de telecomunicación; teletráfico

A process of arrivals and releases of demands for resources in a network.

Note – The unit for the variable traffic is the Erlang (symbol: E).

1.6 **observed traffic**

F: trafic observé

S: tráfico observado

Instantaneous observed traffic is the amount of occupied resources at a given instant. Average observed traffic is the time average of instantaneous observed traffic over a given period.

1.7 **poisson traffic; pure chance traffic**

F: trafic poissonnien: trafic de pur hasard

S: tráfico poissoniano

Traffic that has a Poisson distribution of arrivals.

Note – Poisson traffic has a peakedness factor equal to 1.

1.8 **peakedness factor**

F: facteur d'irrégularité

S: factor de irregularidad

The ratio of variance to mean of a traffic.

1.9 **smooth traffic**

F: trafic régularisé

S: tráfico con distribución uniforme

Traffic that has a peakedness factor less than 1.

1.10 **peaked traffic**

F: trafic survariant

S: tráfico con distribución en pico

Traffic that has a peakedness factor greater than 1.

1.11 **traffic volume**

F: volume de trafic

S: volumen de tráfico

The integral of the instantaneous traffic over a given time interval.

Note 1 – Traffic volume is equal to the sum of the holding times of the resources.

Note 2 – A unit used for traffic volume is the Erlang hour (symbol: E).

1.12 **erlang**

F: erlang

S: erlang

The unit of traffic (symbol: E). In traditional telephony the number of Erlangs is the number of busy resources or the expected number of busy resources under stated conditions.

1.13 **bid**

F: tentative de prise

S: tentativa de toma

A single attempt to obtain the use of a resource of the type under consideration.

Note – In a network management context, the absence of a qualification implies a bid to a circuit group, a route or a destination.

1.14 **seizure**

F: prise

S: toma

A bid that obtains the use of a resource of the type under consideration.

1.15 **idle (state)**

F: libre

S: reposo (estado de); estado libre

Condition of a resource that is free to be seized.

1.16 **busy (state)**

F: occupé

S: ocupado (estado de)

Condition of a resource following its seizure.

1.17 **release**

F: libération

S: liberación

The event which changes the condition of a resource from busy to idle.

1.18 **holding time**

F: durée d'occupation

S: tiempo de ocupación; tiempo de retención

The time between the seizure of a resource and its release.

1.19 **blocked mode of operation**

F: mode d'exploitation avec blocage

S: modo de operación con bloqueo (de llamadas)

A mode of operation in which bids which find no suitable resources idle and accessible are not permitted to wait.

1.20 **delay mode of operation**

F: mode d'exploitation avec attente

S: modo de operación con espera (de llamadas)

A mode of operation in which bids which find no suitable resources idle and accessible are permitted to wait.

1.21 **call congestion**

F: encombrement d'appel

S: congestión de llamadas

The probability that a bid to a particular pool or resources will not result in an immediate seizure.

1.22 **time congestion**

F: congestion temporelle

S: congestión temporal

The proportion of time that a particular pool of resources does not contain any idle resource.

1.23 **waiting time; queuing time**

F: temps de mise en attente

S: tiempo de espera; tiempo de cola

In delay mode of operation, the time interval between the bid for a resource and its seizure.

2 Calls

2.1 **call**

F: appel

S: llamada

A generic term related to the establishment, utilization and release of a connection. Normally a qualifier is necessary to make clear the aspect being considered, e.g. call attempt.

2.2 **call intent**

F: intention d'appel

S: intención de llamada; intento de llamada

The desire to establish a connection to a user.

Note – This would normally be manifested by a call demand. However, demands may be suppressed or delayed by the calling user's expectation of poor Quality of Service performance at a particular time.

2.3 **call demand**

F: demande d'appel

S: demanda de llamada

A call intent that results in a first call attempt.

2.4 **call attempt**

F: tentative d'appel

S: tentativa de llamada

An attempt to achieve a connection to one or more devices attached to a telecommunications network.

Note – At a given point in the network a call attempt is manifested by a single unsuccessful bid, or a successful bid and all subsequent activity related to the establishment of the connection.

2.5 **first call attempt**

F: première tentative d'appel

S: primera tentativa de llamada

The first attempt of a call demand that reaches a given point of the network.

2.6 **repeated call attempt; reattempt**

F: tentative d'appel répétée

S: tentativa de llamada repetida

Any of the call attempts subsequent to a first call attempt related to a given call demand.

Note – Repeated call attempts may be manual, i.e. generated by humans, or automatic, i.e. generated by machines.

2.7 **call string**

F: chaîne d'appel

S: cadena de llamada

All the call attempts related to a single demand.

2.8 **blocked call attempt**

F: tentative d'appel bloquée

S: tentativa de llamada bloqueada

A call attempt that is rejected owing to a lack of resources in the network.

2.9 **abandoned call attempt**

F: tentative d'appel abandonnée

S: tentativa de llamada abandonada

A call attempt aborted by the calling user.

2.10 **successful call attempt; fully routed call attempt**

F: tentative d'appel acheminée

S: tentativa de llamada fructuosa; tentativa de llamada totalmente encaminada

A call attempt that receives intelligible information about the state of the called user.

2.11 **completed call attempt; effective call attempt**

F: tentative d'appel ayant abouti; tentative d'appel efficace

S: tentativa de llamada completada; tentativa de llamada eficaz

A successful call attempt that receives an answer signal.

2.12 **successful call**

F: appel ayant abouti

S: llamada fructuosa

A call that has reached the wanted number and allows the conversation to proceed.

2.13 **completion ratio**

F: taux d'efficacité

S: relación respuesta/toma; tasa de compleción; tasa de eficacia

The ratio of the number of completed call attempts to the total number of call attempts, at a given point of a network.

2.14 **answer seizure ratio (ASR)**

F: taux de prise avec réponse (TPR)

S: tasa de tomas con respuesta (TTR)

On a route or a destination code basis, and during a specified time interval, the ratio of the number of seizures that result in an answer signal, to the total number of seizures.

2.15 **answer bid radio (ABR)**

F: taux de tentatives de prise avec réponse (TTPR)

S: tasa de tentativas de toma con respuesta (TTTR)

On a route or a destination code basis and during a specified time period, the ratio of the number of bids that result in an answer signal, to the total number of bids.

2.16 **calling rate**

F: taux d'appel

S: tasa de llamadas

The number of call attempts at a given point, over a period of time, divided by the duration of the period.

2.17 **dialling-time**

F: durée de numérotation

S: tiempo de marcación

Time interval between the reception of dial tone and the end of dialling of the calling user.

3 Circuits

3.1 circuit

F: circuit (de télécommunication)

S: circuito

A transmission means which allows communication between two points.

3.2 trunk circuit

F: circuit (commuté)

S: circuito (entre centrales); circuito troncal

A circuit terminating in two switching centres.

3.3 one way; unidirectional

F: à sens unique; unidirectionnel

S: en un solo sentido; unidireccional

A qualification applying to traffic or circuits which implies that the establishment of a connection always occurs in one direction.

3.4 two way; bidirectional

F: à double sens; bidirectionnel

S: en ambos sentidos; bidireccional

A qualification applying to traffic or circuits which implies that the establishment of a connection may occur in either direction.

3.5 circuit group

F: faisceau (de circuits)

S: haz de circuitos

A group of circuits which are traffic engineered as a unit.

3.6 circuit subgroup

F: sous-faisceau

S: subhaz de circuitos

A part of a circuit group with similar characteristics (e.g. type of signalling, type of transmission path, etc.).

3.7 first choice circuit group

F: faisceau de premier choix

S: haz de circuitos de primera elección

With respect to a particular traffic relation, the circuit group to which this traffic is first offered.

3.8 high usage circuit group

F: faisceau débordant

S: haz de circuitos de gran utilización

With respect to a particular traffic relation, a circuit group that is traffic engineered to overflow to one or more other circuit groups.

3.9 final circuit group

F: faisceau final

S: haz final de circuitos

With respect to a particular traffic relation, a circuit group from which there is no possibility of overflow to another circuit group within the routing scheme currently in effect.

3.10 fully provided circuit group

F: faisceau totalement fourni

S: haz de circuitos totalmente provisto

With respect to a particular traffic relation, a circuit group which is the first choice circuit group for this traffic and which is traffic engineered as a final circuit group.

4 Grade of service

4.1 grade of service (GOS)

F: qualité d'écoulement du trafic

S: grado de servicio (GDS)

A number of traffic engineering variables used to provide a measure of adequacy of a group of resources under specified conditions; these grade of service variables may be probability of loss, dial tone delay, etc.

Note 1 – The parameter values assigned as objectives for grade of service variables are called grade of service standards.

Note 2 – The values of grade of service parameters achieved under actual conditions are called grade of service results.

4.2 quality of service variable

F: variable de qualité de service

S: variable de calidad de servicio

Any performance variable (such as congestion, delay, etc.) which is perceivable by a user.

Note – For a description of the relations of quality of service factors see Recommendation E.800.

4.3 dial-tone delay

F: durée d'attente de tonalité

S: demora del tono de invitación a marcas; periodo de espera del tono de invitación a marcar

Time interval between off hook and reception of dial tone.

4.4 post-dialling delay

F: attente après numérotation

S: demora después de marcar; periodo de espera después de marcar

Time interval between the end of dialling by the user and the reception by him of the appropriate tone or recorded announcement, or the abandon of the call without tone.

4.5 answer-signal delay

F: délai du signal de réponse

S: demora de la señal de respuesta

Time interval between the establishment of a connection between calling and called users, and the detection of an answer signal at the originating exchange.

4.6 incoming response delay

F: durée de présélection

S: demora de la preselección; duración de la preselección

The interval from the instant when an incoming seizure is recognizable at the incoming side of the exchange to the instant when the proceed to send signal is sent to the preceding exchange by the receiving exchange.

Note – This definition is only applicable in the case of channel associated signalling.

4.7 exchange call set-up delay

F: durée de sélection d'un commutateur

S: demora de establecimiento de la comunicación por una central; tiempo de establecimiento de la comunicación por una central

The interval from the instant when the address information required for setting up a call is received at the incoming side of the exchange to the instant when the seizing signal or the corresponding address information is sent to the subsequent exchange.

4.8 through-connection delay

F: durée d'établissement d'un commutateur

S: demora de transconexión; tiempo de transferencia de una central

The interval from the instant when the information required for setting up a through-connection in an exchange is available for processing in the exchange, to the instant when the switching network through-connection is established and available for communication.

4.9 **internal blocking**

F: blocage interne

S: bloqueo interno

The probability that a connection cannot be made between a given point in a network and any suitable idle resource in an external pool of resources owing to call congestion within the portion of the network being considered.

4.10 **external blocking**

F: blocage externe

S: bloqueo externo

The probability that a connection cannot be made between a given point in a network and any suitable resource in an external pool of resources owing to call congestion within the pool of resources.

5 Traffic engineering

5.1 **busy hour**

F: heure chargée

S: hora cargada

The continuous 1-hour period lying wholly in the time interval concerned for which the traffic or the number of call attempts is greatest.

5.2 **average daily peak hour traffic**

F: moyenne du trafic des heures chargées

S: tráfico medio de las horas punta

The average busy hour traffic of several days; it is usually not related to the same hour each day.

5.3 **time consistent busy hour**

F: heure chargée moyenne

S: hora cargada media repetitiva o sistemática

The 1-hour period starting at the same time each day for which the average traffic of the resource group concerned is greatest over the days under consideration.

5.4 **day to busy hour ratio**

F: rapport du trafic journalier au trafic à l'heure chargée

S: relación del tráfico diario al tráfico en la hora cargada

The ratio of the 24-hour day traffic volume to the busy hour traffic volume.

Note – Busy hour to day ratio is also used.

5.5 **traffic carried**

F: trafic écoulé

S: tráfico cursado

The traffic served by a pool of resources.

5.6 **traffic offered**

F: trafic offert

S: tráfico ofrecido

The traffic that would be carried by an infinitely large pool of resources.

5.7 **effective traffic**

F: trafic efficace

S: tráfico eficaz

The traffic corresponding only to the conversational portion of effective call attempts.

5.8 **overflow traffic**

F: trafic de débordement

S: tráfico de desbordamiento

The part of the traffic offered to a pool of resources which is not carried by that pool of resources.

5.9 **blocked traffic**

F: trafic bloqué

S: tráfico bloqueado

The part of the overflow traffic that is not carried by subsequent pools of resources.

5.10 **lost traffic; abandoned traffic**

F: trafic perdu; trafic abandonné

S: tráfico perdido; tráfico abandonado

That part of the blocked traffic which does not result in reattempts.

5.11 **suppressed traffic**

F: trafic non exprimé; trafic supprimé

S: tráfico suprimido

The traffic that is withheld by users who anticipate a poor quality of service (QOS) performance.

5.12 **origin**

F: origine

S: origen

The location of the calling user. This may be specified to whatever accuracy is necessary.

5.13 **destination**

F: destination

S: destino

The location of the called network termination. This may be specified to whatever accuracy is necessary; in international working, the area or country code is usually sufficient.

5.14 **traffic relation**

F: flux de trafic

S: relación de tráfico

The traffic between a particular origin and a particular destination.

5.15 **traffic matrix**

F: matrice de trafic

S: matriz de tráfico

A structured presentation of the traffic between a number of origins and destinations.

5.16 **originating traffic**

F: trafic de départ

S: tráfico de origen

Traffic generated within the network considered, whatever its destination.

5.17 **terminating traffic**

F: trafic d'arrivée

S: tráfico de destino

Traffic which has its destination within the network considered, whatever its origin.

5.18 **internal traffic**

F: trafic interne

S: tráfico interno

Traffic originating and terminating within the network considered.

5.19 **incoming traffic**

F: trafic entrant

S: tráfico entrante

Traffic entering the network considered, from outside it, whatever its destination.

5.20 **outgoing traffic**

F: trafic sortant

S: tráfico saliente

Traffic leaving the network considered, destined for sinks located outside it, whatever its origin.

5.21 **transit traffic**

F: trafic de transit

S: tráfico de tránsito

Traffic passing through the network considered.

5.22 **traffic distribution imbalance**

F: déséquilibre interne de trafic

S: desequilibrio de la distribución interna de tráfico

Unevenly distributed traffic among similar resources.

5.23 **route**

F: voie d'acheminement

S: ruta

One or more circuit groups providing a connection between switching centres.

5.24 **traffic routing**

F: acheminement de trafic

S: encaminamiento de tráfico

The selection of routes, for a given traffic relation; this term is applicable to the selection of circuit groups by switching systems or operators, or to the planning of routes.

5.25 **call routing**

F: acheminement d'appel

S: encaminamiento de la llamada

The selection of appropriate circuit subgroups or individual circuits for a particular call attempt.

5.26 **alternative route; alternate route**

F: voie d'acheminement détourné

S: ruta alternativa

A second, or subsequent choice route between two switching centres usually consisting of two or more circuit groups in tandem.

5.27 **network cluster**

F: faisceau de faisceaux

S: agrupación de haces

A final circuit group and all the high usage circuit groups which have at least one traffic relation for which the final circuit group is in the last choice route.

5.28 **equivalent random traffic**

F: trafic equivalent

S: tráfico aleatorio equivalente

The theoretical poisson traffic that, when offered to a theoretical circuit group (equivalent random circuit group) produces an overflow traffic with a mean and variance equal to that of a given offered traffic.

Note – The equivalent random traffic and circuit group represent the traffic impact of a more complex arrangement of offered traffics and high usage circuit groups.

SECTION 7

ISDN TRAFFIC ENGINEERING

Recommendation E.700

FRAMEWORK OF THE E.700 SERIES OF RECOMMENDATIONS

1 Introduction

The E.700 Series of Recommendations are intended to provide guidance on traffic engineering for the ISDN. They are organized in sections to deal with various aspects of traffic engineering.

2 General: Recommendations E.700 to E.709

This section deals with general topics applicable throughout the E.700 Series.

3 Traffic modelling: Recommendations E.710 to E.719

This section provides guidance on how to characterize the traffic that will be offered in ISDNs.

4 Grade of Service: Recommendations E.720 to E.729

This section defines Grade of Service (GOS) concepts and parameters that will be significant in ISDNs and sets objectives for these parameters.

5 Dimensioning methods: Recommendations E.730 to E.739

This section provides methods to relate the offered traffic and Grade of Service objectives in order to allocate sufficient resources for planning and design purposes.

6 Traffic measurements: Recommendations E.740 to E.749

This section describes traffic measurement and performance monitoring requirements.

Recommendation E.701

REFERENCE CONNECTIONS FOR TRAFFIC ENGINEERING

1 General

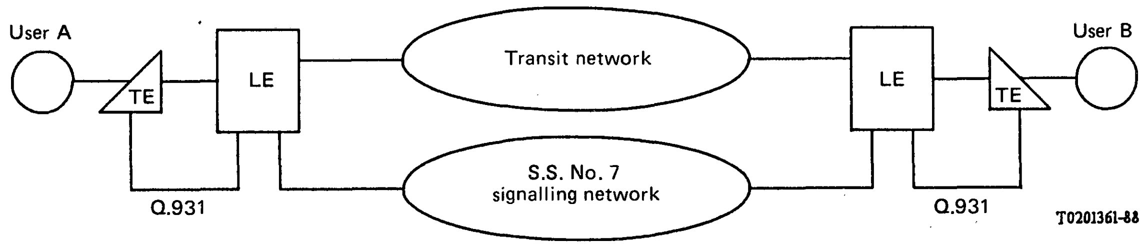
The goal of this Recommendation is to give the E.700 Series of Recommendations a base to define ISDN Grade of Service (GOS) and traffic parameters.

In § 2, two reference connections are defined. Definition of other reference connections is for further study.

2 Reference connections

2.1 *Reference connection for point-to-point circuit switched services*

See Figure 1/E.701.



User A Originating user
User B Terminating user
TE Terminal equipment
LE Local exchange

Note 1 – The transit network may contain zero, one, or more transit exchanges, which may or may not be dedicated ISDN exchanges.

Note 2 – The signalling network may contain zero, one, or more signalling transfer points.

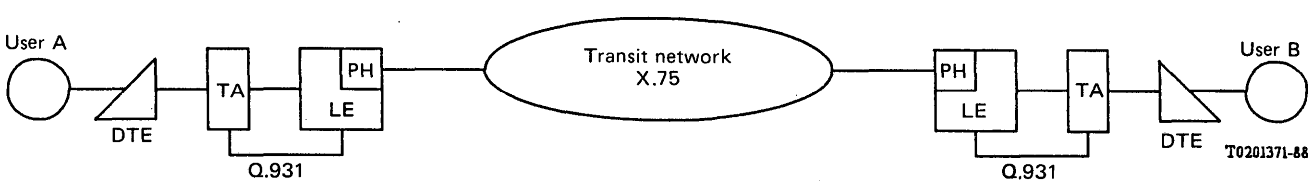
Note 3 – The topology of the signalling network may differ significantly from that of the transit network.

FIGURE 1/E.701

Reference connection for point-to-point circuit switched services

2.2 *Reference connection for point-to-point packet switched services*

See Figure 2/E.701.



User A Originating user
User B Terminating user
DTE Data terminal equipment
TA Terminal adaptor
LE Local exchange
PH Packet handler

Note 1 – The transit network may contain zero, one, or more transit exchanges.

Note 2 – Current CCITT Recommendations support X.25 DTEs using X.31 TAs.

Note 3 – The packet handler may be located outside the local exchange.

FIGURE 2/E.701

Reference connection for point-to-point packet switched services

ISDN TRAFFIC REQUIREMENTS OVERVIEW

1 Introduction

This Recommendation outlines the general consideration in modelling traffic flows in ISDNs. More detailed descriptions for specific services and significant points[1] in the network are given in Recommendations of the E.710 Series as follows:

E.711 – User Demand

E.712 – User Plane Traffic Models

E.713 – Control Plane Traffic Models

E.714 – Management Plane Traffic Models.

Additional Recommendations in this Series will be developed in the future to reflect ISDN developments.

Note – Recommendations E.712 and E.714 are for further study.

2 Context

ISDN concepts, services and networks are described in the Series I Recommendations. The E.710 Series of Recommendations have been developed consistent with the approach. However, the grouping of material in the E.710 Series concentrates on the important aspects from a traffic point of view of ISDN operations in the immediate future.

An important modelling technique used to represent ISDN capabilities is the layered architecture described in Recommendations I.310 and I.320. The E.710 Series have been developed using this approach. At present the E.710 Series concentrates on lower layer (1-3) traffic flows. Higher layer traffic flows are for further study.

The user plane/control plane perspectives described in Recommendation I.320 have been used to provide two separate traffic models in Recommendations E.712 and E.713. Nevertheless, it should be kept in mind that many traffic engineering procedures described in subsequent E.700 Recommendations require incorporating traffic loads from both planes using the same resources. The influence of the management plane is for further study.

The ISDN reference connections used in the E.710 Series are those given in Recommendation E.701.

3 User demand

ISDN users have various needs for information transfer. The user and his terminal equipment transform these needs into a series of call demands for available ISDN services. These transformations involve many functions including coding, and peer-to-peer and inter-layer protocols. The higher layer functions are not analysed in the E.710 Series.

Recommendation E.711 starts from the expression by the user of a call demand to use the ISDN services defined in Recommendations I.230 and I.240. Corresponding traffic variables including the number of attempts per call demand are derived for the relevant attributes of each service.

4 User plane

In the user plane the attributes of some ISDN services give rise to additional traffic parameters beyond those used in telephony. Based on the user models of Recommendation E.711, Recommendation E.712 will be developed in the future to derive traffic models for each basic ISDN service using a common set of parameters that applies to all services.

[1] **significant points** are points in the network where traffic flows and grades of service should be evaluated.

5 Control plane

Each attempt from a terminal will result in signalling messages in the control plane. The number of messages and their length are highly dependent on:

- protocols (S.S. No. 7 and Q.931 digital access);
- call disposition (including user facilities);
- originating and terminating subscriber equipment configurations (e.g. overlap sending).

Control plane traffic models given in Recommendation E.713.

6 Further developments

The preceding sections of this Recommendation and the Recommendations of the E.710 Series concentrate on those services and facilities which will predominate in the first ISDNs to be implemented.

However, the ISDN concepts include the inherent possibility of developing completely new services and functions in response to changing user needs. This will result in new traffic situations that will need to be covered in the E.710 Series. Some of the situations which can be anticipated are:

- the impact of supplementary services, such as call waiting, on control plane traffic;
- new ISDN capabilities such as statistical switching and dynamic allocation;
- connectionless communications;
- the impact of user-to-user signalling;
- the impact of multipresentation type calls, and simultaneous use of different interactive and distribution services, involving multi-slot and multipoint connections, and broadcast mode.

Recommendation E.711

USER DEMAND

1 Introduction

1.1 Traffic offered to layers 1-3 of the ISDN can be modelled by distributions of arrival times and holding times (traffic variables). This Recommendation describes how these traffic variables are related to user demands at higher levels.

2 General structure

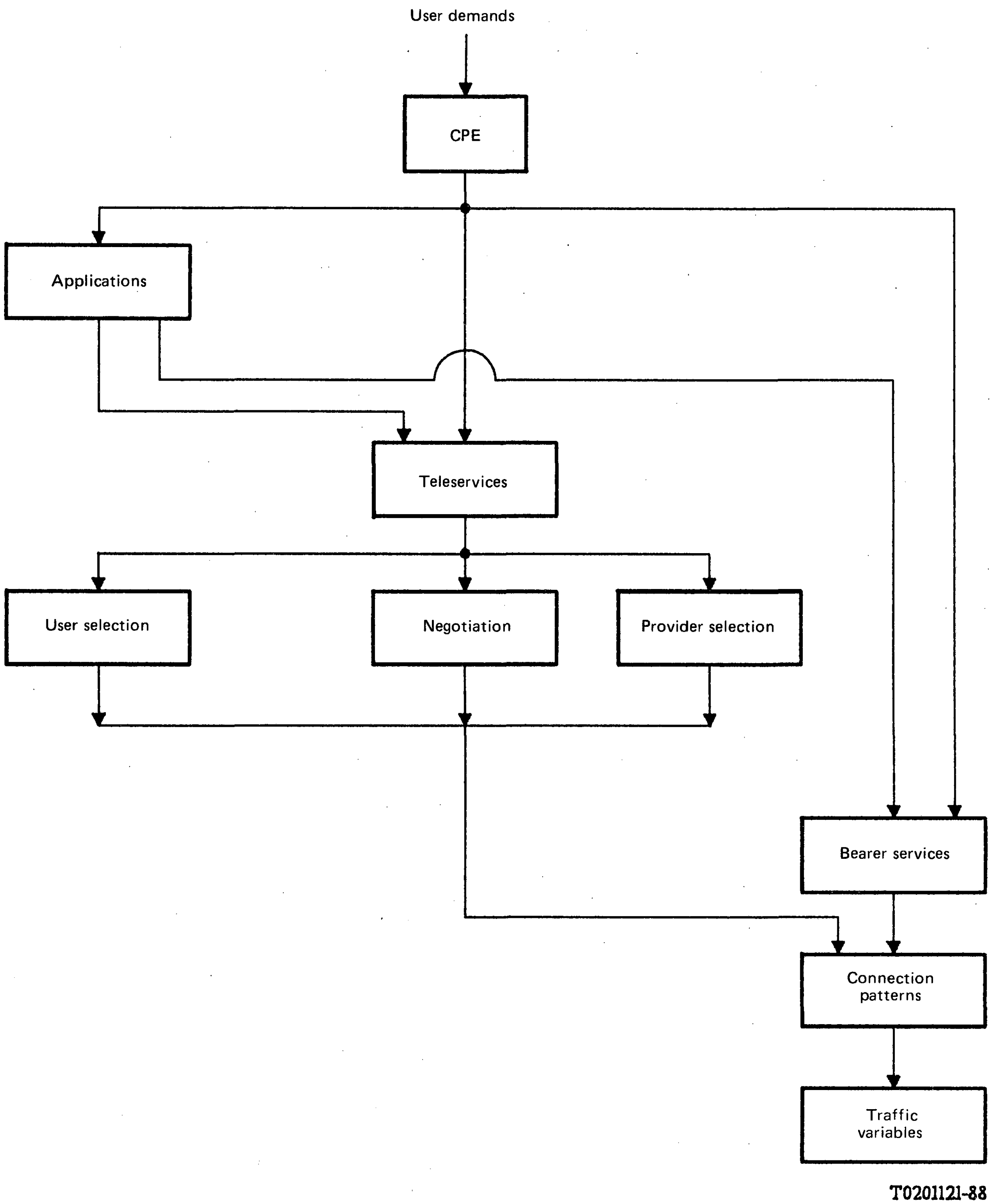
2.1 This section describes the general process by which the distributions of arrival times and holding times (traffic variables) which determine the offered traffic to layers 1-3 may be derived from user demands. The process is illustrated in Figure 1/E.711 and described in detail in Annex B.

2.2 Through the mediation of Customer Premises Equipment (CPE), user demands are translated into sequences of requests for applications, teleservices, and bearer services.

2.3 An **application in an ISDN** is a sequence of teleservice and bearer service requests, predefined in order to satisfy a global communications need.

2.4 A **call pattern** is a specific sequence of events and inter-event times generated by a call demand and modelled by traffic variables as described in § 3 of this Recommendation. Each teleservice class can be modelled by a mix of call patterns, each corresponding to a set of teleservice attributes.

2.5 A **connection pattern** is a specific set of information transfer and general attributes which are significant for traffic engineering. Information transfer and general attributes are described in Recommendation I.210. Each call pattern can be served by one or more connection patterns.



CPE Customer premises equipment

FIGURE 1/E.711

Relation between user demands and traffic variables

2.6 A teleservice has attributes that can be selected by the user, negotiated or selected by the service provider. The result of this selection procedure is a sequence of requests for connection patterns.

2.7 Annex A outlines the relations between user demands, applications, teleservices, bearer services and traffic significant attributes.

2.8 The mix of connection patterns determined by the process in turn determines the distributions of arrival times and holding times.

3 Traffic variables

3.1 Traffic variables are expressed as distributions of arrival times and holding times. For traditional circuit switched services, the shapes of some distributions are such that they can be represented by the mean values. A discussion of traffic variables in the ISDN context is given in the following sections.

3.2 *Call variables*

3.2.1 *Arrival process*

For traditional circuit switched services, the call attempt rate has, for practical purposes, been considered equivalent to the call demand rate. In the ISDN, on the contrary, this equivalence can no longer be assumed. Many teleservices will have attributes such that complex call attempt sequences are generated for each call demand. This will require the introduction of additional considerations such as:

- number of call attempts per call demand;
- number of negotiations per call demand;
- number of call demands requiring reservation.

The entire subject of call attempts sequences requires further study.

3.2.2 *Holding times*

For traditional circuit switched services, call holding time t_1 is the only variable of interest. For reservation services, additional variables are needed to characterize reservation time t_2, completion time t_3 and request time t_4. See Figure 2/E.711. (New holding times require further study.)

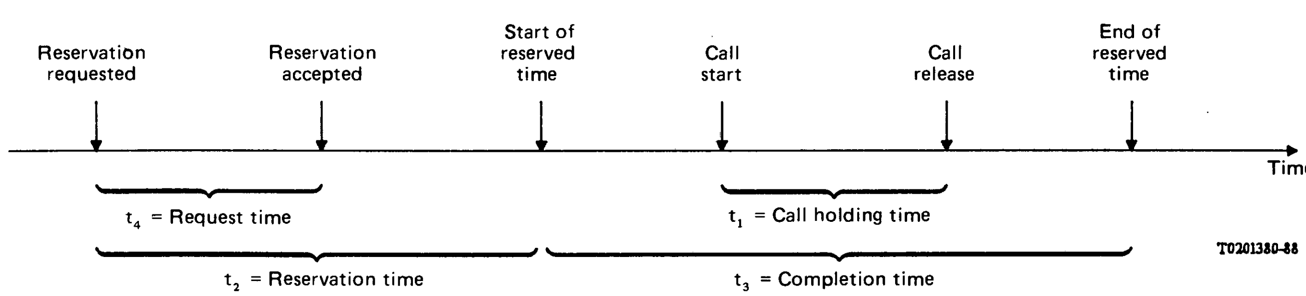


FIGURE 2/E.711

Holding times for reservation services

3.3 *Transaction variables*

Additional information beyond § 3.2 is needed for packet switching services.

For packet-switched services, the information content at the user level during a call may be produced in discrete transactions (intervals during which a user is continuously producing information). This subdivision is significant from a traffic point of view. See Figure 3/E.711.

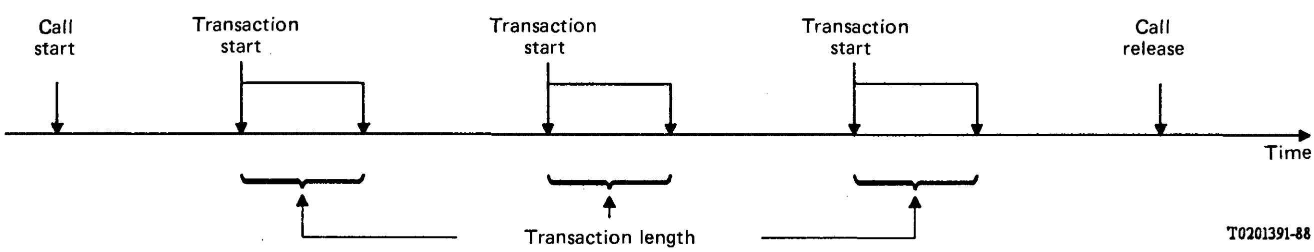
3.3.1 *Arrival process*

The arrival process for transactions within a call is for further study.

3.3.2 *Transaction length*

The transaction length as expressed in bits represents the workload offered by the transaction through the user/network interface. The distribution of transaction lengths is for further study.

Note – For transport purposes, the workload as related to single transactions within a specific call may undergo one or more segmentation stages. The entire subject of workload segmentation is for further study.



Note – Information transfer only occurs during the transactions.

FIGURE 3/E.711

Transactions of a packet-switched service

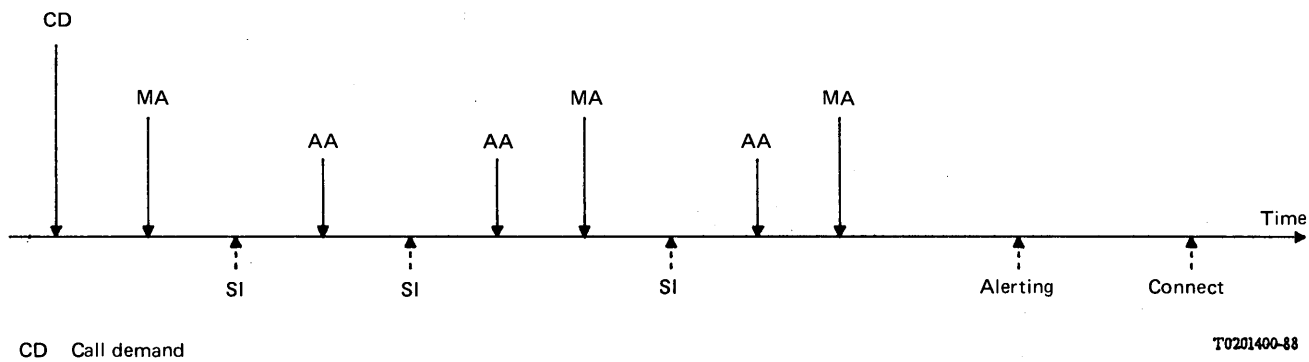
4 Examples

4.1 Traditional telephone service with lost calls cleared is usually characterized by mean arrival rate and mean holding times.

4.2 Telephone service in the ISDN, with a fast signalling system (Signalling System No. 7) and capabilities for automatic repetition, needs the introduction of a supplementary variable, namely the repetition rate, to evaluate the number of call attempts per call demand.

4.3 Personal computer communication using reservation services, associated with the supplementary services of automatic repetition and call waiting, is a teleservice giving rise to a complex call attempt sequence as illustrated in Figures 4/E.711 and 5/E.711.

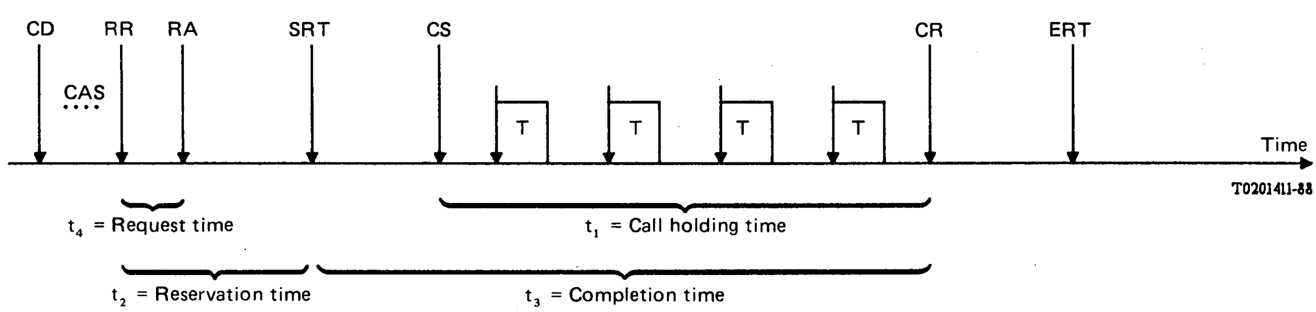
In relating this service to user demand, many additional variables are needed as discussed in § 3. The control and user plane traffics must take into account not only the mean values but also other parameters characterizing the distributions.



CD Call demand
MA Manual attempt
AA Automatic attempt
SI Status indication of network or B-subscriber reason for non-completion

FIGURE 4/E.711

Call attempt sequence



CD – Call demand, manifested by the first call attempt
CAS – Call attempt sequence
RR – Reservation requested
RA – Reservation accepted
SRT – Start of reserved time
CS – Call start
T – Transactions
CR – Call release
ERT – End of reserved time

FIGURE 5/E.711

General call pattern

ANNEX A

(to Recommendation E.711)

Relation between user demands and attributes

A.1 *Introduction*

This Annex provides concrete examples relating user demands (applications, teleservices and bearer services) to attributes which are important for traffic engineering purposes. Tables are provided for illustrative purposes but it must be noted that these are based on a selective summarization of key attributes related to the I.200 Series of Recommendations. Thus they should be only interpreted as illustrations of the process.

A.2 *User demand attributes*

User demands are described by the following attributes:

- user service selections;
- access channels and rates (see Figure A-1/E.711);
- layer 7 to 1 protocols.

A.3 *Application characteristics*

Applications are described by the following characteristics:

- teleservices supporting the application;
- bearer services supporting the application;
- bearer capabilities supporting teleservices and bearer services.

Table A-1/E.711 gives the teleservices recommended in Recommendation I.240 together with the attributes which are important from a traffic engineering point of view. These comprise:

- information transfer mode;
- information transfer rate;
- information transfer capability;
- establishment of communication;
- symmetry;
- communication configuration.

As other teleservices are introduced into ISDN (e.g. electronic shopping) in the future, the traffic engineering attributes may expand (e.g. information handling processes).

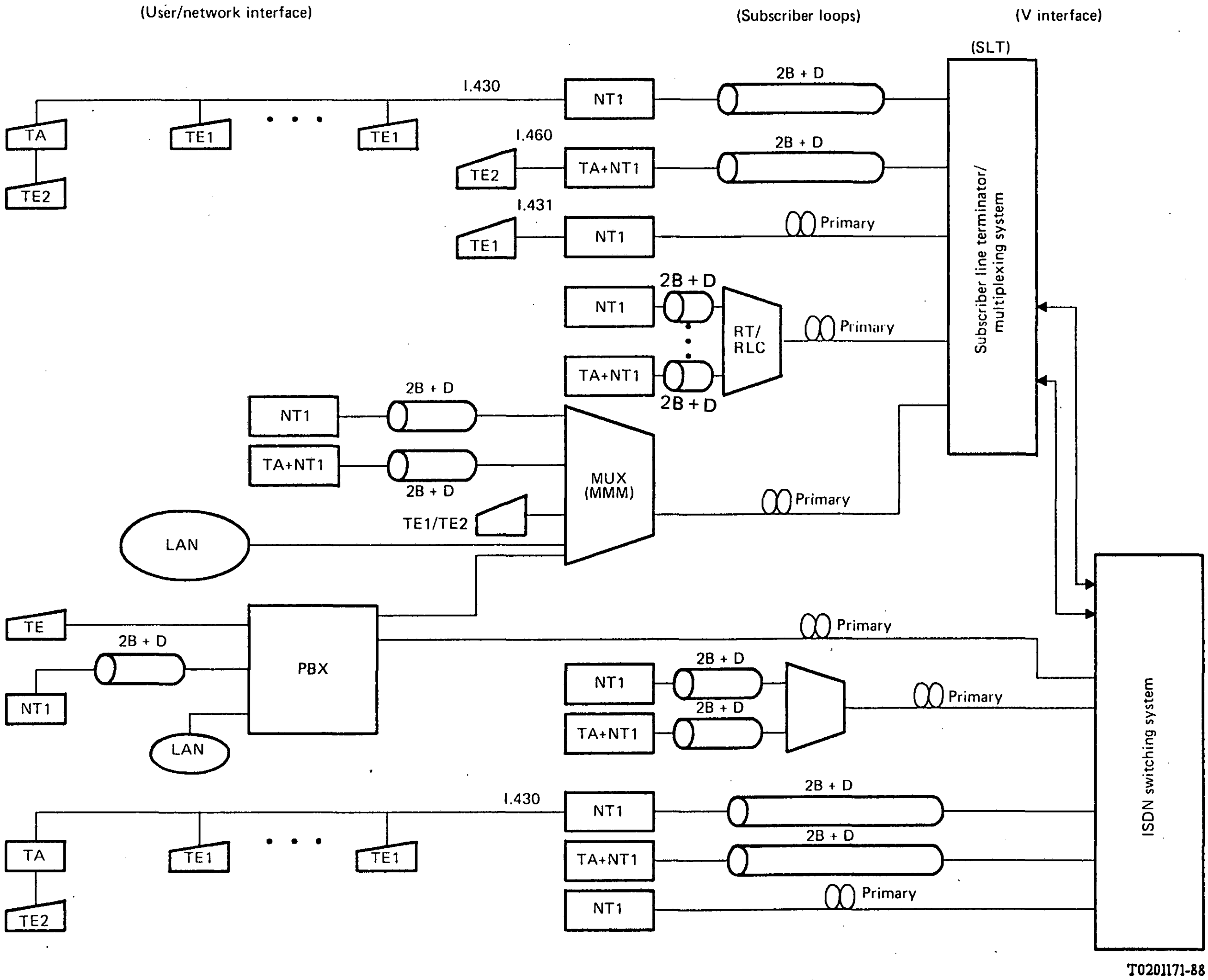
Table A-2/E.711 lists representative bearer services from which those required to support an application may be chosen.

A.4 *Teleservices*

According to Recommendation I.210, a teleservice is the result of one of the following combinations:

- one basic teleservice;
- one basic teleservice plus one or more supplementary services.

Furthermore, a teleservice is implemented using bearer capabilties.



Note – The equipments and channels shown in this diagram are defined in Recommendations I.411 and I.412.

FIGURE A-1/E.711

Example of ISDN user accesses

TABLE A-1/E.711

Examples of teleservices and their attributes

	Information transfer mode										
	Circuit mode [a]										
	Information transfer rate (kbit/s)	Information transfer capability		Establishment of communication		Symmetry			Communication configuration		
		Unrestriction digital	Speech	Demand	Reserved	Unidirectional	Bidirectional		Point-to-point	Multipoint	Broadcast
							Symmetric	Asymmetric			
Telephony	64 (max.)		×	×			×		×	×	
Teletex	64 (max.)	× [b]		×		×			×	×	
Telefax 4	64 (max.)	×		×		×			×		
Mixed mode	64 (max.)	×		×		×			×		
Videotex	64 (max.)	×		×				×	×		

[a] Packet mode is for further study.
[b] Presently this service is provided by 3.1 kHz audio.

TABLE A-2/E.711

Bearer services

Circuit mode (Recommendation I.231)
64 kbit/s, unrestricted, 8 kHz structured
64 kbit/s, 8 kHz structured, usable for speech information transfer
64 kbit/s, 8 kHz structured, usable for 3.1 kHz audio information transfer
Alternate speech / 64 kbit/s unrestricted, 8 kHz structured
2 × 64 kbit/s unrestricted, 8 kHz structured
384 kbit/s unrestricted, 8 kHz structured
1536 kbit/s unrestricted, 8 kHz structured
1920 kbit/s unrestricted, 8 kHz structured
Packet mode (Recommendation I.232)
Virtual call and permanent virtual circuit
Connectionless
User signalling

ANNEX B

(to Recommendation E.711)

Traffic characterization

B.1 *Introduction*

B.1.1 This Annex describes a methodology for relating user demands to the traffic offered to layers 1-3 of the ISDN. The basic approach is to relate the mix of user demands to *call patterns* and *connection patterns.* These latter concepts are defined in §§ 2.4 and 2.5 and repeated below; between them, they contain all of the information needed to derive the distributions of arrival times and holding times.

B.1.2 Call patterns and connection patterns are the means by which the effects of user demands are described as they affect layers 1-3 of the ISDN network.

A call pattern is a specific sequence of events and inter-event times generated by a call demand and modelled by traffic variables as described in § 3 of this Recommendation.

A connection pattern is a specific set of information transfer and general attributes which are significant for traffic engineering. Information transfer and general attributes are described in Recommendation I.210.

Call patterns describe what happens at the user-network interfaces. Connection patterns describe what types of resources are used. Each call pattern can be served by one or more connection patterns.

B.2 *User/customer premises equipment characterization*

B.2.1 *User classes*

The population of users can be divided into user classes characterized by the user selections of applications, teleservices and bearer services, and their rates of occurrence. Each class is associated with a penetration in the population.

B.2.2 *Customer premises equipment (CPE) classes*

The actual application, teleservice and bearer service requests presented to the network as a result of user selections are determined by the user's CPE type. Each user class can be subdivided into CPE classes characterized by the penetration of CPE types in that user class.

B.3 *Application characterization*

For further study.

B.4 *Teleservice characterization*

B.4.1 *Teleservice classes*

The population of teleservices requested by user/CPE combinations may be subdivided into classes defined by the values of attributes significant for traffic engineering.

Teleservices, as defined in Recommendation I.240, are teleservice classes from the traffic point of view.

Of the attributes defined in Recommendation I.210, the following are significant for traffic engineering:

- information transfer mode;
- information transfer rate;
- information transfer capability;
- establishment of communication;
- symmetry;
- communication configuration.

Each combination of attribute values defines a single teleservice class.

B.4.2 *Teleservices*

Within each teleservice class, individual teleservices are defined by values of general attributes which are still under study in Study Group XVIII. Of particular significance for traffic engineering is the attribute "Supplementary services".

B.4.3 *Demands for teleservice classes*

Each user class/CPE class combination is characterized by rates of demand for teleservice classes. This characterization may be represented as shown in Table B-1/E.711. The contents of Table B-1/E.711 must be estimated by statistical studies.

B.4.4 *Teleservice demands*

Combining the concepts of §§ B.4.1 and B.4.2, the total request rate for each teleservice class can be subdivided as shown in Table B-2/E.711.

The entries of Table B-2/E.711 must be estimated by statistical means.

B.4.5 *Call patterns*

For each individual teleservice there is one and only one corresponding call pattern. However, the same call pattern may be representative of several teleservices.

Multiplying the total rates in Table B-1/E.711 by the proportions shown in Table B-2/E.711, rates for each call pattern are obtained as shown in Table B-3/E.711.

TABLE B-1/E.711

Demands for teleservice classes

User class	CPE class (Note 1)	Teleservice class							
		1	2	3	.	.	.	.	.
1	X								
	Y								
	Z								
2	t								
	Z								
. . .	. . .								
Totals									

Note 1 – A given user class will not necessarily use all CPE classes.

Note 2 – Table entries are the rates at which the user/CPE combinations originate requests for each teleservice class.

TABLE B-2/E.711

Demand for individual teleservices

Teleservice class	General attribute combinations							
	1	2	3	.	.	.	.	.
1								
2								
. . .								
Total								

Note – Table entries are the proportions of total requests for each teleservice class for each general attribute combination (defining an individual teleservice). Each row adds to unity.

TABLE B-3/E.711

Call pattern demands

Teleservice class	Call pattern							
	1	2	3	.	.	.	.	.
1								
2								
. . .								
Total								

Note – Table elements are the rates at which each teleservice class creates a demand for each call pattern.

B.5 *Connection pattern characterization*

Each call pattern can be served by one or more connection patterns. A specific connection pattern corresponds to each set having as elements one value for each applicable bearer service attribute.

The breakdown in Table B-4/E.711 of the call patterns on the connection patterns is needed.

TABLE B-4/E.711

Breakdown of the call pattern on the connection patterns

Call pattern	Connection pattern											
	Packet mode								Circuit mode (B-channel)			
	On D-channel				On B-channel							
	1	...	i	...	1	...	j	...	1	...	k	...
CP1												
...												
CPn												
Total												

Note 1 – Table entries are the proportions of the *n*th call pattern served by the different connection patterns.

Note 2 – The total on the columns gives the total rate on each connection pattern.

Note 3 – The sums on the rows may be useful for designing priority classes.

Using Tables B-3/E.711 and B-4/E.711, Table B-5/E.711 can be obtained.

TABLE B-5/E.711

Rate of the call demands requiring a specific connection pattern

Connection patterns	Rate
XP1	
XP2	
. . .	
XPn	
Total	

Bibliography

BONATTI (M.), GIACOBBO SCAVO (G.), ROVERI (A.), VERRI (L.): Terminal exchange access system for NB-ISDN: Key issues for a traffic reference model. *Proc. 12th ITC*, paper 4.1A.3, Turin, 1988.

Recommendation E.713

CONTROL PLANE TRAFFIC MODELS

1 Control plane traffic

For the purposes of teletraffic engineering, the control plane traffic load is assumed to be generated by call attempts on the network. These call attempts are part of the call pattern described in Recommendation E.711.

This Recommendation considers traffic loads at the lower three layers of the CCITT 7-layer reference models (Recommendation I.310 and I.320) described for ISDN in Recommendation Q.931 and in Signalling Sytem No. 7.

The control plane traffic of an ISDN network includes all the control signals sent through the ISDN network. The types of control signals are:

1) signals for user call attempts
   a) to set up the connection paths in the user plane (reservation of time slots for circuit switched connections or control for the virtual calls of packet-switched connections),
   b) to release the connection paths in the user plane,
   c) if required, to order additional communication facilities or change of service by the users during the time of user information transfer,
   d) possibly to send charging information during the time of user information transfer.

2) User-to-user information messages[1)]

Because control plane traffic due to user-to-user messages is left for further study, this Recommendation will consider only signals for user call attempts.

The control plane traffic uses two types of channels in the network:

a) the 16 kbit/s or 64 kbit/s D-channels in the user access, and

b) the 64 kbit/s Signalling System No. 7 channels connecting two different signalling points.

2 Signalling traffic

The end-to-end ISDN signalling traffic depends on the call pattern arrival process defined in Recommendation E.711 and on the signalling protocol.

The basis for the estimation of the signalling traffic is the information given in the Recommendations of the I- and Q-Series dealing with the number and structure of the signals in the D- and Signalling System No. 7 channels for any type of attempt. The total signalling traffic is composed of these signals. The number of signals may be different for each different type of attempt.

3 Estimation of the signalling traffic for a single call attempt

In Figure 1/E.713 the network components supporting the control plane of the ISDN reference connection of Figure 1/E.701 are considered. In each section, a significant point is defined:

DA (D-channel, A user side):	S/T interface at an A user side
DB (D-channel, B user side):	S/T interface at a B user side
CA (S.S. No. 7 channels, A user side):	outgoing side of the local exchange LE(A).
CB (S.S. No. 7 channels, B user side):	incoming side of the local exchange LE(B).

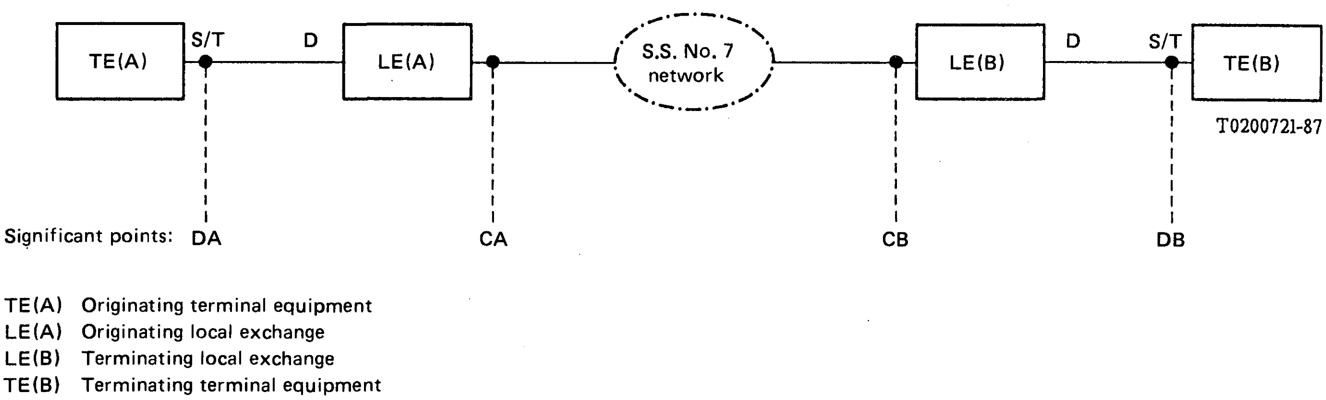


FIGURE 1/E.713

Significant points in the control plane

[1)] The analysis of user-to-user messages in the control plane is left for further study.

The signal flow which is necessary to perform the control functions of a particular call attempt may be represented by a signal flow diagram. It contains all the signals passing the significant points in the control plane for the considered attempt. Figure 2/E.713 shows the basic scheme of this signal flow diagram. The arrows represent the layer 2 signals in the three connection phases: call establishment, user information transfer, and call release.

An example of a signal flow diagram for a successful circuit switched call attempt is given in Annex A.

The signal flow diagram is the basis for the estimation of the amount of signalling traffic caused by the considered attempt using the reference connection. The signalling traffic of a single attempt in a given section of the control plane associated with a significant point can be described by two sets of parameters:

1) the total number of signals passing the significant point in the three call-connection phases in the A-to-B direction and in the B-to-A direction, as in Figure 2/E.713;

2) the length of each signal type passing the significant point in the A-to-B and the B-to-A direction.

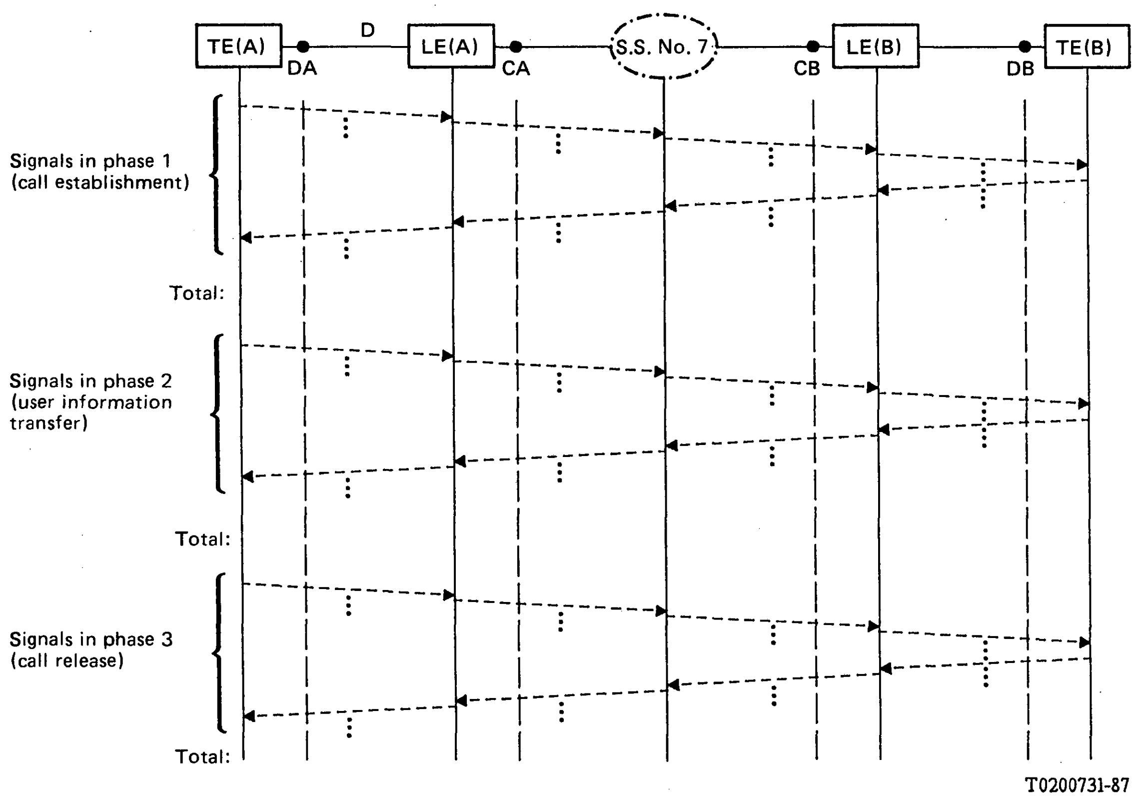


FIGURE 2/E.713

Basic scheme of a signal flow diagram of a call attempt for the reference connection

4 Estimation of the total signalling traffic

The total number of signals in the control plane over a reference period[2] is determined by summing the number of signals caused by call attempts handled in the associated user plane during the reference period. Therefore the estimation of the number of signals is based on the estimation of the amount and types of attempts in the user plane.

In order to estimate the amount of signals, it is necessary to accept a traffic model for the traffic in the user plane assuming the total number of attempts over the reference period and the breakdown of these attempts into the different types of attempts, such as successful call attempts, unsuccessful call attempts and calls to busy tone.

The total traffic load of a section caused by the signals is expressed by the total amount of bits crossing its significant point.

In order to estimate the amount of this traffic load it is necessary to multiply the length of each particular type of signals by the number of signals of each type occuring during the reference period and summing over all types of signals occuring during the reference period.

Since the number and length of the signals do not vary widely for most types of attempts, initially this traffic model will be adequate by taking into consideration only the most frequently experienced types of attempts.

The characteristic of the signalling traffic in a particular section of the control plane will depend on such factors as:

a) the total traffic load caused by layer 2 and 3 signals for the attempts.

b) the distribution of call attempts and release arrivals.

The impact on teletraffic engineering caused by a full characterization of the arrival process is left for further study.

Using Figure 2/E.713, the signalling traffic load at a significant point can be estimated.

If, over the reference period:

- i is the call phase,
- j is the signal type,
- $n_{ij}(u)$ is the average number of signals of type j in call phase i in the A-to-B direction,
- $n_{ij}(d)$ is the average number of signals of type j in call phase i in the B-to-A direction,
- l_j is the length of signal of type j,
- T is the total number of signals types,
- $L(u)$ is the total load in the A-to-B direction,
- $L(d)$ is the total load in the B-to-A direction,

then:

$$L(u) = \sum_{i=1}^{3} \sum_{j=1}^{T} l_j \times n_{ij}(u)$$

$$L(d) = \sum_{i=1}^{3} \sum_{j=1}^{T} l_j \times n_{ij}(d)$$

Each $n_{ij}(u)$ and $n_{ij}(d)$ must be estimated from the number of call attempts and the call attributes in the user plane over the reference period. An example of this procedure is given in Annex A.

[2] The proper reference period to use for dimensioning is for further study.

ANNEX A

(to Recommendation E.713)

Example of procedure for estimating the total signalling traffic in a D-channel

A.1 *Signalling traffic for one call attempt*

A call attempt of the following type is considered:

- effective call attempt,
- circuit switched connection,
- en-bloc sending of dialled information,
- call to an appointed terminal,
- no additional control signals during the information transfer phase,
- installation of data link in the D-channels required for establishment and release of the connection,
- manual answering terminal.

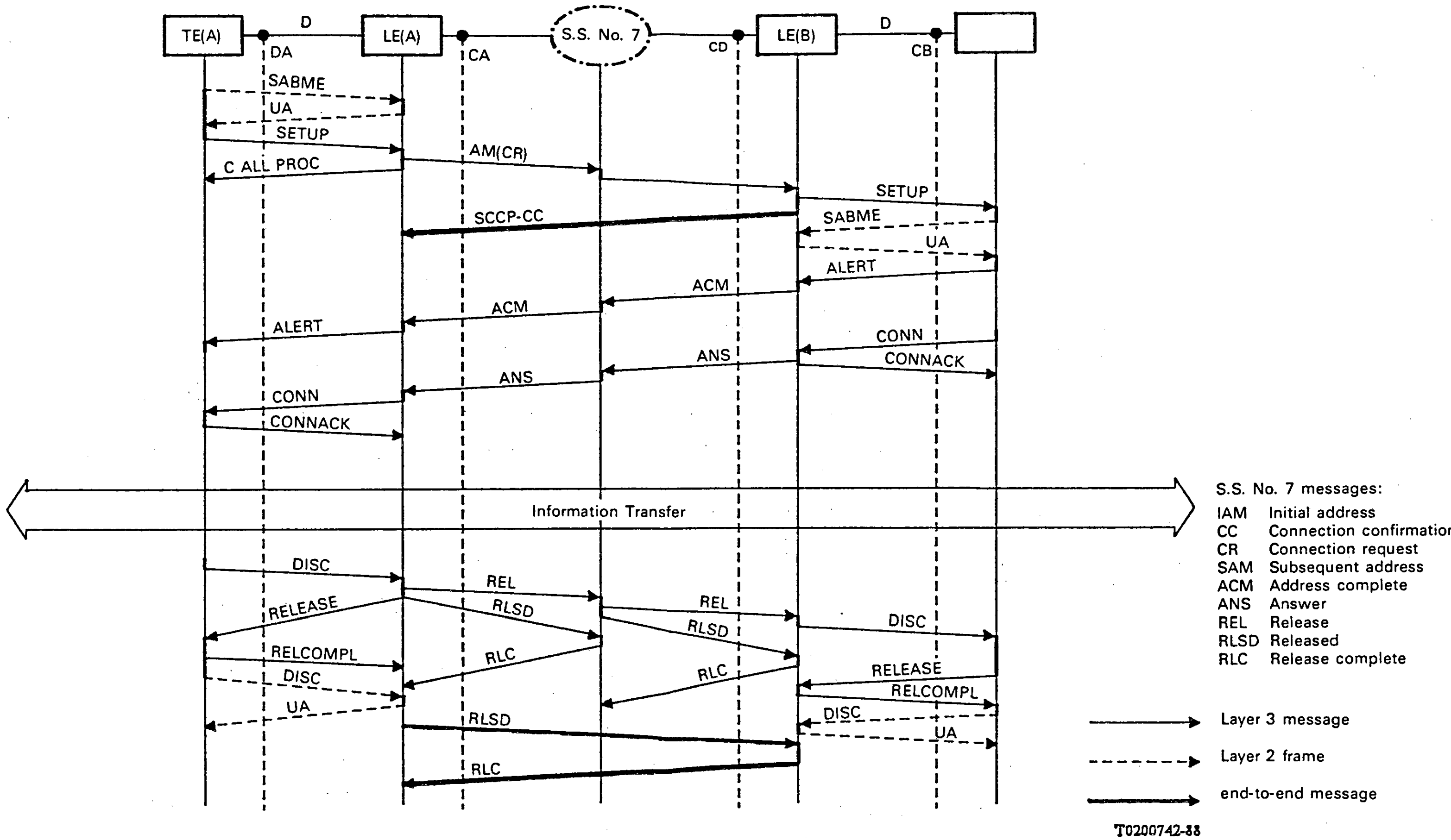
The signal flow diagram for this type of call attempt is given in Figure A-1/E.713 and Figure A-2/E.713. Three kinds of signals are indicated in Figure A-1/E.713:

- layer 3 signals,
- layer 2 signals for the activation and deactivation of the data links,
- end-to-end signals via the S.S. No. 7 network.

Figure A-2/E.713 presents the breakdown of the D-channel signals into layer 2 for the case of multiple terminals on the terminating side. The breakdown of the S.S. No. 7 messages and the total length of signal in the considered call attempt is for further study.

A.2 *Signalling traffic for additional types of call attempts*

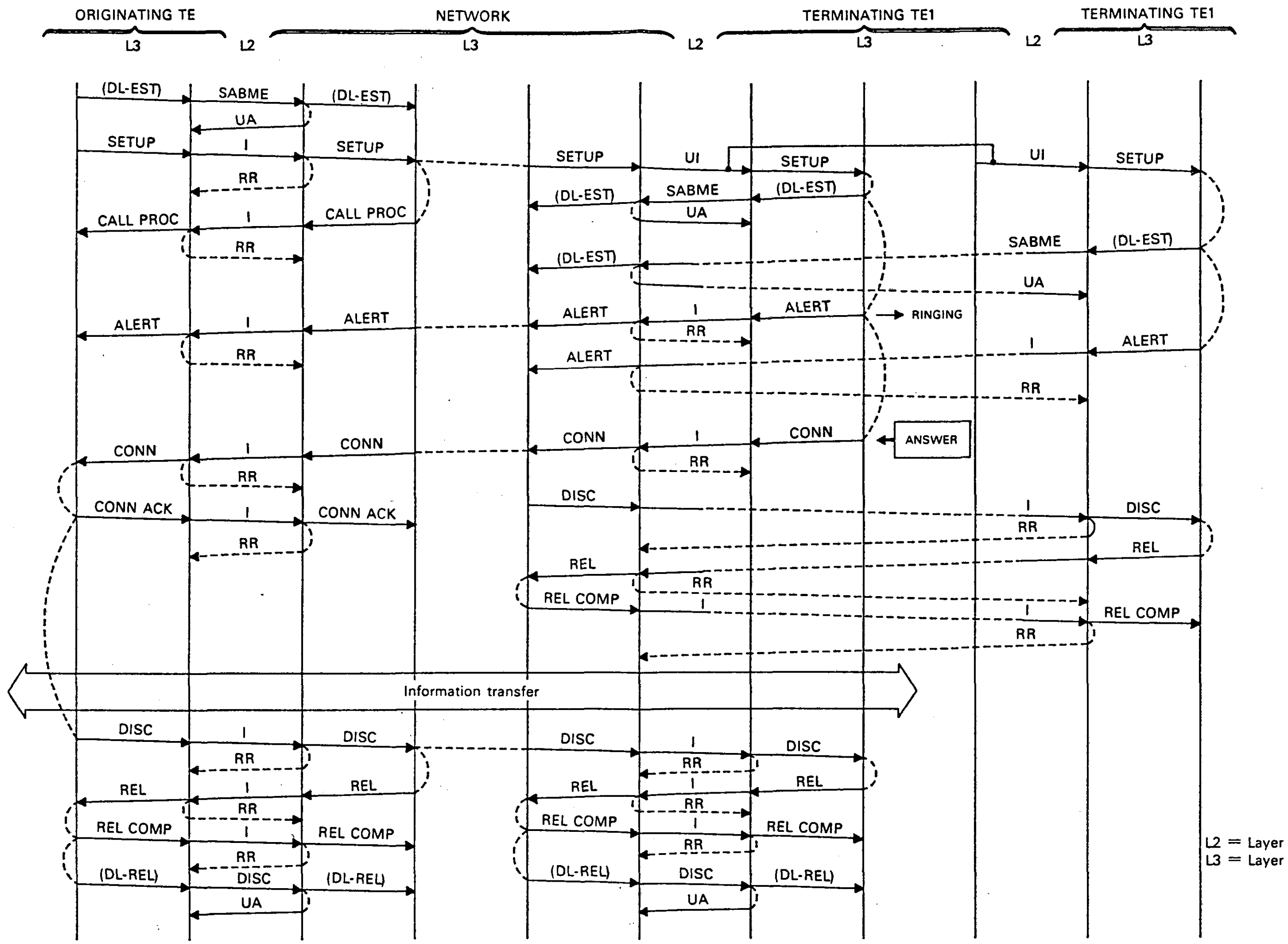
For further study.



Note — See Recommendations I.441 and I.451 for the abbreviations of layer 2 frames and layer 3 messages. In this Recommendation, these are referred to as signals.

FIGURE A-1/E.713

Signal flow diagram for a circuit switched connection with en-bloc sending of dialled information (to appointed terminal)



Note – See Recommendations I.441 and I.451 for the abbreviations of layer 2 frames and layer 3 messages. In this Recommendation, these are referred to as signals.

FIGURE A-2/E.713

Signal flow diagram
Example of Figure A-1/E.713 with layer 2 signals on the D-channels and with multiple terminals on the terminating side for an effective call attempt

ISDN GRADE OF SERVICE CONCEPT

1 Introduction

This Recommendation outlines the general considerations for the ISDN Grade of Service (GOS) concept and provides guidelines for selecting GOS parameters. In this Series of Recommendations, the term GOS always refers to traffic Grade of Service parameters as defined in Recommendation E.600.

ISDN GOS parameters are given in subsequent Recommendations in the E.720 Series.

2 GOS concept

GOS uses a number of traffic engineering parameters to provide a measure of adequacy of plant under specified conditions; these GOS parameters may be expressed as probability of blocking, probability of delay, etc. Blocking and delay are caused by the fact that the traffic handling capacity of a network/network component is finite and that the demand traffic is stochastic by nature.

The users of telecommunication services can experience the effects of GOS parameters depending on their perception of events such as:

1) failure of a call demand or excessive delay to satisfy a call demand;
2) failure of call attempts or excessive delay to satisfy call attempts;
3) failure of automatic re-attempts or excessive delay to satisfy automatic re-attempts.

Events of the first class are always perceived by the user. Events of the other two classes may be perceived by the user depending on the capability of the terminal equipment to transmit signalling information to the calling user.

In all three classes the ability to distinguish GOS depends on having distinct indications of called user conditions and network conditions.

GOS may be distinguished as the user GOS, network GOS and network component GOS as shown in Figure 1/E.720. User GOS relates to user call demands. Network GOS relates to any call attempts including both user generated call attempts and terminal generated automatic reattempts. Network component GOS relates to bids for the utilization of a specific network component including both bids generated by call attempts and bids generated by call attempts and bids generated by internal retrials in the network. Parameters related to user GOS and network component GOS are for further study.

Recommendation E.721 defines network GOS parameters based on any call attempt. Subsequent Recommendations in the E.720 Series will define other GOS parameters. Recommendations in the E.740 series will define traffic measurement and performance monitoring requirements.

User GOS performance effects and other traffic-independent, user-perceived effects such as availability and service integrity contribute to Quality of Service (QOS). Network GOS parameters and their values provide information on the traffic aspects of the QOS.

3 Principles to select ISDN GOS parameters

3.1 *ISDN traffic characteristics*

ISDN has many characteristics different from the existing dedicated networks such as Public Switched Telephone Network (PSTN), Circuit Switched Public Data Network (CSPDN), Packet Switched Public Data Network (PSPDN), etc. The following characteristics are taken into account when defining GOS parameters for ISDN:

- ISDN provides integrated access to a wide variety of telecommunication services through a small set of standardized user-network interfaces.
- Services have heterogeneous traffic demand profiles and diverse performance requirements.

– The traffic streams generated by user demands for bearer services and teleservices utilize layer 1, 2 and 3 resources.

– The configuration and implementation of a user's terminal and its man-machine interface may vary from one service to another service and one user to another user.

– Out-of-band signalling and call control capability, based on D-channel and Signalling System No. 7 (S.S. No. 7) are provided.

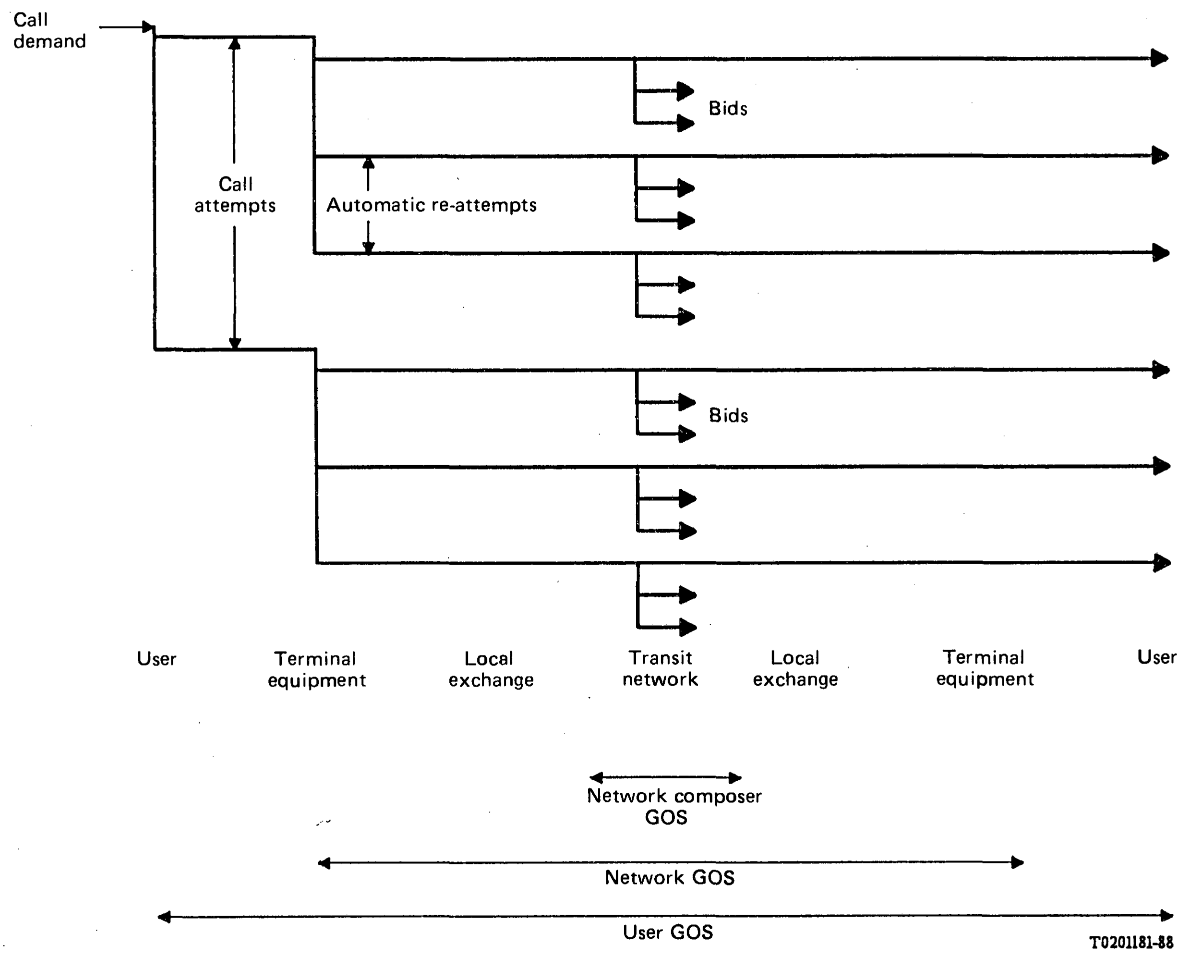


FIGURE 1/E.720

GOS concept

3.2 *Parameter selection principles*

The GOS parameters defined in the E.720 series of Recommendations applies to the first phase of ISDN. The definitions of these parameters may be expanded or additional GOS parameters defined to accomodate future evolution of ISDN architecture and services. Considering the above characteristics in ISDN, the following principles are recommended for selecting GOS parameters in ISDN:

i) A minimum common set of GOS parameters is defined for attempts on layers 1, 2 and 3 – such attempts belonging to different services may or may not share the same out-of-band call set-up and release procedures. (See Note.)

ii) The GOS parameters are defined and specified in such a way that the GOS can be derived at well-defined reference points (traffic significant points).

iii) The GOS parameters should be specified with reference to traffic load conditions in the sense of Recommendation E.500.

iv) Blocking GOS parameters may in the future need to take account of repetitions due to network status, but are presently based on lost-call-cleared assumptions.

v) GOS parameters related to the user plane information transfer phase are for further study.

Note – To assess the performance directly perceived by a user, other additional parameters which are specific to the user's terminal equipment may also be necessary.

3.3 *GOS standard setting principles*

GOS standard setting principles will take into account the current standards for voice and data services so that when the user migrates from a dedicated network to the ISDN, the user does not encounter a marked undesirable contrast.

Recommendation E.721

NETWORK GRADE OF SERVICE PARAMETERS IN ISDN

1 Introduction

This Recommendation proposes network Grade of Service (GOS) parameters for circuit-switched and packet-switched services in ISDN, based on the ISDN Grade of Service concept and guidelines for selecting GOS parameters provided in Recommendation E.720. These parameters are defined assuming that the network and the network components are fully operational.

2 Circuit-switched services

In the current ISDN specifications, call establishment and release for all circuit-switched services provided via B-channels (voice, data, image) will use the out-of-band call control procedures defined by Recommendation Q.931 and S.S. No. 7 (ISUP) signalling protocols. Thus, for the traffic GOS parameters that relate to call establishment and release phases, a single set of parameters can be used for all circuit-switched services provided by the ISDN.

The following four traffic GOS parameters are recommended for circuit-switched calls in ISDN:

1) pre-selection delay (overlap sending),
2) post-selection delay,
3) call release delay, and
4) probability of end-to-end blocking.

The definitions of these traffic GOS parameters are given below. The delay GOS parameters are based on the message flows in Recommendation Q.931 and S.S. No. 7 (ISUP) protocols as indicated, for example, in Figure A-1/E.713.

2.1 *Pre-selection delay (overlap sending)*

pre-selection delay (overlap sending) is defined as the time interval from the instant the SABME message is passed by the calling terminal to the access signalling system until the SETUP ACK message is received by the calling terminal.

2.2 *Post-selection delay*

post-selection delay is defined as the time interval from the instant the INFO message containing the last selection digit (in the case of overlap sending or the SETUP message in the case of en-bloc sending) is passed by the calling terminal to the access signalling system until the first message indicating call disposition is received by the calling terminal.

Note – In the ISDN the called user can choose to delay the sending of the ALERTING signal to the calling user. This definition does not include such user-induced delays.

2.3 *Call release delay*

call release delay is defined as the time interval from the instant the DISCONNECT message is passed by the user terminal which terminated the call to the access signalling system, until the RELEASE message is received by the same terminal (indicating that the terminals can initiate/receive a new call).

2.4 *Probability of end-to-end blocking*

The **probability of end-to-end blocking** is the probability that a call attempt will be unsuccessful due to a lack of network resources.

Note – Resources in the access network are not part of this definition.

3 Packet-switched services

The ISDN user has a choice of two types of packet-switched data services. The B-channel provides 64 kbit/s packet access, while the D-channel can also be used to provide packet data access at 16 kbit/s (64 kbit/s in the case of primary access). For packet-switched services, the current call control procedures are based on X.25 (inband) protocols, except during the initial B- or D-channel set-up between the DTE (Data Terminating Equipment) and the PH (Packet Handler). Thus a certain number of messages will be exchanged between the TE (Terminal Equipment) and the LE (Local Exchange) over the D-channel during the initial establishment phase of a packet-switched call. The messages will have to compete with other signalling (SAPI 0) and data (SAPI 16) traffic on the D-channel.

Thus, the traffic GOS parameters for the current specification of ISDN packet-switched services will have to be based on the Q.931 as well as on the X.25 call control procedures.

The selection and definition of traffic GOS parameters for packet-switched services in ISDN are for further study.

4 Target values for GOS parameters

The target values will be specified at the normal and high loads in the same sense as Recommendation E.500. The delay target values will be specified by the mean and percentile levels for both normal and high loads.

The actual target values are for further study.

PART III

Recommendations E.800 to E.880

QUALITY OF SERVICE; CONCEPTS, MODELS, OBJECTIVES, DEPENDABILITY PLANNING

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

SECTION 1

TERMS AND DEFINITIONS RELATED TO THE QUALITY OF TELECOMMUNICATION SERVICES

Recommendation E.800 [1)]

QUALITY OF SERVICE AND DEPENDABILITY VOCABULARY

CONTENTS

1 Introduction [2)]

A consistent set of terms and definitions is necessary for the development of Recommendations in the important areas of *quality of service* and *network performance* by the numerous Study Groups responsible for the Recommendations. Terminology standardization is also necessary to align the work of the various groups and to avoid confusing the users of Recommendations by the introduction of conflicting terms and definitions. Therefore, this Recommendation sets forth a simple set of terms and definitions relating to the concept of the quality of telecommunications services and *network performance.* These terms and definitions apply to all telecommunications services and all network arrangements used to provide the services.

---

1) Formerly part of Recommendation G.106, *Red Book*, Fascicle III.1

2) Terms printed in italics in the text may be found with their related definitions in Supplement No. 6 or in Recommendation E.600.

The diagram in Figure 1/E.800 is intended to provide an overview of the factors which contribute collectively to the overall *quality of service* as perceived by the *user* of a telecommunication service. The terms in the diagram can be thought of as generally applying either to the quality of service levels actually achieved in practice, to objectives which represent *quality of service* goals to be achieved, or to requirements which reflect design specifications.

The diagram in Figure 1/E.800 is also structured to show that one quality of service factor can depend on a number of others. It is important to note – although it is not explicitly stated in each of the definitions to follow – that the value of a characteristic measure of a particular factor may depend directly on corresponding values of other factors which contribute to it. This necessitates, whenever the value of a measure is given, that all of the conditions having an impact on that value be clearly stated.

An essential aspect of the global evaluation of a service is the opinion of the users of the service. The result of this evaluation expresses the users' degrees of satisfaction. This Recommendation establishes:

1) a general framework for the *quality of service* concept
2) the relationship between *quality of service* and *network performance*
3) a set of measures for these performances.

It is obvious that a service can be used only if it is provided, and it is desirable that the provider have a detailed knowledge about the quality of the offered service. From the provider's viewpoint, *network performance* is a concept by which network characteristics can be defined, measured and controlled to achieve a satisfactory level of service quality. The interests and the viewpoints of users and providers are different, and usually require a compromise between quality and economics.

In the utilization of a *service* the *user* identifies two «bodies»:

1) the "Organization(s)", i.e., the telecommunication Administration, operating company, etc. providing the means and facilities for the access to and the utilization of the *service*;
2) the «network», i.e., the necessary means (terminals[3]), lines, switches, etc.) actually used.

The contribution of the Organization to the *quality of service* is characterized by one performance concept, *service support performance*, as shown in Figure 1/E.800.

The contribution of the network to the *quality of service* is characterized by three performance concepts, which are:

– *service operability performance*, i.e., the ease by which the *service* can be used, including the characteristics of terminal equipment, the intelligibility of tones and messages, etc.;
– *serveability performance*, the ability of a *service* to be obtained – within specified tolerances and other given conditions – when requested by the *user* and continue to be provided for the requested duration. Thus, *serveability performance* describes the response of the network during the establishment, retention and *release* of a service connection;
– *service integrity*, the degree to which a *service* is provided without excessive impairments, once obtained. Thus, *service integrity* is primarily concerned with the level of reproduction of the transmitted signal at the receiving end.

The *serveability performance* is further subdivided into two terms:

– *service accessibility performance*, the ability of a *service* to be obtained – within specified tolerances and other given conditions – when requested by the *user*, further subdivided into (1) *network accessibility*, which is the ability of the *user* to obtain access to the network for a service request, and (2) *connection accessibility*, which is the ability of the network to provide the *user* with a satisfactory connection to the intended *destination*;
– *service retainability performance*, which is the ability of the *service*, once obtained, to continue to be provided under given conditions for a requested period of time. That is, *service retainability performance* covers the proper retention of *connections* and the *release* (disengagement) when requested by the *user*.

Serveability performance is divided into *trafficability performance*, *dependability* and *propagation performance* as shown in Figure 1/E.800. The *trafficability performance* is described in purely teletraffic engineering terms (see Recommendation E.600). The *measures* are expressed in terms of losses and delay times. *Dependability* is the combined aspects of availability, reliability, maintainability and maintenance support performances and relates to the ability of an *item* to be in a state to perform a *required function* (see Supplement No. 6). *Propagation performance* refers to the ability of the transmitting medium to transmit the signal within intended tolerances.

[3] In some countries' terminals are not part of the network and are or may be customer-provided

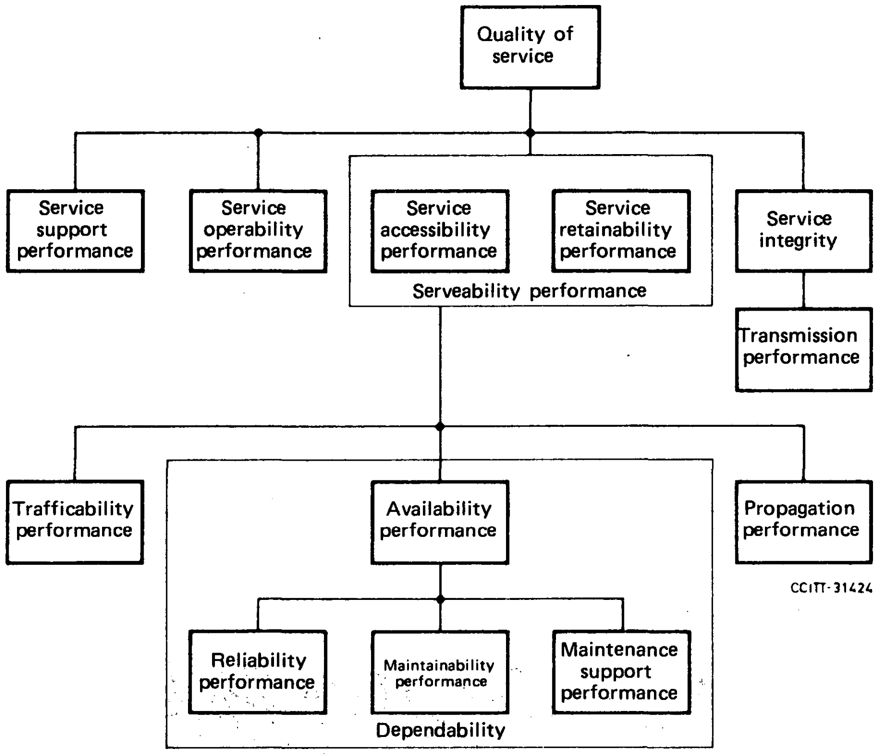


FIGURE 1/E.800

Performance concepts

Measures for all of the above performances may be related to an instant of time (instantaneous, etc.) or expressed as a mean value over a time interval. These and other recommended qualifiers (measure modifiers) are found in Supplement No. 6.

Supplement No. 6 further provides recommended statistical terms and definitions for use in the application of measures related to all performances.

While dependability is used only for a general description in non-quantitative terms, the actual quantification is done under the heading of availability performance, reliability performance, maintainability performance and maintenance support performance.

The most important of these dependability-related measures are found in Supplement No. 6, Part I. The properties expressed by these measures impact the measures related to quality of service and network performance and are thus implicitly characterizations of these performances.

Measures are connected to events (failure, restoration, etc.), states (fault, up state, down state, outage, etc.) or activities (e.g. maintenance), with their time durations.

Part I of Supplement No. 6 provides necessary identification of times, events, states and maintenance activities.

2 Related Recommendations and Supplements

Recommendation E.600:	Terms and definitions of traffic engineering
Supplement No. 6:	Terms and definitions for quality of service, network performance, dependability and trafficability studies.

3 Performances

3.1 *Service related performances*

3101 **quality of service**

F: qualité de service

S: calidad de servicio

The collective effect of service performances which determine the degree of satisfaction of a *user* of the *service.*

Note 1 – The *quality of service* is characterized by the combined aspects of *service support performance, service operability performance, serveability performance, service integrity* and other factors specific to each *service.*

Note 2 – The term "quality of service" is not used to express a degree of excellence in a comparative sense nor is it used in a quantitative sense for technical evaluations. In these cases a qualifying adjective (modifier) shall be used.

3102 **serveability performance**

F: servibilité (d'un service)

S: servibilidad (de un servicio)

The ability of a *service* to be obtained – within specified tolerances and other given conditions – when requested by the *user* and continue to be provided for a requested *duration.*

Note – *Serveability performance* may be subdivided into the *service accessibility performance* and the *service retainability performance.*

3103 **service accessibility performance**

F: accessibilité (d'un service)

S: accesibilidad (de un servicio)

The ability of a *service* to be obtained, within specified tolerances and other given conditions, when requested by the *user.*

Note – This takes into account the transmission tolerance and the combined aspects of *propagation performance, trafficability performance* and *availability performance* of the related systems.

3104 **service retainability performance**

F: continuabilité (d'un service)

S: retenibilidad (de un servicio)

The ability of a *service,* once obtained, to continue to be provided under given conditions for a requested duration.

Note – Generally this depends on the transmission tolerances, the *propagation performance* and *reliability performance* of the related systems. For some services, for example packet switching, this also depends on the *trafficability performance* and the *availability performance* of the related systems.

3105 **service support performance**

F: logistique de service

S: logística del servicio

The ability of an organization to provide a *service* and assist in its utilization.

Note – An example of *service support performance* is the ability to provide assistance in commissioning a basic service, or a supplementary service such as the call waiting service or directory enquiries service.

3106 **service operability performance**

F: facilité d'utilisation (d'un service)

S: facilidad de utilización (de un servicio)

The ability of a *service* to be successfully and easily operated by a *user.*

3107 **service integrity**

F: integrité de service

S: integridad del servicio

The degree to which a *service* is provided without excessive impairments, once obtained.

Note – This *service* is characterized by the *transmission performance* of the system.

3108 **transmission performance**

F: qualité de transmission

S: calidad de transmisión

The level of reproduction of a signal offered to a telecommunications system, under given conditions, when this system is in an *up state.*

3.2 *Item related performances*

3201 **network performance**

F: qualité technique du réseau

S: calidad de funcionamiento de la red

The ability of a network or network portion to provide the functions related to *communications* between *users.*

Note 1 – Network performance contributes to *serveability performance* and *service integrity* (see Figure 2/E.800).

Note 2 – Network performance measures are meaningful to network providers and are quantifiable at boundaries of network portions to which they apply. Quality of service measures are only quantifiable at a service access point.

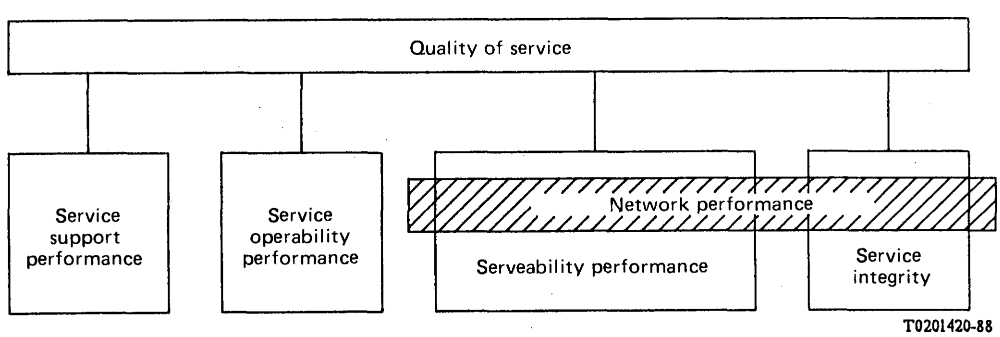


FIGURE 2/E.800

Relationship between quality of service and network performance

3202 **trafficability performance**

F: traficabilité; capacité d'écoulement du trafic

S: aptitud para cursar tráfico

The ability of an *item* to meet a traffic demand of a given size and other characteristics, under given internal conditions.

Note – Given internal conditions refer, for example, to any combination of *faulty* and not *faulty* sub-items.

3203 **capability**

F: capacité; capabilité (d'une entité)

S: capacidad

The ability of an *item* to meet a demand of a given size under given internal conditions.

Note 1 – Internal conditions refer, for example, to any given combination of *faulty* and not *faulty* sub-items.

Note 2 – This is also called *trafficability performance.*

3204 **propagation performance**

F: caractéristiques de propagation

S: característica de propagación

The ability of a propagation medium, in which a wave propagates without artificial guide, to transmit a signal within the given tolerances.

Note – The given tolerances may apply to variations in signal level, noise, interference levels, etc.

3205 **effectiveness (performance)**

F: efficacité

S: efectividad

The ability of an *item* to meet a service demand of a given size.

Note – This ability depends on the combined aspects of the *capability* and the *availability performance* of the *item.*

4 Interruptions

4101 **interruption; break (of service)**

F: interruption; coupure (d'un service)

S: interrupción (de un servicio); corte (de un servicio)

Temporary inability of a *service* to be provided persisting for more than a given *time duration,* characterized by a change beyond given limits in at least one parameter essential for the *service.*

Note 1 – An *interruption* of a *service* may be caused by *disabled states* of the *items* used for the *service* or by external reasons such as high service demand.

Note 2 – An *interruption* of a *service* is generally an *interruption* of the transmission, which may be characterized by an abnormal value of power level, noise level, signal distortion, *error* rate, etc.

4102 **time between interruptions**

F: temps entre interruptions

S: tiempo entre interrupciones

The *time duration* between the end of one *interruption* and the beginning of the next.

4103 **interruption duration**

F: durée d'interruption

S: duración de interrupción

The *time duration* of an *interruption.*

4104 **mean time between interruptions (MTBI)**

F: durée moyenne entre interruptions (DMEI)

S: tiempo medio entre interrupciones

The *expectation* of the *time between interruptions.*

4105 **mean interruption duration (MID)**

F: durée moyenne d'une interruption (DMI)

S: duración media de una interrupción

The *expectation* of the *interruption duration.*

5 Measures of performances

5.1 *Service support performance*

5101 **mean service provisioning time**

F: délai moyen pour la fourniture d'un service

S: tiempo medio de espera (para la prestación de un servicio)

The *expectation* of the *duration* between the *instant of time* a potential *user* requests that an organization provides the necessary means for a *service,* and the *instant of time* when these means are furnished.

5102 **billing error probability**

F: probabilité d'erreur de facturation

S: probabilidad de error de facturación

The *probability* of an *error* when billing a *user* of a *service.*

5103 **incorrect charging or accounting probability**

F: probabilité de taxation erronée

S: probabilidad de tarificación o de contabilidad incorrectas

The *probability* of a *call attempt* receiving incorrect charging or accounting treatment.

5104 **undercharging probability**

F: probabilité de sous-taxation

S: probabilidad de subtarificación

The *probability* that an effective *call* will be undercharged for any reason.

5105 **overcharging probability**

F: probabilité de surtaxation

S: probabilidad de sobretarificación

The *probability* that an effective *call* will be overcharged for any reason.

5106 **billing integrity** (probability)

F: (probabilité de) justesse de facturation

S: integridad de la facturación (probabilidad de)

The *probability* that the billing information presented to a *user* correctly reflects the type, destination and duration of the *call.*

5.2 *Service operability performance*

5201 **service user mistake probability**

F: probabilité d'erreur d'un usager

S: probabilidad de error de un usuario (de un servicio)

Probability of a *mistake* made by a *user* in his attempt to utilize a *service.*

5202 **dialling mistake probability**

F: probabilité d'erreur de numérotation

S: probabilidad de error de marcación

The *probability* that the *user* of a telecommunication network makes dialling *mistakes* during his *call attempts.*

5203 **service user abandonment probability**

F: probabilité d'abandon (d'accès à un service par un usager)

S: probabilidad de abandono de un servicio por un usuario

The *probability* that a *user* abandons the attempt to use a *service.*

Note – Abandonments may be caused by excessive *user* mistake rates, by excessive service access delays, etc.

5204 **call abandonment probability**

F: probabilité d'abandon (d'une tentative d'appel)

S: probabilidad de abandono de una tentativa de llamada

The *probability* that a *user* abandons the *call attempt* to a telecommunication network.

5.3 *Service accessibility performance*

5301 **service accessibility; service access probability**

F: accessibilité (d'un service)

S: accesibilidad de un servicio; probabilidad de acceso a un servicio

The *probability* that a *service* can be obtained within specified tolerances and other given operating conditions when requested by the *user.*

5302 **mean service access delay**

F: durée moyenne d'accès

S: retardo medio de acceso a un servicio; demora media de acceso a un servicio

The *expectation* of the *time duration* between an initial *bid* by the *user* for the acquisition of a *service* and the *instant of time* the user has access to the *service*, the *service* being obtained within specified tolerances and other given operating conditions.

5303 **network accessibility**

F: accessibilité (d'un réseau)

S: accesibilidad (de una red)

The *probability* that the *user* of a *service* after a request receives the proceed-to-select signal within specified conditions.

Note – The proceed-to-select signal is that signal inviting the *user* to select the desired *destination.*

5304 **connection accessibility**

F: accessibilité

S: accesibilidad de una conexión

The *probability* that a *connection* can be established within specified tolerances and other given conditions following receipt by the exchange of a valid code.

5305 **mean access delay**

F: durée moyenne d'accès

S: retardo medio de acceso; demora media de acceso

The *expectation* of the *time duration* between the first *call attempt* made by a *user* of a telecommunication network to reach another *user* or a *service* and the *instant of time* the *user* reaches the wanted other *user* or *service,* within specified tolerances and under given operational conditions.

5306 **p-fractile access delay**

F: quantile-p de la durée d'accès

S: cuantil-p del retardo de acceso; cuantil-p de la demora de acceso

The *p-fractile* value of the *duration* between the *first call attempt* made by a *user* of a telecommunication network to reach another *user* or a *service* and the *instant of time* the *user* reaches the wanted other *user* or *service,* within specified tolerances and under given operational conditions.

5307 **accessibility of a connection to be established**

F: accessibilité d'une communication à établir

S: accesibilidad de una conexión por establecer

The *probability* that a switched *connection* can be established, within specified transmission tolerances, to the correct *destination,* within a given *time interval,* when requested by the *user.*

Note 1 – For user-originated calls, it could express the *probability* of a successful call establishment on the first attempt. For operator-handled calls, it could represent the *probability* of having a satisfactory *connection* established within a given *time duration.*

Note 2 – In general, the tolerances should correspond to a level of *transmission performance* which makes the connection unsatisfactory for *service* such that, for example, a substantial percentage of *users* would abandon the *connection.*

5308 **unacceptable transmission probability**

F: probabilité d'une transmission inacceptable

S: probabilidad de transmisión inaceptable

The *probability* of a *connection* being established with an unacceptable speech path transmission quality.

5309 **no tone probability**

F: probabilité de non tonalité

S: probabilidad de ausencia de tono

The *probability* of a *call attempt* encountering no tone following receipt of a valid code by the exchange.

5310 **misrouting probability**

F: probabilité d'acheminement erroné

S: probabilidad de encaminamiento erróneo

The *probability* of a *call attempt* being misrouted following receipt by the exchange of a valid code.

5.4 *Service retainability performance*

5401 **service retainability**

F: continuabilité (d'un service)

S: retenibilidad (de un servicio)

The *probability* that a *service*, once obtained, will continue to be provided under given conditions for a given *time duration.*

5402 **connection retainability**

F: continuabilité (d'une chaîne de connexion)

S: retenibilidad (de una conexión)

The *probability* that a *connection*, once obtained, will continue to be provided for a *communication* under given conditions for a given *time duration.*

5403 **retainability of an established connection**

F: continuabilité d'une communication établie

S: retenibilidad de una conexión establecida

The *probability* that a switched *connection*, once established, will operate within specified transmission tolerances without *interruption* for a given *time interval.*

5404 **premature release probability; cut-off call probability**

F: probabilité de libération prématurée

S: probabilidad de liberación prematura; probabilidad de corte de una llamada

The *probability* that an established *connection* will be released for a reason other than intentionally by any of the parties involved in the call.

5405 **release failure probability**

F: probabilité de non-libération

S: probabilidad de fallo de liberación

The *probability* that the required *release* of a *connection* will not take place.

5.5 *Serveability performance*

5501 **probability of successful service completion**

F: probabilité d'exécution correcte du service

S: probabilidad de prestación satisfactoria de un servicio

The *probability* that a *connection* can be established, under satisfactory operating conditions, and retained for a given *time interval.*

5.6 *Transmission performance*

5601 **bit error ratio (BER)**

F: taux d'erreur sur les bits (TEB)

S: tasa de errores en los bits; tasa de error en los bits (TEB)

The ratio of the number of bit errors to the total number of bits transmitted in a given *time interval.*

5602 **error free seconds (EFS)**

F: secondes sans erreur (SSE)

S: segundos sin error (SSE)

The ratio of the number of one-second intervals during which no bits are received in error to the total number of one-second intervals in the *time interval.*

Note 1 – The length of the *time interval* needs to be specified.

Note 2 – This ratio is usually expressed as a percentage.

6 Common concepts

The following concepts are used in the definitions of this Recommendation. Others used, such as probability, measure, up state, disabled state, time duration, user and connection may be found in Recommendation E.600 and in Supplement No. 6.

6001 **service**

F: service

S: servicio

A set of functions offered to a *user* by an organization.

6002 **item ; entity**

F: entité; individu

S: elemento; entidad; ítem

Any part, device, subsystem, functional unit, equipment or system that can be individually considered.

Note 1 – An *item* may consist of hardware, software or both, and may also include people, e.g. operators in a telephone operator system.

Note 2 – In French, the term *entité* replaces the term *dispositif* previously used in this meaning, because the term *dispositif* is also the common equivalent for the English term "device".

Note 3 – In French, the term *individu* is used mainly in statistics.

ANNEX A

(to Recommendation E.800)

Alphabetical list of definitions contained in this Recommendation

5307 accessibility of a connection to be established
5102 billing error probability
5106 billing integrity (probability)
5601 bit error ratio
5204 call abandonment probability
3203 capability
5304 connection accessibility
4101 break (of service)
5402 connection retainability
5404 cut-off call probability
5202 dialling mistake probability
3205 effectiveness (performance)
6002 entity
5602 error free seconds (EFS)
5103 incorrect charging or accounting probability
4103 interruption duration
4101 interruption
6002 item
5305 mean access delay
4105 mean interruption duration
5302 mean service access delay
5101 mean service provisioning time
4104 mean time between interruptions
5310 misrouting probability
5303 network accessibility
3201 network performance
5309 no tone probability
5105 overcharging probability
5306 p-fractile access delay
5404 premature release probability
5501 probability of successful service completion
3204 propagation performance
3101 quality of service
5405 release failure probability
5403 retainability of an established connection
3102 serveability performance
6001 service
5301 service access probability
5301 service accessibility
3103 service accessibility performance
3107 service integrity
3106 service operability performance
5401 service retainability
3104 service retainability performance
3105 service support performance
5203 service user abandonment probability
5201 service user mistake probability
4102 time between interruptions
3202 trafficability performance
3108 transmission performance
5308 unacceptable transmission probability
5104 undercharging probability

SECTION 2

MODELS FOR TELECOMMUNICATION SERVICES

Recommendation E.810

MODEL FOR THE SERVEABILITY PERFORMANCE ON A BASIC CALL IN THE TELEPHONE NETWORK[1)]

Introduction

This Recommendation[2)] is one of a set of closely related Recommendations concerned with the accessibility and retainability of telephone services, as listed below.

The CCITT,

considering

(a) that there is a desire to establish overall objectives for the quality of service as perceived by the users;

(b) that such objectives can then be used as a basis for the design, planning, operation and maintenance of telecommunication networks and their component parts;

(c) that Recommendation E.800 contains terms and definitions for the quality of service, the reliability and availability performances and related characteristics of services and networks,

recommends

that the telephone call model given in this Recommendation shall be used by Administrations to design, plan, operate and maintain their networks taking into account the objectives given in Recommendations:

E.830 Models for the allocation of international telephone connection retainability, accessibility and integrity;

E.845 Connection accessibility objective for the international telephone service;

E.850 Connection retainability objective for the international telephone service.

Note – Refer also to the draft Recommendation on interruption objectives which is being studied under Question 39/II.

---

1) Although this Recommendation deals with the telephone service, in principle the model and the decomposition of serveability performance can also be applied to other telecommunication services. The elaboration of this principle is left for further study.

2) Some of the terms in this Recommendation, for example the noun "measure", are used in the sense of their definition given in Recommendation E.800.

1 Model of a basic telephone call and its serveability performance

The following simplified model illustrates the principal phases of a basic telephone call. It also interrelates these phases to the service-related performance concepts and their principal measures as well as to the main causes of failure in the establishment and retention of such a call and its subsequent billing.

The model also indicates where, in this series of phases, user actions or mistakes may influence the call.

2 Comments to the model and its applications

2.1 *Mathematical modelling*

In a simple case of statistical independence, the probabilities may be combined into the following mathematical models:

$$P = (P_{11} \cdot P_{12}) \cdot P_2 \cdot (P_{31} \cdot P_{32}) \cdot P_4$$

to express the probability of a correctly billed revenue-making call and,

$$P = (P_{11} \cdot P_{12}) \cdot P_2 \cdot (P_{31} \cdot P_{32})$$

to express the probability of a successfully completed call.

2.2 *Contributions to causes of call failure*

It is generally recognized that the various parts of a national or international network may be of different importance to the successful completion of the various phases of a call. For example, the network accessibility is mainly determined by the telephone set, the subscriber line and the local exchange; the connection accessibility by the exchanges, transmission network and signalling network used; the billing integrity is dependent on the charging facilities used by the network parts that constitute the connection and the equipment for processing the billing information, etc. In some Administrations, the telephone set is not considered as a part of the network and in that case it is not included in the concept of network performance.

2.3 *Time aspects of measures*

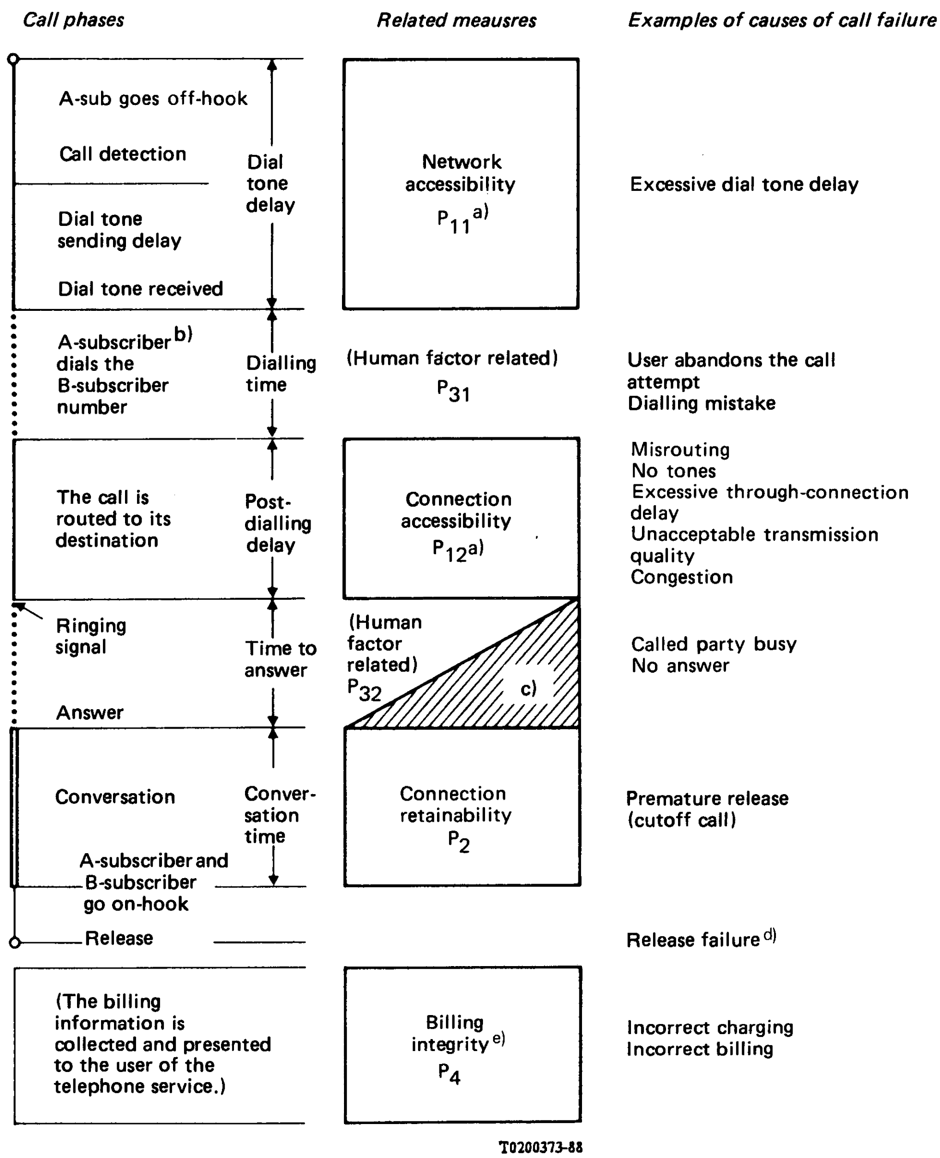
Depending on the intended application of the measures indicated in Figure 1/E.810, it may be appropriate to express these measures as instantaneous values related to a given instant of time or as a mean over a given time interval.

Advice on which variant to use should be given in each specific relevant Recommendation.

2.4 *Space aspects of averages*

The measures as indicated in Figure 1/E.810 could be applied to calls between particular destinations as traffic weighted averages over a number of destinations, etc.

Each relevant Recommendation should clearly specify which alternative(s) to use.



a) Network accessibility and connection accessibility combine into service accessibility.

b) The routing of the call may start before all digits have been received.

c) The shaded area shows that a premature release can occur durine the time to answer.

d) The release of a call is not a separate phase in this model. A release failure may result in network inaccessibility for a new call.

e) The billing integrity has been shown for completeness but is not a part of serveability performance.

FIGURE 1/E.810

Model of the serveability performance on a basic call in the telephone network

MODELS FOR THE ALLOCATION OF INTERNATIONAL TELEPHONE CONNECTION RETAINABILITY, ACCESSIBILITY AND INTEGRITY

Introduction

This Recommendation is one of a set of closely related Recommendations, comprising Recommendations E.810, E.830, E.845, E.850 and E.855 concerned with the accessibility, retainability and integrity of telephone services.

The CCITT,

considering

that there is a need to establish hypothetical reference connection models to allocate overall connection retainability, accessibility and integrity objectives to the component parts of international connections,

recommends

three models for retainability (one of which is for a typical, or average, international connection), and one model for accessibility and integrity.

1 Retainability models

The models are shown respectively, in Figures 1/E.830, 2/E.830 and 3/E.830. As indicated by Figure 1/E.830, the typical connection has two circuits in each of the national systems, and one in the international chain. In the 90th percentile case, there would be three in the national systems and one in the international chain.

2 Number of circuits

The number of circuits in each of the models is based on Table 1/E.830. The entries of this table are based on the information contained in Table 1/G.101.

The mean and model number of national extension circuits are both equal to 2. This applies to both originating and terminating national systems. The mean number of international circuits is 2.1 and the model number is 2.

TABLE 1/E.830

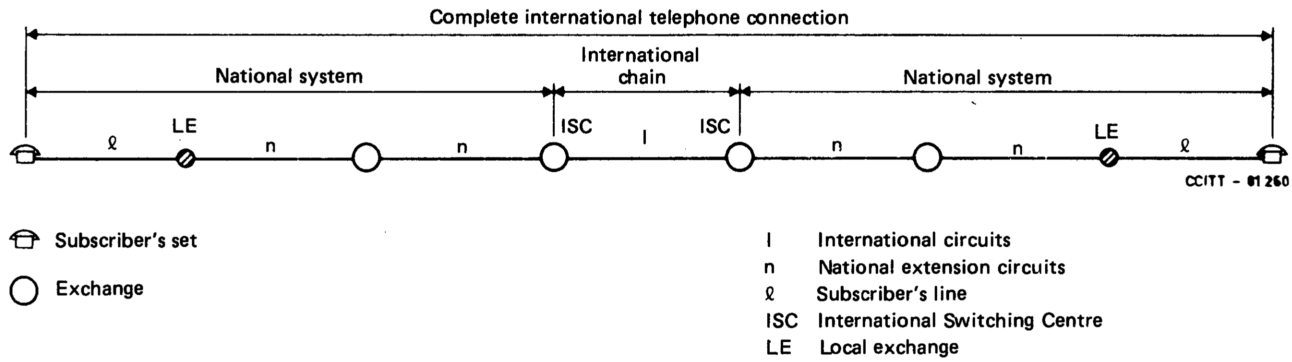
Probabilities of the number of circuits in the two national systems and the international chain (expressed as percentages)

Number of circuits	Originating LE-ISC	International ISC-ISC	Terminating ISC-LE′
1	33.8	95.1	32.9
2	38.9	4.5	39.5
3	20.2	0.3	20.4
4	6.0	–	6.1
5	1.0	–	1.0

LE Local exchange

ISC International switching centre

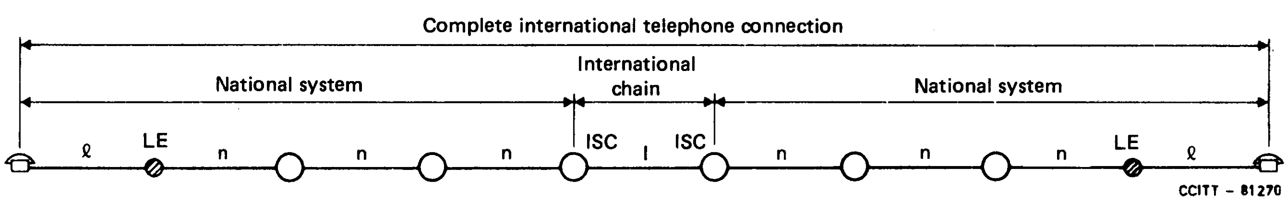
Note – The possibilities of 6 and 7 circuits in the originating national system are 0.005% and 0.0005% respectively. The probabilities of 4, 5 and 6 international circuits are 0.03%, 0.00007% and 0.00009% respectively.



Note – For the purposes of this Recommendation, the international switching centres are considered to be a part of the international chain.

FIGURE 1/E.830

Typical international telephone connection model

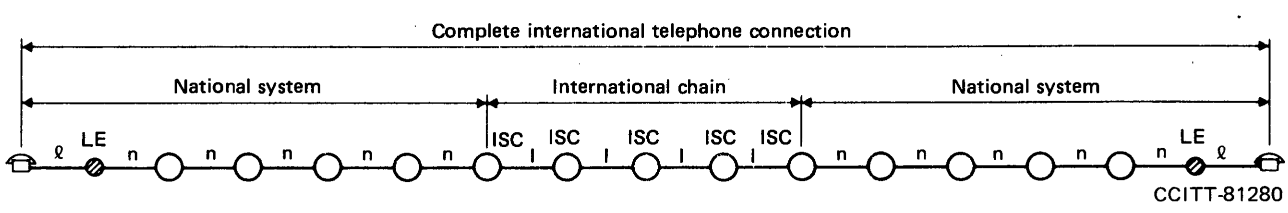


Note 1 – For an explication of legends, see Figure 1/E.830.

Note 2 – For the purposes of this Recommendation, the international switching centres are considered to be part of the international chain.

FIGURE 2/E.830

90th percentile international telephone connection model



Note 1 – For an explanation of legends, see Figure 1/E.830.

Note 2 – For the purposes of this Recommendation, the international switching centres are considered to be part of the international chain.

FIGURE 3/E.830

Possible longest international telephone connection model

3 Accessibility and integrity model

The model to be used for allocation of the connection accessibility and integrity objectives found in Recommendations E.845 and E.855 respectively to the national portions and international chains of international connections is shown in Figure 4/E.830.

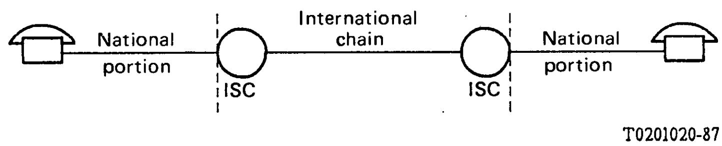


FIGURE 4/E.830

Model for allocating connection accessibility and integrity

SECTION 3

OBJECTIVES FOR QUALITY OF SERVICE AND RELATED CONCEPT OF TELECOMMUNICATION SERVICES

Recommendation E.845 [1]

CONNECTION ACCESSIBILITY OBJECTIVE FOR THE INTERNATIONAL TELEPHONE SERVICE [2]

Introduction

This Recommendation is one of a set of closely related Recommendations, comprising Recommendations E.810, E.830, E.845, E.850 and E.855 concerned with the accessibility, retainability and integrity of telephone services.

Preamble

This Recommendation provides an overall end-to-end connection accessibility (availability) objective for *international* switched telephone service.

Connection accessibility is a component of service accessibility as defined in Recommendation E.800.

This Recommendation contains a measure of connection accessibility, an objective, and an allocation of the objective to the national systems and international chain of international connections. The Recommendation also relates the overall end-to-end performance to the reliability and availability of circuits and exchanges in a way useful for network design purposes.

The objective includes the effects of equipment faults and traffic congestion.

The CCITT,

considering

(a) that connection accessibility is defined in Recommendation E.800;

(b) that customers rank connection inaccessibility as one of the most annoying of call set-up impairments;

(c) that an objective for connection accessibility which takes into account customer opinion about the call set-up phase is consistent with other Recommendations which have recommended an objective for service retainability based, in part, on customer opinion;

---

[1] Formerly G.180, in *Red Book*, Fascicle III.1.

[2] Some of the terms in this Recommendation, for example, the noun "measure", are used in the sense of their definition given in Recommendation E.800.

(d) that connection accessibility will not be constant over time, even for a particular calling and called line pair. One suitable measure is a long-term average network connection failure probability. (Other suitable measures may also be required.);

(e) that the overall objective for connection accessibility should be allocable to the national systems and the international chain of the international connection;

(f) that the objective should take into account the concerns of network planners and system designers, provide useful guidance to both and may be used by Administrations in providing a method for verifying whether or not network performance is acceptable;

(g) that the overall connection accessibility should be controlled by the accessibility performances of individual exchanges and circuits, and that to obtain this control, the overall connection accessibility must be mathematically linked to the equipment availability and reliability,

recommends

1 Measure of connection accessibility

Connection accessibility shall be measured using the long-term average network connection failure probability, which is the complement of the connection access probability as defined in Recommendation E.800.

The network connection failure probability P_{NCF} can be estimated by using the following formula:

$$P_{NCF} = \frac{Q_N}{N}$$

where Q_N is the number of unsuccessful connection access attempts and N is the total number of connection access attempts in some time period (to be determined).

A method for estimating the required call sample size is contained in Annex A.

For purposes of network design, the network connection failure probability, P_{NCF}, can also be calculated using the method outlined in Annex B. Annex C describes how the busy and non-busy hours affect the network connection failure (NCF).

Note 1 – Those unsuccessful connection access attempts reflecting failure of the *network* to work properly, from the user's perspective, are called network connection failure. They are call failures an astute caller can determine and are caused by network faults and congestion. A network connection failure is any valid bid for service which receives one of the following network responses:

1) dial tone returned after dialling is completed;
2) no ring and no answer;
3) all circuits busy signal or announcement;
4) connection to the wrong number (misrouting);
5) double connection.

This list may not be exhaustive.

Note 2 – This definition of network connection failure is based on the response the caller can hear.

Note 3 – There are two generic causes of network connection failures: equipment faults and traffic congestion.

Note 4 – The averaging interval (to be determined) used for estimating the connection failure probability shall include normal and peak hour traffic periods. In the event of exceptionally high traffic demand (public holiday, natural disaster, etc.) failure rates higher than the objective may be tolerated.

Note 5 – The network connection failure probability should be estimated by Administrations in a manner consistent with obtaining, from the Administration's point of view, reasonably accurate estimates.

2 Objective for connection accessibility

Connection accessibility is acceptable if the long-term average connection failure probability, expressed as a percentage, does not exceed a value (overall average for all international calls) of $A\%$ to $B\%$ (values to be determined). Additionally, the long-term average failure probability at any single international homing exchange should never exceed $C\%$ (value to be determined).

Note – Possible values for A, B and C are in the range of 10% to 20%.

3 Allocation of the overall objective to the national systems and international chain

The network connection failure probability objective shall be apportioned as follows:

$X\%$ to the originating national system,

$Y\%$ to the international chain,

$Z\%$ to the terminating national system,

where $X + Y + Z = P$, and P is the overall objective stated in § 2.

Note 1 – The connection access attempt may fail in the national systems or the international chain of the connection.

Note 2 – The objective takes into account all means of "defense" of the network against failure to complete the connection, including alternate routing, if used.

Note 3 – The network connection failure probability of the national systems or international chain is defined as the probability that the call access attempt will fail because of some problem (equipment fault or congestion) in the systems or chain.

Note 4 – Values for X, Y and Z are in the range of 3% to 7%.

ANNEX A

(to Recommendation E.845)

Method for selecting the required call sample size, N

The network connection failure probability shall be estimated by Administrations in a manner consistent with obtaining reasonably accurate estimates.

The number of call access attempts sampled shall be sufficiently large to obtain a good estimate of the probability.

A method of picking a sample size N could be used which could produce a maximum error of measurement, e, (to be determined) with confidence level, α (to be determined).

Recommendation E.850 contains a method for estimating the sample size required to estimate cutoff call probability. This method should be studied for application here.

ANNEX B

(to Recommendation E.845)

Method for relating overall network connection failure probability to the reliability and availability performance of exchanges and circuits

The following equation gives the relationship between the overall network connection failure probability, P_{NCF}, and the probabilities of connection failure in the national systems and international chain of the connection:

$$P_{NCF} = 1 - (1 - P_{OE})(1 - P_I)(1 - P_{TE})$$

where P_{OE} is the probability that the access attempt fails in the originating national system, P_I is the probability of failure in the international chain and P_{TE} is the probability of failure in the terminating national system.

Hypothetical reference connections for the three parts of an international connection are shown in Figure B-1/E.845. The proportion of calls (F_n) which are routed over the parts are also given in the figure. The values are taken from Table 1/G.101.

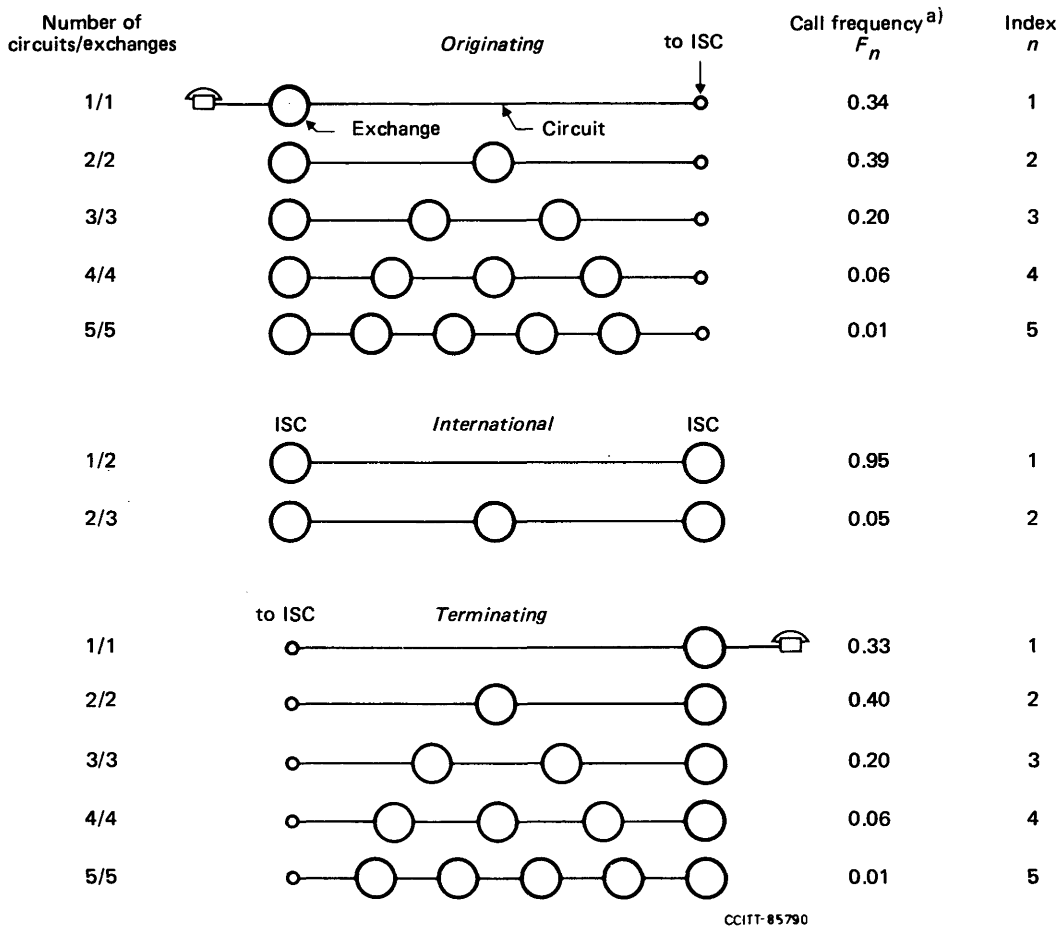
The probability that a connection access attempt fails in either of the parts is given by the following equations:

$$P_{OE} = 1 - \sum_{n=1}^{5} F_n (1 - P_c)^n (1 - P_s)^n$$

$$P_I = 1 - \sum_{n=1}^{2} F_n (1 - P_c')^n (1 - P_s')^{n+1}$$

$$P_{TE} = 1 - \sum_{n=1}^{5} F_n (1 - P_c'')^n (1 - P_s'')^n$$

where n is the number of circuits in a selected part. F_n is the call frequency for an n-circuit system or chain (from Figure B-1/E.845).



a) Values taken from Table 1/G.101.

FIGURE B-1/E.845

Hypothetical reference connections as a function of call frequencies

P_c, P'_c and P''_c are the probabilities that the connection access fails in the originating system, international chain or terminating system circuits, respectively. (It is assumed here for simplicity that all circuits in a system or chain have the same probability of failure. However, this is not a requirement.)

P_s, P'_s and P''_s are the probabilities that the connection access attempt fails in the originating system, international chain (note that ISC is assumed part of the international chain) or terminating system exchanges, respectively. (For simplicity, all exchanges are assumed to have the same failure probability, but this is not a requirement.)

A circuit or exchange can cause a network connection failure for one of three reasons:

1) The call is blocked because of congestion. The probability of blockage is P_{CB} and P_{SB} for circuits and exchanges, respectively.
2) The circuit or exchange fails during the call set-up time. The probability of such a failure is P_{CF} and P_{SF} for circuits and exchanges, respectively.
3) The circuit or exchange is unavailable to arriving calls, so all calls arriving during the downtime fail to be completed. These probabilities are P_{CD} and P_{SD} for circuits and exchanges, respectively.

The probability that a circuit or exchange causes a network connection failure is given by the following equations, respectively:

$$P_C = 1 - (1 - P_{CB})(1 - P_{CF})(1 - P_{CD})$$

$$P_S = 1 - (1 - P_{SB})(1 - P_{SF})(1 - P_{SD})$$

The failure probabilities P_{CF} and P_{SF} can be expressed in terms of the long-term mean failure intensities Z_c and Z_s of circuits and exchanges, respectively, by the following equations:

$$P_{CF} = Z_c \, T_s$$

$$P_{SF} = Z_s \, T_s,$$

where T_s is the long-term average call set-up time.

Similarly, the failure probabilities P_{CD} and P_{SD} can be expressed in terms of the long-term mean accumulated downtime $(MADT)_c$ and $(MADT)_s$ of circuits and exchanges, respectively, by the following equations:

$$P_{CD} = \frac{(MADT)_c \times \alpha_c}{\mathrm{K} \times N}$$

$$P_{SD} = \frac{(MADT)_s \times \alpha_s}{\mathrm{K} \times N}$$

α_c and α_s are the long-term average call arrival rates for circuits and exchanges, respectively, and N is the long-term average number of call attempts (in some interval, such as one year).

K is a constant equal to the number of units of time (minutes or seconds) used to express the downtime in the long-term averaging interval selected (such as a year).

For example, if the downtime is expressed in minutes and the averaging interval is one year, then K = 525 600 min./year.

ANNEX C

(to Recommendation E.845)

Effects of busy hours and non-busy hours on the network connection failure

The two major components of network connection failure (NCF) are the blocking rate due to congestion and connection access attempt failures due to equipment faults. Equipment faults are further divided into major and minor faults. These components affect NCF differently.

C.1 *Influences of faults*

Faults of subsystems in a telephone network may be divided into two categories, according to their influence on network performance. Table C-1/E.845 shows two fault categories: major and minor.

TABLE C-1/E.845

Failure category	Definition	Network components
Major (considerable) influence fault	Fault wherein a connection access attempt encounters a situation such that service degradation of network component(s) lasts for some period of time, owing to large scale failure of equipment, and a subscriber cannot be assured of normal service.	Subscriber line, subscriber terminal [b], exchange, transmission line, service center
Minor (less important) influence fault [a]	Small scale fault wherein a connection access attempt is handled incorrectly and encounters no signal (e.g. dial tone, ring-back tone), no connection, low level speech signal, etc., i.e. less important service degradation is experienced	

[a] Intermittent fault is excluded and its treatment is an unresolved problem.
[b] In some Administrations the subscriber terminal is not considered a network component.

C.2 *Relationship between NCF, congestion and fault*

Congestion-related NCF depends on the traffic offered to a system being considered (a switching system, a network, etc.).

The effects of a minor fault will be considered as so-called white noise where the absolute value is small and fluctuates at random.

The effects of a major (complete) fault depend on the offered traffic volume at the time of fault. If a major fault occurred during busy hours, there would be an extremely high value for NCF. Conversely, a major fault during non-busy hours will merely yield a small NCF, no matter how large the affected system is. This is because the traffic load itself is small. Since it is usually expected that major faults will be very rare, NCF characteristics under major fault conditions are different from those under minor fault conditions which may be daily occurrences.

C.3 *Long-term NCF (averaged throughout a year)*

The long-term NCF concerned with traffic congestion during non-busy hours will be much smaller than that during busy hours. Since both cumulative call failures N_f and total calls offered N_o during non-busy hours are much smaller than those during busy hours, the averaged 24-hour NCF including non-busy and busy hours effects will not be much different from the busy hour NCF.

A major fault can be identified but a minor fault cannot be specified correctly when network operators maintain network equipment. By measuring long-term NCF during non-busy hours, the effect of minor faults can be estimated because NCF during non-busy hours is attributed not to traffic congestion but to minor faults.

C.4 *NCF and busy hour pattern*

In a country (international region) with several standard time zones, there will be several busy hours. In such cases, a connection in the network may include busy and non-busy network components. Thus, an averaged 24-hour NCF would be helpful to administer a network with different time zones.

However, the averaged 24-hour NCF does not seem to be appropriate to administer a network having only one standard time zone because its fault-related term is too small to affect the total NCF, and it might be too late by the time an extraordinary NCF value has been detected. The NCF averaged during non-busy hours would be one measure for monitoring the effect of equipment faults (minor faults) on subscribers, since this will become a major factor during non-busy hours.

Recommendation E.850 [1)]

CONNECTION RETAINABILITY OBJECTIVE FOR THE INTERNATIONAL TELEPHONE SERVICE [2)]

Introduction

This Recommendation is one of a set of Recommendations, comprising Recommendations E.810, E.830, E.845, E.850 and E.855 concerned with the accessibility, retainability and integrity of telephone services.

The CCITT,

considering

(a) that "premature release" is defined in Recommendation E.800 as the event that an established connection will be released for a reason other than intentionally by any of the parties involved in the call;

(b) that premature release is a measure of connection retainability;

(c) that a prematurely released connection is considered high in annoyance as perceived by telephone users;

(d) that the probability of a premature release is a function of network component failure intensity and call holding time;

(e) that the objective should take account of the expectations and tolerances of users to the premature release impairment as well as the capabilities of current technology;

(f) that the objective might not be met at the present time but should be viewed as a long-term goal;

(g) that the objective should take into account the concerns of network planners and system designers, provide useful guidance to each, and it can be used by Administrations in a consistent way to measure connection retainability performance;

(h) that connection retainability is defined in Recommendation E.800,

recommends

1 Definitions

A **prematurely released telephone connection** is known as a cutoff call when the connection is completely broken, or

1) when a single interruption occurs lasting for longer than ten seconds which causes the transmission quality of the connection to be unsuitable for voice communications;
2) when a succession of interruptions occur lasting less than ten seconds where the product of the average duration of each interruption and the frequency of occurrence (i.e., average number of interruptions/seconds) exceeds 0.005.

---

1) Formerly G.181, in *Red Book*, Fascicle III.1.

2) Some of the terms in this Recommendation, for example, the noun "measure", are used in the sense of their definition given in Recommendation E.800.

2 A measure to quantify telephone connection retainability performance

The measure to be used shall be the complement of connection retainability, namely the probability of a prematurely released telephone connection when normalized to a call holding time of one minute (P_r). The estimator of the premature release probability is the premature release call ratio (P_{re}) which is defined as:

$$P_{re} = \frac{1 - \frac{R_N}{N}}{T}$$

where N is the number of telephone calls successfully established in some period of time, T is the mean call holding time in minutes and R_N is the number of telephone calls successfully completed out of such N calls (see Annex A and Annex B).

3 Overall objective for premature release probability

The provisional objective for the normalized premature release probability (P_r) shall be such that the performance is better than the values given below:

for typical international connections:

$$2 \times 10^{-4} \leqslant P_r \leqslant 4 \times 10^{-4},$$

for 90th percentile international connections:

$$4 \times 10^{-4} \leqslant P'_r \leqslant 8 \times 10^{-4}$$

for worst case international connections:

$$8 \times 10^{-4} \leqslant P''_r \leqslant 1.6 \times 10^{-3}.$$

Note 1 – It is intended to establish a single value for P_r, P'_r or P''_r in the future.

Note 2 – The typical 90th percentile and worst case connections mentioned above shall be assumed to be those hypothetical reference connections (HRXs) given in Recommendation E.830.

Note 3 – See Annex B.

4 Allocation of the overall objective

It is desirable, for planning purposes, to allocate the overall objective for a typical connection to the national systems and the international chain of the HRX. The overall objective is given by:

$$P_r = P_{r_{n1}} + P_{r_{n2}} + P_{r_i}$$

where $P_{r_{n1}}$ and $P_{r_{n2}}$ are the premature release probabilities for originating and terminating national systems respectively and P_{r_i} is the premature release probability of the international chain. The allocation of the overall objective to national systems and international chain shall be as follows:

$$P_{r_{n1}} = P_{r_{n2}} = \alpha \, P_{r_i}.$$

Note 1 – α is provisionally recommended as being equal to 2. Thus, for example, if:

$$P_r = 3 \times 10^{-4}$$

then

$$P_{r_{n1}} = P_{r_{n2}} = 1.2 \times 10^{-4}$$

and

$$P_{r_i} = 0.6 \times 10^{-4}.$$

Note 2 – Further allocation of the overall objective to the circuits and exchanges used in a connection might also be desirable.

Note 3 – Objectives for the permissible probability of premature release of an established telephone connection in Integrated Digital Networks (IDNs) and mixed (analogue/digital) networks, due to transit digital or local and combined local/transit exchange malfunctions, are specified in the Recommendations Q.504 or Q.514.

ANNEX A

(to Recommendation E.850)

Relationship between the premature release probability and its estimator

The following relationship exists between the premature release probability normalized to a 1-minute holding time (P_r) and its estimator P_{re}:

$$\lim_{N \to \infty} P_{re} = \lim_{N \to \infty} \left(\frac{1 - \frac{R_N}{N}}{T} \right) = P_r, \text{ if such limit exists.}$$

On the other hand, for the purpose of network design, the probability of a premature release with a mean call holding time of T minutes, $P(Z, T)$, can be expressed using the formula:

$$P(Z, T) = \frac{Z}{Z + T^{-1}}$$

where

$$Z = \sum_{i=1}^{L} Z_i$$

and Z_i is the average number of failures per minute of an i component in the hypothetical connection between two users as shown in Figure A-1/E.850. The connection holding time and the time between failures for the individual components are assumed to be exponentially distributed.

T0201430-88

FIGURE A-1/E.850

Hypothetical connection to estimate the retainability of an established telephone connection

In practice, $Z \ll T^{-1}$ and therefore P_r can be approximated as follows:

$$P_r = P(Z, T)_{T=1} = \frac{Z}{Z+1} \approx \frac{P(Z, T)}{T}.$$

Also, the following relationship exists:

$$\lim_{N \to \infty} \left(1 - \frac{R_N}{N}\right) = P(Z, T).$$

ANNEX B

(to Recommendation E.850)

A method to estimate the premature release probability for an international telephone connection

In this annex, a method is described which can be used to estimate the premature release probability for an international telephone connection.

The method is based on placing end-to-end test calls, whose mean holding time is T in minutes, and observing those which are prematurely released due either to transmission or switching failures, or transmission interruptions lasting longer than ten seconds.

From the results of Annex A, it follows that the simple estimator of P_r is:

$$P_{re} = \frac{1 - \frac{R_N}{N}}{T}.$$

If it can be reasonably assumed that the occurrence or non-occurrence of a premature release for each of the test calls constitutes independent events, then the binomial sampling theory can be used to derive confidence intervals for P_r, and to determine minimum sample sizes (N).

In particular, it would be required that N be chosen such that:

$$P_r\{\,|(R_N/N) - P_r T| \leqslant e\,P_r T/100\,\} \geqslant a/100$$

where e is the estimation error in percent, and a is the confidence level in percent. Writing $P = P_r \times T$, it follows from the central limit theorem that, for large N,

$$\frac{(eNP)}{100} \,/\, \left[NP(1 - P)\right]^{1/2} \geqslant Z_a \qquad \text{(B-1)}$$

where Z_a is the root of the equation:

$$(2/\pi)^{1/2} \int_0^{Z_a} \exp(-1/2\,y^2)\,\mathrm{d}y = a/100.$$

Neglecting terms of order P^2, the inequality (B-1) becomes:

$$N \geqslant (100\ Z_a/e)^2/P \qquad \text{(B-2)}$$

In this last formula, P is generally not known. As an example, however, if we have to verify that P is in conformity with the overall objectives of typical connections (see § 3), such that P is in the order of 3×10^{-4}, then a choice of a = 90% and e = 40% would lead to $N \geqslant 56\,720$.

Similar calculations based on varying assumptions are reproduced in Figure B-1/E.850.

Based on these results, it is proposed that for an average holding time of T = 1 min, N = 60 000. For other values of T (in minutes), $N = 60\,000/T$.

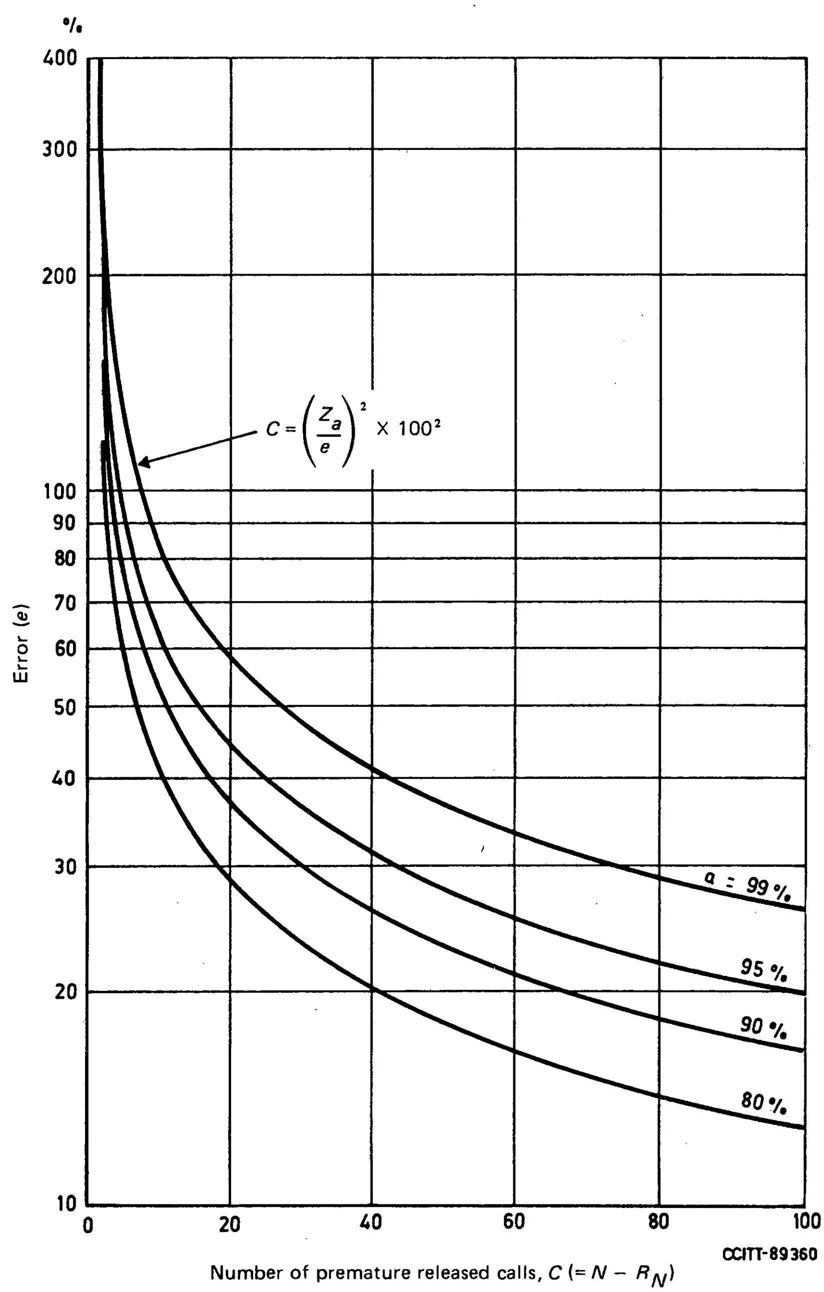


FIGURE B-1/E.850

Relative precision in estimating P_r from large samples when C/N = 0.1

Bibliography

TORTORELLA (M.): The Bell System Technical Journal, *Cutoff calls and telephone equipment reliability*, Vol. 60, No. 8, pp. 1861-1890, October 1981.

Recommendation E.855

CONNECTION INTEGRITY OBJECTIVE FOR INTERNATIONAL TELEPHONE SERVICE

Introduction

This Recommendation is one of a set of closely related Recommendations comprising Recommendations E.810, E.830, E.845, E.850 and E.855 concerned with the accessibility, retainability and integrity of telecommunication services, specially telephone services.

The CCITT,

considering

(a) that users of the telephone service can perceive the speech loss due to transmission interruptions with durations shorter than 10 seconds;

Note – Transmission interruptions with durations longer than or equal to 10 seconds in a conversation phase are not tolerable by telephone users (Annex A). Such transmission interruptions are considered as a premature release of the connection as defined in Recommendation E.850.

(b) that speech loss causing transmission interruptions are caused by a change beyond given limits for a given period of time in one or more parameters, e.g. power level, noise level, signal-to-noise ratio, bit error ratio, etc.;

(c) that the objective should take into account the expectations of the users for quality of voice communications as well as the capabilities of current technologies;

(d) that the objective should take into account the concerns of network planners and system planners, provide useful guidance to each and that it can be used by Administrations in a consistent way to measure transmission interruptions;

(e) that the objective should be in conformity with other Recommendations;

(f) the definition of *interruption* as given in Recommendation E.800,

recommends

1 Definitions

1.1 connection integrity for telephone service

The degree to which an established telephone connection is offered without excessive transmission interruptions.

1.2 mean time between interruptions (MTBI)

The expectation of the time between interruptions. The time between interruptions is the time duration between the end of one interruption and the beginning of the next.

1.3 mean interruption duration (MID)

The expectation of interruption duration.

1.4 transmission interruption

Temporary inability of the user-to-user transmission path to be provided persisting for less than 10 seconds (maximum duration) and more than another given time duration (or minimum duration) characterized by a reduction below a certain threshold in received signal power level. The minimum duration of transmission interruption and the minimum power threshold are for further study. Transmission interruptions caused by changes beyond certain thresholds of other parameters essential to connection integrity e.g. noise level, signal-to-distortion ratio, are for further study.

2 A measure to quantify telephone connection integrity performance

The measure to be used shall be the complement of connection integrity, namely the probability of speech loss, P_i, which is tolerable to telephone users due to transmission interruptions with durations shorter than 10 s. The estimator of the speech loss probability, P_{ie}, is the ratio of accumulated transmission interruption duration to the total observation period of time.

$$P_{ie} = \sum_{i=1}^{N} TD_i / T$$

where T is the observation time and TD_i is the time duration of the ith transmission interruption of N transmission interruptions measured during T (see Annex B).

Note – There are two major parameters: time between interruptions (or frequency) and duration to specify characteristics of transmission interruptions. Those parameters should be easy to observe from the practical point of view. Actually it seems very difficult to measure very short duration of transmission interruptions in analogue networks and to separate interruptions from burst errors in digital networks.

3 Overall objective for speech loss probability

The provisional objective for P_i shall be such that the performance is better than the value given below:

$$P_i = x \text{ (to be defined with further study)}$$

Note – A percentage of speech loss of less than 0.5% due to transmission interruptions with durations shorter than 10 s (Annex C) in a conversation phase is assumed to be tolerable to telephone users.

4 Allocation of the overall objective

From a practical point of view, instead of P_i, the value $\frac{P_i}{1 - P_i}$ should be allocated to various network components.

The method to allocate this value is for further study.

ANNEX A

(to Recommendation E.855)

Tolerability of telephone user to transmission interruptions lasting several seconds or more

A.1 *Measure*

The time interval between the start of transmission interruption occurring in the middle of a conversation and the abandoning of the disturbed call either by a calling or a called party is used as a measure to assess or evaluate the tolerability of telephone users.

A.2 *Measuring method*

Fifty intra-office calls were selected at random and deliberately interrupted by callers soon after the calls had been established, and time intervals between the start of transmission interruption and the release of calls by the called party were measured.

A.3 *Results of tests*

The distribution of the durations of interruptions which forced the telephone users to give up their established calls is depicted in Figure A-1/E.855. The distribution curve is well approximated by an exponential distribution function with the mean value of 17.26 sec.

The figure shows that 50% of users released the established calls when the interruption lasted longer than 11.96 sec.

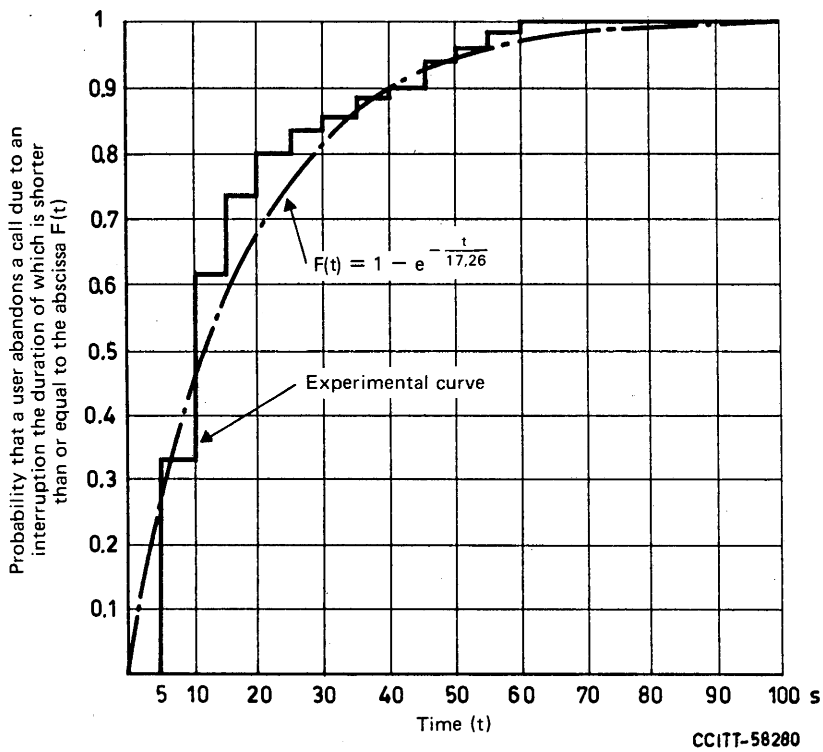


FIGURE A-1/E.855

Tolerability to transmission interruptions

ANNEX B

(to Recommendation E.855)

Relationship between speech loss probability and its estimation

The following relationship exists between the speech loss probability (P_i) and its estimator (P_{ie}):

$$\lim_{T \to \infty} P_{ie} = \lim_{T \to \infty} \sum_{k=1}^{K} \frac{TD_k}{T} = P_i$$

if such limit exists and where T is the observation period of time and TD_k is the duration of the kth transmission interruption of K transmission interruptions over T.

It should be noted that there is also the following relationship:

$$P_i = \frac{\rho}{1 + \rho}, \ \rho = \sum_{i=1}^{L} \frac{MID_i}{MTBI_i}$$

where, MID_i is the mean time duration of transmission interruption caused by the ith component of a telephone connection and $MTBI_i$ is the mean time between interruptions caused by the ith component of the connection, under the assumption that the transmission interruption duration and the time between transmission interruptions are exponentially distributed (see also Figure B-1/E.855).

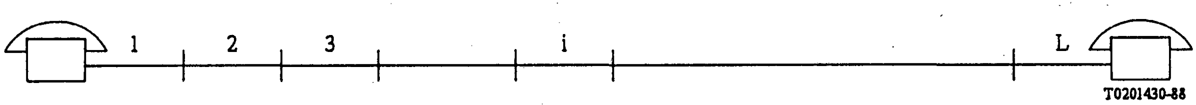


FIGURE B-1/E.855

Hypothetical connection to estimate the speech loss probability of an established telephone connection

ANNEX C

(to Recommendation E.855)

Quality of speech impaired by short interruptions

C.1 *Measure*

The subjective opinion is used as a measure to assess or evaluate speech quality impaired by short interruptions with durations shorter than 1 s.

C.2 *Measuring method*

Recommendation P.77 was applied for this subjective evaluation with five grade opinion scores (Excellent = 4, Good = 3, Satisfactory = 2, Poor = 1 and Unacceptable = 0). The test procedure was comprised of a 40-second text tape recorded in Japanese spoken by a female, which was listened to by 20 test subjects through indoor test circuits with a transmission interruption generator.

C.3 *Test results*

The relationship between frequency and duration of transmission interruptions for a given Mean Opinion Score (MOS) is depicted in Figure C-1/E.855.

The dotted line in this figure shows the locus of "frequency × duration = 0.5%" which is considered as a permissible limit of the freeze out rate or the percentage of speech loss for designing Digital Speech Interpolation (DSI) and Time Assignment Speech Interpolation (TASI) equipment.

Note – The product of frequency and duration of the short interruption is identical to $P[= MID/(MTBI + MID)]$ in Annex B.

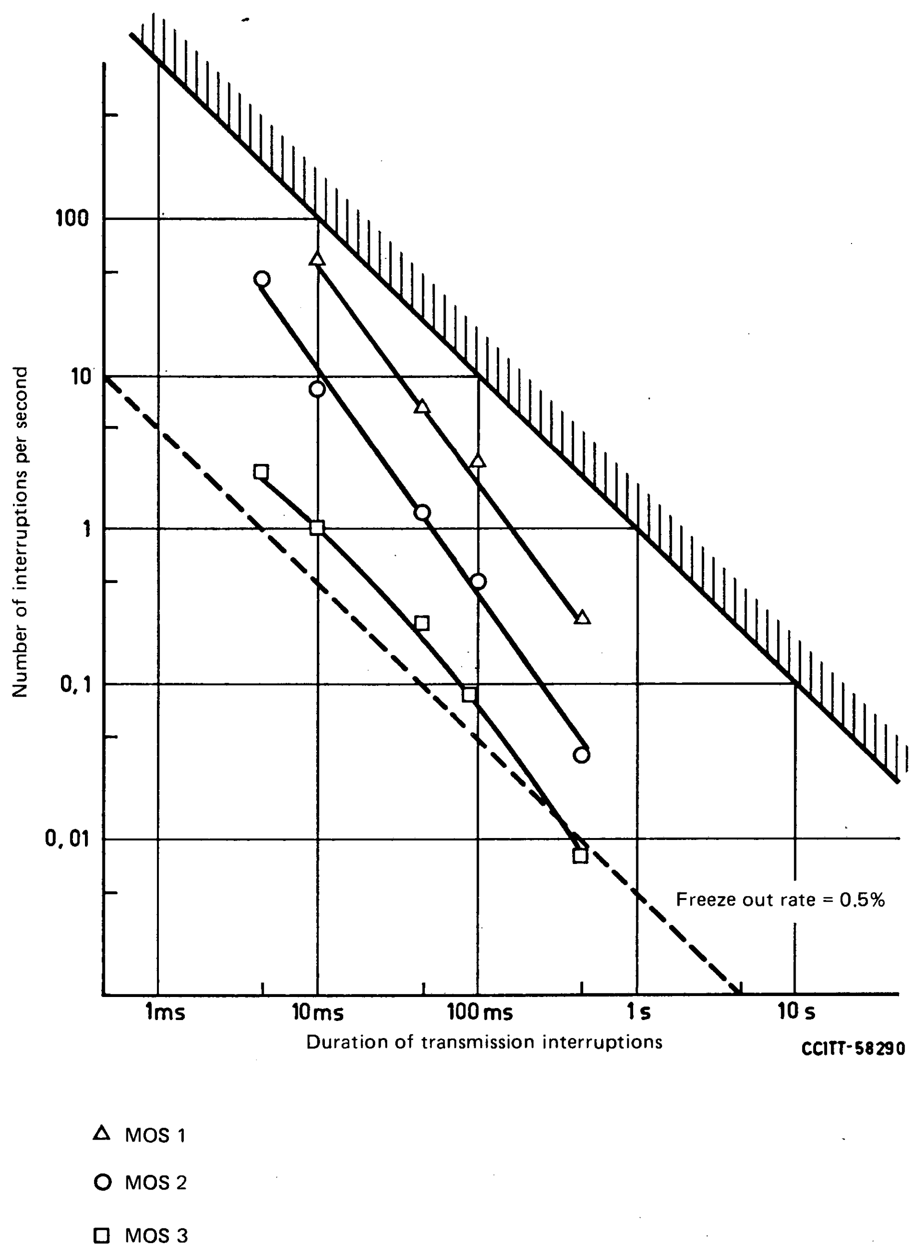


FIGURE C-1/E.855

Speech impairment due to transmission interruptions

SECTION 4

USE OF QUALITY OF SERVICE OBJECTIVES FOR PLANNING OF TELECOMMUNICATION NETWORKS

Recommendation E.862

DEPENDABILITY PLANNING OF TELECOMMUNICATION NETWORKS

Introduction

This Recommendation is concerned with models and methods for dependability planning, operation and maintenance of telecommunication networks, and the application of these methods to the various services in the international network.

The CCITT,

considering

(a) that economy is often an important aspect of dependability planning;

(b) that the ability of achieving a certain level of dependability differs between network providers;

(c) that network providers often operate in a competitive environment;

(d) that Recommendations E.845, E.850 and E.855 establish objectives for serveability performance;

(e) that objectives for dependability performance are deducible from Recommendations Q.504, Q.514, and X.134 to X.140;

(f) that these objectives have been established in an intuitive manner rather than based on analysis of user needs;

(g) that there exists no unambiguous way of implementing these objectives in planning;

(h) that there is a need of establishing a method for dimensioning and allocating dependability in the telecommunication network;

(i) that terms and definitions relevant to concepts used for dependability may be found in Recommendation E.800,

recommends

that the procedures defined in this Recommendation shall be used by Administrations to plan, design, operate and maintain their networks.

1 General

Dependability planning may be accomplished by using essentially two different methods.

Intuitive method

The level of dependability is determined by making a synthesis of objectives and procedures presently used. It is a pragmatic method in absence of an analytical method or in the case when necessary data for a thorough analysis is not available.

This method reflects the present status, but is inconsistent in achieving what Administrations actually want to attain: the most economic level of dependability taking into account customer needs and inconvenience.

Analytical method

The analytical method is based on principles defining the object of dependability planning. The principles are realized through a quantitative model. The level of dependability is deduced by applying the model, taking into account all relevant factors in each planning case.

- Basic principle: The main object of dependability planning is to find a balance between the customers' needs for dependability and their demand for low costs.
- Model: Fault consequences are expressed in terms of money and are included as additional cost factors in planning and cost-optimization. The cost factor reflects the customers' experience of faults in the network, quantified in terms of money, as well as the Administration's costs for lost traffic revenue and corrective maintenance.
- Application: The Administration is provided with a method to integrate dependability as a natural part of planning, taking local information from the actual planning case into account. This method enables the preparation of simplified planning rules.

The application of the analytical method gives, economically, the best-balanced level of dependability, seen from the customer's point of view. This reduces the risk of customer complaints and loss of business to competitors as well as the risk of unnecessary investments. It is therefore considered to be the best general way of planning dependability for the Administration, as well as for the customers.

Recommendations for operational dependability objectives are needed in order to discover impairments and to check and compare dependability performance in the national and international network. Experience from the application of the analytical method may give reason to revise existing Recommendations.

2 Generic measures for dependability planning

The dependability is described by measures defining the availability performance, the reliability performance and the maintainability performance of the network and its constituent parts as well as the maintenance support performance (for the maintenance of the network). The recommended measures are:

Availability performance

Mean accumulated down time

Reliability performance

Mean failure intensity

Maintainability performance

Mean undetected fault time

Mean time to restoration

Mean active repair time

Maintenance support performance

Mean administrative delay

Mean logistic delay

Note – The definitions of these measures are given in Recommendation E.800 and Supplement No. 6.

3 Characteristics of network faults

The faults occurring in the telecommunication network are characterized mainly by their impact on the service provided by the network, i.e. by the traffic disturbance they cause. Important measures determining the traffic disturbance due to a fault are:

Duration of the fault (mean down time), T in hours (h)

Mean traffic intensity affected by the fault, A in Erlangs (E)

Mean probability of congestion during the fault, P

The seriousness of a fault also depends on how the customers experience the fault, and on the Administration's loss of revenue. In order to express this fact, the value of a unit of traffic volume (Eh) disturbed by the fault is quantified in economic terms.

Measure: the economic valuation of affected traffic volume is c (monetary units per Eh).

A number of factors may influence this variable such as:

- the category of customers and services affected,
- the degree of congestion or transmission disturbance during the fault,
- the duration of the fault,
- the accessibility to alternative communication means for the affected customers,
- time of day, week or year when the fault is in effect,
- how often faults have occurred in the past.

Additionally, the Administration's costs for corrective maintenance also contribute to the assessment of fault consequences.

Measure: the maintenance cost per fault is c_m (monetary units per fault).

4 Planning for economic optimum

4.1 *Economic dimensioning and allocation method*

Mathematically expressed, the main principle of dependability planning is to find actions that minimize the total cost of the network:

$$\min \left\{ C_I + C_m \cdot d + C_t \cdot d + \dots \right\}$$

where

C_I is the investment costs to achieve a certain degree of dependability,

C_m is the expected annual costs for corrective maintenance,

C_t is the expected annual traffic disturbance costs,

d is the discount factor for calculating the present value of the annual cost over the lifetime of the investment.

C_t reflects the annoyance caused by faults and should be regarded as the basic service parameter which dimensions and allocates dependability in the network under given conditions.

An action is optimal if the following two conditions are met:

1) The benefit of the action (e.g. lower traffic disturbance cost) is larger than the cost, i.e. the action is profitable.
2) The action is the best in the sense that the ratio benefit/cost is maximal. There are no alternative actions that give a higher profit.

The method points out a profit seen from the customer's point of view, i.e. the actions are not necessarily profitable for the Administration in the short run. Rates and charges might therefore have to be increased to finance the actions. However, satisfying the customer's needs is recommended as the generally most profitable policy for the Administration in the long run.

This method is applicable for planning all parts of the national and international network and for dimensioning the dependability of network components and the level of the maintenance support. It may be used in short term planning as well as in long term optimization and strategic planning.

The method does not become out of date with technological advances, changes in cost structure etc. Dependability is converted to one clear-cut measure (money) which makes it easier to evaluate actions to promote dependability and to compare and choose between different alternatives.

4.2 *A model for traffic disturbance costs*

The annual traffic disturbance cost is obtained by multiplying the disturbed traffic volume (lost, delayed or affected by transmission impairments) by the monetary valuation of disturbed traffic volume c and the mean failure intensity z which gives:

$$C_t = PAzTc$$

where

- T is the duration of the state of increased congestion or transmission disturbance due to the fault, mainly the down time. Congestion due to traffic overload after the fault has been repaired might however also have to be included.
- A is the intensity of offered traffic.
- P is the portion of the offered traffic volume over the time T, delayed or lost.
- z is the mean failure intensity.
- c is the monetary valuation of disturbed traffic volume. c may be dependent on any number of factors, i.e. $c = c(P, T, A, \dots)$.

Assuming traffic variations, $A(t)$, and consequently variations of congestion, $P[A(t)] = P(t)$, then A and P are calculated as follows:

$$P = \frac{\int_0^T A(t)\,P(t)\,dt}{\int_0^T A(t)\,dt} \qquad A = \frac{1}{T}\int_0^T A(t)\,dt$$

Normally it is not possible to predict the instant of time when a failure will occur. In this case A is a long time average incorporating yearly variations and long time trends. P is calculated by using an average traffic profile. Recommendations E.506, E.510 and E.520 to E.523 deal with methods for traffic calculations.

4.3 *Economic assessment of disturbed traffic volume, c*

The factor c reflects the level of ambition of an Administration or Operating Company in dependability planning. A high valuation of c will give a high level of dependability and vice versa. The values used by an Administration are related to the society's dependence on telecommunications which in turn might be dependent on standard of living, national economy, price level, etc. The establishment of c on the national level is therefore a national matter.

However, it is recommended that c should reflect the combined experience of the Administration and the customer, i.e. it should consist of:

1) the Administration's loss of revenue due to traffic not recurring after the fault,
2) an assessment of the average customer's economic loss due to a unit of traffic volume (Eh) being affected by a fault,
3) a symbolic price tag reflecting the annoyance experienced by the average customer.

The sum of 2) and 3) should reflect the price the average customer is willing to pay to avoid one Erlang-hour of offered traffic, delayed or lost due to a fault. The result will then be a level of dependability the customers are satisfied with and prepared to pay for.

Administrations are recommended to make their own investigations among their customers in order to determine the values to be used for planning. Annex B gives an example of such an investigation.

If this is not possible, rough estimates may be obtained from information about actions taken in the present network. The cost of an action is compared to the amount of traffic it saves. Actions that intuitively are regarded as reasonable give a lower limit of c, actions that obviously are unreasonable give an upper limit. The values derived this way are then used in optimization under the assumption that they are valid also for planning the future network.

If c is not possible to estimate at all, the method may still be used to find an optimum allocation of a given amount of resources. The level of dependability attained in this case does however not necessarily satisfy the customers.

4.4 *Planning procedure*

Traffic disturbance costs are included as additional cost-factors in economical calculations for planning, thus integrating dependability as a natural part of planning.

The procedure of dependability planning is performed in four steps:

Step 1: Plan a network attaining functional and capacity requirements.

The starting point is a network planned and dimensioned in order to comply with the functional and capacity requirements, but without special consideration of dependability (zero-alternative). The second step is to identify what changes may be necessary to promote dependability.

Step 2: Search for actions to promote dependability.

There is a need for actions to promote dependability if traffic disturbance costs are high or if the actions can be taken at a low cost. A non-exhaustive list from which actions could be identified is given below:

- Protection of equipment in order to prevent failures
- Choice of reliable and maintainable equipment
- Modernization and reinvestment of worn out equipment
- Redundancy
- Overdimensioning
- Increase in maintenance support
- Network management actions to reduce fault effects.

Step 3: Analyse the actions.

Express improvements in terms of changes in traffic disturbance and maintenance costs ($\Delta C_t + \Delta C_m$) for each action. It is only necessary to calculate costs that differ between the alternatives. Annex A gives examples of dependability models for network design, maintenance support planning and for determining requirements for network components.

Compare $\Delta C_t + \Delta C_m$ to the increased investment cost (ΔC_I) for each action, e.g. by the present value method.

Choose the best set of actions, i.e. which gives the lowest total cost.

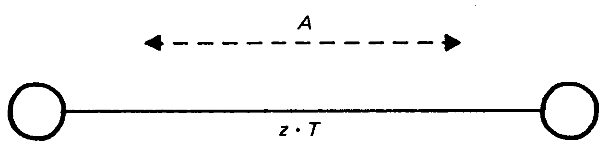
Step 4: Check that minimum requirements are complied with.

A minimum service level may be stipulated by governmental regulations, by CCITT Recommendations, for commercial or for other reasons. The establishment of any minimum requirements on the national level is a national matter. For planning of the international network the Administration is recommended to check if dependability objectives deducible from existing CCITT Recommendations are met. If not, the reasons for non-compliance should be examined more closely. If it is justified, the level of dependability should be adjusted.

4.4.1 *Numerical example based on the above*

Step 1: Network planned without special consideration of dependability.

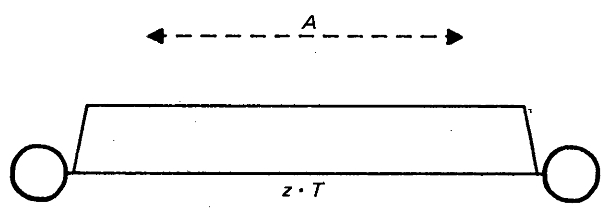
The network studied is the trunk between two exchanges.



T0201480-88

Step 2: Search for actions to promote dependability.

The action considered is to introduce a physically redundant cable. It is assumed to be dimensioned to carry the whole traffic load, i.e. a single failure will not disturb the traffic.



T0201490-88

Step 3: Analyse the action.

Assumptions

Failure intensity	z	= 0.1 failures/year
Mean down time	T	= 24 h
Mean offered traffic	A	= 100 E
Congestion	P P	= 1 (without redundancy) = 0 (with redundancy)
Monetary valuation of disturbed traffic volume	c	= 400 monetary units/Eh
Discount factor (lifetime 25 years, interest 5% per year)	d	= 14
Maintenance cost per failure	c_m	= 1000 monetary units/failure
Cost of redundant cable	C_I	= 400 000 monetary units

Calculations

Traffic disturbance costs for network without redundancy:

$$C_t = P \cdot A \cdot z \cdot T \cdot c = (1)\ (100)\ (0.1)\ (24)\ (400) = 96\,000 \text{ per year}$$

$$\text{Present value } C_t d = (96\,000)(14) = 1\,344\,000$$

Traffic disturbance costs for network with redundancy (the possibility of simultaneous faults is negligible):

$$C_t = 0$$

Change in traffic disturbance costs:

$$\Delta C_t d = 0 - 1\,344\,000 = -1\,344\,000$$

Maintenance costs without redundancy:

$$C_m = zc_m = (0.1)(1000) = 100 \text{ per year}$$
$$\text{Present value } C_m d = (100)(14) = 1400$$

Maintenance costs with redundancy:

$$C_m = 2zc_m = (2)(0.1)(1000) = 200 \text{ per year}$$
$$\text{Present value } C_m d = (200)(14) = 2800$$

Change in maintenance costs:

$$\Delta C_m d = 2800 - 1400 = 1400$$

Cost reduction:

$$\Delta C_t d + \Delta C_m d = -1\,344\,000 + 1400 = -1\,342\,600$$

Change in total cost:

$$\Delta C_I + \Delta C_m d + \Delta C_t d = 400\,000 - 1\,342\,600 = -942\,600$$

Conclusion

Since $\Delta C_I + \Delta C_m d + \Delta C_t d < 0$, the action is profitable. Whether or not it is optimal depends on whether there are alternative actions that are more profitable.

Step 4: Check minimum requirements

Any additional actions to meet governmental requirements (for defence reasons, emergency, etc.) should be taken.

5 Applications to the international network

5.1 *Value of c for international traffic (for further study)*

In order to dimension and allocate dependability to different parts of the international network a uniform way of evaluating affected traffic should be established. It is recommended that the following values (c_i) be used as a guide in the planning of the international network

$$c_i = x_i \, SDR : s/Eh \qquad \text{(values to be determined)}$$

The values refer to a particular reference year. Price increase due to inflation, society's increasing dependence on telecommunication etc., should be taken into account.

5.2 *Planning recommendations (for further study)*

When values of c have been established, it is possible to make economic dependability analyses of the international network. These studies may be done in a similar manner and using partly the same data as for cost studies of charging and accounting.

The object of the studies is to arrive at planning recommendations, e.g. for the amount of redundancy, maintenance support, etc., in different parts of the international network.

5.3 *Operational objectives for dependability (for further study)*

The result of the economical dependability analysis of the international network is presented in terms of reliability, maintainability and maintenance support performances of different parts of the network. This will help Administrations monitoring and checking their networks to discover impairments, misplanning, etc.

ANNEX A

(to Recommendation E.862)

Simplified models for dependability planning

A.1 *General*

The object of this annex is to show simple examples of how different models of dependability may be used to calculate traffic disturbance costs and how the calculations can be used in planning. A list of actions is given in § 4.4. The applications may be divided into:

- Network planning (§§ A.2 and A.3)
- Dimensioning dependability of network components (§ A.4)
- Maintenance support planning (§ A.5).

A.2 *Example: Redundancy*

The traffic disturbance cost of a redundancy consisting of two independent items as shown in Figure A-1/E.862 is:

$$C_t = P_1 z_1 T_1 Ac(P_1) + P_2 z_2 T_2 Ac(P_2) + z_1 z_2 T_1 T_2 Ac(1)/8760$$

where

P_1 is the average congestion when item 1 is faulty,

P_2 is the average congestion when item 2 is faulty.

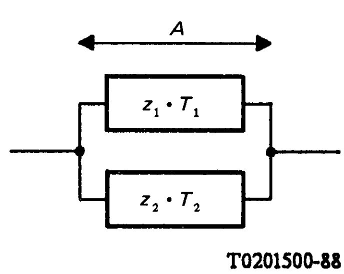


FIGURE A-1/E.862

A simple case is when the two items are identical and each can carry the whole traffic load (see Figure A-2/E.862), then:

$$C_t = z^2 T^2 Ac(1)/8760.$$

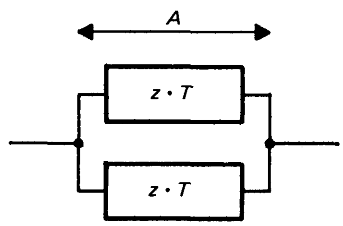


FIGURE A-2/E.862

By installing a redundant item, the traffic disturbance costs are reduced by

$$\Delta C_t = zTAc(1) - z^2 T^2 Ac(1)/8760.$$

The second term is often negligible, thus ΔC_t may be approximated by $\Delta C_t = zTAc(1)$.

A.3 *Example: Optimal dimensioning for diversified routes*

The problem is to determine the optimal number of channels, N_1 and N_2 respectively, for which the two redundant routes should be dimensioned, see Figure A-3/E.862.

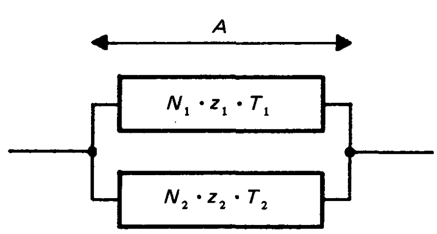


FIGURE A-3/E.862

Denote C_N to be the cost per channel. The optimal allocation of channels each way is found by solving

$$\min_{N_1,\, N_2} \left\{ (N_1 \cdot C_{N1} + N_2 \cdot C_{N2}) + (P_1 \cdot A \cdot z_1 \cdot T_1 \cdot C(P_1) + P_2 \cdot A \cdot z_2 \cdot T_2 \cdot C(P_2)) \cdot d \right\}$$

This implies an overdimensioning in the fault free condition. The benefit of this is not included in this formula. The effect of simultaneous faults does not influence the optimization.

A.4 *Example: Optimal testing time*

Assume that the failure intensity $z(t)$ after a certain operation time (t) is given by

$$z(t) = z_0 + z\mathrm{e}^{-bt}$$

where

$z_0 + z$ is the failure intensity at $t = 0$,

z_0 is the constant failure intensity after the early failure period,

b is the factor determining the decrease in failure intensity during the early failure period.

By testing, faults may be corrected before causing traffic disturbance and maintenance costs. Assume that:

$c_m + ATc$ are the maintenance and traffic disturbance costs per fault,

C is the cost per year of testing.

The optimal testing time (t') is found by solving

$$\min_t \left\{tC + \frac{z}{b}\,\mathrm{e}^{-bt}\,(c_m + ATc)\right\}$$

where

$\frac{z}{b}\,\mathrm{e}^{-bt}$ is the additional number of faults occurring in operation as a function of the testing time.

Optimal test time: $t' = \frac{1}{b} \ln \frac{z(c_m + ATc)}{C}$.

A.5 *Example: Optimal number of maintenance units*

Mean delay $w(N)$ as a function of the number of maintenance men (N) may in some cases be mathematically expressed by using queuing theory. The simplest case is if the times between failures and repair times are exponentially distributed (an $M/M/N$ queue model). $w(N)$ is obtained by calculating:

$$w(N) = \left[\frac{(z/\mu)^N \cdot \mu}{(N-1)\,!\,(N\mu - z)^2}\right] \Big/ \left[\sum_{k=0}^{N-1} \frac{1}{k!}\left(\frac{z}{\mu}\right)^k + \frac{1}{N!}\left(\frac{z}{\mu}\right)^N\left(\frac{N}{N\mu - z}\right)\right]$$

where

N is the number of maintenance units,

z is the intensity of failures,

$w(N)$ is the mean delay as a function of N,

A is the affected traffic intensity,

c is the valuation of affected traffic volume,

μ is the repair rate.

The model may be refined by taking into account classes of priority. It is also possible to let faults of a higher priority interrupt assignments with a lower priority.

If C_N is the annual cost per maintenance unit, the optimal number of maintenance units is obtained by solving:

$$\min_N \left\{NC_N + zw(N)Ac\right\}$$

ANNEX B

(to Recommendation E.862)

Example of an investigation to assess the monetary valuation of disturbed traffic volume, *c*

B.1 The aim is to arrive at cost data to assess *c*. Different customer groups and their monetary valuation of total and partial failures with respect to typical traffic relations and different services is studied. Investigations are carried out among residential and business customers based on the following assumptions:

a) The customers are affected by telecommunication interruptions in mainly two ways: in terms of annoyance and in terms of direct costs.

b) For residential customers, annoyance is likely to predominate. For business customers, the direct cost may be important.

c) Both costs and annoyance increase by the duration of the interruptions and the amount of traffic disturbed.

d) As a natural consequence of the great variations in dependence on telecommunications there is a great variation of costs and annoyance.

e) Residential customers are not able to quantify their annoyance in monetary terms. Faults on the home telephone mostly result in irritation, and not in direct costs (except in the case of long-time faults).

B.2 *Complete faults*

B.2.1 *Business traffic*

Companies chosen at random are asked to answer the following question: "What is the estimated approximative cost of a total interruption of the telephone or data service in connection with down times of 5 minutes, 1 hour, 4 hours, 8 hours, 24 hours and 3 days?"

Companies with experience of a specific fault are asked the question: "What was the estimated cost of the fault just experienced?"

An estimate of the affected traffic intensity in connection with total interruptions can be made on the basis of the number of exchange lines and the number of data terminals for communication of each company, together with information on how trunks are dimensioned and measurements on the calling intensity of various customer classes.

On the basis of a stated cost, *c* is estimated according to the formula:

$$c = \frac{\text{(cost stated by the customer)}}{\text{(mean traffic intensity) (down time)}}$$

Average values of *c* for telephony and data traffic are calculated for different trades by means of a market profile (distribution of workplaces by trade).

B.2.2 *Residential customers*

Group discussions on interruptions can be held in order to arrive at a reasonable valuation. If there is little willingness to pay for increased dependability a relatively low value of *c* is assigned.

B.3 *Partial faults*

A partial interruption of a traffic relation results in costs for the customer mainly in the form of delays to commerce. By using a calculated hourly salary this cost is estimated for business customers. On the basis of information about the amount of business and household traffic, an average value of *c* for traffic disturbed by partial faults is obtained.

B.4 *Results*

Table B-1/E.862 gives a few examples of figures derived by the Swedish Administration. The figures have been used in various planning cases. The Administration's loss of revenue is included in these figures. The cost figures and exchange rate relate to 1986-01-01 [1 SEK (Swedish Krona) ≈ 0.1 USD (US Dollar)].

TABLE B-1/E.862

Economic assessment of prevented communication (c)		
Field of application	Class of failure	
	Complete fault ($P = 1$)	Partial fault ($P < 0.5$)
Business customers with a large portion of data traffic	1000 SEK/Eh	250 SEK/Eh
Used in the long distance network	400 SEK/Eh	100 SEK/Eh
Customers in a sparsely populated area. High cost for alternative communication	200 SEK/Eh	50 SEK/Eh
An average value for areas with mostly residential customers	100 SEK/Eh	25 SEK/Eh
Residential area where it is easy to reach essential services. Low costs for alternative communication	30 SEK/Eh	10 SEK/Eh

SECTION 5

FIELD DATA COLLECTION, ANALYSES AND EVALUATION ON THE PERFORMANCE OF EQUIPMENT, NETWORKS AND SERVICES

Recommendation E.880

FIELD DATA COLLECTION AND EVALUATION ON THE PERFORMANCE OF EQUIPMENT, NETWORKS AND SERVICES

1 Introduction

This Recommendation provides guidelines for the collection of field data relating to dependability. It covers general aspects with an overview of sources, measures and information that may be involved when collecting field data. It is anticipated that specific practical needs of operation, maintenance and planning staff, in applying these guidelines, will be dealt with in a handbook under preparation.

The Recommendation emphasizes that meaningful data must include the data on successes (operation without failures) as well as data on failures and faults. In other words, this Recommendation is not intended to be only a failure reporting guideline.

It is applicable, without any restriction, to different items ranging from components to systems and networks (including hardware, software and people).

Terms and definitions used are in line with Recommendation E.800.

2 Scope

It is the intention of this Recommendation to provide guidelines for setting up data collection and reporting schemes which can be applied either during monitoring of samples of items or on a more widespread basis on almost all items (of the same type) by large operational and maintenance organizations.

It is considered that, if such guidelines are followed, accuracy and completeness of reporting are enforced and the quality of the monitored items and their parts can be improved on a medium- to long-term basis. Moreover, such an effort will facilitate the interchange of information between user and providers.

No recommendations are made on how to organize maintenance support. It is nevertheless understood that some items are repaired on site, others only replaced on site and possibly repaired at centralized facilities. Field data may be obtained at each of those stages.

In order to obtain maximum efficiency from the collection of data, it is suggested that the programmes of reporting, analysis and dissemination of results be closely co-ordinated.

The items considered may either have been designed, manufactured, or installed and may be operated by the same organization or by different organizations. This Recommendation applies to all possible cases of provider-user relations.

3 The need for data collection

Any data collection scheme must aim to provide the information required to enable the correct decisions to be taken in order to reach specified objectives; these objectives should be well defined and documented at the outset.

The specific objectives of the field data collection and presentation are as follows:

a) to provide for a survey of the actual performance level of the items monitored for information to management, operation and planning, maintenance support, training of personnel, etc.;

b) to indicate a possible need for the improvement of:
   - items already installed and in operation, or
   - further items to be delivered;

c) to compare the specified or predicted characteristics of the item(s) with the actual field performance;

d) to improve future designs;

e) to improve predictions (data bases and procedures);

f) to inform the provider about the performance of items on a regular or on a single occasion basis;

g) to have a common reporting basis.

4 Sources and means of data collection

In the following, the various information sources are described and the methods for systematically collecting information are outlined.

4.1 *Sources of data*

The following sources of data may generally be available:
- maintenance activities;
- repair activities (on site, repair and/or complaint centre);
- performance observation activities (e.g. anomaly reports, traffic measurements);
- existing information (e.g. stocklist, installation list, modifications, a regularly updated data base for configuration control purposes).

4.2 *Means of collecting data*

It is not intended to recommend any particular format for the recording medium (e.g. paper based or computer data base), however it should be recognized that early consideration of the format is necessary and important in setting up an effective data collection scheme and also aids subsequent successful processing.

Frequently the recording of data will be by manual means but automated and interactive data collection systems may be also considered. The advantages to be gained from holding data in a form suitable for processing by an electronic data processing system include easy and accurate updating of information and the possibility of performing new extended analyses.

Data may be collected by one or several of the following reporting means.

4.2.1 *Operation reporting*

Data reporting should be supported by information on the use of the items. Where systems are in operation for the reporting of all failures, it is necessary to collect data on the use of the whole population of items (the total number of similar items under observation).

4.2.2 *Failure reporting*

At any level, failure reporting is dependent on the fault coverage test resources used at the considered level: cases such as "fault not found" or "right when tested" should be clearly mentioned.

Failure reporting should cover all failures that have been observed. They should also contain sufficient information to identify failures. Failures considered to be attributable to any maintenance action should be so noted.

The failure reporting should be sufficiently comprehensive to cover the requirements of detailed investigation of an individual failure and the resulting fault. Where economic reasons or lack of resources make it undesirable to collect all of the failure data indicated, it may be desirable to agree upon a shortened form of report which can be used to collect limited data on all relevant failures, with an option to call for the full report in specific cases.

4.2.3 *Maintenance reporting*

The maintenance report should contain all information relevant to the manual or automatic action taken to restore the item.

When there is need to distinguish between corrective maintenance and preventive maintenance reporting, if no replacements or repairs are made, the action can be classified as a preventive maintenance report. If a preventive maintenance action results in a replacement or repair, the report may be treated as a corrective maintenance report even though the item has in fact not failed in operation.

4.3 *Storage, updating and checking procedures*

Independently of the structure chosen for the data storage, data should be checked at the time of input so as to ensure validity.

It is evident that every data bank needs an in-depth study appropriate to its specific requirements, in order to define the most suitable method of data checking, error correction, and updating.

5 **List of dependability measures**

The selection of the data to be collected is very dependent on the kind of performance measures to be evaluated/estimated.

Field data reporting may have to be limited by economic necessity to the minimum necessary to meet the requirement, whilst recognizing that collection systems should be capable of future expansion.

It is likely that certain data may be needed for more than one purpose, and careful consideration can therefore result in the most cost-effective data collection scheme.

The dependability measures that might be taken into consideration are listed as follows.

5.1 *Reliability performance*

Failure rate
Failure intensity
Replacement intensity
Mean operating time between failures
[.][1] Up time.

5.2 *Maintainability performance*

5.2.1 *Time related performance*

[.][1] Down time
[.][1] Technical delay
[.][1] Fault localization time
[.][1] Fault correction time
[.][1] Restart time
[.][1] Checkout time
[.][1] Repair time
[.][1] Active corrective maintenance time.

[1] [.] indicates according to specific applications a mean value or a fractile.

5.2.2 *Probabilities*

Probability of fault coverage
Probability of false alarm
Probability of fault nondetection
Probability of alarm detection
Probability of a failure being localized within a given number of replaceable units.

5.3 *Maintenance support performance*

5.3.1 *Time related performance*

[.][2] Logistic time
[.][2] Administrative delay.

5.3.2 *Probabilities*

Spare parts shortage probability
Test resource shortage probability
Human resource shortage probability.

5.4 *Availability performance*

Steady state availability
[.][2] Accumulated down time.

6 Data required

Consideration of the foregoing objectives defines the need for a system which provides for the collection of documented data covering:

a) the identity of items or population of items under observation;
b) operational conditions;
c) maintenance support conditions;
d) performance monitoring.

For each individual item, sufficient information has to be recorded to clearly identify the item itself and its operating environment.

Depending on the item under consideration (e.g. equipment, printed circuit board, component, personnel), and on the depth and kind of analysis of collected data, the necessary item identification data shall be used, on a case by case basis.

The item identification should also allow the analysis of the relationships between the items for which data is collected.

In relation to the particular analysis to be done, some items may be considered as equivalent, therefore separate small items need not to be collected in such cases.

The following information may be needed and could be collected or will be available from existing sources:

- type of item
- manufacture/provider
- item configuration
- individual No. or serial No.
- date of manufacture
- supplier
- delivery date

[2] [.] indicates according to specific applications a mean value or a fractile.

- installer (company)
- installation date
- customer (name)
- site (geographic)
- system.

Consideration should be given to possible limitations due to non-completeness of collected data or possible difficulties in data collection or particular assumptions made for the collection itself.

The choice of the kinds of data to be collected and the design of the related collection procedure depend on many factors, some of which are:

- the required end-result;
- the diversity of components or systems;
- the duration of the data collection project;
- the data handling method (manual or computer based);
- a sufficient knowledge of the capability to collect the required quantity of information and the accessibility to data to be gathered.

6.1 *Number of items to be considered*

The number of items to be considered depends mainly on the characteristics to be dealt with, the statistical aspect of the evaluation to be made and the cost involved.

6.2 *Information on items being considered*

6.2.1 *Operating conditions*

6.2.1.1 *Environment classes*

a) Fixed (outdoors, indoors, underground, undersea, off-shore, etc.);

b) Portable (item specially built for easy transportation by one man only);

c) Mobile (in motor vehicle, in ship, in aircraft);

d) Other (specify).

6.2.1.2 *Specific environment data*

a) Climatic conditions

- weather-protected,
- not weather protected,
- air temperature,
- air pressure,
- humidity;

b) Electrical environment (EMC);

c) Mechanical conditions (vibration, shocks, bumps);

d) Mechanically active substances (sand, dust, etc.);

e) Chemically active substances;

f) Biological conditions.

6.2.1.3 *Mode of operation*

a) Continuous;

b) Intermittent (give cycle);

c) Stand-by;

d) Single operation (e.g. one shot devices);

e) Storage.

6.2.1.4 *Load conditions*

a) Overload;

b) Other (specified).

6.3 *Failure and fault description*

- Fault recognition: symptoms and indications, fault detected, fault not detected, false alarm.
- Item fault mode (identification of functions affected).
- Failure causes:
  a) Inherent to item under observation;
  b) Misuse failure;
  c) Induced by maintenance or administrative action;
  d) External to item under observation;
  e) Secondary (caused by related item);
  f) Other.

  In cases where the failure immediately follows a period of transport, storage or stand-by, the relevant conditions shall be stated.
- Fault consequences
- List (identification) and physical location of faulty replaced parts:
  a) quantity of suspected replaceable items;
  b) quantity of replaced items.
- Fault evidence and documentation (printouts, photograph, etc.).
- Action taken: Replacement, repair, adjustment, modification, lubrication, etc.
- Active maintenance time (diagnostic + repair + tests + ...).
- Downtime, including, where applicable:
  - undetected fault time,
  - fault localization time,
  - reconfiguration time[3],
  - technical delay,
  - logistic delay,
  - administrative delay,
  - fault correction time,
  - checkout time,
  - restart time.

6.4 *Maintenance support data:*

- spare resources shortage,
- test resources shortage,
- resources shortage.

7 Data presentation for evaluation

When collected data is offered for subsequent evaluation by using approximate statistical methods, all conditions for their correct use and understanding should be clearly stated.

[3] Time required for automatic reconfiguration (if manual operations are needed, they are integrated into technical delay).

These conditions should encompass the purpose of the data gathering especially with respect to type and variation of the data chosen. Information on the circumstances should also be provided such as when (e.g. busy hour conditions), where (e.g. geographic considerations) and for how long the collection took place. Specific situations, which may limit the data application and use, should be indicated, e.g. difficulties encountered, particular assumptions made, non-completeness.

Considerations should also be given to the form of presentation: where appropriate, a condensed form (e.g. diagrams, histograms) may prevail over a detailed raw data presentation.

8 Statistical methods for data treatment

In most cases the need for data treatment appears in connection with one of the following activities:

- estimation,
- compliance evaluation,
- monitoring of performances,
- comparison of performances.

For each performance of interest, estimations, hypothesis tests, control charts and comparison techniques are used for evaluating.

The application of a given statistical procedure usually requires the fulfilment of some general conditions and assumptions which have to be carefully investigated. Some of these preliminary investigations relate directly to the properties and the characteristics of the (stochastic) process generating the collected data, some other relate to the distribution underlying the collected data.

Both preliminary investigations and data treatment may require statistical procedures not dealt with in this Recommendation. International organizations other than CCITT, e.g. IEC, have produced valuable material in this field [1].

Reference

[1] International Electrotechnical Commission – Catalogue of Publications, Ed. 1987.

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

PART IV

SUPPLEMENTS TO THE SERIES E RECOMMENDATIONS RELATING TO TELEPHONE NETWORK MANAGEMENT AND TRAFFIC ENGINEERING

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

TABLE OF THE ERLANG FORMULA

Table of the Erlang loss formula
(Erlang No. 1 formula, also called Erlang B formula)

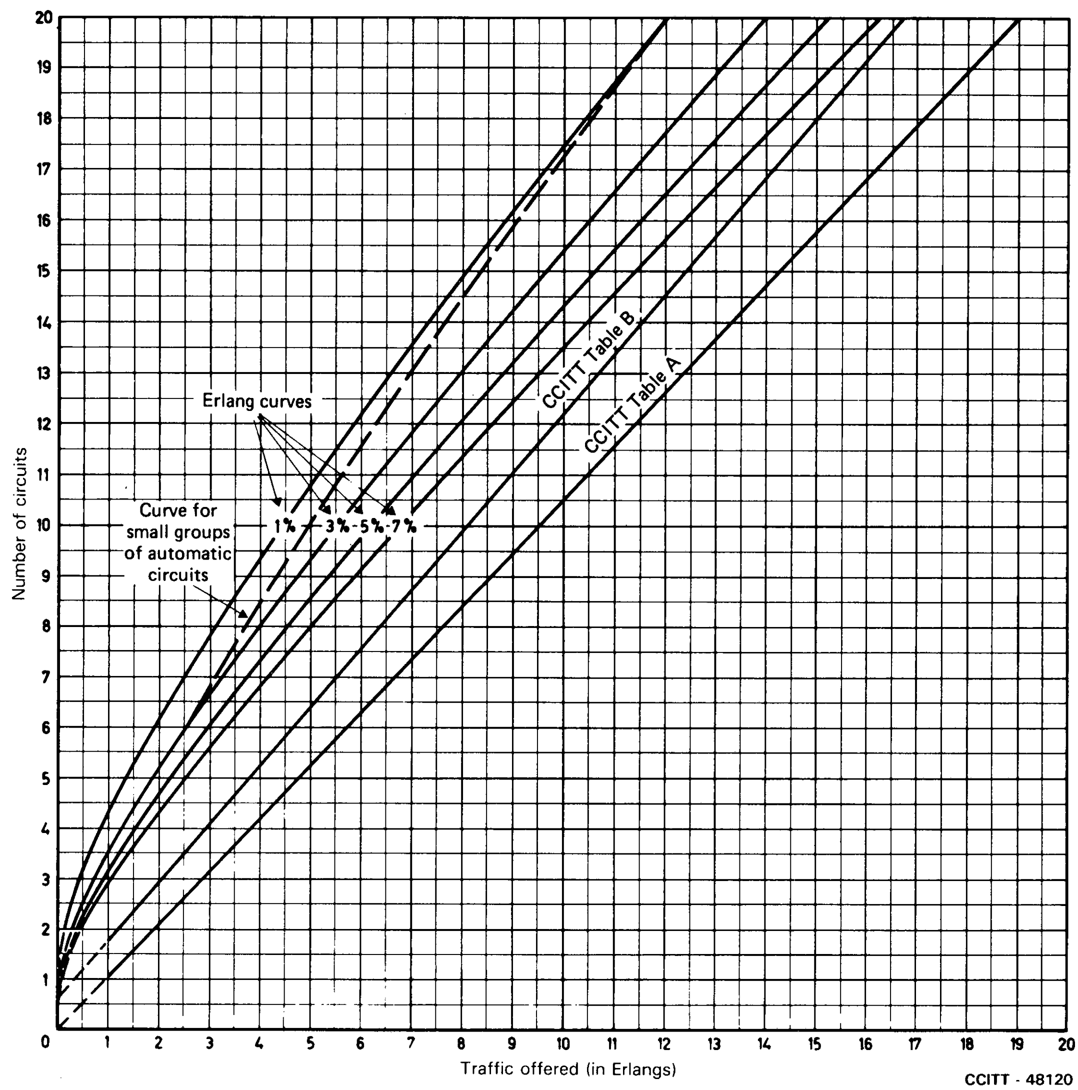
Loss probabilities: 1%, 3%, 5%, 7%.

Let p = the loss probability
y = the traffic offered (in Erlangs)
n = the number of circuits

Formula: $$E_{1,n}(y) = p = \frac{\frac{y^n}{n!}}{1 + \frac{y}{1} + \frac{y^2}{2!} + \ldots + \frac{y^n}{n!}}$$

n	$p = 1\%$	$p = 3\%$	$p = 5\%$	$p = 7\%$	n	$p = 1\%$	$p = 3\%$	$p = 5\%$	$p = 7\%$
1	0.01	0.03	0.05	0.08	51	38.80	42.89	45.53	47.72
2	0.15	0.28	0.38	0.47	52	39.70	43.85	46.53	48.76
3	0.46	0.72	0.90	1.06	53	40.60	44.81	47.53	49.79
4	0.87	1.26	1.53	1.75	54	41.50	45.78	48.54	50.83
5	1.36	1.88	2.22	2.50	55	42.41	46.74	49.54	51.86
6	1.91	2.54	2.96	3.30	56	43.31	47.70	50.54	52.90
7	2.50	3.25	3.74	4.14	57	44.22	48.67	51.55	53.94
8	3.13	3.99	4.54	5.00	58	45.13	49.63	52.55	54.98
9	3.78	4.75	5.37	5.88	59	46.04	50.60	53.56	56.02
10	4.46	5.53	6.22	6.78	60	46.95	51.57	54.57	57.06
11	5.16	6.33	7.08	7.69	61	47.86	52.54	55.57	58.10
12	5.88	7.14	7.95	8.61	62	48.77	53.51	56.58	59.14
13	6.61	7.97	8.84	9.54	63	49.69	54.48	57.59	60.18
14	7.35	8.80	9.73	10.48	64	50.60	55.45	58.60	61.22
15	8.11	9.65	10.63	11.43	65	51.52	56.42	59.61	62.27
16	8.88	10.51	11.54	12.39	66	52.44	57.39	60.62	63.31
17	9.65	11.37	12.46	13.35	67	53.35	58.37	61.63	64.35
18	10.44	12.24	13.39	14.32	68	54.27	59.34	62.64	65.40
19	11.23	13.11	14.31	15.29	69	55.19	60.32	63.65	66.44
20	12.03	14.00	15.25	16.27	70	56.11	61.29	64.67	67.49
21	12.84	14.89	16.19	17.25	71	57.03	62.27	65.68	68.53
22	13.65	15.78	17.13	18.24	72	57.96	63.24	66.69	69.58
23	14.47	16.68	18.08	19.23	73	58.88	64.22	67.71	70.62
24	15.29	17.58	19.03	20.22	74	59.80	65.20	68.72	71.67
25	16.13	18.48	19.99	21.21	75	60.73	66.18	69.74	72.72
26	16.96	19.39	20.94	22.21	76	61.65	67.16	70.75	73.77
27	17.80	20.31	21.90	23.21	77	62.58	68.14	71.77	74.81
28	18.64	21.22	22.87	24.22	78	63.51	69.12	72.79	75.86
29	19.49	22.14	23.83	25.22	79	64.43	70.10	73.80	76.91
30	20.34	23.06	24.80	26.23	80	65.36	71.08	74.82	77.96
31	21.19	23.99	25.77	27.24	81	66.29	72.06	75.84	79.01
32	22.05	24.91	26.75	28.25	82	67.22	73.04	76.86	80.06
33	22.91	25.84	27.72	29.26	83	68.15	74.02	77.87	81.11
34	23.77	26.78	28.70	30.28	84	69.08	75.01	78.89	82.16
35	24.64	27.71	29.68	31.29	85	70.02	75.99	79.91	83.21
36	25.51	28.65	30.66	32.31	86	70.95	76.97	80.93	84.26
37	26.38	29.59	31.64	33.33	87	71.88	77.96	81.95	85.31
38	27.25	30.53	32.62	34.35	88	72.81	78.94	82.97	86.36
39	28.13	31.47	33.61	35.37	89	73.75	79.93	83.99	87.41
40	29.01	32.41	34.60	36.40	90	74.68	80.91	85.01	88.46
41	29.89	33.36	35.58	37.42	91	75.62	81.90	86.04	89.52
42	30.77	34.30	36.57	38.45	92	76.56	82.89	87.06	90.57
43	31.66	35.25	37.57	39.47	93	77.49	83.87	88.08	91.62
44	32.54	36.20	38.56	40.50	94	78.43	84.86	89.10	92.67
45	33.43	37.16	39.55	41.53	95	79.37	85.85	90.12	93.73
46	34.32	38.11	40.54	42.56	96	80.31	86.84	91.15	94.78
47	35.22	39.06	41.54	43.59	97	81.24	87.83	92.17	95.83
48	36.11	40.02	42.54	44.62	98	82.18	88.82	93.19	96.89
49	37.00	40.98	43.53	45.65	99	83.12	89.80	94.22	97.94
50	37.90	41.93	44.53	46.69	100	84.06	90.79	95.24	98.99

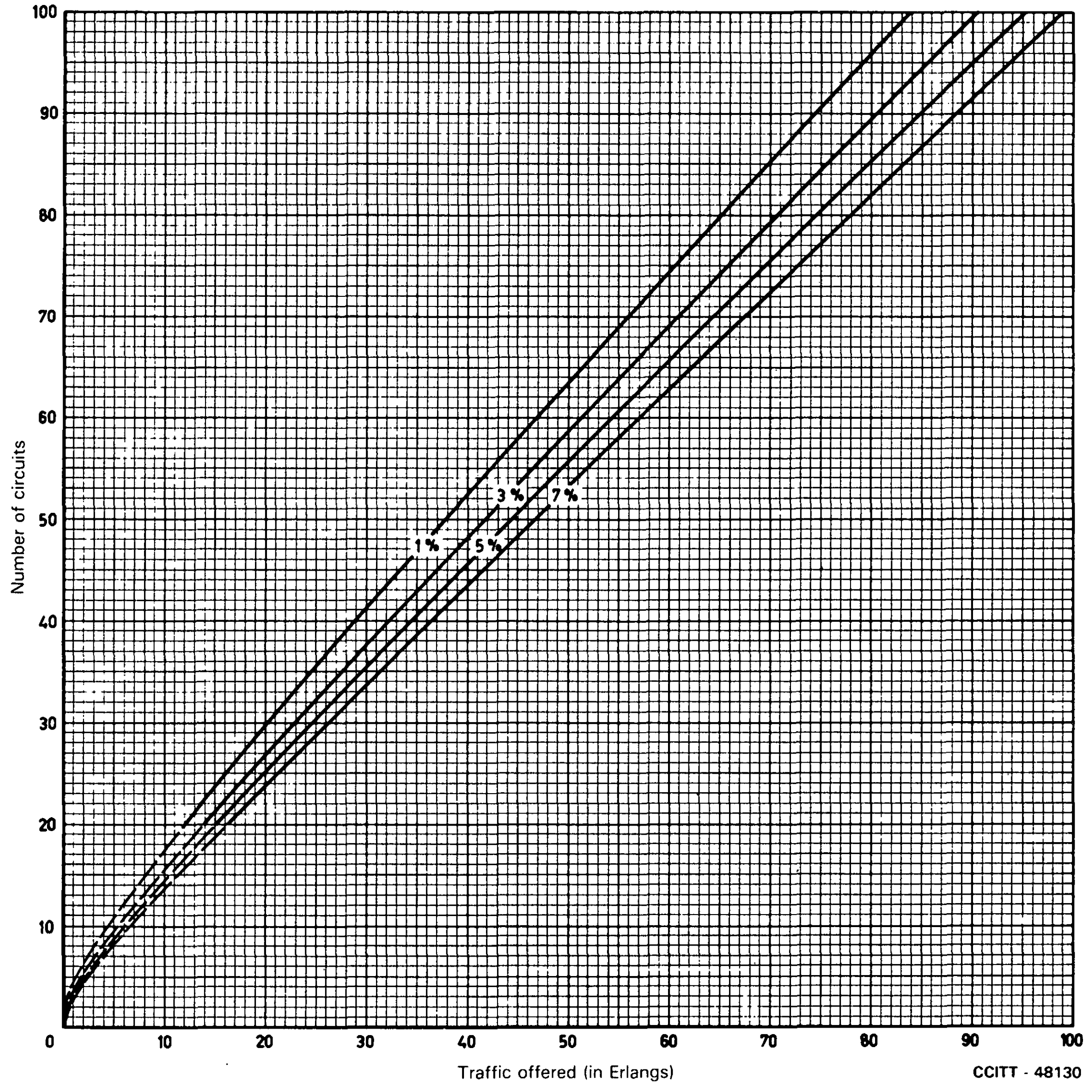
CURVES SHOWING THE RELATION BETWEEN THE TRAFFIC OFFERED AND THE NUMBER OF CIRCUITS REQUIRED



Relation between the traffic (in Erlangs) offered and the number of circuits required in the case of:
— the curves A and B of Table 1/E.510;
— the Erlang formula (p = 1 %, 3 %, 5 % and 7 %);
— the curve for small groups of automatic circuits (see Annex A to Recommendation E.520).

FIGURE 1

Number of circuits between 1 and 20



Relation between the traffic (in Erlangs) offered and the number of circuits required in the case of the Erlang formula for (p = 1 %, 3 %, 5 % and 7 %).

FIGURE 2

Number of circuits between 1 and 100

Supplement No. 3

INFORMATION ON TRAFFIC ROUTING IN THE INTERNATIONAL NETWORK

(Results from study in 1973-1976 of Question 11/XIII
concerning actual connections of international telephone calls

(For the text of this Supplement, see Supplement No. 7,
Volume II.2, *Orange Book*, Geneva, 1976)

Supplement No. 4

USE OF COMPUTERS FOR NETWORK PLANNING AND CIRCUIT GROUP DIMENSIONING

(For the text of this Supplement, see Supplement No. 8, Volume II.2, *Orange Book*, Geneva, 1976)

Supplement No. 5

TELETRAFFIC IMPLICATIONS FOR INTERNATIONAL SWITCHING AND OPERATIONAL PROCEDURES RESULTING FROM A FAILURE OF A TRANSMISSION FACILITY

1 Very considerable changes have occurred in the international network over the past decade. These changes have arisen mainly from:

- the growth in the number of long-distance routes;
- the growth in the number of circuits forming individual long-distance routes;
- the world-wide introduction of international automatic operation;
- technological developments associated with all aspects of the international network: switching unit design, transmission facility design, and routing and operational strategies;
- the integration into the international automatic service of the more isolated geographical areas and of centres having low-capacity international switching units.

2 The resultant multiplicity of circumstances and situations arising within the international network is now such that it is no longer possible to specify one single criterion for initiating corrective action to counter the loss of a transmission facility. Indeed, the failure of the whole, or part, of a transmission facility may manifest itself in a different manner to each of several Administrations affected by the failure.

3 Among the many aspects of international switching and operational procedures which can influence the degree of curtailment of service arising from a transmission-facility failure, i.e. which can reduce the ability of part of the international network to carry its designed traffic load successfully, the following are specifically stressed (their order has no particular significance):

- the introduction of fully automatic international operation, which means that the control of the network, formerly completely operator-controlled, now depends directly on subscribers' habits;
- the number of routes that could be affected by failure and their proportion of the total routes on the switching unit to which they are directly connected: the range can be from one whole route to a few circuits in each of many routes, depending on the method of allocating circuits to transmission facilities;
- the influence of any route, for which no alternative transmission facility exists, on the performance of the international switching unit to which it is connected;
- the effect on the grade of service of the switching unit itself due to the loss of a complete route or routes, or parts of several routes, directly connected to it;
- the methods of limiting the effect of failure on service by action within the switching unit or at preceding international or national switching unit, e.g. by code blocking or recorded announcements;

- the cause of the failure, and thus the possible restoration time, relative to the 24-hour traffic profile;
- the effect of a failure on overflow and automatic alternative routing strategies;
- the use of diversity of international switching units;
- the use of diversity of international transmission facilities.

4 Attention is also drawn to four major factors of maintaining continuity of service:
- reliability,
- diversity,
- network management, and
- any redundancy specifically provided to allow restoration of service.

5 Clearly, no practical transmission facility provided will give 100 per cent reliability, so it is inevitable that the other three factors will be involved to varying degrees in maintaining service. The interaction of these four factors will depend largely on the emphasis placed upon each of them by each Administration, thus reinforcing the view that the degree of corrective action that can be taken will depend considerably upon the investment policy (in materials and equipment) and forward-planning objectives of individual Administrations.

6 With respect to diversity, it is recommended that Administrations give consideration to the provision of an adequate number of paths for a particular route, with an adequate level of independence between the paths. Such independence could reduce the effect of a breakdown or other adverse event by confining it, as far as possible, to only one of the paths used by that route.

7 For the further assistance of Administrations in their study of those teletraffic aspects of international switching and operational procedure which influence the degree of curtailment of service and which arise from a transmission facility failure, these factors are included in Question 23/II related to continuity of service, accepted for study during the 1985-1988 Study Period.

Supplement No. 6

TERMS AND DEFINITIONS FOR QUALITY OF SERVICE, NETWORK PERFORMANCE, DEPENDABILITY AND TRAFFICABILITY STUDIES

CONTENTS

PART I

Dependability Vocabulary

1 Introduction

A consistent set of terms and definitions is required for the development and use of the various Recommendations addressing aspects relating the specifications, planning, data collection, analysis and evaluation of the dependability, including availability performance, reliability performance, maintainability performance and maintenance support performance. These needs address the whole range of operational and maintenance aspects as applicable to telecommunication networks, exchanges, transmission routes and channels, etc., independent of the types of services supported, including aspects of providing the necessary maintenance.

The vocabulary provided by this Supplement is also of importance in the handling of concepts related to quality of service and network performance. The conceptual relationships between quality of service and network performance to dependability and other item-related performances is shown in Figure 1. Further information on these performances is found in Recommendation E.800.

The vocabulary also provides the necessary linking to the performances of components and building blocks (hardware and software) of networks at any level of indenture (ref. definition 6019).

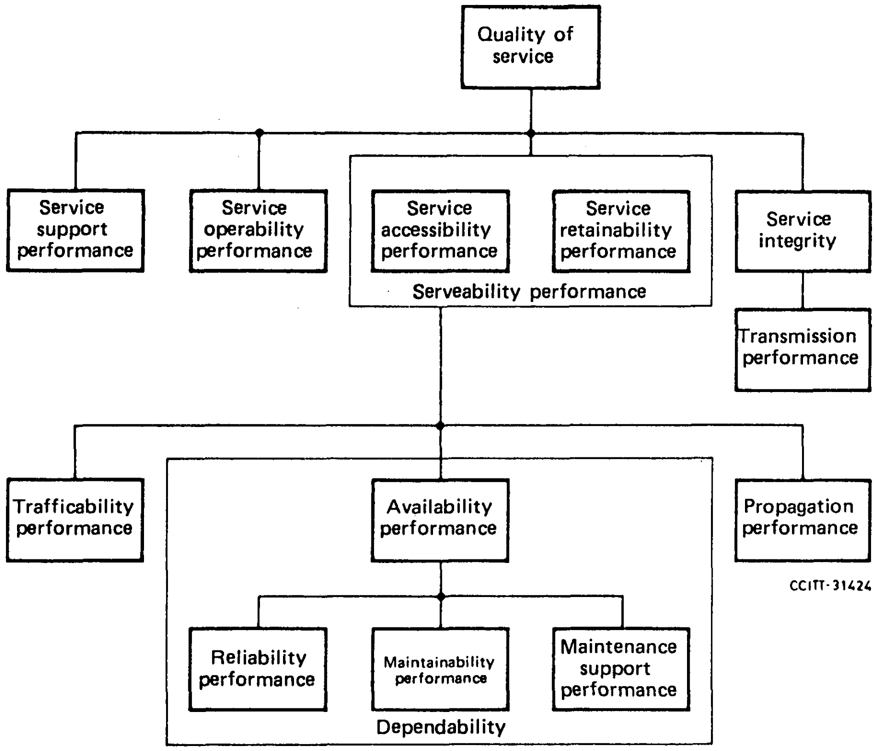


FIGURE 1

Dependability and its relation to other performance concepts

For the quantitative application and determination of measures, the terms and definitions of Part II also need to be consulted.

2 Related Recommendations

E.600 – Terms and definitions of traffic engineering

E.800 – Quality of service and dependability vocabulary.

3 Basic concepts

3001 **item ; entity**

F: entité; individu

S: elemento; entidad; ítem

Any part, device, subsystem, functional unit, equipment or system that can be individually considered.

Note 1 – An *item* may consist of hardware, software or both, and may also include people, e.g. operators in a telephone operator system.

Note 2 – In French, the term *entité* replaces the term *dispositif* previously used in this meaning, because the term *dispositif* is also the common equivalent for the English term "device".

Note 3 – In French, the term *individu* is used mainly in statistics.

3002 **repaired item**

F: entité réparée

S: elemento reparado

A repairable *item* which is in fact repaired after a *failure.*

3003 **non-repaired item**

F: entité non réparée

S: elemento no reparado

An *item* which is not repaired after a *failure.*

Note – A *non-repaired item* may be repairable or not.

3005 **required function**

F: fonction requise

S: función requerida

A function or a combination of functions of an *item* which is considered necessary for the provisioning of a given *service.*

3006 **functional mode**

F: mode de fonctionnement

S: modo de funcionamiento

A subset of the whole set of possible functions of an *item.*

3007 **instant of time**

F: instant

S: instante (de tiempo)

A single point on a time scale.

Note – The time scale may be continuous as calendar time, or discrete, e.g. number of use cycles.

3008 **time interval**

F: intervalle de temps

S: intervalo de tiempo

All *instants of time* between two given *instants of time.*

3009 **(time) duration**

F: durée

S: duración

The difference between the end points of a *time interval.*

3010 **accumulated time**

F: durée cumulée

S: tiempo acumulado

The sum of *time durations* characterized by given conditions over a given *time interval.*

3011 **measure** (as applied in the study of reliability performance and related areas)

F: caractéristique (probabilité); mesure (en fiabilité et domaines connexes)

S: medida (aplicada en estudios de fiabilidad y de aspectos conexos)

A function or quantity used to describe a *random variable* or a *random process.*

Note – For a *random variable,* examples of *measures* are the *distribution function* and the *mean.*

3012 **operation**

F: exploitation

S: explotación; operación

Combination of all technical and corresponding administrative actions intended to allow an *item* to perform a *required function,* recognizing necessary adaptation to changes in external conditions.

Note – By external conditions are understood, for example, service demand and environmental conditions.

3013 **modification** (of an item)

F: modification (d'une entité)

S: modificación (de un elemento)

The combination of all technical and corresponding administrative actions intended to alter the *capability* of an *item* by changing, adding or deleting one or more *required functions.*

4 Performances

4001 **dependability**

F: sûreté de fonctionnement

S: seguridad de funcionamiento

The collective term used to describe the *availability performance* and its influencing factors: *reliability performance, maintainability performance* and *maintenance support performance.*

Note – *Dependability* is used only for general descriptions in non-quantitative terms.

4002 **availability performance**

F: disponibilité

S: disponibilidad

The ability of an *item* to be in a state to perform a *required function* at a given *instant of time* or at any *instant of time* within a given *time interval,* assuming that the external resources, if required, are provided.

Note 1 – This ability depends on the combined aspects of the *reliability performance,* the *maintainability performance* and the *maintenance support performance* of an *item.*

Note 2 – In the definition of the *item* the external resources required must be delineated.

Note 3 – The term *availability* is used as an *availability performance measure.*

Note 4 – Warning: the term *availability* has occasionally been used in connection with the term *item,* but with an implied meaning of *item* being entirely different from that of this Supplement.

4003 **reliability performance**

F: fiabilité

S: fiabilidad

The ability of an *item* to perform a *required function* under given conditions for a given *time interval.*

Note 1 – It is generally assumed that the *item* is in a state to perform this *required function* at the beginning of the *time interval.*

Note 2 – The term *reliability* is used as a *measure* of *reliability performance.*

4004 **maintainability performance**

F: maintenabilité

S: mantenibilidad

The ability of an *item* under stated conditions of use, to be retained in, or restored to, a state in which it can perform a *required function,* when *maintenance* is performed under given conditions and using stated procedures and resources.

Note – The term *maintainability* is used as a *measure* of *maintainability performance.*

4005 **maintenance support performance**

F: logistique de maintenance

S: logística de mantenimiento

The ability of a maintenance organization, under given conditions, to provide upon demand the resources required to maintain an *item,* under a given *maintenance policy.*

Note – The given conditions are related to the *item* itself and to the conditions under which the *item* is used and maintained.

4006 **durability**

F: durabilité

S: durabilidad

The ability of an *item* to remain in a condition where it can perform a *required function* under stated conditions of use and *maintenance* until a limiting state is reached.

Note – A limiting state of an *item* may be characterized by the end of the *useful life,* unsuitability for any economic or technological reasons, etc.

5 Events and states

5.1 *Defects*

5101 **defect**

F: défaut

S: defecto

Any departure of a characteristic of an *item* from requirements.

Note 1 – The requirements may or may not be expressed in the form of a specification.

Note 2 – A defect may or may not affect the ability of an *item* to perform a *required function.*

5102 **bug**

F: erreur de programmation; bogue

S: error de programación

A software *defect* caused by a *mistake.*

5103 **critical defect**

F: défaut critique

S: defecto crítico

A *defect* that is assessed likely to result in injury to persons or significant material damage.

5104 **non-critical defect**

F: défaut non critique

S: defecto no crítico

A *defect* other than a *critical defect.*

5105 **major defect**

F: défaut majeur

S: defecto mayor

A *defect* that is likely to result in a *failure* or to reduce materially the usability of the *item* for its intended purpose.

5106 **minor defect; imperfection**

F: défaut mineur; imperfection

S: defecto menor; imperfección

A defect other than a *major defect.*

5107 **defective; defective item**

F: défectueux; entité défectueuse

S: defectuoso; elemento defectuoso

An *item* which contains one or more *defects.*

5108 **critical defective item**

F: défectueux critique

S: elemento defectuoso crítico

An *item* which contains one or more *critical defect.*

5109 **major defective item**

F: défectueux majeur

S: elemento defectuoso mayor

An *item* which contains one or more *major defects.*

5110 **minor defective item**

F: défectueux mineur

S: elemento defectuoso menor

An *item* which contains one or more *minor defects* but no *major defects.*

5111 **design defect**

F: défaut de conception

S: defecto de diseño

A *defect* due to an inadequate design of an *item.*

5112 **manufacturing defect**

F: défaut de fabrication

S: defecto de fabricación

A *defect* due to nonconformance in manufacture to the design of an *item* or to specified manufacturing processes.

5.2 *Failures*

5201 **failure**

F: défaillance

S: fallo

The termination of the ability of an *item* to perform a *required function.*

Note – After *failure* the *item* has a fault.

5202 **critical failure**

F: défaillance critique

S: fallo crítico

A *failure* which is assessed likely to result in injury to persons or significant material damage.

5203 **non-critical failure**

F: défaillance non critique

S: fallo no crítico

A *failure* other than a *critical failure.*

5204 **misuse failure**

F: défaillance par mauvaise utilisation

S: fallo por uso incorrecto

A *failure* due to induced stresses during use which are beyond the stated capabilities of the *item.*

5205 **mishandling failure**

F: défaillance par fausse manœuvre

S: fallo por manejo incorrecto

A *failure* caused by incorrect handling or lack of care of the *item.*

5206 **(inherent) weakness failure**

F: défaillance par fragilité (inhérente)

S: fallo por fragilidad (inherente)

A *failure* due to a weakness inherent in the *item* itself when subjected to stresses within the stated capabilities of the *item.*

5207 **design failure**

F: défaillance de conception

S: fallo de diseño

A *failure* due to a *design defect.*

5208 **manufacturing failure**

F: défaillance de fabrication

S: fallo de fabricación

A *failure* due to a *manufacturing defect.*

5209 **ageing failure; wearout failure**

F: défaillance par vieillissement; défaillance par usure

S: fallo por envejecimiento; fallo por desgaste

A *failure* whose probability of occurrence increases with the passage of time, as a result of processes inherent in the *item.*

5210 **sudden failure**

F: défaillance soudaine

S: fallo repentino

A *failure* that could not be anticipated by prior examination or monitoring.

5211 **gradual failure; degradation failure; drift failure**

F: défaillance progressive; dégradation; défaillance par dérive

S: fallo gradual; fallo por degradación; fallo por deriva

A *failure* due to a gradual change in time of given characteristics of an *item* and that could be anticipated by prior examination or monitoring.

Note – A *gradual failure* can sometimes be avoided by *preventive maintenance.*

5212 **cataleptic failure; catastrophic failure** (deprecated)

F: défaillance cataleptique

S: fallo cataléptico; fallo catastrófico (desaconsejado)

A *sudden failure* which results in a *complete fault.*

5213 **relevant failure**

F: défaillance pertinente; défaillance à prendre en compte

S: fallo pertinente; fallo relevante

A *failure* to be included in interpreting test or operational results or in calculating the value of a *reliability performance measure.*

Note – The criteria for the inclusion should be stated.

5214 **non-relevant failure**

F: défaillance non pertinente; défaillance à ne pas prendre en compte

S: fallo no pertinente; fallo irrelevante

A *failure* to be excluded in interpreting test or operational results or in calculating the value of a *reliability performance measure.*

Note – The criteria for the exclusion should be stated.

5215 **primary failure**

F: défaillance primaire

S: fallo primario

A *failure* of an *item*, not caused either directly or indirectly by the *failure* or the *fault* of another *item*.

5216 **secondary failure**

F: défaillance secondaire

S: fallo secundario

A *failure* of an *item*, caused either directly or indirectly by the *failure* or the *fault* of another *item*.

5217 **failure cause**

F: cause de défaillance

S: causa de fallo

The circumstances during design, manufacture or use which have led to a *failure*.

5218 **failure mechanism**

F: mécanisme de défaillance

S: mecanismo de fallo

The physical, chemical or other process which has led to a *failure*.

5219 **systematic failure; reproducible failure; deterministic failure**

F: défaillance systématique; défaillance reproductible

S: fallo sistemático; fallo reproducible; fallo determinístico

A *failure* related in a deterministic way to a certain cause, which can only be eliminated by a *modification* of the design or manufacturing process, operational procedures, documentation or other relevant factors.

Note 1 – *Corrective maintenance* without *modification* will usually not eliminate the *failure cause*.

Note 2 – A *systematic failure* can be induced at will by simulating the *failure cause*.

5.3 *Faults*

5301 **fault**

F: panne; dérangement

S: avería

The inability of an *item* to perform a *required function*, excluding that inability due to *preventive maintenance*, lack of external resources or planned actions.

Note – A *fault* is often the result of a *failure* of the *item* itself, but may exist without prior *failure*.

5302 **critical fault**

F: panne critique

S: avería crítica

A *fault* which is assessed likely to result in injury to persons or significant damage to material.

5303 **non-critical fault**

F: panne non critique

S: avería no crítica

A *fault*, other than a *critical fault.*

5304 **major fault**

F: panne majeure

S: avería mayor

A *fault* which affects a function considered to be of major importance.

5305 **minor fault**

F: panne mineure

S: avería menor

A *fault* other than a *major fault.*

5306 **misuse fault**

F: panne par mauvaise utilisation

S: avería por uso incorrecto

A *fault* due to induced stresses during use which are beyond the stated capabilities of the *item.*

5307 **mishandling fault**

F: panne par fausse manœuvre

S: avería por manejo incorrecto

A *fault* caused by incorrect handling or lack of care of the *item.*

5308 **(inherent) weakness fault**

F: panne par fragilité (inhérente)

S: avería por fragilidad (inherente)

A *fault* due to a weakness inherent in the *item* itself when subjected to stresses within the stated capabilities of the *item.*

5309 **design fault**

F: panne de conception

S: avería de diseño

A *fault* due to a *design defect.*

5310 **manufacturing fault**

F: panne de fabrication

S: avería de fabricación

A *fault* due to a *manufacturing defect.*

5311 **ageing fault; wearout fault**

F: panne par vieillissement; panne par usure

S: avería por envejecimiento; avería por desgaste

A *fault* resulting from an *ageing failure.*

5312 **programme-sensitive fault**

F: panne mise en évidence par le programme

S: avería dependiente del programa

A *fault* that is revealed as a result of the execution of some particular sequence of instructions.

5313 **data-sensitive fault**

F: panne mise en évidence par les données

S: avería dependiente de los datos

A *fault* that is revealed as a result of the processing of a particular pattern of data.

5314 **complete fault; function preventing fault**

F: panne complète

S: avería completa

A *fault* characterized by complete inability to perform all *required functions* of an *item.*

Note – The criteria for a *complete fault* have to be stated.

5315 **partial fault**

F: panne partielle

S: avería parcial

A *fault* of an *item* other than a *complete fault.*

5316 **persistent fault; permanent fault; solid fault**

F: panne permanente

S: avería permanente

A *fault* of an *item* that persists until an action of *corrective maintenance* is performed.

5317 **intermittent fault; volatile fault; transient fault**

F: panne intermittente; panne temporaire

S: avería intermitente; avería transitoria

A fault of an *item* which persists for a limited *time duration* following which the *item* recovers the ability to perform a *required function* without being subjected to any action of *corrective maintenance.*

Note – Such a *fault* is often recurrent.

5318 **determinate fault**

F: panne franche

S: avería clara; avería determinable

For an *item,* which produces a response as a result of an action, a *fault* for which the response is the same for all actions.

5319 **indeterminate fault**

F: panne indéterminée

S: avería indeterminable

For an *item,* which produces a response as a result of an action, a *fault* such that the *error* affecting the response depends on the action applied.

Note – An example would be a *data-sensitive fault.*

5320 **latent fault**

F: panne latente

S: avería latente

An existing *fault* that has not yet been recognized.

5321 **systematic fault**

F: panne systématique

S: avería sistemática

A *fault* resulting from a *systematic failure.*

5322 **fault mode; failure mode** (deprecated)

F: mode de panne; mode de défaillance (terme déconseillé)

S: modo de avería; modo de fallo (desaconsejado)

One of the possible states of a *faulty item*, for a given *required function.*

5323 **faulty**

F: en panne

S: averiado

Property of having a *fault.*

5.4 *Errors and mistakes*

5401 **error**

F: erreur

S: error

A discrepancy between a computed, observed or measured value or condition and the true, specified or theoretically correct value or condition.

Note – An *error* can be caused by a *faulty item*, e.g. a computing *error* made by a *faulty* computer equipment.

5402 **execution error; generated error**

F: erreur d'exécution

S: error de ejecución

Error produced during the operation of a *faulty item.*

5403 **interaction error** (man-machine)

F: erreur d'interaction (homme-machine)

S: error de interacción (hombre-máquina)

An *error* in the response of an *item* caused by a *mistake* during its use.

5404 **propagated error**

F: erreur propagée

S: error propagado

An *error* in the response to erroneous data input to a non-faulty *item.*

5405 **mistake; error** (deprecated in this sense)

F: erreur (humaine); faute

S: equivocación; error (desaconsejado en este sentido)

A human action that produces an unintended result.

5.5 *Item related states* (see also Figure 2)

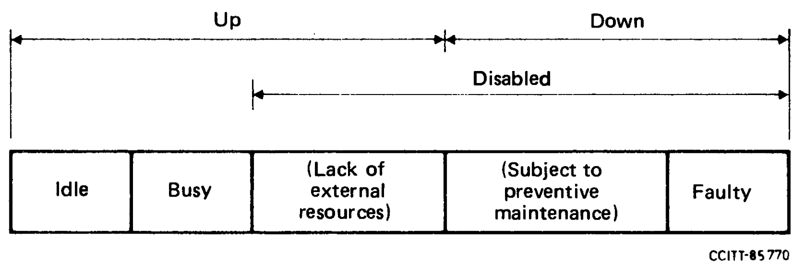


FIGURE 2

Classification of item states

5501 **operating state**

F: (état de) fonctionnement

S: estado de funcionamiento; estado operacional

The state when an *item* is performing a *required function.*

5502 **non-operating state**

F: (état de) non-fonctionnement

S: estado de no funcionamiento

The state when an *item* is not performing a *required function.*

5503 **standby state**

F: (état d') attente

S: estado de espera (en reserva)

A non-operating *up state* during the *required time.*

5504 **idle state; free state**

F: état vacant; état libre

S: estado de reposo; estado libre

A non-operating *up state* during *non-required time.*

5505 **disabled state; outage**

F: état d'incapicité

S: estado de incapacidad

A state of an *item* characterized by its inability to perform a *required function,* for any reason.

5506 **external disabled state**

F: état d'incapacité externe

S: estado de incapacidad externa

That subset of the *disabled state* when the *item* is in an *up state,* but lacks required external resources.

5507 **down state; internal disabled state**

F: état d'indisponibilité; état d'incapacité interne

S: estado de indisponibilidad; estado de incapacidad interna

A state of an *item* characterized by a *fault* or by a possible inability to perform a *required function* during *preventive maintenance.*

Note – This state relates to *availability performance.*

5508 **up state**

F: état de disponibilité

S: estado de disponibilidad

A state of an *item* characterized by the fact that it can perform a *required function,* assuming that the external resources, if required, are provided.

Note – This state relates to *availability performance.*

5509 **busy state**

F: état occupé; occupation

S: estado de ocupación; estado de ocupado

The state of an *item* in which it performs a *required function* for a user and for that reason is not accessible by other users.

5510 **critical state**

F: état critique

S: estado crítico

A state of an *item* assessed likely to result in injury to persons or significant material damage.

Note – A *critical state* may be the result of a *critical fault,* but not necessarily.

6 Maintenance

6001 **maintenance philosophy**

F: philosophie de maintenance

S: filosofía de mantenimiento

A system of underlying principles for the organization and execution of the *maintenance.*

6002 **maintenance policy**

F: politique de maintenance

S: política de mantenimiento

A description of the interrelationship between the *maintenance echelons,* the *indenture levels* and the *levels of maintenance* to be applied for the *maintenance* of an *item.*

6003 **maintenance**

F: maintenance

S: mantenimiento

The combination of all technical and corresponding administrative actions, including supervision actions, intended to retain an *item* in, or restore it to, a state in which it can perform a *required function.*

6004 **preventive maintenance**

F: maintenance préventive; entretien

S: mantenimiento preventivo

The *maintenance* carried out at predetermined intervals or according to prescribed criteria and intended to reduce the *probability* of *failure* or the degradation of the functioning of an *item.*

6005 **corrective maintenance; repair**

F: maintenance corrective; réparation; dépannage

S: mantenimiento correctivo; reparación

The *maintenance* carried out after *fault recognition* and intended to restore an *item* to a state in which it can perform a *required function.*

6006 **deferred maintenance**

F: maintenance différée

S: mantenimiento diferido

Such *corrective maintenance* which is not immediately initiated after a *fault recognition* but is delayed in accordance with given maintenance rules.

6007 **scheduled maintenance**

F: maintenance programmée; entretien systématique

S: mantenimiento programado

The *preventive maintenance* carried out in accordance with an established time schedule.

6008 **unscheduled maintenance**

F: maintenance non programmée

S: mantenimiento no programado

The *maintenance* carried out, not in accordance with an established time schedule, but, for example, after reception of an indication regarding the state of an *item.*

6009 **on-site maintenance; in situ maintenance; field maintenance**

F: maintenance in situ

S: mantenimiento local; mantenimiento sobre el terreno

Maintenance performed at the premises where the *item* is used.

6010 **off-site maintenance**

F: maintenance déportée

S: mantenimiento no local

Maintenance performed at a place different from where the *item* is used.

Note – An example is the *repair* of a sub-item at a maintenance centre.

6011 **remote maintenance**

F: télémaintenance

S: mantenimiento remoto; telemantenimiento

Maintenance of an *item* performed without physical access of the personnel to the *item.*

6012 **automatic maintenance**

F: maintenance automatique

S: mantenimiento automático

Maintenance accomplished without human intervention.

6013 **function-affecting maintenance**

F: maintenance affectant les fonctions

S: mantenimiento que afecta a la función

A *maintenance action* that affects one or more of the *required functions* of a maintained *item.*

Note – Function-affecting maintenance is divided into *function-preventing maintenance* and *function-degrading maintenance.*

6014 **function-preventing maintenance**

F: maintenance-arrêt; maintenance empêchant l'accomplissement des fonctions

S: mantenimiento con discontinuidad de funciones

A *maintenance action* that prevents a maintained *item* from performing a *required function* by causing complete loss of all the functions.

6015 **function-degrading maintenance**

F: maintenance avec dégradation; maintenance dégradant les fonctions

S: mantenimiento con degradación de funciones

A *maintenance action* that affects one or more of the *required functions* of a maintained *item,* but not to such extent as to cause complete loss of all the functions.

6016 **function-permitting maintenance**

F: maintenance en fonctionnement; maintenance en exploitation

S: mantenimiento sin discontinuidad de funciones

A *maintenance action* that does not affect any of the *required functions* of a maintained *item.*

6017 **level of maintenance**

F: niveau de maintenance

S: nivel de mantenimiento

The *maintenance action* to be carried out at a specified *indenture level.*

Note – Examples of a *maintenance action* are replacing a component, a printed circuit board, a subsystem, etc.

6018 **maintenance echelon; line of maintenance**

F: échelon de maintenance

S: escalón de mantenimiento; línea de mantenimiento

The position in an organization where specified *levels of maintenance* are to be carried out on an *item.*

Note 1 – Examples of *maintenance echelons* are: field, repair shop, manufacturer.

Note 2 – The *maintenance echelon* is characterized by the skill of the personnel, the facilities available, the location, etc.

6019 **indenture level** (for maintenance)

F: niveau d'intervention (pour la maintenance)

S: nivel de intervención (para el mantenimiento)

A level of subdivision of an *item* from the point of view of a *maintenance action.*

Note 1 – Examples of *indenture levels* could be a subsystem, a circuit board, a component.

Note 2 – The *indenture level* depends on the complexity of the item's construction, the accessibility to sub-items, skill level of maintenance personnel, test equipment facilities, safety considerations, etc.

6020 **elementary maintenance activity**

F: opération élémentaire de maintenance

S: acción elemental de mantenimiento

The unit of work into which a maintenance activity may be broken down at a given *indenture level.*

6021 **maintenance action; maintenance task**

F: opération de maintenance; tâche de maintenance

S: acción de mantenimiento; tarea de mantenimiento

A sequence of *elementary maintenance activities* carried out for a given purpose.

Note – Examples are *fault diagnosis, fault localization* and *function check-out* or combinations thereof.

6022 **supervision**

F: surveillance; supervision

S: supervisión

Activity, performed either manually or automatically, intended to observe the state of an *item.*

Note – Automatic *supervision* may be performed internally or externally to the *item.*

6023 **controlled maintenance**

F: maintenance dirigée

S: mantenimiento dirigido

A method to sustain a desired *quality of service* by the systematic application of analysis techniques using centralized supervisory facilities and/or sampling to minimize *preventive maintenance* and to reduce *corrective maintenance.*

6024 **fault recognition**

F: détection (de panne)

S: detección (de una avería)

The event when a *fault* is recognized.

6025 **fault diagnosis**

F: diagnostic (de panne)

S: diagnóstico (de una avería)

Actions taken for *fault recognition, fault localization* and cause identification.

6026 **fault localization; fault location** (deprecated in this sense)

F: localisation de panne

S: localización (de una avería)

Actions taken to identify the *faulty* sub-item or sub-items at the appropriate *indenture level.*

6027 **fault correction**

F: correction (de panne)

S: corrección (de una avería)

Actions taken after *fault localization* intended to restore the ability of the *faulty item* to perform a *required function.*

6028 **function check-out**

F: vérification (de fonctionnement)

S: verificación de funcionamiento

Actions taken after *fault correction* to verify that the *item* has recovered its ability to perform the *required function.*

6029 **restoration; recovery**

F: rétablissement

S: restablecimiento; restauración

That event when the *item* regains the ability to perform a *required function* after a *fault.*

6030 **maintenance entity**

F: cellule de maintenance

S: célula de mantenimiento; entidad de mantenimiento

A sub-item of a given *item* defined with the intention that an alarm – caused by a *fault* in that sub-item – will be unambiguously referrable to the sub-item.

7 **Time concepts** (see also Figure 3)

7.1 *Maintenance related times*

7101 **maintenance time**

F: temps de maintenance

S: tiempo de mantenimiento

The *time interval* during which a *maintenance action* is performed on an *item* either manually or automatically, including *technical delays* and *logistic delays.*

Note – Maintenance may be carried out while the *item* is performing a *required function.*

7102 **maintenance man-hours (MMH)**

F: durée équivalente de maintenance

S: duración equivalente de mantenimiento; horas-hombre de mantenimiento

The accumulated durations of the *maintenance times,* expressed in hours, used by all maintenance personnel for a given type of *maintenance action* or over a given *time interval.*

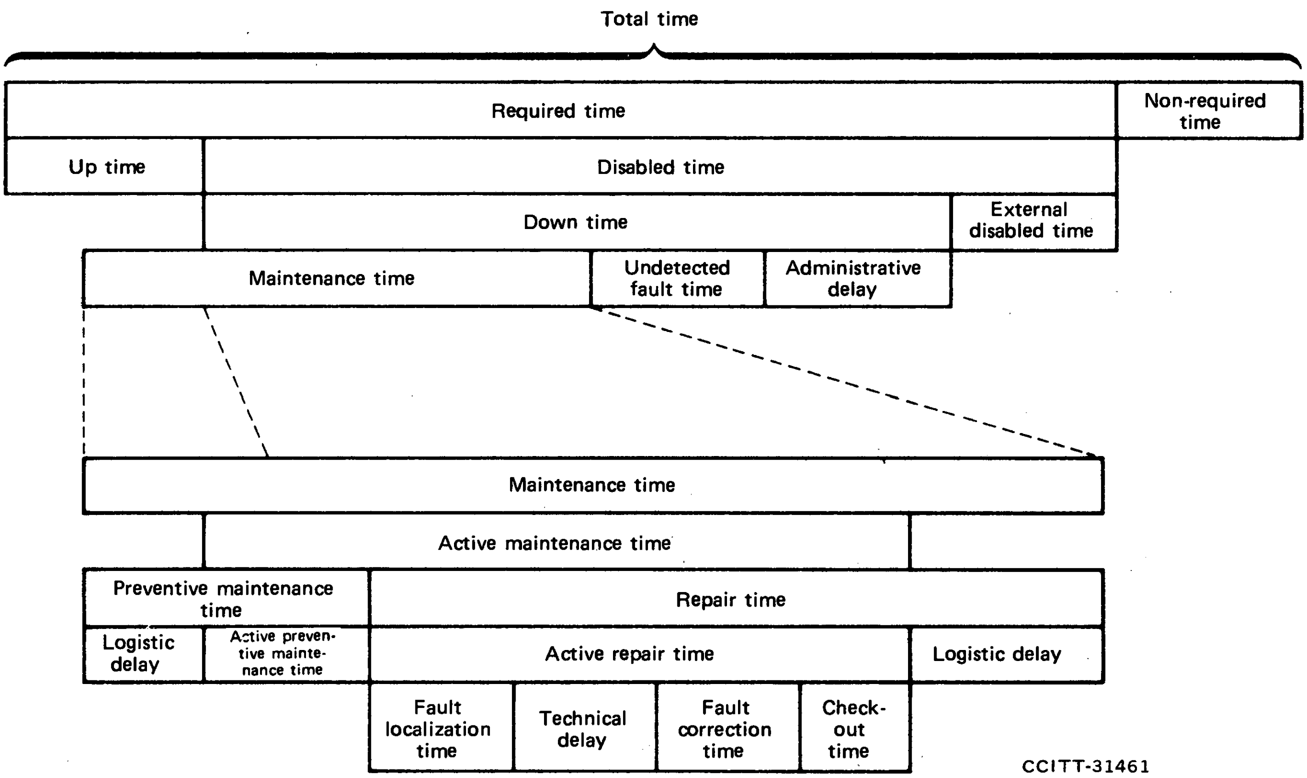


FIGURE 3

Time diagram

7103 **active maintenance time**

F: temps de maintenance active

S: tiempo de mantenimiento activo

That part of the *maintenance time* during which a *maintenance action* is performed on an *item*, either automatically or manually, excluding *logistic delays.*

Note – Active maintenance may be carried out while the *item* is performing a *required function.*

7104 **preventive maintenance time**

F: temps de maintenance préventive

S: tiempo de mantenimiento preventivo

That part of the *maintenance time* during which *preventive maintenance* is performed on an *item*, including *technical delays* and *logistic delays* inherent in *preventive maintenance.*

7105 **repair time; corrective maintenance time**

F: temps de réparation; temps de maintenance corrective

S: tiempo de reparación; tiempo de mantenimiento correctivo

That part of the *maintenance time* during which *corrective maintenance* is performed on an *item*, including *technical delays* and *logistic delays* inherent in *corrective maintenance.*

7106 **active preventive maintenance time**

F: temps de maintenance préventive active

S: tiempo de mantenimiento preventivo activo

That part of the *active maintenance time* during which actions of *preventive maintenance* are performed on an *item.*

7107 **active repair time; active corrective maintenance time**

F: temps de réparation active; temps de maintenance corrective active

S: tiempo de reparación activo; tiempo de mantenimiento correctivo activo

That part of the *active maintenance time* during which actions of *corrective maintenance* are performed on an *item.*

7108 **undetected fault time**

F: temps de non-détection de panne

S: tiempo de no detección de una avería

The *time interval* between a *failure* and recognition of the resulting *fault.*

7109 **administrative delay** (for corrective maintenance)

F: délai administratif (pour la maintenance corrective)

S: retardo administrativo (para el mantenimiento correctivo); *demora administrativa*

The *accumulated time* during which an action of *corrective maintenance* on a *faulty item* is not performed due to administrative reasons.

7110 **logistic delay**

F: délai logistique

S: retardo logístico; demora logística

That *accumulated time* during which a *maintenance action* cannot be performed due to the necessity to acquire *maintenance* resources, excluding any *administrative delay.*

Note – Logistic delays can be due to, e.g. travelling to unattended installations, awaiting the arrival of spare parts, specialists or test equipment.

7111 **fault correction time**

F: temps de correction de panne

S: tiempo de corrección de una avería

That part of *active repair time* during which *fault correction* is performed.

7112 **technical delay**

F: délai technique

S: retardo técnico; demora técnica

The *accumulated time* necessary to perform auxiliary technical actions associated with the *maintenance action* itself.

7113 **check-out time**

F: temps de vérification (du fonctionnement)

S: tiempo de verificación (de funcionamiento)

That part of *active repair time* during which *function check-out* is performed.

7114 **fault localization time; fault location time** (deprecated)

F: temps de localisation (de panne)

S: tiempo de localización de una avería

That part of *active repair time* during which *fault localization* is performed.

7.2 *Item-state related times*

7201 **operating time**

F: temps de fonctionnement

S: tiempo de funcionamiento

The *time interval* during which an *item* is an *operating state.*

7202 **non-operating time**

F: temps de non-fonctionnement

S: tiempo de no funcionamiento

The *time interval* during which an *item* is in a *non-operating state.*

7203 **required time**

F: période requise

S: periodo requerido

The *time interval* during which the user requires the *item* to be in a condition to perform a *required function.*

7204 **non-required time**

F: période non requise

S: periodo no requerido

The *time interval* during which the user does not require the *item* to be in a condition to perform a *required function.*

7205 **stand-by time**

F: période d'attente

S: tiempo de espera (en reserva)

The *time interval* during which an *item* is in a *stand-by state.*

7206 **idle time; free time**

F: période vacante; temps mort; temps libre

S: tiempo de reposo; tiempo muerto; tiempo libre

The *time interval* during which an *item* is in a *free state.*

7207 **disabled time**

F: temps d'incapacité

S: tiempo de incapacidad

The *time interval* during which an *item* is in a *disabled state.*

7208 **down time**

F: temps d'indisponibilité

S: tiempo de indisponibilidad

The *time interval* during which an *item* is in a *down state.*

7209 **accumulated down time**

F: durée cumulée d'indisponibilité

S: tiempo de indisponibilidad acumulado

The sum of the duration of *down times* over a given *time interval.*

7210 **external disabled time; external loss time**

F: temps d'incapacité externe

S: tiempo de incapacidad externa

The *time interval* during which an *item* is in an *external disabled state.*

7211 **up time**

F: temps de disponibilité; temps de bon fonctionnement

S: tiempo de disponibilidad

The *time interval* during which an *item* is in an *up state.*

7.3 *Time concepts related to reliability performance*

7301 **time to first failure**

F: durée de fonctionnement avant la première défaillance

S: tiempo hasta el primer fallo

Total *time duration* of the *operating time* of an *item* from the *instant of time* it is first put in an *up state,* until *failure.*

7302 **time to failure**

F: durée de fonctionnement avant défaillance

S: tiempo hasta el fallo

Total *time duration* of the *operating time* of an *item,* from the *instant of time* it goes from a *down state* to an *up state,* after a corrective *maintenance action,* until the next *failure.*

7303 **time between failures**

F: temps entre défaillances

S: tiempo entre fallos

The *time duration* between two successive *failures* of a *repaired item.*

Note 1 – Those parts of *non-operating time* which are included must be identified.

Note 2 – In some applications only the *up time* is considered.

7304 **time to restoration; time to recovery**

F: temps de panne

S: tiempo de avería

The *time interval* during which an *item* is in a *down state* due to a *failure.*

7305 **useful life**

F: (durée de) vie utile

S: vida útil

Under given conditions, the *time interval* beginning at a given *instant of time*, and ending when the *failure intensity* becomes unacceptable or when the *item* is considered unrepairable as a result of a *fault.*

7306 **early failure period**

F: période initiale de défaillance

S: periodo de fallos inicial

That possible early period in the life of an *item*, beginning at a given *instant of time* and during which the *instantaneous failure intensity* for a *repaired item* or the *instantaneous failure rate* for a *non-repaired item* decreases rapidly.

Note – In any particular case, it is necessary to explain what is meant by "decreases rapidly".

7307 **constant failure intensity period**

F: période d'intensité constante de défaillance

S: periodo de intensidad de fallos constante

That possible period in the life of a *repaired item* during which the *failure intensity* is approximately constant.

Note – In any particular case it is necessary to explain what is meant by "approximately constant".

7308 **constant failure rate period**

F: période de densité constante de défaillance; période de taux constant de défaillance

S: periodo de tasa de fallos constante

That possible period in the life of a *non-repaired item* during which the *failure rate* is approximately constant.

Note – In any particular case it is necessary to explain what is meant by "approximately constant".

7309 **wear-out failure period**

F: période de défaillance par vieillissement; période de défaillance par usure

S: periodo de fallos por envejecimiento

That possible later period in the life of an *item* during which the *instantaneous failure intensity* for a *repaired item* or the *instantaneous failure rate* for a *non-repaired item* increases rapidly.

Note – In any particular case it is necessary to explain what is meant by "increases rapidly".

8 Measures of performances

8.1 *Availability performance*

8101 **instantaneous availability; pointwise availability; A(t)** (symbol)

F: disponibilité (instantanée), A(t) (symbole)

S: disponibilidad instantánea, A(t) (símbolo)

The *probability* that an *item* is in an *up state* at a given *instant of time, t.*

Note – In French the term *disponibilité* is also used to denote the performance quantified by this *probability.*

8102 **instantaneous unavailability; pointwise unavailability; U(t)** (symbol)

F: indisponibilité (instantanée), U(t) (symbole)

S: indisponibilidad instantánea, U(t) (símbolo)

The probability that an *item* is in a *down state* at a given *instant of time, t.*

8103 **mean availability, $\overline{A}(t_1, t_2)$** (symbol)

F: disponibilité moyenne, $\overline{A}(t_1, t_2)$ (symbole)

S: disponibilidad media, $\overline{A}(t_1, t_2)$ (símbolo)

The normalized integral of the *instantaneous availability* in a given *time interval* (t_1, t_2).

Note – The *mean availability* is related to the *instantaneous availability* as

$$\overline{A}\,(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} A(t)\,\mathrm{d}t.$$

8104 **mean unavailability, $\overline{U}(t_1, t_2)$** (symbol)

F: indisponibilité moyenne, $\overline{U}(t_1, t_2)$ (symbole)

S: indisponibilidad media, $\overline{U}(t_1, t_2)$ (símbolo)

The normalized integral of the *instantaneous unavailability* in a stated *time interval* (t_1, t_2).

Note – The *mean unavailability* is related to the *instantaneous unavailability* as

$$\overline{U}\,(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} U(t)\,\mathrm{d}t.$$

8105 **(asymptotic) availability; (steady-state) availability; A** (symbol)

F: disponibilité asymptotique; disponibilité, A (symbole)

S: disponibilidad (asintótica); disponibilidad (en régimen permanente); A (símbolo)

The limit, if this exists, of the *instantaneous availability* when the time tends to infinity.

Note – Under certain conditions, for instance constant failure rate and constant repair rate, the *asymptotic availability* may be expressed as:

$$A = \frac{MUT}{MUT + MDT}$$

where

MDT is the *mean down time*

MUT is the *mean up time.*

8106 **asymptotic unavailability, U** (symbol)

F: indisponibilité asymptotique, U (symbole)

S: indisponibilidad asintótica, U (símbolo)

The limit, if this exists, of the *instantaneous unavailability* when the time tends to infinity.

Note – Under certain conditions, for instance constant failure rate and constant repair rate, the *asymptotic unavailability* may be expressed as:

$$U = \frac{MDT}{MDT + MUT}$$

where

MDT is the *mean down time*

MUT is the *mean up time.*

8107 **asymptotic mean availability, $\overline{\mathbf{A}}$** (symbol)

F: disponibilité moyenne asymptotique, $\overline{A}$ (symbole)

S: disponibilidad media asintótica, $\overline{A}$ (símbolo)

The limit, if this exists, of the *mean availability* over a *time interval* (t_1, t_2) when t_2 tends to infinity.

Note 1 – The *asymptotic mean availability* is related to the *mean availability* as

$$\overline{A} = \lim_{t_2 \to \infty} \overline{A}(t_1, t_2)$$

Note 2 – When such a limit exists it is not dependent on t_1.

8108 **asymptotic mean unavailability, $\overline{\mathbf{U}}$** (symbol)

F: indisponibilité moyenne asymptotique, $\overline{U}$ (symbole)

S: indisponibilidad media asintótica, $\overline{U}$ (símbolo)

The limit, if this exists, of the *mean unavailability* over a *time interval* (t_1, t_2) when t_2 tends to infinity.

Note 1 – The *asymptotic mean unavailability* is related to the *mean unavailability* as

$$\overline{U} = \lim_{t_2 \to \infty} \overline{U}(t_1, t_2)$$

Note 2 – When such a limit exists it is not dependent on t_1.

8109 **mean up time (MUT)**

F: temps moyen de disponibilité; durée moyenne de disponibilité (TMD)

S: tiempo medio de disponibilidad (TMD)

The *expectation* of the *up time.*

8110 **mean accumulated down time (MADT)**

F: durée cumulée moyenne d'indisponibilité

S: tiempo medio acumulado de indisponibilidad (TMAI)

The *expectation* of the *accumulated down time.*

8111 **instantaneous availability of a leased circuit**

F: disponibilité instantanée d'un circuit loué

S: disponibilidad instantánea de un circuito arrendado

The *probability* that, under stated operating conditions, a leased circuit can perform a *required function* when requested by the subscriber.

8.2 *Reliability performance*

8201 **reliability, R** (symbol)

F: fiabilité, R (symbole)

S: fiabilidad, R (símbolo)

The *probability* that an *item* can perform a *required function* under stated conditions for a given *time interval.*

Note 1 – It is generally assumed that the *item* is in a state to perform this *required function* at the beginning of the *time interval.*

Note 2 – In French, the term *fiabilité* is also used to denote the performance quantified by this *probability.*

8202 **(instantaneous) failure rate, λ(t)** (symbol)

F: densité (temporelle) (instantanée) de défaillance; taux (instantané) de défaillance, λ(t) (symbole)

S: tasa (instantánea) de fallos, λ(t) (símbolo)

The limit, if this exists, of the ratio of the conditional *probability* that the *time to failure, T,* of an *item* falls within a given *time interval,* $(t, t+\Delta t)$, to the length of this interval, Δt, when Δt tends to zero, given that the *item* is in a state to perform a *required function* at the beginning of the *time interval.*

Note – The *instantaneous failure rate* is expressed by formula as:

$$\lambda(t) = \lim_{\Delta(t) \to 0+} \frac{Pr(t < T \leqslant t + \Delta t \mid T > t)}{\Delta t}$$

where T is the *instant of time* of *failure.*

The formula is also applicable if T denotes the *time to failure.*

8203 **mean failure rate, $\overline{\lambda}(t_1, t_2)$** (symbol)

F: taux moyen de défaillance; densité (temporelle) moyenne de défaillance, $\overline{\lambda}(t_1, t_2)$ (symbole)

S: tasa media de fallos, $\overline{\lambda}(t_1, t_2)$ (símbolo)

The normalized integral of the *instantaneous failure rate* over a given *time interval,* (t_1, t_2).

Note – The *mean failure rate* relates to *instantaneous failure rate* as

$$\overline{\lambda}(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \lambda(t)\, \mathrm{d}t.$$

8204 **(instantaneous) failure intensity, z(t)** (symbol)

F: intensité (instantanée) de défaillance, z(t) (symbole)

S: intensitad (instantánea) de fallos, z(t) (símbolo)

The limit, if this exists, of the ratio of the mean number of *failures* of a *repaired item* in a *time interval,* $(t, t + \Delta t)$, to the length of this interval, Δt, when the length of the *time interval* tends to zero.

Note – The *instantaneous failure intensity* is expressed by formula as:

$$z(t) = \lim_{\Delta t \to 0+} \frac{E[N(t + \Delta t) - N(t)]}{\Delta t}$$

where $N(t)$ is the number of *failures* in the *time interval* $(0, t)$.

8205 **mean failure intensity, $\bar{z}(t_1, t_2)$** (symbol)

F: intensité moyenne de défaillance, $\bar{z}(t_1, t_2)$ (symbole)

S: intensidad media de fallos, $\bar{z}(t_1, t_2)$ (símbolo)

The normalized integral of the *instantaneous failure intensity* over a given *time interval* (t_1, t_2).

Note – The *mean failure intensity* is related to *instantaneous failure intensity* as:

$$\bar{z}(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} z(t)\, dt.$$

8206 **mean time to first failure (MTTFF)**

F: durée moyenne de fonctionnement avant la première défaillance (MTTFF)

S: tiempo medio hasta el primer fallo (MTTFF)

The *expectation* of the *time to first failure.*

8207 **mean time to failure (MTTF)**

F: durée moyenne de fonctionnement avant défaillance (MTTF)

S: tiempo medio hasta el fallo (MTTF)

The *expectation* of the *time to failure.*

8208 **mean time between failures (MTBF)**

F: moyenne des temps entre défaillances (MTBF)

S: tiempo medio entre fallos (MTBF)

The *expectation* of the *time between failures.*

8209 **failure rate acceleration factor**

F: facteur d'accélération de la densité de défaillance; facteur d'accélération du taux de défaillance

S: factor de aceleración de la tasa de fallos

The ratio of the accelerated testing *failure rate* to the *failure rate* under stated reference test conditions.

Note – Both *failure rates* refer to the same time period in the life of the tested *items.*

8210 **failure intensity acceleration factor**

F: facteur d'accélération de l'intensité de défaillance

S: factor de aceleración de la intensidad de fallos

In a *time interval* of given *duration,* whose beginning is specified by a fixed age of a *repaired item,* the ratio of the number of *failures* obtained under two different sets of stress conditions.

8.3 *Maintainability performance*

8301 **maintainability**

F: maintenabilité

S: mantenibilidad

The *probability* that a given active *maintenance action*, for an *item* under given conditions of use can be carried out within a stated *time interval*, when the *maintenance* is performed under stated conditions and using stated procedures and resources.

Note – In French the term *maintenabilité* is also to denote the performance quantified by this *probability.*

8302 **(instantaneous) repair rate, $\mu(t)$** (symbol)

F: densité (temporelle) (instantanée), de réparation, μ (t) (symbole)

S: tasa (instantánea) de reparaciones, $\mu(t)$ (símbolo)

The limit, if this exists, of the ratio of the conditional *probability* that the corrective *maintenance action* terminates in a *time interval*, $(t, t+\Delta t)$ to the length of this *time interval*, when Δt tends to zero, given that the action had not terminated at the beginning of the *time interval.*

Note – The *instantaneous repair rate* is expressed by formula as:

$$\mu(t) = \lim_{\Delta(t)\to 0} \frac{Pr(t < T \leqslant t + \Delta t | T > t)}{\Delta t}$$

where T is the *instant of time* of *restoration.*

T may also represent the *time to restoration.*

8303 **mean repair rate, $\bar{\mu}(t_1, t_2)$** (symbol)

F: densité (temporelle) moyenne de réparation, $\bar{\mu}(t_1, t_2)$ (symbole)

S: tasa media de reparaciones, $\bar{\mu}(t_1, t_2)$ (símbolo)

The normalized integral of the *instantaneous repair rate* over a given *time interval* (t_1, t_2).

Note – The *mean repair rate* is related to *instantaneous repair rate* as:

$$\bar{\mu}(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \mu(t)\, dt.$$

8304 **mean maintenance man-hours**

F: durée moyenne équivalente de maintenance

S: duración media equivalente de mantenimiento; media de horas-hombre de mantenimiento

The *expectation* of the *maintenance man-hours.*

8305 **mean down time (MDT)**

F: temps moyen d'indisponibilité; durée moyenne d'indisponibilité (TMI)

S: tiempo medio de indisponibilidad (TMI)

The *expectation* of the *down time.*

8306 **mean repair time (MRT)**

F: durée moyenne de réparation

S: tiempo medio de reparación

The *expectation* of the *repair time.*

8307 **p-fractile repair time**

F: quantile-p de la durée de réparation

S: cuantil-p del tiempo de reparación

The *p-fractile* value of the *repair time.*

8308 **mean active repair time (MART)**

F: durée moyenne de réparation active

S: tiempo medio de reparación activa

The *expectation* of the *active repair time.*

8309 **p-fractile active repair time**

F: quantile-p de la durée de réparation active

S: cuantil-p del tiempo de reparación activa

The *p-fractile* value of the *active repair time.*

8310 **mean time to restoration (MTTR); mean time to recovery; mean time to repair** (deprecated)

F: durée moyenne de panne; moyenne des temps pour la tâche de réparation (MTTR)

S: tiempo medio hasta el restablecimiento (MTTR)

The *expectation* of the *time to restoration.*

8311 **fault coverage**

F: couverture de pannes

S: cobertura de averías

The proportion of *faults* of an *item* that can be recognized under given conditions.

8312 **repair coverage**

F: couverture des réparations

S: cobertura de reparaciones

The proportion of *faults* of an *item* that can be successfully removed.

8.4 *Maintenance support performance*

8401 **mean administrative delay (MAD)**

F: durée moyenne du délai administratif

S: retardo medio administrativo; demora media administrativa

The *expectation* of the *administrative delay.*

8402 **p-fractile administrative delay**

F: quantile-p du délai administratif

S: cuantil-p del retardo administrativo; cuantil-p de la demora administrativa

The *p-fractile* value of the *administrative delay.*

8403 **mean logistic delay (MLD)**

F: durée moyenne du délai logistique

S: retardo medio logístico; demora media logística

The *expectation* of the *logistic delay.*

8404 **p-fractile logistic delay**

F: quantile-p du délai logistique

S: cuantil-p del retardo logístico; cuantil-p de la demora logística

The *p-fractile* value of the *logistic delay.*

9 Test, data, design and analysis

9.1 *Test concepts*

9101 **test**

F: essai

S: prueba

An experiment made in order to measure or classify a *characteristic.*

9102 **compliance test**

F: essai de conformité

S: prueba de conformidad

A *test* used to show whether or not a *characteristic* of an *item* complies with the stated requirements.

9103 **determination test**

F: essai de détermination

S: prueba de determinación

A *test* used to establish the value of a *characteristic.*

9104 **laboratory test**

F: essai en laboratoire

S: prueba de laboratorio

A *compliance test* or a *determination test* made under prescribed and controlled conditions which may or may not simulate field conditions.

9105 **field test**

F: essai dans des conditions d'exploitation

S: prueba en condiciones de explotación; prueba en condiciones reales

A *compliance test* or *determination test* made in the field where operating, environmental, maintenance and measurement conditions are recorded.

9106 **endurance test**

F: essai d'endurance

S: prueba de resistencia

A *test* carried out over a *time interval* to investigate how the properties of an *item* are affected by the application of stated stresses and by their *time duration.*

9107 **accelerated test**

F: essai accéléré

S: prueba acelerada

A *test* in which the applied stress level is chosen to exceed that stated in the reference conditions in order to shorten the *time duration* required to observe the stress response of the *item,* or to magnify the responses in a given *time duration.*

Note – To be valid, an *accelerated test* shall not alter the basic *fault modes* and *failure mechanisms,* or their relative prevalence.

9108 **step stress test**

F: essai sous contrainte échelonnée

S: prueba de esfuerzo escalonado

A *test* consisting of several stress levels applied sequentially for periods of equal *time duration* to an *item*, in such a way that during each *time interval* a stated stress level is applied and the stress level is increased from one *time interval* to the next.

9109 **screening test**

F: essai de sélection

S: prueba de selección

A *test*, or combination of *tests*, intended to remove or detect unsatisfactory *items* or those likely to exhibit early *failures.*

9110 **time acceleration factor**

F: facteur d'accélération temporelle

S: factor de aceleración temporal

The ratio between the *time durations* necessary to obtain the same stated number of *failures* or degradations in two equal size samples under two different sets of stress conditions involving the same *failure mechanisms* and *fault modes* and their relative prevalence.

Note – One of the two sets of stress conditions should be a reference set.

9111 **maintainability verification**

F: vérification de la maintenabilité

S: verificación de la mantenibilidad

A procedure applied for the purpose of determining whether the requirements for *maintainability performance measures* for an *item* has been achieved or not.

Note – The procedures may range from analysis of appropriate data to a *maintainability demonstration.*

9112 **maintainability demonstration**

F: vérification expérimentale de maintenabilité

S: demostración de la mantenibilidad

A *maintainability verification* performed as a *compliance test.*

9.2 *Data concepts*

9201 **observed data**

F: valeur observée; donnée observée

S: datos observados; valores observados

Values related to an *item* or a process obtained by direct observation.

Note – Values referred to could be events, *time instants, time intervals,* etc.

9202 **test data**

F: données d'essai

S: datos de prueba

Observed data obtained during *tests.*

9203 **field data**

F: donnée d'exploitation

S: datos de explotación

Observed data obtained during field operation.

9204 **reference data**

F: valeur de référence; données de référence

S: datos de referencia; valores de referencia

Data, which by general agreement may be used for *prediction* and/or comparison with *observed data.*

9.3 *Design concepts*

9301 **redundancy**

F: redondance

S: redundancia

In an *item*, the existence of more than one means for performing a *required function.*

9302 **active redundancy**

F: redondance active

S: redundancia activa

That *redundancy* wherein all means for performing a *required function* are intended to operate simultaneously.

9303 **standby redundancy**

F: redondance en attente; redondance passive; redondance en secours

S: redundancia pasiva; redundancia de reserva

That *redundancy* wherein one means for performing a *required function* is intended to operate, while the alternative means are inoperative until needed.

9304 **fail safe**

F: protégé contre défaillances (critique); à sûreté intégrée

S: prevención de fallos

A designed property of an *item* which prevents its *failures* from resulting in *critical faults.*

9305 **fault tolerance**

F: tolérance aux pannes

S: tolerancia a las averías

The attribute of an *item* that makes it able to perform a *required function* in the presence of certain given sub-item *faults.*

9306 **fault masking**

F: masquage de panne

S: enmascaramiento de avería

The condition in which a *fault* exists in a sub-item of an *item* but cannot be recognized because of a feature of the *item* or because of another *fault* of the sub-item or of another sub-item.

9.4 *Analysis concepts*

9401 **prediction**

F: prévision; prédiction

S: previsión; predicción

1) The process of computation used to obtain (a) *predicted* value(s) of a quantity.

2) The *predicted* value(s) of a quantity.

9402 **reliability model**

F: modèle de fiabilité

S: modelo de fiabilidad

A mathematical model used for *prediction* or *estimation* of *reliability measures* of an *item* or for similar purposes.

9403 **fault modes and effects analysis (FMEA)**

F: analyse des modes de panne et de leurs effets (AMDE)

S: análisis de los modos de avería y de sus efectos (AMAE)

A qualitative method of *reliability* analysis which involves the study of the *fault modes* which can exist in every sub-item of the *item* and the determination of the effects of each *fault mode* on other sub-items of the *item* and on the *required functions* of the *item.*

9404 **fault modes, effects and criticality analysis (FMECA)**

F: analyse des modes de panne, de leurs effets et de leur criticité (AMDEC)

S: análisis de los modos de avería, sus efectos y su criticidad (AMAEC)

Fault modes and effect analysis together with a consideration of the *probability* of occurrence and a ranking of the seriousness of the *fault.*

9405 **fault tree analysis (FTA)**

F: analyse par arbre de panne

S: análisis en árbol de averías

An analysis to determine which *fault modes* of the sub-items or external events, or combinations thereof, may result in a stated *fault mode* of the *item,* resulting in a *fault tree.*

9406 **stress analysis**

F: analyse de contraintes

S: análisis de esfuerzos

A quantitative or qualitative determination of the physical, chemical or other stresses an *item* is subjected to under given use conditions.

9407 **reliability block diagram**

F: diagramme de fiabilité

S: diagrama de bloques de fiabilidad

Block diagram showing, for one or more *functional modes* of a complex *item,* how *faults* of the sub-items represented by the blocks, or combinations thereof, result in a *fault* of the *item.*

9408 **fault tree**

F: arbre de panne

S: árbol de averías

A logic diagram showing which *fault modes* of sub-items or external events, or combinations thereof, result in a given *fault mode* of the *item.*

9409 **state-transition diagram**

F: diagramme de transition

S: diagrama de transición de estados

A diagram showing the set of possible states of an *item* and the possible one step transitions between these states.

9410 **stress model**

F: modèle de contraintes

S: modelo de esfuerzos

A mathematical model which describes how a *reliability performance measure* of an *item* varies as a function of the applied stresses.

9411 **fault analysis**

F: analyse des pannes

S: análisis de averías

The logical, systematic examination of an *item* or its diagram(s) to identify and analyse the *probability,* causes and consequences of potential and real *faults.*

9412 **maintainability model**

F: modèle de maintenabilité

S: modelo de mantenibilidad

A mathematical model used for *prediction* or *estimation* of *maintainability performance measures* of an *item* or for similar purposes.

Note – An example is the *maintenance tree.*

9413 **maintainability prediction**

F: prévision de maintenabilité; prédiction de maintenabilité

S: previsión de la mantenibilidad; predicción de la mantenibilidad

An activity performed with the intention to forecast the numerical values of a *maintainability performance measure* of an *item,* taking into account the *maintainability performance* and *reliability performance measures* of its sub-items, under given operational and maintenance conditions.

9414 **maintenance tree**

F: arbre de maintenance

S: árbol de mantenimiento

A logic diagram showing the pertinent alternative sequences of *elementary maintenance activities* to be performed on an *item* and the conditions for their choice.

9415 **maintainability allocation; maintainability apportionment**

F: répartition de la maintenabilité

S: distribución de la mantenibilidad; asignación de la mantenibilidad

A procedure applied during the design of an *item* intended to apportion the requirements for *maintainability performance measures* for an *item* to its sub-items according to given criteria.

9.5 *Improvement processes*

9501 **learning process**

F: apprentissage

S: aprendizaje

Growth in experience and familiarity by personnel with design or constructional techniques, which reduces the risk of future *mistakes.*

9502 **burn-in**

F: rodage

S: rodaje

A process of *reliability improvement* of hardware, employing operation of every *item* in a prescribed environment, with successive *fault correction*, replacement or removal at every *failure*, during the steeply falling *failure intensity* period within the *early failure period.*

9503 **reliability growth**

F: croissance de la fiabilité

S: crecimiento de la fiabilidad; incremento de la fiabilidad

A condition characterized by a progressive improvement of a *reliability performance measure* of an *item*, or population of similar *items*, with time.

Note – A growth can result either from active improvement or from *burn-in.*

9504 **reliability improvement**

F: amélioration de fiabilité

S: mejora de la fiabilidad

A process undertaken with the deliberate intention of promoting *reliability growth* by the elimination of *systematic faults.*

9505 **maintainability programme**

F: programme de maintenabilité

S: programa de mantenibilidad

A detailed plan, including the human and material resources, procedures, tasks and responsibilities during the life of an *item*, intended to determine the fulfilment of the requirements for *maintainability performance measures* for an *item* and facilitate the planning of the *maintenance.*

10 Measure modifiers

1001 **true ...**

F: ... vrai

S: ... verdadero

The ideal value which characterizes a quantity perfectly defined under the conditions which exist at the moment when that quantity is observed, or the subject of a determination.

Note – This value could be arrived at only if all causes of measurement *error* were eliminated.

1002 **predicted ...**

F: ... prédit; ... prévu

S: ... previsto; ... predicho

The numerical value assigned to a quantity, before the quantity is actually observable, computed on the basis of earlier observed or estimated values of the same quantity or of other quantities using a mathematical model.

1003 **extrapolated ...**

F: ... extrapolé

S: ... extrapolado

The *predicted* value based on *estimated* values for one or a set of conditions, intended to apply to other conditions such as time, *maintenance* and environmental conditions.

1004 **estimated ...**

F: ... estimé

S: ... estimado

The value obtained as the result of an *estimation.*

Note – The result may be expressed either as a single numerical value, a point estimate, or as a *confidence interval.*

1005 **intrinsic ... ; inherent ...**

F: ... intrinsèque; ... inhérent

S: ... intrínseco; ... inherente

Value of a *measure* determined when maintenance and operational conditions are assumed to be ideal.

1006 **operational ...**

F: ... opérationnel

S: ... operacional

Value determined under given operational conditions.

1007 **mean ... ; average ...** (deprecated)

F: ... moyen (adjectif)

S: ... medio (adjetivo); promedio (desaconsejado)

1) The value obtained as the *expectation* of a *random variable.*

2) The normalized integral of a time dependant quantity.

1008 **p-fractile ...**

F: ... quantile-p

S: cuantil-p de ...

The value obtained as the *p-fractile* of the distribution of a *random variable.*

1009 **instantaneous ...**

F: ... instantané

S: ... instantáneo

The value of a *measure* determined for a given *instant of time.*

PART II

Statistical vocabulary

1 Introduction

The quantitative applications of measures for quality of service, network performance, dependability and trafficability performance requires a fundamental set of statistical concepts.

This Part provides the terms and definitions for such applications.

2 Terms and definitions

2001 **characteristic**

F: caractère (statistique)

S: característica

A property which helps to differentiate between the individuals of a given population.

Note – The differentiation may be either quantitative (by variables) or qualitative (by attributes).

2002 **probability**

F: probabilité

S: probabilidad

For practical reasons, it may be considered that, whenever the conditions of a *test* can be reproduced, the *probability* $Pr(E)$ of an event E occurring is the value around which the occurrence frequency of the latter oscillates and towards which it tends when the number of tests are indefinitely increased.

Note – The concept of *probability* may be introduced in either of two forms, depending on whether it is intended to designate a degree of belief or whether it is considered as the limit value of a frequency. In both cases, its introduction requires that some precautions be taken which cannot be developed within the context of an International Standard and for which users should refer to specialized literature.

2003 **random variable; variate**

F: variable aléatoire

S: variable aleatoria

A variable which may take any of the values of a specified set of values and with which is associated a probability distribution.

Note – A *random variable* which may take only isolated values is said to be "discrete". A *random variable* which may take all the values of a finite or infinite interval is said to be "continuous".

2004 **random process**

F: processus aléatoire; processus stochastique

S: proceso aleatorio; proceso estocástico

A collection of time-dependent *random variables* where the values are governed by a given set of multivariate distributions for all combinations of the *random variables.*

2005 **distribution function**

F: fonction de répartition

S: función de distribución

A function giving, for every value x, the *probability* that the *random variable* X is less than or equal to x:

$$F(x) = Pr(X \leqslant x).$$

2006 **probability density function**

F: densité de probabilité

S: función densidad de probabilidad

The derivative, if this exists, of the *distribution function*:

$$f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}X}.$$

2007 **p-fractile; p-quantile** (of a probability distribution)

F: quantile d'ordre p; quantile-p (d'une loi de probabilité)

S: cuantil-p; cuantil de orden p (de una ley de distribución de probabilidades)

If p is a number between 0 and 1, the *p-fractile* is the value of the *random variable* for which the *distribution function* equals p or "jumps" from a value less than or equal to p to a value greater than p.

Note – It is possible that the *distribution function* equals p throughout the interval between consecutive possible values of the variate. In this case, any value in this interval may be considered as the *p-fractile.*

2008 **expectation** (of a random variable); **mean** (of a random variable)

F: espérance mathématique (d'une variable aléatoire); *moyenne* (d'une variable aléatoire)

S: esperanza matemática (de una variable aleatoria); *media* (de una variable aleatoria)

a) For a discrete *random variable X* taking the values x_i with the *probabilites* p_i,

$$E(X) = \Sigma\, p_i x_i$$

the sum being extended over all the values x_i which can be taken by X.

b) For a continuous *random variable X* having the probability density function $f(x)$,

$$E(X) = \int x f(x)\,\mathrm{d}x$$

the integral being extended over all values of the interval of variation of X.

Note 1 – No distinction is made between the *expectation* of a *random variable* and that of a *probability* distribution.

Note 2 – The term *mean* is also used with other meanings, for example as the normalized integral over a *time interval.*

2009 **variance** (of a random variable)

F: variance (d'une variable alêatoire)

S: varianza (de una variable aleatoria)

The *expectation* of the square of the difference between a *random variable* and the *expectation* of this variable.

2010 **standard deviation,** δ (symbol)

F: écart-type, δ (symbole)

S: desviación típica, δ (símbolo)

The positive square root of the *variance.*

2011 **observed value** (in statistics)

F: valeur observée (en statistique)

S: valor observado (en estadística)

The value of a *characteristic* determined as the result of an observation or *test.*

2012 **relative frequency**

F: fréquence (statistique)

S: frecuencia relativa

The ratio of the number of times a particular value, or a value falling within a given class, is observed to the total number of observations.

2013 **statistical test**

F: test (statistique)

S: prueba estadística

A procedure that is intended to decide whether a hypothesis about the distribution of one or more populations should be rejected or not rejected (accepted).

Note 1 – The decision taken is a result of the value of an appropriate *statistic* or *statistics,* calculated from values observed in samples taken from the populations under consideration. As the value of the *statistic* is subject to random variations, there is some risk of *error* when the decision is taken.

Note 2 – It is important to note that, generally speaking, a *test* assumes *a priori* that certain assumptions are fulfilled (for example, assumption of independence of the observations, assumption of normality, etc.). These assumptions serve as a basis of the *test.*

2014 **one-sided test**

F: test unilatéral

S: prueba unilateral

A *statistical test* in which the *statistic* used is one-dimensional and the *critical region* is the set of values lower than, or the set of values greater than, a given number.

2015 **two-sided test**

F: test bilatéral

S: prueba bilateral

A *statistical test* in which the *statistic* used is one-dimensional and in which the *critical region* is the set of values lower than a first given number and the set of values greater than a second given number.

2016 **null hypothesis,** $\mathbf{H_0}$ (symbol)

F: hypothèse nulle, H_0 (symbole)

S: hipótesis nula, H_0 (símbolo)

The hypothesis to be rejected or not rejected (accepted) at the outcome of the *statistical test.*

2017 **alternative hypothesis,** $\mathbf{H_1}$ (symbol)

F: hypothèse alternative, H_1 (symbole)

S: hipótesis alternativa, H_1 (símbolo)

The hypothesis, usually composite, which is opposed to the *null hypothesis.*

2018 **critical region**

F: région critique

S: región crítica

The set of possible values of the *statistic* used such that, if the value of the *statistic* which results from the *observed values* belongs to the set, the *null hypothesis* will be rejected, whereas it will not be rejected (accepted) if the opposite is the case.

2019 **critical values**

F: valeurs critiques

S: valores críticos

The given value(s) which limit the *critical region.*

2020 **error of the first kind**

F: erreur de première espèce

S: error de primera clase

The *error* committed in rejecting the *null hypothesis,* because the *statistic* takes a value which belongs to the *critical region,* when the *null hypothesis* is true.

2021 **type I risk**

F: risque de première espèce

S: riesgo de tipo I

The *probability* of committing the *error of the first kind,* which varies according to the real situation (within the framework of the *null hypothesis*). Its maximum value is the *significance level* of the *statistical test.*

2022 **error of the second kind**

F: erreur de seconde espèce

S: error de segunda clase

The *error* committed in failing to reject (accept) the *null hypothesis* (because the value of the *statistic* does not belong to the *critical region*), when the *null hypothesis* is not true (the *alternative hypothesis* therefore being true).

2023 **type II risk**

F: risque de seconde espèce

S: riesgo de tipo II

The *probability,* designated β, of committing the *error of the second kind.* Its value depends on the real situation and can only be calculated if the *alternative hypothesis* is adequately specified.

2024 **operating characteristic curve; OC curve** (for a statistical test plan)

F: courbe d'efficacité (d'un plan de test)

S: curva característica de funcionamiento (para un plan de prueba estadística)

A curve showing, for a given *statistical test* plan, the *probability of acceptance* as a function of the actual value of a given *measure.*

2025 **producer's risk (point)**

F: (point du) risque du fournisseur

S: (punto de) riesgo del proveedor

A point on the *operating characteristic curve* corresponding to some predetermined and usually low *probability of rejection.*

2026 **consumer's risk (point)**

F: (point du) risque du client

S: (punto de) riesgo del consumidor

A point on the *operating characteristic curve* corresponding to a predetermined and usually low *probability of acceptance.*

2027 **power of the test**

F: puissance du test

S: potencia de la prueba

The *probability* of not committing the *error of the second kind,* equal to $1 - \alpha$, and thus the *probability* of rejecting the *null hypothesis* when this hypothesis is false.

2028 **significance level** (of a statistical test), α (symbol)

F: niveau de signification (d'un test); *seuil de signification, α* (symbole)

S: nivel de significación (de una prueba estadística); *umbral de significación, α* (símbolo)

The given value which limits the *probability* of the *null hypothesis* being rejected, if the *null hypothesis* is true.

Note – The *critical region* is determined in such a way that if the *null hypothesis* is true, the *probability* of this *null hypothesis* being rejected should be not more than this given value.

2029 **probability of acceptance**

F: probabilité d'acceptation

S: probabilidad de aceptación

The *probability* that an *item* will be accepted by a given *statistical test* plan.

2030 **probability of rejection**

F: probabilité de rejet

S: probabilidad de rechazo

The *probability* that an *item* will not be accepted by a given *statistical test* plan.

2031 **confidence interval**

F: intervalle de confiance

S: intervalo de confianza

The random interval limited by two *statistics* or by a single *statistic,* such that the *probability* that a parameter to be estimated is covered by this interval is equal to a given value $1 - \alpha$, where α is the *significance level.*

2032 **statistical tolerance interval**

F: intervalle statistique de dispersion

S: intervalo estadístico de tolerancia

A random interval limited by two *statistics* or by a single *statistic*, such that the *probability* that a fraction of the population, equal to or greater than a given value between 0 and 1, is covered by this interval is equal to a given value $1 - \alpha$, where α is the *significance level.*

2033 **confidence limit**

F: limite de confiance

S: límite de confianza

Each of the limits of a two-sided *confidence interval*, or the single limit of a one-sided *confidence interval.*

2034 **estimation**

F: estimation (de paramètres)

S: estimación (de parámetros)

The operation made for the purpose of assigning, from the observed values in a sample, numerical values to the parameters of the distribution chosen as the statistical model of the population from which this sample is taken.

2035 **estimate**

F: estimation

S: estimación

The result of an *estimation.*

Note – This result may be expressed either as a single numerical value (point estimation) or as a *confidence interval.*

2036 **estimator**

F: estimateur

S: estimador

A *statistic* intended to estimate a population parameter.

2037 **confidence coefficient; confidence level**

F: niveau de confiance

S: coeficiente de confianza; nivel de confianza

The value $1 - \alpha$ of the *probability* associated with a *confidence interval* or a *statistical tolerance interval*, where α is the *significance level.*

2038 **statistic**

F: statistique

S: estadístico

A function of the *observed values* derived from a sample.

2039 **acceptable level** (of a measure)

F: niveau acceptable (d'une caractéristique)

S: nivel aceptable (de una medida)

A level for a *measure* of a given performance which in a *test* plan corresponds to a specified but relatively high *probability of acceptance.*

ANNEX A

(to Supplement 6)

Alphabetical list of definitions contained in this Supplement

ANNEX B

(to Supplement 6, Part I)

Relations between defect, failure and fault concepts

TABLE B-1

Defect	Failure	Fault
Critical defect	Critical failure	Critical fault
Non-critical defect	Non-critical failure	Non-critical fault
Major defect	–	Major fault
Minor defect	–	Minor fault
–	Misuse failure	Misuse fault
–	Mishandling failure	Mishandling fault
–	Inherent weakness failure	Inherent weakness fault
Design defect	Design failure	Design fault
Manufacturing defect	Manufacturing failure	Manufacturing fault
–	Ageing failure	Ageing fault
–	Sudden failure	–
–	Gradual failure	–
–	Cataleptic failure	–
–	Relevant failure	–
–	Non-relevant failure	–
–	Primary failure	–
–	Secondary failure	–
–	Failure cause	–
–	Failure mechanism	–
–	–	Programme-sensitive fault
–	–	Data-sensitive fault
–	–	Complete fault
–	–	Partial fault
–	–	Persistent fault
–	–	Intermittent fault
–	–	Fault mode
–	–	Determinate fault
–	–	Indeterminate fault
–	–	Latent fault
–	Systematic failure	Systematic fault
Bug	–	–

ANNEX C

(to Supplement 6, Part I)

List of recommended symbols and abbreviations

α	Significance level
β	Type II risk
$\lambda(t)$	Instantaneous failure rate
$\bar{\lambda}(t_1, t_2)$	Mean failure rate [in time interval (t_1, t_2)]
$\mu(t)$	Instantaneous repair rate
$\bar{\mu}(t_1, t_2)$	Mean repair rate [in time interval (t_1, t_2)]
δ	Standard deviation
A	Asymptotic availability
$A(t)$	Instantaneous availability
$\bar{A}$	Asymptotic mean availability
$\bar{A}(t_1, t_2)$	Mean availability [in time interval (t_1, t_2)]
ASR	Answer seizure ratio
$E(X)$	Mean (of X)
$f(x)$	Probability density function
$F(x)$	Distribution function
FMEA	Fault modes and effect analysis
FMECA	Fault modes, effects and criticality analysis
FTA	Fault tree analysis
H_0	Null hypothesis
H_1	Alternative hypothesis
MAD	Mean administrative delay
MADT	Mean accumulated down time
MART	Mean active repair time
MDT	Mean down time
MLD	Mean logistic delay
MMH	Maintenance man-hours
MRT	Mean repair time
MTBF	Mean time between failures
MTTF	Mean time to failure
MTTFF	Mean time to first failure
MTTR	Mean time to restoration
MUT	Mean up time
$N(t_1, t_2)$	Number of failures [in time interval (t_1, t_2)]
R	Reliability
U	Asymptotic unavailability
$U(t)$	Instantaneous unavailability
$\bar{U}$	Asymptotic mean unavailability
$\bar{U}(t_1, t_2)$	Mean unavailability [in time interval (t_1, t_2)]
$z(t)$	Instantaneous failure intensity
$\bar{z}(t_1, t_2)$	Mean failure intensity [in time interval (t_1, t_2)]

GUIDE FOR EVALUATING AND IMPLEMENTING ALTERNATE ROUTING NETWORKS

A systematic procedure consisting of a number of distinct steps is used for the evaluation of alternate routing networks.

These steps are given in the flowchart of Annex A and are provided as guidance. Administrations may wish to expand, delete or change the order of these steps to meet circumstances.

The steps may be grouped into the following six processes:

- Identification of alternate route.
- Preliminary screening.
- Data gathering.
- Evaluation.
- Implementation.
- Monitoring.

1 Identification of alternate route

A terminal Administration selects an alternate route.

A tentative agreement is reached with the opposite terminal Administration to use the selected alternate route, and both terminal Administrations reach tentative agreement with the transit Administration to explore the use of its network as an alternate route.

If no tentative agreements are reached, another alternate route is selected or if none is available, the procedure is abandoned.

2 Preliminary screening

Using available data, the organizational elements of the terminal Administrations responsible for transmission, routing and call completion, analyse the feasibility of utilizing the alternate route.

If an objection is raised, another alternate route is selected or, if none is available, the procedure is abandoned.

3 Data gathering

A questionnaire is issued to all Administrations involved to obtain additional information before an evaluation is made of the proposed alternate route.

The questionnaire can include requests for transmission, routing, call completion rates, traffic profiles, circuit costs, and transit charges.

If there is no response to the questionnaire or if the information provided indicates that the alternate route is unsuitable, another alternate route is selected or, if none is available, the procedure is abandoned.

4 Evaluation

The alternate routing network is dimensioned according to Recommendation E.522.

If additional circuits are required on the alternate route, and the required increment exceeds the available capacity, another alternate route is selected. If no other alternate route is available Administrations may choose to retain the selected alternate route and accept a cost disadvantage.

5 Implementation

Final negotiations are carried out and approval of all Administrations involved in the alternate route network is sought.

The negotiations would include the reporting procedure and responsibility for recording traffic overflowing to the alternate route.

If final agreement cannot be reached, another alternate route is selected or, if none is available, the procedure is abandoned.

6 Monitoring

Traffic volumes and performance data for the alternate route are recorded and exchanged at regular intervals.

ANNEX A

(to Supplement No. 7)

Flowchart of evaluation and implementation procedure for alternate routing networks

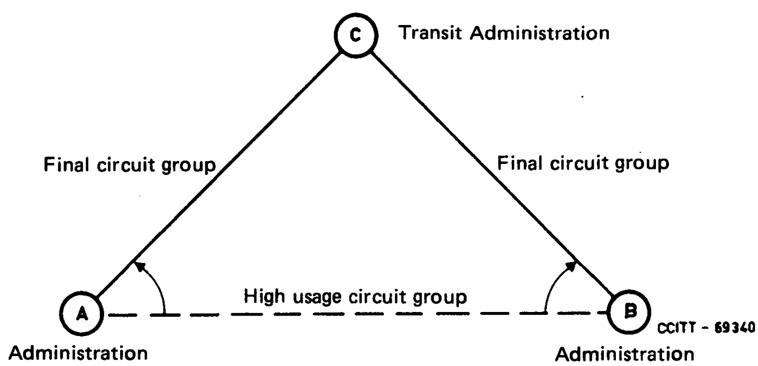
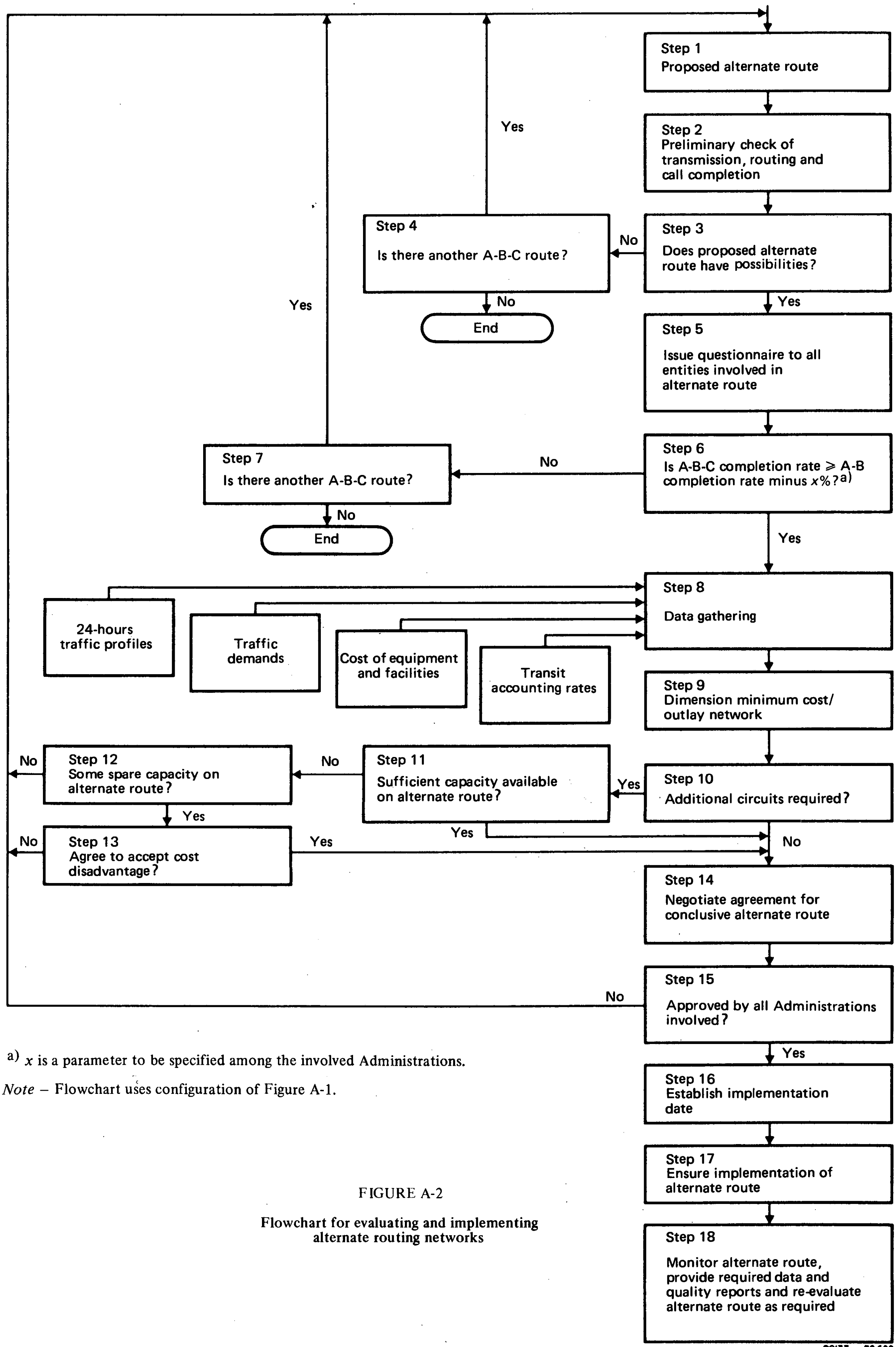


FIGURE A-1

Alternate routing network



a) x is a parameter to be specified among the involved Administrations.

Note – Flowchart uses configuration of Figure A-1.

FIGURE A-2

Flowchart for evaluating and implementing alternate routing networks

ISBN 92-61-03271-0