

国际电信联盟

ITU-T

国际电信联盟
电信标准化部门

E.806

(06/2019)

E系列：综合网络运行、电话业务、业务运行和人为因素
电信业务质量：概念、模型、指标和可靠性规划 – 与电信业务质量相关的术语和定义

测量活动、监测系统和抽样方法，以监测移动网络中的服务质量

ITU-T E.806建议书



国际操作	
定义	E.100-E.103
有关主管部门的一般规定	E.104-E.119
有关用户的一般规定	E.120-E.139
国际电话业务的操作	E.140-E.159
国际电话业务的编号方案	E.160-E.169
国际选路方案	E.170-E.179
用于国内信令系统的信令音	E.180-E.189
国际电话业务的编号方案	E.190-E.199
水上移动业务和公众陆地移动业务	E.200-E.229
国际电话业务中与计费 and 结算有关的操作规定	
国际电话业务的计费	E.230-E.249
为结算目的对呼叫时长的测量和记录	E.260-E.269
利用国际电话网作非话应用	
概述	E.300-E.319
传真电报	E.320-E.329
有关用户的ISDN规定	E.330-E.349
国际选路方案	E.350-E.399
网络管理	
国际业务统计	E.400-E.404
国际网络管理	E.405-E.419
国际电话业务质量检测	E.420-E.489
业务工程	
话务的测量和记录	E.490-E.505
业务预测	E.506-E.509
确定人工操作的电路数量	E.510-E.519
确定自动和半自动操作的电路数量	E.520-E.539
服务等级	E.540-E.599
定义	E.600-E.649
IP网络的业务工程	E.650-E.699
ISDN业务工程	E.700-E.749
移动网络业务工程	E.750-E.799
电信业务质量：概念、模型、指标和可靠性规划	
与电信业务质量相关的术语和定义	E.800-E.809
电信业务的模型	E.810-E.844
电信业务的业务质量指标和相关概念	E.845-E.859
业务质量指标在电信网络规划设计中的使用	E.860-E.879
设备、网络和业务的性能的现场数据收集和评估	E.880-E.899
其他	E.900-E.999
国际操作	
国际电话业务的编号方案	E.1100-E.1199
网络管理	
国际网络管理	E.4100-E.4199

如果需要进一步了解细目，请查阅ITU-T建议书清单。

ITU-T E.806建议书

测量活动、监测系统和抽样方法， 以监测移动网络中的服务质量

摘要

ITU-T E.806建议书描述了测量移动网络服务质量（QoS）最佳实践的基线框架。它提供了测量活动、监测系统特性和要求、后处理一般建议以及监测移动电子服务的抽样方法的高级概述。

本建议书技术中立，但可能会根据所测量的服务陈述不同的要求。

历史沿革

版本	建议书	批准时间	研究组	唯一ID*
1.0	ITU-T E.806	2019-06-29	12	11.1002/1000/13924

关键词

关键性能指标（KPI）、测量活动、移动网络、监测系统、服务质量（QoS）、抽样方法。

* 欲查阅建议书，请在您的网络浏览器地址域键入URL <http://handle.itu.int/>，随后输入建议书的唯一识别码，例如，<http://handle.itu.int/11.1002/1000/11830-en>。

前言

国际电信联盟（ITU）是从事电信和信息通信技术（ICT）领域工作的联合国专门机构。国际电信联盟电信标准化部门（ITU-T）是国际电信联盟的常设机构，负责研究技术、操作和资费问题，并且为在世界范围内实现电信标准化，发表有关上述研究项目的建议书。

每四年一届的世界电信标准化全会（WTSA）确定ITU-T各研究组的研究课题，再由各研究组制定有关这些课题的建议书。

WTSA第1号决议规定了批准ITU-T建议书须遵循的程序。

属ITU-T研究范围的某些信息技术领域的必要标准，是与国际标准化组织（ISO）和国际电工技术委员会（IEC）合作制定的。

注

本建议书为简要而使用的“主管部门”一词，既指电信主管部门，又指经认可的运营机构。

遵守本建议书的规定是以自愿为基础的，但建议书可能包含某些强制性条款（以确保例如互操作性或适用性等），只有满足所有强制性条款的规定，才能达到遵守建议书的目的。“须”或“必须”等其他一些强制性用语及其否定形式被用于表达特定要求。使用此类用语不表示要求任何一方遵守本建议书。

知识产权

国际电联提请注意：本建议书的应用或实施可能涉及使用已申报的知识产权。国际电联对无论是其成员还是建议书制定程序之外的其他机构提出的有关已申报的知识产权的证据、有效性或适用性不表示意见。

至本建议书批准之日止，国际电联尚未收到实施本建议书可能需要的受专利保护的知识产权的通知。但需要提醒实施者注意的是，这可能并非最新信息，因此，特大力提倡他们通过下列网址查询电信标准化局（TSB）的专利数据库：<http://www.itu.int/ITU-T/ipr/>。

© 国际电联 2020

版权所有。未经国际电联事先书面许可，不得以任何手段复制本出版物的任何部分。

目录

页码

1	范围	1
2	参考文献	1
3	定义	2
3.1	他处定义的术语	2
3.2	本建议书中定义的术语	2
4	缩略语和首字母缩写词	3
5	惯例	3
6	监测服务质量的测量活动	3
6.1	测量环境	4
6.2	测试方法	4
6.3	在移动网络上测量服务质量参数的指南	6
7	监测系统特性和要求	10
8	对于后处理的一般建议	11
9	抽样方法	12
附件A	– 获取代表性结果的统计指南	14
A.1	用于监测国家层面服务质量的代表性样本	13
A.2	假设检验	15
参考书目	17

ITU-T E.806建议书

测量活动、监测系统和抽样方法， 以监测移动网络中的服务质量

1 范围

本建议书描述了测量全行业服务质量（QoS）的最佳实践的基线框架，涵盖了移动网络QoS测量活动、监测系统的特性和要求、后处理方案，以及供在国家层面监测QoS的监管机构、测试设备供应商、提供网络测量数据的公司、数据分析师和服务提供商使用的抽样方法。

2 参考文献

下列ITU-T建议书和其他参考文献的条款，在本建议书中的引用而构成本建议书的条款。在出版时，所指出的版本是有效的。所有的建议书和其他参考文献均会得到修订，本建议书的使用者应查证是否有可能使用下列建议书或其他参考文献的最新版本。当前有效的ITU-T建议书清单定期出版。本建议书引用的文件自成一体时不具备建议书的地位。

- [ITU-T E.802] ITU-T E.802建议书（2007年），确定和应用QoS参数的框架和方法。
- [ITU-T E.804] ITU-T E.804建议书（2014年），移动网络流行业务的服务质量问题。
- [ITU-T E.807] ITU-T E.807建议书（2014年），用于处理蜂窝移动语音业务呼叫的以用户为中心的参数的定义、相关测量方法和指导目标。
- [ITU-T E.840] ITU-T E.840建议书（2018年），用于端对端网络性能基准评分和排名的统计框架。
- [ITU-T G.1031] ITU-T G.1031建议书（2014年），网络浏览中的体验质量（QoE）因素。
- [ITU-T P.863.1] ITU-T P.863.1建议书（2019年），ITU-T P.863建议书应用指南。
- [ITU-T P.1401] ITU-T P.1401建议书（2012年），客观质量预测模型的静态评估、验资和比较的方法、矩阵和流程。
- [ITU-T Y.1540] ITU-T Y.1540建议书（2016年），网际协议数据通信业务—IP分组传送和可用性性能参数。
- [ITU-T Y.1545.1] ITU-T Y.1545.1建议书（2017年），监测互联网协议网络服务质量的框架。
- [ETSI TR 125 942] ETSI Technical Report 125 942, V15.0.0 (2018), *Universal mobile telecommunications system (UMTS); Radio frequency (RF) system scenarios*.
- [ETSI TR 138 900] ETSI Technical Report 138 900, V15.0.0 (2018), *LTE; 5G; Study on channel model for frequency spectrum above 6 GHz*.
- [IETF RFC 7799] IETF RFC 7799 (2016), *Active and passive metrics and methods (with hybrid types in-between)*.

3 定义

3.1 他处定义的术语

本建议书使用下列他处定义的术语：

3.1.1 A方 (A-party) [ITU-T E.804]：一个连接的发起方（也称：移动发端，MO）或者在直接事务处理中，发起事务的一方（呼叫方）。

注 – 在存储转发事务处理中，发送内容的一方。

3.1.2 B方 (B-party) [ITU-T E.804]：在直接事务处理中，事务处理中终止或对应的一方。

注 – 在存储转发事务处理中，接收内容的一方。

3.1.3 置信系数 (confidence coefficient)；置信水平 (confidence level) [b-ITU-T E.800]：与置信区间或统计容限间隔相关的概率值。

3.1.4 下载 (download) [ITU-T E.800]：服务器或主机向个人计算机或设备传送数据或程序。

3.1.5 端对端质量 (end-to-end quality) [b-ITU-T E.800]：与通信系统（包括所有终端设备）性能相关的质量。

3.1.6 衡量单位 (measure) [b-ITU-T E.800]：可表示一参数的单位。

3.1.7 网络运营商 (network operator) [ITU-T E.804]：为提供公共电信服务而提供网络的组织。

3.1.8 网络性能 (network performance) [b-ITU-T E.800]：网络或部分网络提供用户之间通信功能的能力。

3.1.9 探测器 (probe) [ITU-T Y.1545.1]：是一个使用探测信息包来收集测量数据的端点测试工具。

3.1.10 服务质量 (quality of service) (QoS) [b-ITU-T E.800]：一种电信业务的特性总和，表明其满足明示和暗示业务用户需求的能力。

3.1.11 流媒体 (streaming) (多媒体服务) [ITU-T E.800]：数据包流形式传送的多媒体数据（通常为语音、文本、视频和音频的组合），软件应用在数据包到达时对其进行解释和渲染。

注1 – 流媒体是传送多媒体数据的技术。

注2 – 流媒体可能是实时的，也可能不是实时的。

3.2 本建议书中定义的术语

本建议书定义了以下术语：

3.2.1 众包数据收集 (crowdsourced data collection)：一种从大量终端用户装置收集主动或被动的服务质量测量数据的方法。

3.2.2 驾驶测试 (drive testing)：一种使用配备网络装置检测设备的车辆来进行检测的方法。

3.2.3 监测系统 (monitoring system)：一个能够对网络性能进行评估的工具或设备。

3.2.4 监测系统用户 (monitoring system user) : 可以是监管机构、测试设备供应商、提供网络测量数据的公司、数据分析师和服务提供商。

3.2.5 无人值守探测器 (unattended probe) : 可被安装在室内或者户外地点或者车辆中的远程管理的移动或固定监测系统 (智能手机、单sim卡装置、双sim卡装置、虚拟sim卡装置)。

3.2.6 步行测试 (walk testing) : 一种使用装置、手持测试设备的测试方式, 通常在无法使用车辆的时候采用。

4 缩略语和首字母缩写词

本建议书使用下列缩略语和首字母缩写词:

2G	第二代
3G	第三代
4G	第四代
BS	基站
BTS	基站收发信台
GIS	地理信息系统
GPS	全球定位系统
HTTP	超文本传输协议
KPI	关键性能指标
MMS	多媒体信息服务
MNO	移动网络运营商
OMC	户外测量活动
QoS	服务质量
SINR	信号与干扰噪声比
SMS	短信息服务
UPS	不间断电源

5 惯例

无。

6 监测服务质量的测量活动

QoS测量活动旨在收集信息, 以便从关键性能指标 (KPI) 的角度描述移动网络QoS的特性, 进行基准测试, 并评估对于现行立法承诺的遵守情况。

与地理足迹、时间框架和测试频率相关的QoS测量的预定范围主要取决于测量活动的目的和测量环境。

6.1 测量环境

对于移动网络来说，测量环境可被分为两个主要类别：室内和户外。测量可以同时包括室内和户外、静止和移动方案，从而覆盖终端用户使用其移动服务的所有不同环境。目前，高比例的移动流量产生于室内环境；因此，建议除了户外之外，还要测量室内QoS性能。

6.1.1 室内测试

以下建议的方法可帮助监测系统用户辨别室内QoS状态：

- 步行测试；
- 无人值守探测器；
- 众包数据收集。

6.1.2 户外测试

以下建议的方法可帮助监测系统用户辨别户外QoS状态：

- 驾驶/步行测试；
- 无人值守探测器；
- 众包数据收集。

第6.2条提供了对于监测系统用户可用来进行室内和户外测试的四种不同测量方法的高层次描述。

6.2 测试方法

6.2.1 步行测试

火车站、地铁站台、机场、体育馆、商场、大学校园和步行区等地点正在成为无线通信的热点。在户外测量中，步行测试通常在无法使用车辆的时候采用。在为多个运营商进行基准测试或测试涵盖多项服务和无线接入技术时，多装置、手持测试设备提供了一种方便的工作方法。

但是，由于物理参数（例如空间维度、设备重量和技术要求（比如天线隔离））的原因，步行测试方法在样本数量、移动网络运营商（MNO）和服务方面存在一些限制，最终影响抽样精度。建议在发起步行测试进行测量之前采取以下措施：

- 1) 绘制建筑物布局蓝图（如果是室内地点的话）。
- 2) 计划并起草一份目标地点或位置的清单。
- 3) 每个服务或MNO的目标样本大小取决于设备可检测到的智能手机的数量。但是，样本大小和分布必须根据考虑的变量类型和所针对的统计表述等来选择。例如，覆盖更广大的区域或测量更长时间可以提供更具统计代表性的结果。参考附件A和[ITU-T E.840]。
- 4) 采取和执行一般的常规程序来检查硬件或软件功能。
- 5) 考虑使用轻型、便携的控制单元来报告实时结果或测量设备状态。
- 6) 须禁用自动网络选择功能，设定为网内（on-net）功能。

6.2.2 驾驶测试

驾驶测试测量活动要预先规划好流程，以满足活动的目标和范围。目标服务、人口分布和社会经济因素等会施加不同的活动设计参数。从本质上来说，成功的驾驶测试活动须考虑样本大小，从而使测量的数据能够代表目标区域的人口。

为了执行驾驶测试测量活动，监测系统用户须考虑以下建议：

- 1) 必须对样本大小进行选择，从而使结果能够代表该研究区域的移动网络行为。为此，监测系统用户必须定义一个抽样方法。可参考附件A和[ITU-T E.802]。
- 2) 规划需要初始无线电覆盖足迹或MNO/技术。
- 3) 用于执行测量的天线须被安装在人类平均高度上。
- 4) 若进行的是监管机构基准测试，须对所有接入技术和所有MNO同时进行随机测量。
- 5) 须建立测量轮廓（技术参考见[ITU-T E.804]）。
- 6) 如果驾驶测试被用于推导一个广阔区域的KPI，则路线须覆盖人类活动区域，避免路线重复。
- 7) 对于在移动中评估的服务，车辆速度的设置必须考虑一台装置被安放在一个固定点，同时另一台装置处于运动中。
- 8) 须禁用自动网络选择功能，设定为网内（on-net）功能。
- 9) 在活动开始之前，须研究目标区域的人口分布，以确保：
 - a. 收集的样本考虑到住宅和商业集中度；
 - b. 考虑到次区域的发展因数。
- 10) 须优先考虑营业日和白天的测试间隔。

6.2.3 无人值守探测器

无人值守探测器可提供接近实时和历史端到端QoS性能，可被用于收集有助于探测QoS下降情况的精细数据。建议监测系统用户考虑以下操作和设计建议来部署使用探测器进行的测量活动。

- 1 无人值守探测器须被安全地安装在所选择的地点或车辆上，并配有电源供应；地点或车辆的选择取决于测量目的。
- 2 室内无人值守探测器须被安装在具备适当无线电覆盖条件和高网络使用的地方。取决于测量范围，如果监测的是峰值网络性能，可安装在尽可能具备最佳覆盖的位置。如果无人值守探测器要监测的是最小无线电覆盖，可以考虑不具备最佳覆盖条件的位置。
- 3 所有无人值守探测器须在全球定位系统（GPS）（或其他适合的全球导航卫星系统）的连接下操作。
- 4 每个目标位置的无人值守探测器的数量取决于抽样要求、需要评估的服务或运营商数量、用户密度和正在进行的活动（音乐会、体育比赛等）。

已知高人类活动地点的网络使用可发生很大变化，在虑及高效利用测量资源的需要时，建议监测系统用户使用可管理的无人值守探测器，一旦某个地点为认为是活跃热点时可远程激活。

6.2.4 众包数据收集

众包数据收集活动可被用于获取只有收集到来自不同终端用户具有代表性数量的样本才有意义的QoS测量数据。为了对结果有更深入了解，可使用具备数据传送功能的终端用户设备来检查环境状况。建议测量结果只有在对于各测量目标来说具有代表性的情况下才被用于得出结论。必须虑及数据间隙限制。众包数据收集解决方案必须符合国家数据保护立法，保证没有个人数据会被不当处理。

与传统的（模拟收集）QoS样本收集不同，众包样本收集频率可以是连续的，因此该测量可以成为24小时×7天QoS测量值的来源。

众包数据收集大致可以分为两类，即主动和被动测量（见[IETF RFC 7799]）：

注 – [ITU-T E.804]中探讨了可以用来收集关于网络QoS的连续信息的另一个来源。监管机构也许可以访问网络性能计数器信息（取决于各国的法律框架）。根据数字保护立法，并且如果双方都有意愿，运营商和监管机构可达成如前所述的保密协议。

6.2.4.1 主动测量

主动数据测量通常是数据速度测试和由终端用户发起或进行脚本编辑的各应用的特定测试。主动方式测量会创建人工流量来确定网络/应用能力。

在使用主动众包方法测量网络的容量、能力或性能时，建议监测系统用户根据位于网内和网外（例如，在最近的互联网交换点）的目标服务器采用自动安排脚本来进行QoS测量。同时，必须考虑服务器的可用带宽和链路的提供，以防测量受到由服务器或测试链路降级带来的贻误。

不仅如此，对于监管机构开发的、可能被集成进运营商SIM卡或独立的可下载应用中的众包应用来说，避免将所有终端用户的测量安排在同一时间以防止可能出现的堵塞十分重要。

6.2.4.2 被动测量

被动数据测量要么是应用特定，要么与应用无关，不需要任何形式的终端用户干涉。被动测量不向网络注入人工流量或测试负荷，旨在基于终端用户的活动来测量QoS。

这一方式移除了对服务器和链路的依赖，可利用主机移动应用程序（例如，交通应用、政务应用、MNO应用）来提供测量数据点的具有代表性的地理分布。

建议监测系统用户考虑来自唯一识别符的更多样本（终端用户），因为它们为监测系统用户提供更高的网络数据准确度。

6.3 在移动网络上测量服务质量参数的指南

本节提供了关于监测一些相关参数的测量活动的进一步信息。本节列出的参数可通过第6.2节描述的任何方法进行测量，仅可作为示例考虑；更多参考见[ITU-T E.804]。

6.3.1 无线电覆盖

覆盖测量是在给定地理位置为每个无线电接入网络技术分别提取接收到的信号电平。这些测量使用频率扫描仪自动执行。

建议监管机构考虑使用驾驶测试活动之前提交、获取或生成的MNO数据的目标区域的覆盖地图。覆盖地图反映MNO提供服务的地区。从这个意义上来说，监管机构可以使用这些地图来规划测量活动和选择感兴趣的区域。例如，监管机构可以选择测量投诉率最高的区域，或者，为了进行比较，选择所有MNO都有覆盖的区域。还可以在测量活动中对覆盖地图进行验证。对于4G及以上网络，服务覆盖是要测量的另一个相关参数。覆盖和服务覆盖的测量方法可参考[b-ECC Report 103]、[b-ECC Report 256]、[b-ECC Report 118]、[b-ECC Report 231]和[b-ECC Rec (12) 03]。

监测系统用户可使用不同方式来反映和可视化移动网络覆盖。第6.3.1.1和6.3.1.2条描述了可被独立或结合使用的两种方法，取决于监测系统用户的需要和可用的信息。

6.3.1.1 理论或分析模型测量方法

本方法主要基于以用于预测网络覆盖和性能的模拟工具为基础的数学和统计学计算的表述方法。这一测量中涉及的一些参数包括基站（BS）位置、信号与干扰噪声比（SINR）、BS功率、服务SINR阈值、室内覆盖信号的衰减和每个服务的天线信息或敏感度。

如果关于MNO基础设施的信息可用，那么监测系统用户可使用此类模型。模型仅可作为参考，因为模型是基于统计预测的。更多关于传播模型的信息，见[ETSI TR 125 942]、[ETSI TR 138 900]和它们的参考文献。

6.3.1.2 地理信息系统

地理信息系统（GIS）使高级映射和空间分析在地图上形成可视化地理信息。有了这类工具，使用空间位置集成不同类型的数据层成为可能，因为这是MNO的覆盖服务地图。此类信息层可以是：

- MNO的基站收发信台（BTS）层；
- QoS测量的输出结果（按比例缩放的样本）；
- 众包系统的具体覆盖或网络指标；
- 人口和管理层；
- 地形和海拔层。

不仅如此，GIS生成层可以很容易通过网站进行共享，可以使用网页浏览器、手机和平板电脑访问。

由于这个原因，建议监测系统用户接入与GIS工具兼容的覆盖地图，为终端用户提供明确牢靠的信息，并且信息要以清晰、友好和可理解的方式呈现。

通过使用GIS系统，监测系统用户可运行分析，从而确定，例如要测量的地理区域、为驾驶测试规划线路和进行人口覆盖分析等。GIS分析的示例见[b-ECC Report 103]、[b-ECC Report 256]、[b-ECC Report 118]、[b-ECC Report 231]和[b-ECC Rec (12) 03]。

GIS系统应能够执行并、交、叠、面定义（点、多边形），生成热图和专题图。此外，须允许运行不同分辨率的网格分析来表征地理区域，以显示KPI的变动性和允许对人口和杂波图进行分析。

6.3.2 语音通话测量

语音服务测量包括发起一系列自动选择无线电接入技术的通话尝试。通话尝试由不同场景生成，可以是移动到移动、固定到移动，或移动到固定。适合语音测量的样本必须根据场景来选择。信息见[ITU-T P.863.1]和[ITU-T E.807]。

图1中的主叫语音通话测试的时间线方案仅为示例，并未规定任何规格。除了无线网络特性之外，相关的时间间隔参数和整体窗口测试时间也会影响样本的总获取数。

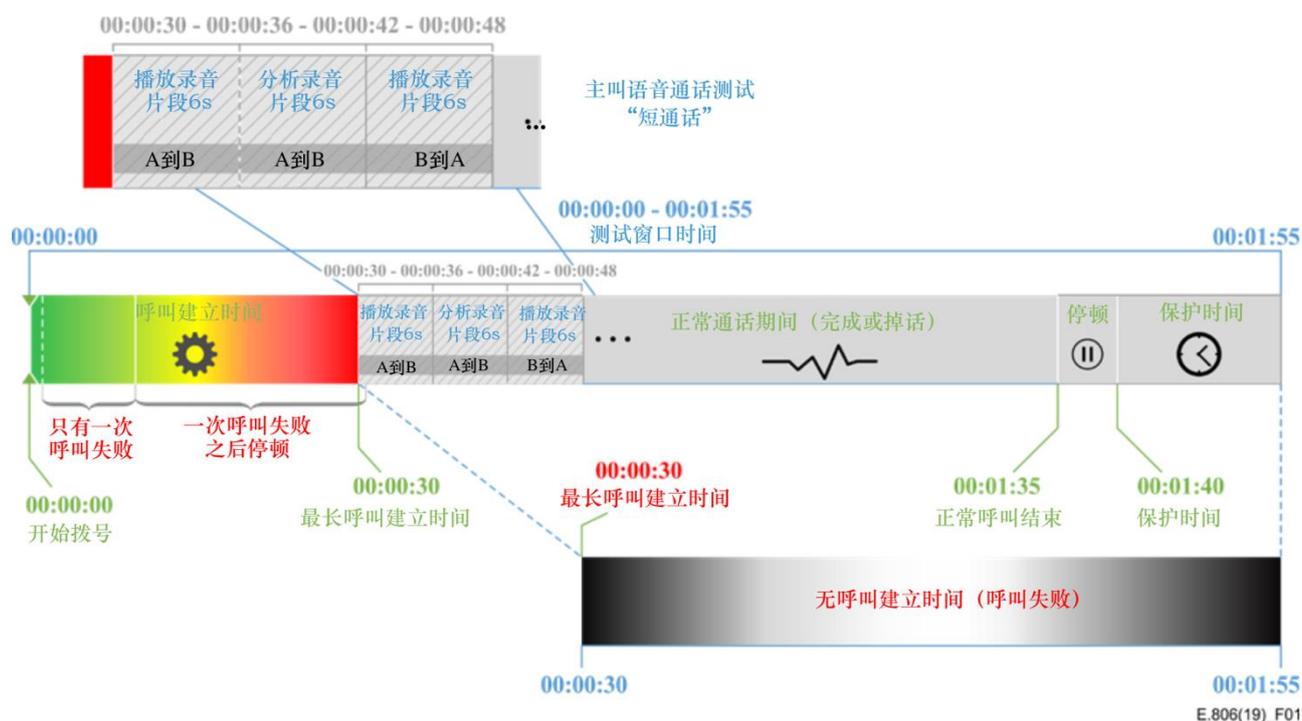


图1 – 呼叫测量时间线

描述的方案表示了一个语音通话测试程序的示例；虽然正常通话时长可能不一，取决于测试目的（短通话、长通话），但对于测量测试期间面对的每个案例来说，几乎所有其他属性都是相同的。

此外，每个国家行为的具体特征都应被考虑到，因为这会影响到方案设计。例如，如果根据统计，A国的平均通话时间为2分钟，而B国为4分钟，那么每个掉话的使用分钟数就更高。

在测试开始点之前，呼叫方（A）开始给预先设定的被叫方（B）拨打电话，在这时，允许使用固定持续时间窗口作为建立与B方连接的最大呼叫建立时间，在这期间测试在等待B方接听测试呼叫。如果B方未接听，或者呼叫方在测试通话建立阶段出现问题，测试通话流程进入“无呼叫建立时间”，测试通话将被标记为“呼叫失败”。

6.3.3 短信息和多媒体信息测量

短信息服务（SMS）和多媒体信息服务（MMS）测量可以以当终端用户的移动终端设备不断变更接入技术时，不强制移动终端设备使用特定的接入技术来模拟类似场景的情况下执行测量。

测量方式是从一个模拟移动用户的移动探测器向一个模拟属于同一运营商的另一个移动用户的固定探测器发送一条由固定数量的字母数字字符构成的SMS或固定大小的MMS。如果送达时间短于设定的最长时间，则认为该SMS/MMS已被接收。

图2展示了一个用于SMS测量的参考时间线图。



图2 – 短信息服务测量时间线

语音服务和SMS/MMS的测量可以同时进行，从而更好地管理时间资源。

6.3.4 宽带数据测量

宽带数据测量活动应在不同的监测系统上进行平衡，以减少窗口测试框架和增加样本数量；这可以使测试地图区域的覆盖面积更广。

场景设置取决于被测试的服务应用。图3显示了对数据业务进行评估的一个时间线图示例。更多信息见[ITU-T Y.1540]、[ITU-T Y.1545.1]。

注 – 当测量涉及具体的服务器资源、用户设备或无人值守探测器，或者多个网络以获取想要的内容或测试服务器，则无法区分它们对于结果的作用和移动网络性能。见[ITU-T G.1031]第6条此类对结果产生作用的影响因素的示例清单。

第6.3.4.1到6.3.4.3条是宽带数据测量的示例。

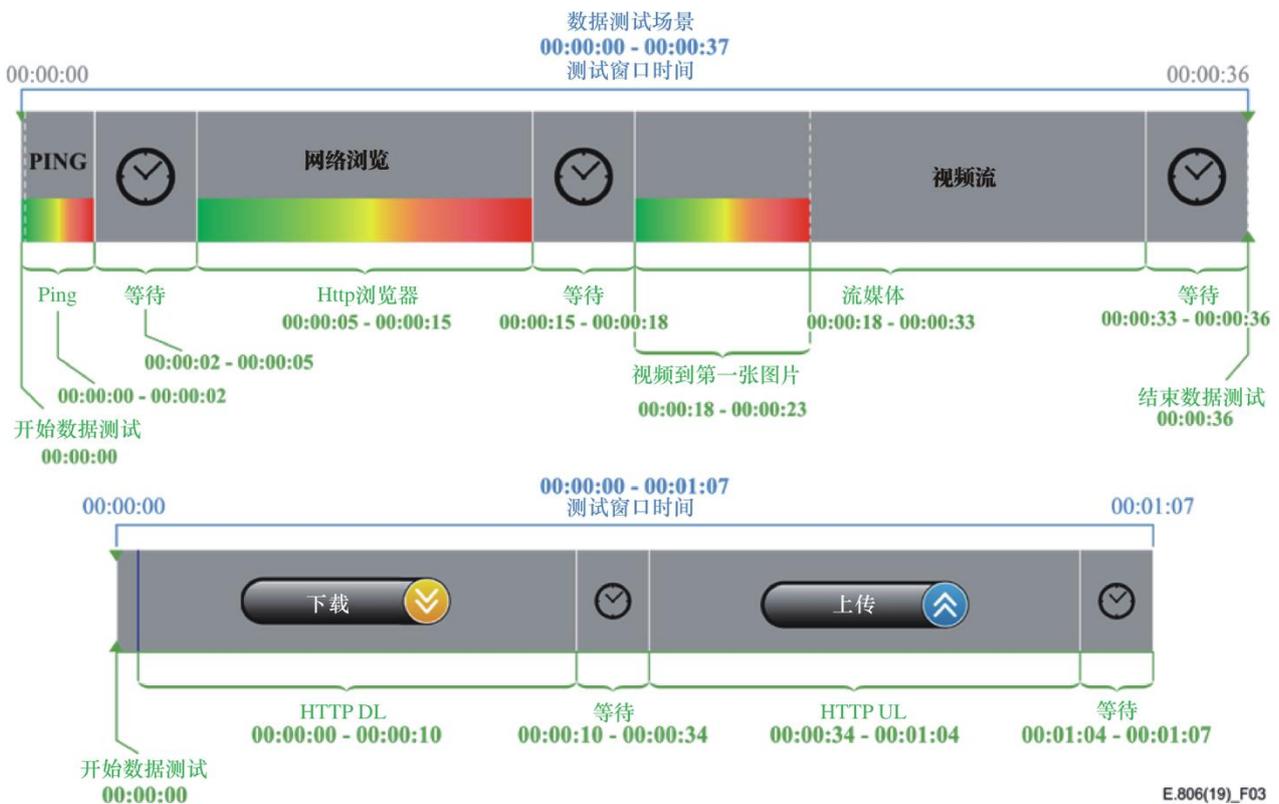


图3 – 宽带数据测量示例

6.3.4.1 网络浏览测量

与网络浏览相关的测量的特点之一是根据页面加载时间以及页面下载的整体成功率来测量导航性能。平均固定时段的网络浏览测试从移动设备或模拟移动用户的探测器在静止或移动环境中向一个预先确定的建议的网站集合发起。访问尝试在不同的测试站点之间随机分布。

6.3.4.2 文件传输测量

文件传输测量方法是发送和接收由一个固定地点的移动设备或模拟移动用户的探测器和一个配备专门用于该测量的资源的服务器之间的上行链路和下行链路信道传输的一个固定大小的文件。可以测量下载会话成功率、平均下载时间或平均上传时间等KPI（但对两个方向允许的完成时间都存在限制）。

6.3.4.3 音频或视频流测量

对于音频或视频流服务，访问寄放在流服务器上的音频或视频文件集的成功率可通过从一个固定地点的模拟移动用户的移动探测器向预先设定的一系列流服务器发起的平均固定时间段测试来测量。此外，在初始视频加载期间，暂停事件出现的频率和时长以及整体通量可被用作流性能的指标。

访问尝试在不同的测试服务器之间随机分布。

7 监测系统特性和要求

监测系统须能够评估网络性能下降如何对终端用户的服务质量造成影响，从而对网络覆盖性能、容量和端到端QoS提供指示。

以下监测系统特性适用于第6.2条描述的除众包数据收集之外的所有测量方法，除非另有说明。

- 1) 能够同时测量所有接入技术。
- 2) 如有需要，能够同时测量语音、SMS和数据等。
- 3) 须能够在运动中、静止地点或结合两种形式进行测量。
- 4) 须自动存储测量数据，并有额外的外部存储作为备份。
- 5) 须装有GPS（或其他适用的全球导航卫星系统）来辨别和记录每个测量的地理位置和速度。
- 6) 须通过考虑到了所有适用法规认证要求的维护计划来保持最佳操作状态。例如，应对扫描设备进行校准。校准须在测量活动时期内有效。
- 7) 天线必须被放置在合适的距离，以避免干扰。
- 8) 须能够对所有接入技术进行测量，因此需要更新以满足技术演进的需要。
- 9) 须允许对测量脚本模板进行设置并重复使用。
- 10) 须使用与终端用户在MNO商店通常购买的移动设备类似的、经认证的移动设备。
- 11) 须支持所有接入技术的语音编解码器。
- 12) 须能够提供视觉或声音警报，从而对测量设备的任何故障发出通知；在这种情形下做出的测量须被舍弃。

- 13) 测量设备须提供记录每个测量的日期、时间和地理参考位置的功能（作为附加功能或使用附加设备）。
- 14) 最理想的是，测量设备至少能够提供测量活动开始和结束、中断、设备故障和异常情况（例如非典型的用户集中度、故意损坏、干扰或自然灾害）的记录。
- 15) 须记录任何可能潜在地影响结果并且对MNO网络没有帮助的情况。
- 16) 上述提到的测量设备记录的所有活动须与日志文件和在这些时间间隔内进行的测量相互关联，并且须在后处理阶段舍弃。
- 17) 监测系统须生成加密的日志文件。
- 18) 无人值守探测器须配备自监测功能（例如，温度感应器或足够的电力供应水平）以检测异常操作条件。须具备相应的后处理来确保数据质量，并检测和排除数据假象。
- 19) 无人值守探测器须具备能够对抗外部电源供应中断的鲁棒性（例如，配有供电缓冲来确保安全的关闭程序和可靠的开启程序）。如果无人值守探测器对此不支持，须考虑不间断电源（UPS），以确保稳定的操作条件。这对于避免日志文件损坏十分重要。
- 20) 无人值守探测器和后端基础设施须提供适当的监测功能，取决于测量数据的传输方式（例如，通过监测的网络）和想要生成的信息类型。例如，无人值守探测器可以被设置为定期传送测量的数据，或者每隔一段时间发送状态报告，如果此类信息逾期，或者测量数据表明选择的网络参数在预先定义的范围之外，则后端程序会生成警报。

注 – 众包数据收集监测系统的特性需要进一步研究。

8 对于后处理的一般建议

对结果进行后处理的第一步是确定考虑提供移动通信网络的电子业务的无线电覆盖限制。必须为每个技术（例如2G、3G、4G）分别设置这一限制。这些限制将决定在导出最终结果时哪些测试将被纳入考虑。还必须规定何时必须将一个测量纳入结果的考虑中。这类例子有：如果在测量期间，有一段时间不符合无线电覆盖限制，那么，要将测试纳入结果中的话，时间段应为多久？

所有监测系统用户在对所有测量方法进行后处理行动期间应执行以下一般建议。

- 1) 为了检查在测量期间发生的事件，信令信息须被纳入考虑。测量结果的后处理须考虑和管理在测量活动期间获取的所有日志文件。建议将所有软件错误考虑在内。在测量活动中收集到的最终测量数量将对结果的估计误差造成影响。可基于样本大小计算估计误差。如果误差大于预期，建议对目标区域进行重复测量。附件A和[ITU-T E.802]、[ITU-T E.840]和[ITU-T P.1401]提供了更多信息。
- 2) 执行QoS测量的监测系统用户须检查测量数据的完整性，并清洗数据。异常样本会由于多种原因产生：例如，由于设备或软件故障或失灵导致设备告警，网元升级、安全事件或维护。
- 3) 须对所有从测试设备中收集到的日志文件进行检查，看是否有遗失或者损坏。
- 4) 设定被接受的测量活动中收集到的虚假或错误样本的百分比阈值至关重要。除了众包数据收集方法之外，确定是否需要重新测量以维护结果的准确性和完整性的需要十分重要。

9 抽样方法

测量可在室内和户外场景之间分布，取决于每个电信市场用户的密度或行为。测量的数量必须以统计上的相对准确度小于所确定的最大值的方式来确定，并且置信水平至少为95%。[ITU-T E.802]中提供了更多信息。

可用来获得具有代表性的样本的一个“两步走”方法涉及的步骤是分层和简单随机抽样。

分层法可被用于计算在一个测量活动中要覆盖的地理区域的数量（例如城市、市镇或区），以获得代表国家层面的网络QoS结果。从这个意义上来说，这是一个在无法测量一个广阔区域（例如，国土面积广大的国家）时获取有代表性的结果的有用工具。

之后，可使用简单随机抽样来计算通过抽样方法选择出的每个地理区域要执行的测量数量。附件A提供了更多详情。

重要的是要考虑从每个测量活动中获得的结果，以推导网络在性能方面的行为。可从这些结果中计算平均值和标准差值，并与统计公式共同使用以监测QoS参数。

为执行对广阔区域的测量，建议监测系统用户将基域分成几个部分，根据确定哪些地方更有必要进行更多测量（取决于测量活动目的）的标准，为每一部分分配一个权重。这些标准的一些示例包括人口权重、电信密度、流量密度和终端用户密度；为了确定这些标准，十分建议监测系统用户从官方来源获取最新数据。

基准性能结果的发布在本建议书范围之外。不过，[ITU-T E.840]中提供了旨在对网络性能进行评分和排名的统计框架。

附件A

获取代表性结果的统计指南

（本附件是本建议书不可分割的组成部分。）

监测系统用户可根据在测量活动中应用的抽样方法，并使用遵照基于假设检验的流程的统计推断来获取有代表性的网络性能QoS结果。

A.1 用于监测国家层面服务质量的代表性样本

建议想要通过户外测量活动（OMC）在国家层面监测QoS的监测系统用户使用的统计建模方法基于一个“两步走”模型：

- 分层随机抽样；
- 简单随机抽样。

第一步是选择该国要测量的地理区域（在此情况下，地理区域指的是城市、市镇、村镇或道路），第二步是为第一步中选择出的每个区域确定样本大小（例如，在每个城市进行多少次语音通话）。在每个区域内的测量分布可根据测量目的进行选择。例如，如果不需要进行区分，测量可以均匀分布，或者可以加权，给城市区域更多重要性。

A.1.1 分层随机抽样

从统计的角度来看，分层法被用于产生一个在估计误差方面比一个相同大小的简单随机样本单独产生的误差更小的限制范围。在监测QoS时，分层法是一个可被用作减少测量区域数量但仍可获得具有代表性结果的工具。例如，开车穿越地域辽阔的国家不太可能，所以可以将该地区分成各个小组（层），在每组中测量一个样本。

从这个意义上来说，分层法允许将地理区域分成同质或非同质的组或者层。基于这些组对于被测KPI具有直接影响共同特征对这些组进行定义。同质层的例子包括但不限于具备类似人口或基站密度、流量需求、干扰电平、站间距离等的地理区域（城市或村镇）等。然后，通过从该国的地理区域中选择非重叠的群来构造层（例如，如果根据人口密度来选择层，则该国的聚居区可被分成城市或农村，每个聚居区被分给一个群，即一个聚居区不可以同时既是城市又是农村）。最后，每个测量层中的地理区域必须是随机选择的，以确保结果的有效性。

以下是分层抽样的统计方程，包括一个国家的城市和农村聚居区的示例。

在活动期间要测量的地理区域的数量可从以下方程中推导出（一个待测国家的聚居区数量）：

$$n = \frac{(\sum_{i=1}^L N_i \sigma_i)^2}{N^2 D + \sum_{i=1}^L N_i \sigma_i^2}$$

其中：

L 是层的总数（用于划分国家的各群；如果划分为城市和农村，L = 2）

σ_i 是*i*层的预计标准差；

N_i 是每个层中的地理区域数量（被划分为城市或农村的聚居区数量）；

N 是地理区域的总数（一个国家聚居区的总数）；

$D = \frac{B^2}{4}$, 其中 B 是估算误差的限制范围

$$n_i = n \left(\frac{N_i \sigma_i}{\sum_{i=1}^L N_i \sigma_i} \right) \quad i = 1, 2, 3$$

其中

n_i 是 i 层待测地理区域的数量（对于 $i = 1$ （城市），待测城市聚居区的数量为 n_i ）；

n 是待测地理区域的总数（待测城市和农村聚居区的数量）。

A.1.2 简单随机抽样

在选择了该国的待测地理区域数量之后，第二步是采用简单随机抽样来确定之前选择的每个地理区域的样本大小，即，需要进行多少次语音通话、数据测量或发送多少SMS。

在第二步数据建模中，简单随机抽样被用于确定在定义的置信水平上和估计误差范围内测量某一KPI所需的事件数量。

若研究的参数是一个平均值，则按照以下方程计算样本大小

$$m_i = \frac{z_{1-\alpha/2}^2}{a^2} \cdot \left(\frac{\sigma_i}{\bar{x}_i} \right)^2$$

其中：

m_i 是 i 层的样本大小；

$z_{1-\alpha/2}$ 是标准正态分布的百分位 $1 - \alpha/2$ ；

$1 - \alpha$ 是置信水平；

a 是估计误差的限制范围；

\bar{x}_i 是 i 层观察参数的平均值；

σ_i 和 \bar{x}_i 从以前的测量活动中计算得出。

若研究的参数是一个比例，则每个KPI的样本大小须使用以下方程计算：

$$n \geq \frac{k^2 * P * (1 - P)}{d^2}$$

其中：

n 是样本大小；

k 是置信水平；

P 是比例目标值（研究的KPI的阈值）；

d 是最大可接受估计误差（实际值 P 与其估计值 p 之间的差异）。

对于每个层，与从测量活动中获得的 i 层对应的标准差和 σ_i 平均值 \bar{x}_i 可通过以下方程计算：

$$\sigma_i^2 = \frac{\sum_{k=1}^{n_i} (x_k - \bar{x}_i)^2}{n_i - 1}$$
$$\bar{x}_i = \frac{\sum_{k=1}^{n_i} x_k}{n_i}$$

$$x_k = \frac{\sum_{j=1}^{m_i} p_j}{m_i}$$

其中：

σ_i^2 是*i*层观察参数的方差；

x_k 是与*i*层对应的每个*n_i*样本的平均值；

P_j 是在大小为*m_i*的样本范围内测量数据*i*的值。

为了确定运营商是否达到某一阈值，建议对所有层使用基于测量结果的假设检验。

A.2 假设检验

假设检验是一个统计推断工具，用于确定从选择的样本中获得的结果从统计角度来看是否有意义，表明MNO遵循最小QoS阈值。

为实施统计假设检验，确定测量活动结果的加权值的平均和标准偏差：

相对于总人口数*N*，基于每*N_i*层的人口定义权重：

$$w_i = N_i / N$$

以下方程提供了每个*i*层基于平均值的KPI平均和标准偏差的加权值：

$$\bar{x} = \sum_{i=1}^L w_i \bar{x}_i$$

$$\sigma^2 = \sum_{i=1}^L w_i^2 \sigma_{\bar{x}_i}^2$$

在计算了测量结果的平均和标准偏差之后，假设检验的第一步是确定一个零假设和一个备择假设。之后，统计假设检验确定零假设是否正确（有一定概率），或者备择假设是否被接受。

考虑到测量结果，假设检验可以以已确定的总平均值和KPI阈值为基础。零假设（ H_0 ）表明MNO遵循最低KPI阈值（ μ ）；与此同时，备择假设（ H_a ）表明MNO未遵守该阈值。统计假设检验显著性水平为 α 。

因此，假设可被表达为：

$$H_0: x_{st} \geq \mu$$

$$H_a: x_{st} < \mu$$

根据从测量活动的测量样本中获得的数据，统计假设检验可确定应该接受还是拒绝零假设。

下面是一个示例，显示当监管机构以比例的形式确立KPI阈值时，如何应用假设检验。关于统计评估的不同度量指标和流程的更多信息和参考，见[ITU-T E.840]、[ITU-T P.1401]及其参考。

通过比较检验统计量(x_{st})来做出决定。检验统计量可从临界值 $z_{1-\alpha}$ 中计算得出：

$$x_{st} = \frac{\bar{x} - \mu}{\sigma}$$

其中：

μ 是监管机构确立的KPI阈值（平均值）；

$$X = \frac{p_{st} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

其中：

p 是监管机构确立的KPI阈值（比例）；

$Z_{1-\alpha}$ 是与显著性水平为 α 的标准正态分布相对应的z值；

n 是从每个QoS参数的测量活动中获得的样本总数（考虑所有层）。

最后，如果检验统计量(x_{st})大于或等于临界值 $z_{1-\alpha}$ ，那么，从统计上来说，没有足够的信息来拒绝显著水平为 α 的零假设（这意味着MNO遵循监管机构确立的阈值，无须施加任何惩罚）；否则，零假设被拒绝，备择假设被接受（MNO未遵循监管机构确立的阈值）。

参考书目

- [b-ITU-T E.800] ITU-T E.800建议书（2008年），有关服务质量的术语定义。
- [b-ECC Report 103] Electronic Communications Committee within the European Conference of Postal and Telecommunications Administrations (CEPT) Report 103 (2007), *UMTS coverage measurements*.
- [b-ECC Report 256] Electronic Communications Committee within the European Conference of Postal and Telecommunications Administrations (CEPT) Report 256 (2016), *LTE coverage measurements*.
- [b-ECC Report 118] Electronic Communications Committee within the European Conference of Postal and Telecommunications Administrations (CEPT) Report 118 (2008), *Monitoring methodology to assess the performance of GSM networks*.
- [b-ECC Report 231] Electronic Communications Committee within the European Conference of Postal and Telecommunications Administrations (CEPT) Report 231 (2015), *Mobile coverage obligations*.
- [b-ECC Rec (12) 03] Electronic Communications Committee within the European Conference of Postal and Telecommunications Administrations (CEPT) Recommendation (12) 03 (2013), *Determination of the radiated power through field strength measurements in the frequency range from 400 MHz to 6000 MHz*.

ITU-T 系列建议书

- 系列 A ITU-T 工作的组织
- 系列 D 资费及结算原则和国际电信/ICT 的经济和政策问题
- 系列 E 综合网络运行、电话业务、业务运行和人为因素**
- 系列 F 非话电信业务
- 系列 G 传输系统和媒介、数字系统和网络
- 系列 H 视听及多媒体系统
- 系列 I 综合业务数字网
- 系列 J 有线网络和电视、声音节目及其他多媒体信号的传输
- 系列 K 干扰的防护
- 系列 L 环境与 ICT、气候变化、电子废物、节能；线缆和外部设备的其他组件的建设、安装和保护
- 系列 M 电信管理，包括 TMN 和网络维护
- 系列 N 维护：国际声音节目和电视传输电路
- 系列 O 测量设备的技术规范
- 系列 P 电话传输质量、电话设施及本地线路网络
- 系列 Q 交换和信令，以及相关的测量和测试
- 系列 R 电报传输
- 系列 S 电报业务终端设备
- 系列 T 远程信息处理业务的终端设备
- 系列 U 电报交换
- 系列 V 电话网上的数据通信
- 系列 X 数据网、开放系统通信和安全性
- 系列 Y 全球信息基础设施、互联网的协议问题、下一代网络、物联网和智慧城市
- 系列 Z 用于电信系统的语言和一般软件问题