HELLENIC TELECOMMUNICATIONS & POST COMMISSION

TELECOMMUNICATION STANDARDIZATION SECTOR OF ITU

ITU-T

E.806 (06/2019)

SERIES E: OVERALL NETWORK OPERATION, TELEPHONE SERVICE, SERVICE OPERATION AND HUMAN FACTORS
Quality of telecommunication services: concepts, models, objectives and dependability planning – Terms and definitions related to the quality of telecommunication services

Measurement campaigns, monitoring systems and sampling methodologies to monitor the quality of service in mobile networks

Dr. Stavroula Bouzouki

Head of Regional Office of Hellenic Telecommunication & Post Commission-EETT

ITU Workshop on Telecommunication Service Quality as Enabler of the Digital Economy

Singapore, 19-21 August 2019

Scope

This Recommendation describes a baseline framework of:

- best practices for measuring quality of service (QoS) throughout the industry
- mobile network QoS measurement campaigns
- characteristics and requirements for monitoring systems
- post-processing scenarios
- sampling methodologies

used by ***regulators***, ***test equipment vendors***, ***companies*** that deliver network measurements, ***data analysts and service providers*** to monitor QoS on a national level.

Measurement campaigns to monitor QoS

QoS measurement campaigns aim to gather information to characterize the QoS of mobile networks in terms of:

- key performance indicators (KPIs)
- perform benchmarking
- evaluate conformance to existing legislation commitments.

The intended scope of QoS measurements ***related*** to geographical footprint, timeframe and testing frequency ***depends*** mainly on the purpose of the measurement campaign and the measurement environment

NOTE: The result produced by testing for each of these purposes is different

Measurement environment

For mobile networks, measurement environments can be divided into two main categories:

- ✓ Indoor
- ✓ Outdoor

Measurements can include both indoor and outdoor, ***stationary and mobile scenarios*** thus covering all the different environments in which end users make use of their mobile service.

Currently, there is a high rate of mobile traffic generated from indoor environments; hence, it is advisable to measure indoor QoS performance, in addition to outdoor.

Measurement methodologies

Indoor testing

- walk testing
- unattended probes
- crowdsourced data collection

Outdoor testing

- drive testing
- walk testing
- unattended probes
- crowdsourced data collection

Walk testing (1/2)

Where: Locations such as train stations, underground train platforms, airports, sports stadiums, shopping malls, university campuses and pedestrian zones are becoming hotspots for wireless communication.

In outdoor measurements, walk testing is normally employed where ***vehicles cannot be used***.

Measurement setup: Multi-device, carry-on testing equipment provides a convenient way to work when benchmarking multiple operators or testing coverage for multiple services and radio access technologies.

Walk testing measurement campaigns have some ***limitations*** in terms of:

1. number of samples and of mobile network operator (MNO)

2. service due to physical parameters such as spatial dimensions, equipment weight and technical requirements, like antenna isolation, which ultimately can influence sample resolution.

Walk testing (2/2)

The following actions are recommended before initiating a walk testing measurement.

1 Draft blueprints of building layouts, in the case of an indoor location.

2 Schedule and draft a list of targeted spots or locations.

3 The targeted sample size per service or MNO depends on the number of smartphones that can be accommodated by the equipment. However, the sample size and distribution should be chosen according to, among others, the type of variable under consideration and the statistical representation being targeted.

For example, covering a wider area or measuring over a longer period could provide more statistically representative results. See Annex A and [ITU-T E.840] for references.

4 Adopt and conduct a general routine procedure to check hardware or software functionality.

5 Consider a light portable control unit to report real-time results or status of the measurement equipment.

6 The automatic network selection feature should be blocked and set to the on-net feature.

Drive testing (1/3)

Drive testing measurement campaigns involve preplanning procedures to satisfy the goal and scope of the campaign. Targeted services, demographic distribution and socio-economic factors can impose different campaign design parameters. Essentially, a successful drive testing campaign shall consider a sample size so that the measured data are representative of the targeted region's population.

Drive testing (2/3)

In order to perform drive testing measurement campaigns, users should consider the following recommendations:

1 The sample size should be chosen so that the results are representative of the behaviour of the mobile networks in the area under study. For this purpose, monitoring system users must define a sampling methodology. References can be found in Annex A and [ITU-T E.802].

2 An initial radio coverage footprint or MNO/technology is required for planning.

3 The antennas used to perform the measurements should be installed at average human height.

4 In the case of a regulator's benchmarking, measurements should be performed randomly, for all access technologies and for all MNOs simultaneously.

5 Measurement profiles should be established (technical references can be found in [ITU-T E.804]).

6 The routes should cover areas with human activity, avoiding route repetition if drive testing is used to derive KPIs for a wide area.

Drive testing (3/3)

7 For services that are evaluated in movement, the speed of the vehicle should be set considering that one device can be located at a fixed point while another device will be in motion.

8 The automatic network selection feature should be blocked and set to the on-net feature.

9 The targeted region's population distribution should be studied prior to the campaign to guarantee that:

a. the collected sample may consider residential and commercial concentrations;

b. development factors of sub-regions may be considered.

10 Business days and daytime testing intervals should be considered referentially.

Unattended probes

Unattended probes may provide ***near real time and historical*** end-to-end QoS performance and can be used to collect granular data that can help to detect QoS degradations

Users should consider the following operating and design recommendations to deploy a probe-based measurement campaign.

1 Unattended probes should be installed ***securely in chosen locations*** or vehicles with power supply; the choice of these locations or vehicles depends on the measurement objective.

2 Indoor unattended probes should be installed in locations with ***adequate radio coverage conditions*** and include high network usage. Depending on the measurement scope, this may be a location with the best possible coverage if peak network performance is monitored. If unattended probes are intended to monitor minimum radio coverage, a location with non-optimal coverage conditions can be considered.

3 All unattended probes should ***operate under a global positioning system*** (GPS) (or another suitable global navigation satellite system) connection.

4 The ***number*** of unattended probes per targeted location ***depends on the sampling requirements***, the ***number of services or operat***ors to be evaluated, ***user density*** and ***occurring events*** (concerts, sports tournaments, etc.).

Crowdsourced data collection (1/2)

Crowdsourced data collection campaigns can be used to obtain QoS measurement data that can only be meaningful if a representative number of samples from different end users are gathered.

It is recommended that measurements results be ***used for any conclusion*** only if they are ***representative for the respective measurement goa***l. Data gap limitations have to be taken into account.

Crowdsourced data collection solutions must be in compliance with national data protection legislation, guaranteeing that no personal data is mishandled.

The sample collection frequency can be continuous unlike the conventional (simulating collection) QoS sample collection, therefore crowd-sourced measurements can be a source of 24 hours by 7 days QoS measurements.

Crowdsourced data collection can be largely categorized into two categories namely, ***active and passive measurements*** (see [IETF RFC 7799]):

NOTE – Alternatives sources to collect continuous information on network QoS are discussed in [ITU-T E.804]. Depending on the legal framework in each country, regulators may have access to network performance counters information. In accordance to data protection legislation and where both parties so desire, operators and regulators may want to enter into non-disclosure agreements as per the foregoing.

Crowdsourced data collection (2/2)

Active Measurements

Active data measurements are typically data speed tests and application-specific tests that are end user initiated or scripted. The active approach measurement creates artificial traffic to determine the network/application capability.

Passive measurements

Passive data measurements are either application-specific or application-agnostic and do not require any form of end user intervention.

Passive measurements do not inject artificial traffic or test payload into the network and aim at measuring QoS based on the end users' activity

Guidelines for measuring quality of service parameters on mobile networks

This section provides further details on the measurement campaigns for monitoring some relevant parameters. The parameters listed here can be measured by any of the methodologies described before and should be considered as examples only; more references can be found in [ITU-T E.804].

Radio coverage (1/2)

The monitoring of the radio coverage can be performed with two approaches:

- Through measurement
- Through theoretical calculations

The measurement results can be used as verification of the calculated results

The coverage measurements consist of taking the ***received signal level*** at a given geographical position for each radio access network ***technology separately***.

These measurements are carried out automatically using a frequency scanner.

References for measurement methodologies for coverage and service coverage can be found in [b-ECC Report 103], [b-ECC Report 256], [b-ECC Report 118], [b-ECC Report 231] and [b-ECC Rec (12) 03]

Radio coverage (2/2)

Theoretical or analytical modelling measurement methodology

This method is mostly based on formulation of mathematical and statistical computations based on simulation tools for predicting network coverage and performance.

A monitoring system user could use such models if information about the MNO infrastructure is available and should be used only as a reference since it is based on statistical predictions. For more information on propagation models, see [ETSI TR 125 942], [ETSI TR 138 900] and the references therein.

Using GIS systems, a monitoring system user may run analysis to determine, for example, the geographical regions to measure, plan routes for drive testing, and perform population coverage analysis.

Examples of GIS analysis are given in [b-ECC Report 103], [b-ECC Report 256], [b-ECC Report 118], [b-ECC Report 231] and [b-ECC Rec (12) 03].

A GIS system should be able to perform union, intersection, overlapping and areal definition (points, polygons), generate heat maps and thematic maps. Moreover, it should allow running grid analysis in different resolutions to characterize geographical regions, to show the variability of a KPI and allow the analysis of population and clutter maps

Voice call measurement

Voice service measurements include the launch of a series of call attempts for automatic selection of radio access technologies. Call attempts are generated from different scenarios, can be mobile to mobile, fixed line to mobile, or mobile to fixed line. An appropriate sample for voice measurements should be chosen according to the scenario. Information can be found in [ITU-T P.863.1] and [ITU-T E.807].

Additionally, specific characteristics per country behaviour should be considered as it can impact the scenario design, for example, if statistically the average call time for country A is 2 min while country B is 4 min, the minutes of usage per dropped call would be higher.

Short and multimedia message measurement

Short message service (SMS) and multimedia messaging service (MMS) measurements can be executed without forcing the mobile terminal equipment to a particular access technology to simulate a similar scenario when the end user mobile terminal equipment is continually changing access technology.

Measurements consist of sending an SMS with a fixed number of alphanumeric characters and fixed size for MMS from a mobile probe simulating a mobile subscriber to a fixed one simulating another mobile subscriber belonging to the same operator. The SMS/MMS is considered as received if the delivery time is less than the maximum time established.

Broadband data measurement

Broadband data measurement campaigns should be balanced on different monitoring systems to reduce the window test frame and increase the number of samples; this will enable greater coverage of the testing map area.

The setup for the scenarios will depend on the service application being measured.

More information can be found in [ITU-T Y.1540], [ITU-T Y.1545.1].

NOTE – When measurements involve specific server resources, user equipment or unattended probes, or multiple networks to reach the desired content or test server, then their contribution to results cannot be distinguished from mobile network performance. See clause 6 of [ITU-T G.1031] for an example list of such influence factors contributing to results.

Characteristics and requirements for monitoring systems (1/3)

Monitoring systems should be able to ***perform an evaluation*** on how ***degraded*** network ***performance*** affects ***service quality to the end user***. It therefore gives an indication of network coverage performance, capacity and end-to-end QoS.

Characteristics and requirements for monitoring systems (2/3)

The following monitoring systems characteristics apply to all measurement, unless otherwise stated, except for crowdsourced data collection.

- Ability to execute simultaneous measurements in all access technologies.
- Ability to execute simultaneous measurements, e.g., voice, SMS and data, if needed.
- It should be possible to carry out the measurements in movement, at static locations or combining both modalities.
- It should automatically store data from measurements and should count with additional external storage for backup.
- It should have a GPS (or another suitable global navigation satellite system) to identify and register the geographical location and speed for each measurement.
- It should be kept in optimal operation conditions through maintenance programmes considering all applicable regulation certification requirements.
- Antennas must be placed at an appropriate distance to avoid interference.
- It should be able to execute measurements for all access technologies and therefore be updated to meet the needs of technological evolution.
- It should allow the setup and reuse of measurement script templates.
- It should use certified mobile equipment similar to those regularly purchased in MNO shops by end users.
- It should support voice codecs for all access technologies.

Characteristics and requirements for monitoring systems (3/3)

- It should be able to produce visual or audio alarms to notify any failure in the measuring equipment
- Measuring equipment should provide a feature, as an added functionality or additional equipment, to register the day, hour and geo-referenced position of each measurement
- It is desirable that the measurement equipment provide the ability to register, at least, the start and the end of the measurement campaign, breaks, equipment failures and extraordinary situations such as atypical end user concentration, vandalism, interference or natural disasters
- Any situation that could potentially affect the results and is not attributed to the network of the MNO should be registered
- All activities registered by the measurement equipment, mentioned in the foregoing, should be correlated with the log files and measurements taken during these time intervals and should be discarded during the post-processing stage
- Monitoring systems should generate encrypted log files.
- Unattended probes should be equipped with self-monitoring functionality (e.g., temperature sensors or adequate power supply level) to detect unusual operating conditions
- Unattended probes should be robust against loss of external power supply, e.g., have power supply buffering to ensure secure turn-off processes and reliable turn-on processes
- Depending on the way measured data are transferred (e.g., via the monitored network), and the type of information that is expected to be created, unattended probes and back-end infrastructure shall provide adequate monitoring functionality

NOTE – The characteristics for crowdsourced data collection monitoring systems need further study.

General recommendations for post processing (1/2)

The first step in the post processing of results is to ***determine the radio coverage limit*** for which the electronic service of a mobile communications network is considered to be provided.

This limit should be set for each technology (e.g., 2G, 3G, 4G) separately.

These limits will determine which tests will be taken into consideration when exporting the final results.

Also, rules should be laid down as to when a measurement should be taken into account in the results. Such rules are, for example, if during a measurement there is a period of time that does not meet the radio coverage limit, how long should that period be for the test to be taken into account in the results.

General recommendations for post processing (2/2)

The following general recommendations should be practised by all monitoring system users during post-processing action for all measurement methodologies.

- In order to examine the events that occurred during the measurements, signalling information should be taken into account. Post processing of measurements results should consider and manage all log files obtained during the measurement campaign. It is recommended that all software errors be considered. The final number of measurements gathered in the measurement campaign will have an impact on the estimation error of the results. The estimation error can be calculated based on the sample size. If it is larger than desired, it is recommended that the measurement campaign for the target area be repeated. More information is provided in Annex A and [ITU-T E.802], [ITU-T E.840] and [ITU-T P.1401].
- Users conducting QoS measurements should check measurement data integrity and cleanse data. Outlier samples can be generated for many reasons, e.g., equipment alarms due to equipment or software failure or malfunction, network element upgrade, incidents or maintenance.
- All log files gathered from the testing equipment should be checked, if missing or corrupted.
- It is essential to set the accepted percentage threshold for fake or false samples collected in a measurement campaign.

Sampling methodologies (1/2)

The number of measurements must be determined in such a way that the statistical relative accuracy is less than the maximum value established and for at least a 95% level of confidence. More information is provided in [ITU-T E.802].

A two-step methodology that can be used to obtain representative samples involves stratification and simple random sampling.

- Stratification can be used to calculate the number of geographical areas (e.g., cities, municipalities or districts) to be covered during a measurement campaign to get results that represent the network QoS at a national level. In this sense, stratification is a tool useful to obtain representative results when it is not possible to measure a large area (e.g., large countries).
- Simple random sampling can then be used to calculate the number of measurements to perform at each of the geographical areas that were selected through sampling methods. More details are provided in Annex A.

Sampling methodologies (2/2)

To perform measurements over a large area, monitoring system users are advised to:

1. divide the base zone into parts, and
2. allocate a weight to each part in accordance with criteria that will identify where it is more relevant to carry out a greater number of measurements, depending on the purpose of the measurement campaign.

Some examples of criteria are:

✓ population weight
✓ telecommunications density
✓ traffic density
✓ end user density

In order to define these criteria, it is highly advisable that monitoring system users obtain updated data from an official source.

Publication of benchmarking performance results lies outside the scope of this Recommendation; however, a statistical framework aimed to score and rank network performance can be found in [ITU-T E.840].

References

[ITU-T E.802] Recommendation ITU-T E.802 (2007), *Framework and methodologies for the determination and application of QoS parameters*.

[ITU-T E.804] Recommendation ITU-T E.804 (2014), *Quality of service aspects for popular services in mobile networks.*

[ITU-T E.807] Recommendation ITU-T E.807 (2014), *Definitions, associated measurement methods and guidance targets of user-centric parameters for call handling in cellular mobile voice service.*

[ITU-T E.840] Recommendation ITU-T E.840 (2018), *Statistical framework for end-to-end network performance benchmark scoring and ranking.*

[ITU-T G.1031] Recommendation ITU-T G.1031 (2014), *QoE factors in web-browsing*.

[ITU-T P.863.1] Recommendation ITU-T P.863.1 (2019), *Application guide for Recommendation ITU-T P.863*.

[ITU-T P.1401] Recommendation ITU-T P.1401 (2012), *Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models*.

[ITU-T Y.1540] Recommendation ITU-T Y.1540 (2016), *Internet protocol data communication service – IP packet transfer and availability performance parameters*

[ITU-T Y.1545.1] Recommendation ITU-T Y.1545.1 (2017), *Framework for monitoring the quality of service of IP network services*.

[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)*.

Annex A
Statistical guidelines to obtain representative results

(This annex forms an integral part of this Recommendation)

A monitoring system user can obtain representative results of the QoS performance of the network based on sampling applied during measurement campaigns together with the use of statistical inference by following a procedure based on hypothesis testing.

Bibliography

- [b-ITU-T E.800] Recommendation ITU-T E.800 (2008), *Definitions of terms related to quality of service*.
- [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*.