International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.1545.2

(07/2022)

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS, NEXT-GENERATION NETWORKS, INTERNET OF THINGS AND SMART CITIES

Internet protocol aspects – Quality of service and network performance

---

Quality of service metrics for continuity of performance of packet data-based services

Recommendation ITU-T Y.1545.2

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS, NEXT-GENERATION NETWORKS, INTERNET OF THINGS AND SMART CITIES

GLOBAL INFORMATION INFRASTRUCTURE	
General	Y.100–Y.199
Services, applications and middleware	Y.200–Y.299
Network aspects	Y.300–Y.399
Interfaces and protocols	Y.400–Y.499
Numbering, addressing and naming	Y.500–Y.599
Operation, administration and maintenance	Y.600–Y.699
Security	Y.700–Y.799
Performances	Y.800–Y.899
INTERNET PROTOCOL ASPECTS	
General	Y.1000–Y.1099
Services and applications	Y.1100–Y.1199
Architecture, access, network capabilities and resource management	Y.1200–Y.1299
Transport	Y.1300–Y.1399
Interworking	Y.1400–Y.1499
Quality of service and network performance	**Y.1500–Y.1599**
Signalling	Y.1600–Y.1699
Operation, administration and maintenance	Y.1700–Y.1799
Charging	Y.1800–Y.1899
IPTV over NGN	Y.1900–Y.1999
NEXT GENERATION NETWORKS	
Frameworks and functional architecture models	Y.2000–Y.2099
Quality of Service and performance	Y.2100–Y.2199
Service aspects: Service capabilities and service architecture	Y.2200–Y.2249
Service aspects: Interoperability of services and networks in NGN	Y.2250–Y.2299
Enhancements to NGN	Y.2300–Y.2399
Network management	Y.2400–Y.2499
Computing power networks	Y.2500–Y.2599
Packet-based Networks	Y.2600–Y.2699
Security	Y.2700–Y.2799
Generalized mobility	Y.2800–Y.2899
Carrier grade open environment	Y.2900–Y.2999
FUTURE NETWORKS	Y.3000–Y.3499
CLOUD COMPUTING	Y.3500–Y.3599
BIG DATA	Y.3600–Y.3799
QUANTUM KEY DISTRIBUTION NETWORKS	Y.3800–Y.3999
INTERNET OF THINGS AND SMART CITIES AND COMMUNITIES	
General	Y.4000–Y.4049
Definitions and terminologies	Y.4050–Y.4099
Requirements and use cases	Y.4100–Y.4249
Infrastructure, connectivity and networks	Y.4250–Y.4399
Frameworks, architectures and protocols	Y.4400–Y.4549
Services, applications, computation and data processing	Y.4550–Y.4699
Management, control and performance	Y.4700–Y.4799
Identification and security	Y.4800–Y.4899
Evaluation and assessment	Y.4900–Y.4999

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

Recommendation ITU-T Y.1545.2

Quality of service metrics for continuity of performance of packet data-based services

Summary

Recommendation ITU-T Y.1545.2 specifies spatially resolved metrics for packet data-based services and a methodology for their computation, using the same conceptual framework as Recommendation ITU-T G.1034. The metrics cover the quality of service (QoS) and quality of experience (QoE) aspects of a wide range of applications used in motion, i.e., during travel. A taxonomy of application categories is provided, taking into account their absolute data rate (DR) requirements and, in particular, their sensitivity against temporary drops in available DRs, which are caused by motion through an environment that is characterized by spatial variation of network performance, i.e., available DR or latency. The methodology establishes ways to create a description of such spatial distributions of performance, termed route profiles and use route profiles to create predictions of QoS and QoE of application usage. It also establishes a new entity to describe local network performance that provides an abstraction and thereby a versatile way to express performance requirements.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T Y.1545.2	2022-07-29	12	11.1002/1000/15013

Keywords

End-to-end, mobile communication, packet data services, QoE, QoS, spatially resolved information.

---

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

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

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

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

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

NOTE

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

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

INTELLECTUAL PROPERTY RIGHTS

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

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Introduction

[ITU-T G.1034] specifies metrics for mobile telephony communication during rail travel, based on actual measurement data from telephony tests on trains.

[ITU-T G.1034] introduces two metrics: local drop probability (LDP) describes local call stability; and call completion probability (CCP) predicts expected call stability for telephone calls of selected length.

In many services using packet-data connectivity, continuity of performance is also key to perceived quality. The range of services that suffer from insufficient data rate (DR) – the DR zero in case of complete loss of connectivity being just an extreme case – starts with real-time services such as video conferencing or gaming where even short dips of DRs cause drastic degradations in quality of service (QoS). The case of video streaming is a bit more complex than for adaptive streaming, no direct relation between bit rate changes and QoS change may exist, and in any case, buffering can conceal insufficient DRs for a few seconds. Even then, however, perceived quality will suffer during longer periods of poor packet data performance.

The categorization of application builds upon, and extends, the terminology and categories established in [ITU-T G.1010], and uses the definition framework for hand-over processes provided by [ITU-T Y.1546].

NOTE –The term 'service' is used as the subject of a QoS-related view. However, be aware that the subject of user experience is actually an application that uses primary wireless packet data services. For completeness, mobile networks are not the only means to provide such primary services. Connectivity can also be provided by, for example, a network of wireless fidelity (Wi-Fi) hotspots with respective hand-over functionality, or a Wi-Fi hotspot in a vehicle that is connected via single or multiple modems to either public mobile networks or trackside infrastructure.

Fifth generation is expected to be an enabler for higher bitrate packet data-based service offerings. Correspondingly, continuity of performance will increase in importance wherever use cases include mobility. In particular, autonomous driving may lead to a large boost in usage of such services, as passengers on board of autonomous vehicles will significantly add to demand. The proposed open-model framework allows ease of addition of QoS and quality of experience predictions for new services and is also open to the integration of multi-level quality rating scales, as well as to the addition of more dimensions of performance indicators, e.g., round-trip times or packet loss.

As in the case of telephony on-board vehicles moving along a predefined track, user experience can vary widely depending on the actual distribution of DRs along the route. Conventional key performance indicators, such as the mean DR, aggregated over a longer period of time or a larger area, can only give a very rough estimate of the actual QoS. In analogy to telephony, exploiting the special conditions in the case of movement along predefined routes allows the derivation of enhanced metrics, again with a solid foundation in actually measured data. To also support this view, the concept of a local expected data rate is introduced, computed from an abstraction of the various methods of sample pre-selection and aggregation used.

This Recommendation builds on the concepts and definitions given in [ITU-T G.1034] and establishes a set of respective metrics for QoS of services depending on continuity of packet data performance. It also builds upon [ITU-T G.1010] and evolves the taxonomy further to reflect ongoing developments in popular use cases.

Recommendation ITU-T Y.1545.2

Quality of service metrics for continuity of performance of packet data-based services

1 Scope

This Recommendation describes an extension of the metrics provided in [ITU-T G.1034] to cover packet data applications and services. It builds upon actual measurement data to create spatially resolved information. Applications are categorized depending on dependency of their quality of service (QoS) on temporal continuity. By using typical travel speeds, spatial and temporal criteria are interlinked, allowing for assessment of the fitness of a given route for particular types of applications.

2 References

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

[ITU-T G.1010] Recommendation ITU-T G.1010 (2001), *End-user multimedia QoS categories*.

[ITU-T G.1034] Recommendation ITU-T G.1034 (2020), *Quality of experience metrics for mobile telephony communication during rail travel*.

[ITU-T Y.1546] Recommendation ITU-T Y.1546 (2014), *Hand-over performance among multiple access networks*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 call completion probability (CCP) [ITU-T G.1034]: The probability that a telephone call started at a given location can be completed as intended (i.e., without being dropped).

3.1.2 geographical unit (GU) [ITU-T G.1034]: A segment of road or railway track, or a square or rectangular shaped area, with given coordinates on a map. Used to aggregate measurement data based on their geographical coordinates.

3.1.3 local drop probability (LDP) [ITU-T G.1034]: An indicator, computed from drive test data, to indicate the call-drop probability for a given geographical unit (GU).

3.1.4 quality of experience (QoE) [b-ITU-T P.10]: The degree of delight or annoyance of the user of an application or service.

3.1.5 quality of service (QoS) [b-ITU-T E.800]: Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service.

3.1.6 virtual call [ITU-T G.1034]: A concept using the local drop probability (LDP) values in a route profile to compute the call completion probability (CCP) for a call of given duration.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 aggregated continuity of service indicator (ACOSI): A service quality indicator for a given service type with given characteristics, aggregated over a route or journey. See the derivation in clause 10.

3.2.2 generic continuity of service indicator (GCOSI): An indicator expressing the degradation of the quality of service for a given type of service with given characteristics, during a single usage of that service which starts at a given geographical unit. See the derivation in clause 10.

3.2.3 local expected data rate (LEDR): Data rate expected at a given geographical unit, calculated from actually measured data rates. See the derivation in clause 9.

3.2.4 virtual transaction (vTA): A transaction used to model the user experience with respect to continuity of service in a given spatial sequence of network performance.

NOTE – This concept is similar to that of virtual call in [ITU-T G.1034].

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AR	Augmented Reality
CCP	Call Completion Probability
GCOSI	Generic Continuity Of Service Indicator
GNSS	Global Navigation Satellite System
GU	Geographical Unit
KPI	Key Performance Indicator
LDP	Local Drop Probability
LEDR	Local Expected Data Rate
MOS	Mean Opinion Score
QoE	Quality of Experience
QoS	Quality of Service
RTT	Round-Trip Time
STT	Single-Trip Time
VR	Virtual Reality
Wi-Fi	Wireless Fidelity

5 Conventions

NOTE – The following conventions are the same as defined in clause 5.2 of [ITU-T G.1034].

5.1 Geographical entities

For the purpose of this Recommendation, the following applies:

Track

Physical road or rail element.

Routes and nodes

A transportation network is understood as a directed graph consisting of routes and nodes. A node can be of two general types: stations, and general junctions in the network. The distinction is made on a functional level where in station-type nodes, passengers can enter or leave vehicles.

NOTE – For the sake of simplicity, text of this Recommendation may make reference to railway tracks and related elements of a railway network only. It is understood that similar considerations also apply in the case of road networks.

A route may subsume several tracks in cases where multiple tracks are running in spatial vicinity. The distinction whether physical tracks are treated as separate routes is made on a functional level, with respect to mobile network properties. If the spatial separation is such that different network conditions have to be assumed, different routes need to be used. This may lead to the introduction of additional nodes, i.e., location where it is assumed that relevant mobile network properties are becoming different.

Special case: If a given route is systematically being driven through with significantly different velocities, it can be represented by different data objects of type route, designated by corresponding attributes. An example would be a station where part of the vehicles stop, and other vehicles drive through. This distinction is required for subsequent combination of routes to describe end user experience.

A route may exclude the immediate physical vicinity of nodes. In that case e.g., a station itself is not part of a route. This distinction is made to avoid zero-velocity effects or to ensure modularity and does not affect the general concept.

Geographical unit

A geographical unit (GU) is the smallest unit in the representation of routes (typical dimension: 100 to 200 m). A route is equivalent to a set of GU arranged in proper order.

A GU is a uniquely identified data object. If geographical positions of routes are available, a GU is typically a one-dimensional object, i.e., a logical segment of a route. Otherwise, a GU can also be a two-dimensional entity associated with a given rectangular area, typically a square. A way to create respective data object information is to use a grid overlaid on a map and record the sequence of grid tiles corresponding to the route.

To express a position along a given route, a respective relative index of a list of the route's component GU can be used.

If the representation of a transportation network is overlap-free, a GU also is unambiguously associated with one (centre) or two (start and end point of associated track segment) geographical location(s).

Journey

Instance of an end-to-end travel use case. A journey has a start and an end point which are station-type nodes in the transportation network. A journey is a set of routes arranged in proper order. On a data-evaluation level, a journey can be understood as the union of the GUs of its component routes.

Figure 1 provides an overview of the geographical entities described in this clause.

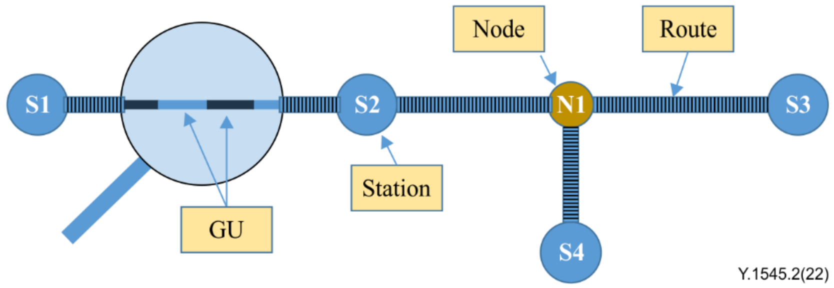


Figure 1 – Nomenclature for entities used in this Recommendation

6 General considerations for continuity of performance metrics

In the case of telephony, the underlying model is straightforward. For a single call, the outcome is binary: the call drops or it does not. CCP, the metric defined for call stability in [ITU-T G.1034] reflects this. There is a direct relationship between the LDP and CCP: CCP uses values provided by LDP.

A similar model for a packet data key performance indicator (KPI) needs a more generalized view. In the case of most if not all packet data-based services, quality of experience (QoE) degradation is a mix of hard and soft criteria, i.e., a gradual function of available data rate (DR). Moreover, there may be interaction between the end points of the services. For instance, adaptive video rates (e.g., dynamic adaptive streaming over hypertext transfer protocol) used in video streaming react to available DRs to reduce the risk of stalling. Such techniques create additional, non-linear dynamics.

Last but not least, some services use buffers (again, as in video streaming). In effect, this means that locations that have been passed along a journey may have lasting effects for subsequent QoS. An example is a buffer that runs low in low-DR regions. If, and how fast, the buffer is filled up again when higher DRs are available, depends on the actual implementation of the service.

A stepwise approach towards a generic model starts with considerations of spatial and temporal scales.

Every use case, seen from a QoE perspective, has its own time scale for performance-related effects. For instance, the time scale for interactive audio exchange using voice over Internet protocol (telephony or conferencing) will be of the order of 0.1 s. Performance degradations of underlying packet-data services larger than a few hundred milliseconds will start to produce perceivable degradations, ranging from slight annoyance towards total practical unusability of the service. This applies for the audio part of video conferencing; the time scale for video transmissions itself can be assumed to be longer. If an image freezes for a second, this may create irritation or annoyance; the basic function is, however, still maintained.

In video streaming, adaptive bitrates and buffering can conceal temporary impairments. In the case of buffering, the tolerance time scale is determined by the temporal depth of the buffer, i.e., typically a few seconds.

NOTE 1 – This Recommendation does not make assumptions about quantitative dimensions of impairments, or the impact of temporal structures (e.g., repeated smaller impairments versus more severe impairments occurring with a lower frequency), and refers readers to respective studies.

The general context of continuity of service-related KPI is motion, and this motion takes place in a spatial pattern of mobile network performance. The temporal pattern experienced during the usage of a service is therefore the result of a spatial pattern transformed into time by the velocity of motion. Typical dimensions of spatial patterns are, in effect, the result of both physics on the radio level (considered to be "microscopic"), and mobile network topology ("macroscopic"). In addition, performance is also influenced by shared-media effects, i.e., load, which creates another layer of effects.

Mobility may also lead to hand-over processes between different access networks, e.g., switching between a mobile network and a wireless fidelity (Wi-Fi) network. [ITU-T Y.1546] deals directly with the relevant aspects and also offers further elements of a conceptual framework for underlying processes that is also useful in a general sense.

Beside the intrinsic temporal and spatial scales, there is also the practical question of measurement accuracy, which usually depends on global navigation satellite system (GNSS) accuracy.

All in all, we assume that a useful spatial scale is 100 m. By combining this with typical travel speeds, the matrix in Table 1 shows resulting typical time scales for continuity of performance-related considerations.

Table 1 – Typical time scales for continuity of service-related considerations

Dimension (m)	**Speed (km/h)**						
	40	**80**	**120**	**160**	**200**	**240**	**280**
	Time (s)						
100	9	4.5	3	2.3	1.8	1.5	1.3
200	18	9	6	4.5	3.6	3	2.6
400	36	18	12	9	7.2	6	5.1
600	54	27	18	13.5	10.8	9	7.7
800	72	36	24	18	14.4	12	10.3
1 000	90	45	30	22.5	18	15	12.9

Table 1 shows that for services in the near-real time category, even a single 100 m track segment, traversed by a high-speed train, can lead to severe (temporary) impairments in QoE if the DR in this place is insufficient.

For streaming video or other services using buffers, the tolerance is higher. A single segment with an insufficient DR is usually not critical for QoE. However, if the DR is insufficient for a sequence of segments, this will also lead to impairments.

Another factor influencing QoE is the spatial frequency of DR distributions.

NOTE 2 – The concept of spatial frequency was originally introduced in optical information processing. A spatial frequency is the rate at which a spatial property is changing.

Assuming a spatially fixed distribution of high and low DRs (which is surely an idealization), traversing this pattern at a given speed creates a corresponding pattern of QoE-related behaviour. As in the case of telephony, user experience for packet data-based services depends on the actual distribution pattern, and calculating only average values for DR aggregated over an entire route or larger periods of time cannot provide differentiation in expected QoE.

So far, the considerations have only looked at DR. For assessment of fitness of a particular service, in particular those that are interactive, near-real time, latency is also a central factor. Adding a respective indicator, i.e., some flavour of single-trip time (STT) or round-trip-time (RTT), can easily be done within the given concept. For the time being, and for ease of reading, we assume that a good DR is a sufficiently applicable proxy or approximation for good STT or RTT.

7 Taxonomy of service types versus robustness against spatiotemporal performance degradations

Following the considerations in clause 6, typical packet data-based services can be categorized into four groups. This categorization also builds upon and integrates the taxonomy provided by [ITU-T G.1010].

NOTE – Note that the current edition of ITU-T G.1010 dates back to 2001, which limits its direct usability as a foundation. Many applications that are considered "popular" today did not exist at this time, e.g., augmented reality/virtual reality (AR/VR); usage habits and expectations have also changed. Also, in many cases applications have become considerably more "data-intensive" (which limits e.g., the applicability of Table I.1 of [ITU-T G.1010]). For instance, sending an e-mail today cannot simply be considered to be "fire and forget" as a typical mail client shows a progress bar, and users may want to wait for confirmation that the mail has actually been sent successfully.

With respect to [ITU-T G.1010], this Recommendation identifies three key parameters impacting user QoE: Delay, delay variation, and information loss. These effects are in effect coupled by the basic transport protocol. For instance, when a transmission control protocol is used, there is no loss of data due to the retransmission mechanism, but such retransmissions cause delay and delay variation instead. With the user datagram protocol, packet loss can occur directly.

However, in the case of mobile networks, there are additional mechanisms of error detection and retransmission at lower layers. In summary, for the purpose of the following taxonomy which focuses on user perception, the summary term "performance degradation" is used, which subsumes different underlying effects.

a) Services with near-real time or high-bandwidth performance requirements (audio/video telephony and conferencing; many types of AR, VR and gaming); termed, for brevity, type A. For the given spatial resolution (100 m) and speeds (50 km/h to 250 km/h) associated with rail travel, their QoE will be affected by performance degradations even in a single GU. In the ITU-T G.1010 taxonomy, this class would be of type "interactive".

b) Services (termed type B; ITU-T G.1010 categories "responsive" or "timely") with high performance requirements but using buffering (typically flavours of video streaming, or more generally non-real time multimedia/AR/VR/gaming; also, the spectrum of "enterprise mobility" applications would fall into this category): Those will be robust against drops in performance on the equivalent spatiotemporal scale of some seconds. Their QoE will, however, also be affected if too many consecutive GUs show poor DRs.

c) Services (termed type C; ITU-T G.1010 category "timely" to "non-critical") that still have some degree of interactivity, but with – in today's view – moderate low or moderate performance requirements, such as web browsing, or, in a wider sense, e-commerce): These services have a certain degree of robustness against temporary performance degradations. QoE will, however, also be affected if those degradations are too severe or last too long.

d) Services (termed type D) of the "store and forward" or "background activity" type (ITU-T G.1010 category "non-critical"). Their QoE will practically not suffer from temporally limited performance degradations. They are listed for completeness here.

It is understood that a classification of services into these types interacts and scales with the assumed spatiotemporal resolution and can become finer grained with corresponding accuracy (which is in essence not only a function of GNSS accuracy, but even more a function of available data density).

The type classification depends on the spatial scale. With a scale of 10 m instead of 100 m, the typical time associated with a segment traversed at 160 km/h becomes approx. 230 ms. A service which is type A on a 100 m scale would then be classified as type B on a 10 m scale.

8 Layered approach to metrics

In analogy to the metrics for telephony specified in [ITU-T G.1034], a layered approach is used.

The first layer is a local packet-data performance indicator (analogous to LDP as specified in [ITU-T G.1034]) that is calculated from actually measured data.

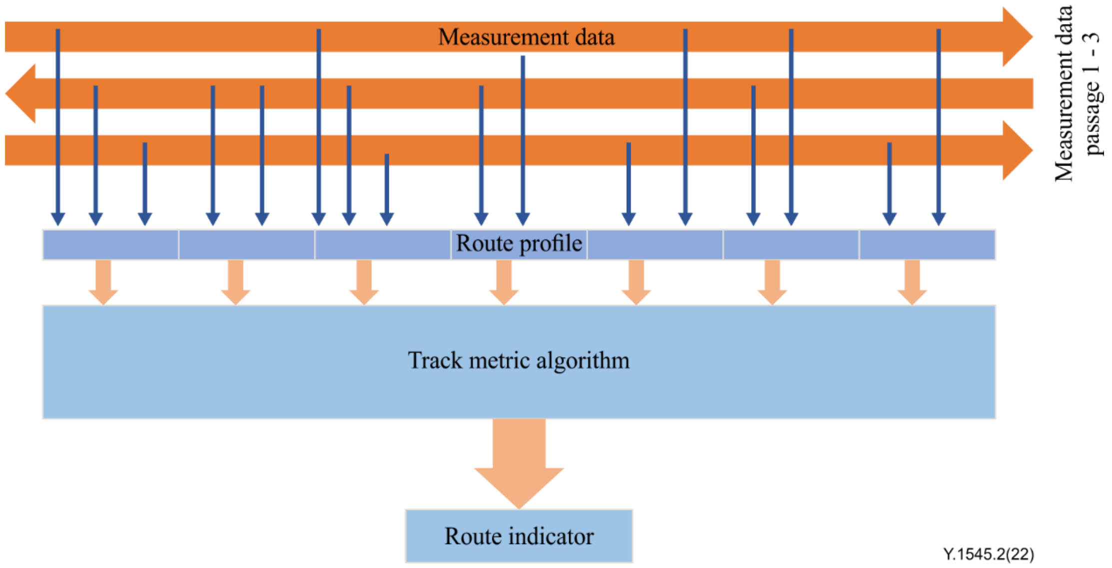


Figure 2 – Symbolic description of spatially resolved data processing (as in [ITU-T G.1034])

Measurement data taken on subsequent passages (drive tests) on a route are aggregated into the corresponding GUs.

The second layer uses this indicator to calculate a KPI that describes the QoS for typical service use cases, by analogy to CCP as specified in [ITU-T G.1034]. As in the case of CCP, the metric is designed and able to deliver predictions on perceived quality for the entire spectrum of use cases.

While the parameter space for telephony is just one-dimensional (the length of a telephone call), the packet data metric supports a two-dimensional parameter space, i.e., the duration of a service usage, and the type of the use case, e.g., video streaming, audio or video conferencing, gaming.

NOTE – In extension, the concept would also support a wider parametric model. In this case, the local performance indicators would be extended to include both DR and latency. The stage two modelling would then deliver a fitness-for-service figure based on required characteristics in the DR-latency space. Using e.g., concepts of fuzzy logic (membership functions), this allows definitions in the sense of a traffic-light fitness model, such as "fully fit", "borderline case", and "completely unfit".

9 Local expected data rate

Typically, KPIs of type DR are calculated by aggregating primary data values. However, there are many different ways to do this, e.g., arithmetic mean, median values or percentiles. Also, calculation algorithms typically include rules for selection of data items to be included or excluded, so the overall process is in many cases more complex than just applying a standard mathematical transformation.

In order to enhance modularity and versatility, it is useful to provide an abstraction, termed local expected data rate (LEDR). This indicator describes the DR expected for a given GU.

NOTE 1 – Based on the considerations in clause 7 and with reference to [ITU-T G.1010], it is assumed that DR acts as a suitable proxy for factors impacting the user experience. This Recommendation provides all necessary foundation to extend the methodology by direct indicators for delay, delay variation or corresponding derived indicators.

A simple implementation of LEDR uses a constant scaling factor to the averaged DR. In practice, algorithms such as *n*-percentiles are useful. This corresponds to an end-user perspective where the indicated value is met or exceeded with a given probability. For instance, using the P_{30} value (30% percentile) from measured data samples would express the expectation that the given DR is met or exceeded in 70% of the usages.

NOTE 2 – The calculation of a meaningful and valid LEDR depends on DR, with the basic calculation of DR as transferred data volume divided by time interval. Packet data transfer has a bursty nature; if the time interval is too short, artefacts in the form of spikes in values can occur. It is assumed that the time interval is chosen reasonably to avoid such overshoots, see also clause A.2.1 of [b-ITU-T Y.1540]. In the context of this Recommendation, the default value of 1 s as specified in Annex B of [b-ITU-T Y.1540] is assumed to provide sufficient smoothing, while being consistent with assumed spatial resolution (see Table 1), and the potential to adjust if a higher spatial resolution is required.

10 Generic continuity of service indicator

Ideally, the specification of a metric for continuity of service covers all categories specified in clause 9. This is achieved by using a two-step approach, based on the LEDR indicator as specified in clause 9. For that purpose, LEDR is assumed to have the highest possible spatial resolution.

For ease of reading, it is assumed, without loss of generality, that this resolution is 100 m.

In the first step, all locations along a route are identified where the requirement for continuity of service is not met. This step creates the first-stage figure of merit denoted $\boldsymbol{I}_{\mathrm{GCOS}}(x)$. This quantity is a vector, i.e., each position along a route has its generic continuity of service indicator (GCOSI) value.

In another simplification for the sake of easy explanation, it is further assumed that this outcome is binary, i.e., there is only "yes" or "no". In clause 12, the expansion towards a multi-valued approach is described.

A particular service is categorized by assigning it a threshold value *P* for the minimum performance required to work free of degradations, and the threshold number *L* of consecutive segments at which the performance degradation sets in. The broad categorization shown in clause 9 is therefore to a set of values for a particular service, as explained by the following examples for selection of *L*. The DR threshold *P* is, in all cases, set to the DR requirements of the respective service.

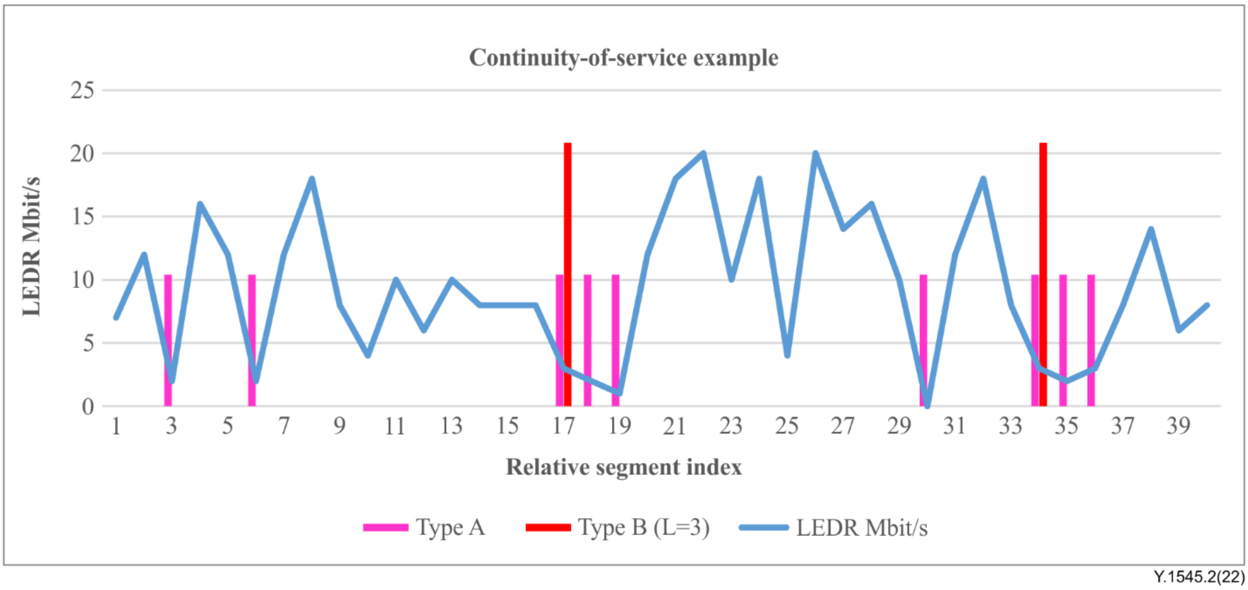
– A real-time service of type *A* has an *L* value of 1: Even a single 100 m-segment of poor network performance leads to a degradation of user experience.

– A service of type B will have an *L* value of 2 or 3.

  NOTE – When a mapping of segment dimension to time per segment is made, the assumed buffer length of this service can be used directly).

– A service of type C will have its *L* set to an even higher value.

The processing algorithm itself works by analogy to the CCP algorithm by assuming a virtual transaction (vTA) having a given duration. It moves along an LEDR profile of a route, indexed by the relative position *x*, and keeps a counter value *z*. As long as the end of the vTA is not reached, and, from *x*, a sequence of L LEDR elements with values below *P* is encountered, *z* is incremented by the number of segments for which LEDR is below *P*, and *x* is moved to the next position after the block.



Key

Bar type	**Likely**
A	QoE degradations for a type A service
B	QoE degradations for a type B service with sequence length $L = 3$

Bar height is set differently for better visual differentiation

Figure 3 – Example for continuity of service metrics

The end of the vTA is determined by the average speeds for respective segments. If the end of a vTA is reached within a sequence of degraded segments, only those segments are counted that are inside the duration of the vTA. The GCOSI value is then determined from the time fraction of degraded segments to the overall duration of the vTA and is assigned to the starting position *x*. This processing step therefore generates a vector $\boldsymbol{I}_{GCOS}(x)$ for the respective route.

The processing algorithm for $\boldsymbol{I}_{GCOS}(p)$, i.e., the value for the specific starting position *p* along a route:

– proceed, with *x* initially set to *p*, along the LEDR(*x*) profile until a given relative position is reached;

– count the segments which belong to a sequence of at least *L* units with a performance below the DR threshold *P*;

– assign this counter value to the starting segment at index *p*.

In the second step, these $\boldsymbol{I}_{GCOS}(x)$ values are combined into a single aggregated continuity of service indicator (ACOSI) number. As mentioned, this indicator is analogous to CCP but has one more dimension, namely the specific service it deals with. So, the full denomination would be an ACOSI (service, usage time per transaction).

If an LEDR has more than two possible values, the model can easily be adapted by corresponding aggregation algorithms.

In words, an ACOSI describes the user experience for a given application and a given usage time per instance of usage.

NOTE – The algorithm for ACOSI is seen as a specific instance of a whole class of algorithms for the computation of that indicator. The given algorithm, which just counts or averages local GCOSI values, is easy to implement and serves as a useful starting point for such study. Refinements and alternatives are a matter of further study. It is assumed that subjective testing, with a particular focus on effects of the frequency and the distribution pattern of service degradations, will play a valuable role in that evolution.

11 Special case: Macroscopic continuity

In extension to mobility management in a single mobile network, there are cases where mobility includes handover processes between different access networks, such as the hand-over between a local Wi-Fi network and a mobile network or between two different Wi-Fi access networks. Practical examples include passengers boarding a train and traversing a station with local Wi-Fi or riding a bus that leaves or enters local Wi-Fi coverage. In today's mobile devices, there are also selectable network selection strategies that may cause similar effects when traversing areas with different network coverage.

Hand-over processes between access networks are covered by [ITU-T Y.1546], which provides a conceptual framework and performance indicators for underlying processes. From the nature of these processes, it is expected that the time scale of related effects is in the range of seconds. With respect to the service-type taxonomy, this would be the domain of service type B.

Macroscopic continuity is considered to be a special case because the range of applications involved or affected may be different from that related to a mobility single network (apart from cases where a passenger traverses a region with very volatile single-network coverage).

For instance, in the aforementioned case of a passenger boarding a train through a train station, the passenger will typically be on foot, which means that the range of likely use cases within that period is different from that used during typical travel through a single network. For instance, usage of VR or video conferencing is not compatible with this kind of motion. On the other hand, the likelihood of other usages such as AR (in gaming as well as in other contexts) may even be higher.

12 Multi-level quality ratings

So far, this Recommendation works with the assumption of a binary classification of service degradation, i.e., local network performance is either sufficient or not. A refined version of such an assessment would use a multi-valued figure of merit and a mapping relation between performance and expected QoE, e.g., on a mean opinion score (MOS) or other scale with discrete levels.

Table 2 provides an example for such a mapping relation; in this case, a symbolic representation is used. Usage of a five-point MOS, a numerical or alphabetical school mark scale, are other examples.

Table 2 – Example performance mapping for a data application requiring a nominal fixed data rate and tolerance threshold values *x* and *y*

Available data rate as % of DR	Quality rating
≥100	+++
<100 and ≥100 − *x*	++
<100 − *x* and ≥100 − *y*	−−
<100 − *y*	−−−

Reading examples:

– if the available DR is below 100% but within a tolerance margin (*x*), the rating is ++;

– if the DR is below a critical threshold (*y*), the rating is −−−.

NOTE – Table 2 has been deliberately designed for a high level of abstraction. It shall not propose any concrete numbers or target values.

A mapping table is just the most generic solution; it is assumed that there are cases where a closed mathematical formula would also work, e.g., a sigmoid function or a proportionality with the *n*th root of the DR. Regardless of the variant that is chosen, it will need to be backed by solid data (e.g., from subjective testing) or at least a plausible model.

Also, the methodology provided in this Recommendation is open to the extension of such a mapping relation to a more-dimensional input, such as adding latency or RTT as a second dimension for value ranges. In a further extension, other ways of categorization established elsewhere are available, such as the concepts of fuzzy logic using membership functions.

Bibliography

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

[b-ITU-T P.10] Recommendation ITU-T P.10/G.100 (2017), *Vocabulary for performance, quality of service and quality of experience*.

[b-ITU-T Y.1540] Recommendation ITU-T Y.1540 (2019), *Internet protocol data communication service – IP packet transfer and availability performance parameters*.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	**Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities**
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2022