International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.1028.1
(02/2019)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

Multimedia Quality of Service and performance – Generic and user-related aspects

---

End-to-end quality of service for video telephony over 4G mobile networks

Recommendation ITU-T G.1028.1

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS	G.200–G.299
INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES	G.300–G.399
GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES	G.400–G.449
COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY	G.450–G.499
TRANSMISSION MEDIA AND OPTICAL SYSTEMS CHARACTERISTICS	G.600–G.699
DIGITAL TERMINAL EQUIPMENTS	G.700–G.799
DIGITAL NETWORKS	G.800–G.899
DIGITAL SECTIONS AND DIGITAL LINE SYSTEM	G.900–G.999
MULTIMEDIA QUALITY OF SERVICE AND PERFORMANCE – GENERIC AND USER-RELATED ASPECTS	**G.1000–G.1999**
TRANSMISSION MEDIA CHARACTERISTICS	G.6000–G.6999
DATA OVER TRANSPORT – GENERIC ASPECTS	G.7000–G.7999
PACKET OVER TRANSPORT ASPECTS	G.8000–G.8999
ACCESS NETWORKS	G.9000–G.9999

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

Recommendation ITU-T G.1028.1

End-to-end quality of service for video telephony over 4G mobile networks

Summary

Recommendation ITU-T G.1028.1 provides guidelines concerning key aspects impacting end-to-end performance of carrier-grade (in opposition to over-the-top (OTT) approaches, which are outside of the scope of this Recommendation) conversational video services over long-term evolution (LTE) networks, also known as video-telephony over LTE (ViLTE), as defined by Global System for Mobile communications Association (GSMA). It identifies the preconditions for an optimally operating ViLTE network and provides remedial measures that operators can leverage to address the associated impact of quality of service (QoS) degradations in the LTE network.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T G.1028.1	2019-02-06	12	11.1002/1000/13831

Keywords

LTE, QoS, quality of service, video, video telephony, ViLTE, 4G.

* 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, 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 TSB patent database at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Introduction

Mobile broadband operators, facing a competitive broadband market, are obliged to redefine their business models to enhance revenue-generating streams. This has necessitated a deployment shift to converged IP-based technology platforms and high-throughput access network technologies that deliver high-quality triple-play services (telephony, internet and video streaming) to consumers whose expectation for improved user experience continues to remain insatiable. In this perspective, video-telephony services over 4G networks (i.e., long-term evolution (LTE)) present an opportunity for operators to offer new value-added services to their customers and convince them to remain faithful. There still exists ongoing research work by academics, systems developers and standards organizations; all attempting to help fill the knowledge gap for successful commercial video-telephony over LTE (ViLTE) deployment worldwide.

Recommendation ITU-T G.1028.1

End-to-end quality of service for video telephony over 4G mobile networks

1 Scope

This Recommendation covers end-to-end quality of service (QoS) requirements for video-telephony over long-term evolution (LTE) (ViLTE) network segments (see [b-GSMA IR.94]), budget allocation considerations for different service architecture scenarios, QoS parameterization for regulatory compliance, impact assessment of some relevant operating conditions on identified service parameters as well as a diagnostic strategy for QoS degradations in ViLTE. The intention of this Recommendation is to serve as a reference guide for LTE operators and regulators.

This Recommendation is a complement to [ITU-T G.1028]. All voice-related aspects of ViLTE are exactly similar to those for voice over LTE (VoLTE), and therefore covered by [ITU-T G.1028], and thus, they are not repeated in this Recommendation.

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.1011] Recommendation ITU-T G.1011 (2016), *Reference guide to quality of experience assessment methodologies*.

[ITU-T G.1028] Recommendation ITU-T G.1028 (2016), *End-to-end quality of service for voice over 4G mobile networks*.

[ITU-T G.1070] Recommendation ITU-T G.1070 (2018), *Opinion model for video-telephony applications*.

[ITU-T G.1071] Recommendation ITU-T G.1071 (2016), *Opinion model for network planning of video and audio streaming applications*.

[ITU-T H.264] Recommendation ITU-T H.264 (2017), *Advanced video coding for generic audiovisual services*.

[ITU-T H.265] Recommendation ITU-T H.265 (2018), *High efficiency video coding*.

[ITU-T J.144] Recommendation ITU-T J.144 (2004), *Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference*.

[ITU-T J.246] Recommendation ITU-T J.246 (2008), *Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference*.

[ITU-T J.247] Recommendation ITU-T J.247 (2008), *Objective perceptual multimedia video quality measurement in the presence of a full reference*.

[ITU-T J.249] Recommendation ITU-T J.249 (2010), *Perceptual video quality measurement techniques for digital cable television in the presence of a reduced reference*.

[ITU-T J.341] Recommendation ITU-T J.341 (2016), *Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference*.

[ITU-T J.342] Recommendation ITU-T J.342 (2011), *Objective multimedia video quality measurement of HDTV for digital cable television in the presence of a reduced reference signal*.

[ITU-T J.343.1] Recommendation ITU-T J.343.1 (2014), *Hybrid-NRe objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of encrypted bitstream data*.

[ITU-T J.343.2] Recommendation ITU-T J.343.2 (2014), *Hybrid-NR objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of non-encrypted bitstream data*.

[ITU-T J.343.3] Recommendation ITU-T J.343.3 (2014), *Hybrid-RRe objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a reduced reference signal and encrypted bitstream data*.

[ITU-T J.343.4] Recommendation ITU-T J.343.4 (2014), *Hybrid-RR objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a reduced reference signal and non-encrypted bitstream data*.

[ITU-T J.343.5] Recommendation ITU-T J.343.5 (2014), *Hybrid-FRe objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a full reference signal and encrypted bitstream data*.

[ITU-T J.343.6] Recommendation ITU-T J.343.6 (2014), *Hybrid-FR objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a full reference signal and non-encrypted bitstream data*.

[ITU-T P.863] Recommendation ITU-T P.863 (2018), *Perceptual objective listening quality prediction*.

[ITU-T P.1201] Recommendation ITU-T P.1201 (2012), *Parametric non-intrusive assessment of audiovisual media streaming quality*.

[ITU-T P.1201.1] Recommendation ITU-T P.1201.1 (2012), *Parametric non-intrusive assessment of audiovisual media streaming quality – Lower resolution application area*.

[ITU-T P.1201.2] Recommendation ITU-T P.1201.2 (2012), *Parametric non-intrusive assessment of audiovisual media streaming quality – Higher resolution application area*.

[ITU-T P.1202.1] Recommendation ITU-T P.1202.1 (2012), *Parametric non-intrusive bitstream assessment of video media streaming quality – Lower resolution application area*.

[ITU-T P.1202.2] Recommendation ITU-T P.1202.2 (2013), *Parametric non-intrusive bitstream assessment of video media streaming quality – Higher resolution application area*.

[ETSI TS 122 105] ETSI TS 122 105 v15.0.0 (2018-07), *Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Services and service capabilities (3GPP TS 22.105 version 15.0.0 Release 15)*.

[ETSI TS 123 203] ETSI TS 123 203 v15.4.0 (2018-09), *Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System*

(UMTS); LTE; Policy and charging control architecture (3GPP TS 23.203 version 15.4.0 Release 15).

[ETSI TS 126 114] ETSI TS 126 114 v15.4.0 (2018-10), *Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction (3GPP TS 26.114 version 15.4.0 Release 15*.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

3G Third Generation of radio access network

4G Fourth Generation of radio access network

AEC Acoustic Echo Control

AGC Automatic Gain Control

AMR-WB Adaptive Multi-Rate Wideband

AS Application Server

ATCF Access Transfer Control Function

ATGW Access Transfer Gateway

BE Best Effort

BGCF Border Gateway Control Function

BSC Base Station Controller

BTS Base Transceiver Station

CIF Common Intermediate Format

CS Circuit Switched

CSFB Circuit Switched Fallback

DL Downlink

DRB Data Radio Bearer

DRX Discontinuous Reception

DSCP Differentiated Services Code Point

DTMF Dual-Tone Multi-Frequency

EF Expedited Forwarding

eMSC Enhanced MSC

e-NodeB Enhanced Node B

EPC Evolved Packet Core

E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
GBR	Guaranteed Bit Rate
GERAN	GSM/Edge Radio Access Network
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
GSMA	GSM Association
GTP	GPRS Tunnelling Protocol
GW	Gateway
HARQ	Hybrid Automatic-Repeat-Request
HD	High Definition
HSS	Home Subscriber Server
HVGA	Half Video Graphics Array
IBCF	Interconnection Border Control Function
I-CSCF	Interrogating Call Session Control Function
IMS	IP Multimedia Subsystem
LTE	Long-Term Evolution
MBR	Maximum Bit Rate
MGCF	Media Gateway Controller Function
MGW	Media Gateway
M-LWDF	Modified Largest Weighted Delay First
MME	Mobility Management Entity
MOS	Mean Opinion Score
MOS-LQ	Mean Opinion Score – Listening Quality
MRFC	Multimedia Resource Function Controller
MRFP	Multimedia Resource Function Processor
MSC	Mobile Switching Centre
MSCS	MSC Server
MTSI	Multimedia Telephony Service for IMS
NB	Narrowband
NGN	Next Generation Network
NR	Noise Reduction
OFDMA	Orthogonal Frequency-Division Multiple Access
OT	Third Operator
OTT	Over-The-Top
PCC	Policy and Charging Control
PCEF	Policy and Charging Enforcement Function
PCRF	Policy and Charging Rule Function

P-CSCF	Proxy Call Session Control Function
PDA	Personal Digital Assistant
PDCP	Packet Data Convergence Protocol
PDD	Post Dialling Delay
PF	Proportionality Fair
P-GW	Packet Data Network Gateway
PLF	Packet Loss Fair
PSTN	Public Switched Telephone Network
QCI	QoS Class Identifier
QCIF	Quarter Common Intermediate Format
QoS	Quality of Service
QVGA	Quarter Video Graphics Array
RACH	Random Access Channel
RLC	Radio Link Control
RNC	Radio Network Controller
RoHC	Robust Header Compression
RRC	Radio Resource Control
RSRP	Reference Signal Received Power
RTCP	Real-time Transport Control Protocol
RTP	Real-time Transport Protocol
S-CSCF	Serving Call Session Control Function
SD	Standard Definition
SDP	Session Description Protocol
S-GW	Serving Gateway
SIP	Session Initiation Protocol
SRB	Signalling Radio Bearer
SRVCC	Single Radio Voice Call Continuity
TAS	Telephony Application Server
TrGW	Trunking Gateway
TTI	Transmission Time Interval
UDP	User Datagram Protocol
UE	User Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunication System
UTRAN	UMTS Terrestrial Radio Access Network
ViLTE	Video-telephony over LTE
VGA	Video Graphics Array

VoLTE Voice over LTE

VT Video Telephony

WB Wideband

5 Conventions

None.

6 Brief introduction on video-telephony over LTE and assumptions

This Recommendation considers some key assumptions in respect of the IP multimedia subsystem (IMS) profile for video as defined by the Global System for Mobile communications Association (GSMA) in [b-GSMA IR.94] and the multimedia telephony service for IMS (MTSI) media handling procedures (video part only) defined by 3GPP in [ETSI TS 126 114].

– To deploy ViLTE, VoLTE is required as a prerequisite. Voice aspects and network service architecture of ViLTE are adequately addressed in [ITU-T G.1028];
– To support a video call, the user equipment (UE) transmits its video capability to the LTE network. The video call request encapsulates the video media with real-time transport protocol (RTP) on user datagram protocol (UDP) (RTP/UDP);
– RTP is the media protocol for the transmission of realtime audio or video streams. Different than VoLTE, packet data network gateway (P-GW) and serving gateway (S-GW) establish two bearers for a video call: one for voice and one for video;
– ViLTE uses the mandatory ITU-T H.264 codecs or preferably the optional (ITU-T H.265 Main tier level 3.1 codecs) to encode and decode the video stream with trade-off consideration for the optimization of both the bit rate and video signal quality;
– The ITU-T H.264/ITU-T H.265 codec delivers superior quality as compared to the low bit ITU-T H.263 codec that is used in third generation (3G) conversational video calls;
– Video resolution and coding rate is likely to adapt, during a call, to network conditions such as a reduction in downlink bandwidth. The real-time transport control protocol (RTCP) is used for the communication of capacities between the UE and the IMS entities inside the network during a call, thus triggering adaptation;
– ViLTE uses the same control plane protocol as VoLTE, namely session initiation protocol (SIP);
– The IMS core network along with the applicable application server (AS) performs the call control;
– ViLTE video calls are allocated appropriate quality of service (QoS) to differentiate and prioritize such delay and jitter sensitive conversational traffic from other streaming video traffic that is not as delay or jitter sensitive;
– The mechanism used is called QoS class identifier (QCI). The ViLTE bearer traffic is typically allocated QCI-2, and the SIP-based IMS signalling QCI-5;
– During ViLTE sessions, video capable devices often ensure lip-synchronization across audio and video components, a phenomenon that is characterized by the sending of timing information to each other;
– Call handling in ViLTE provides communicating devices with options to turn off video at any time during the call and continue with voice only;
– Conversational video calling services may be carried out in either simplex or duplex mode;

– Video streams can be changed from one mode to another by the sending of a re-INVITE request with a session description protocol (SDP) offer using appropriate media descriptors (e.g., sendrecv, sendonly, recvonly).

Table 1 – Standardized QCI characteristics for use in ViLTE [ETSI TS 123 203]

QCI	Resource type	Priority level	Packet delay budget	Packet error rate	Service type
1	Guaranteed bit rate (GBR)	2	100 ms	1/100	Conversational voice
2		4	150 ms	1/1000	Conversational video (live streaming)
5	Non-GBR	1	100 ms	1/1000000	IMS signalling

7 ViLTE network architecture

The network architecture for ViLTE is similar to the one for VoLTE (see [ITU-T G.1028]).

Figure 1 (taken from [ITU-T G.1028]) shows the overall network architecture for ViLTE services.

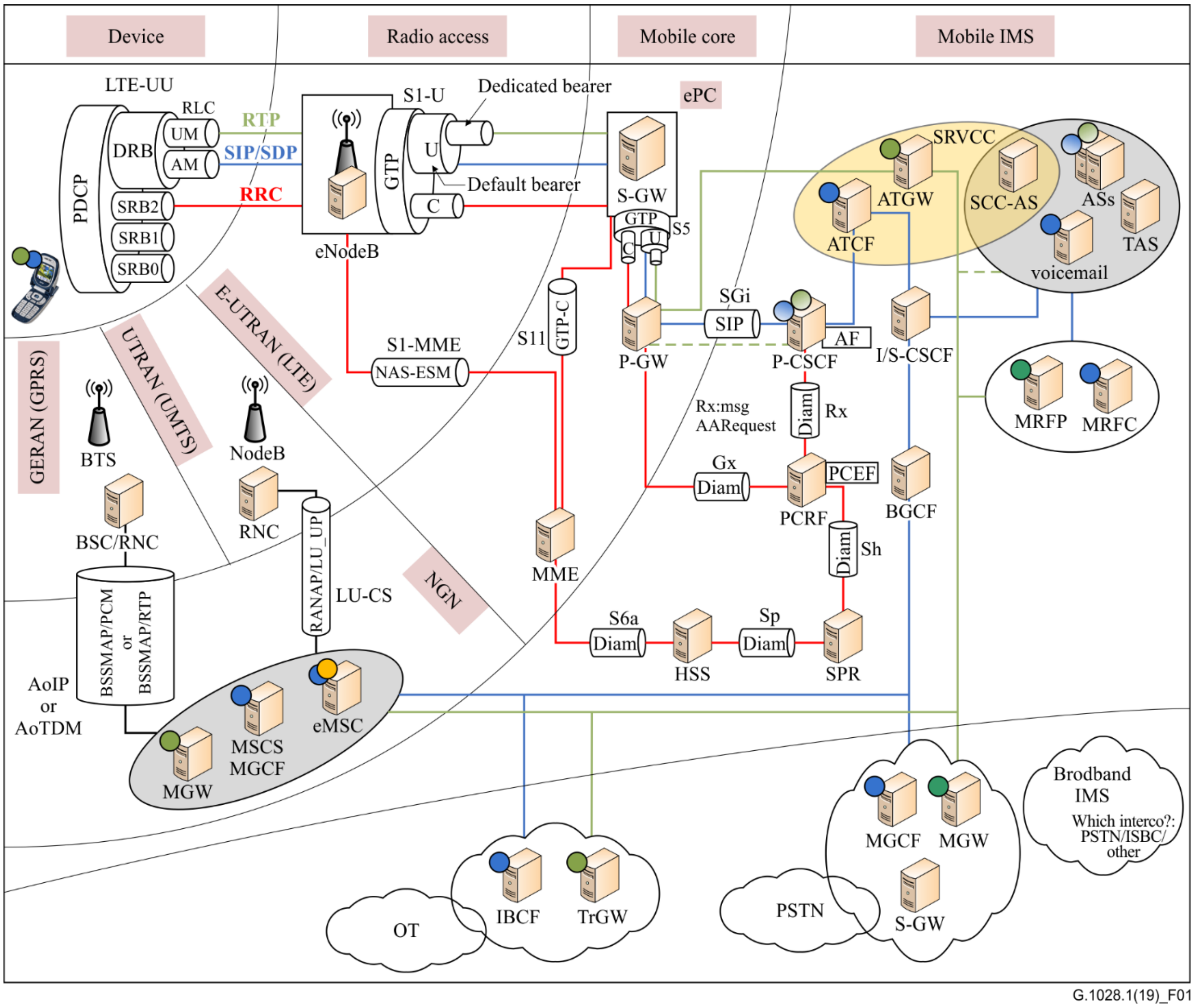


Figure 1 – Overall network architecture for ViLTE services

8 QoS requirements for ViLTE – Segmented approach

8.1 Overview of QoS issues experienced by end-users

ViLTE is a relatively new service, and not enough data is available yet to understand global QoS perceived by customers and how big the weight of the different dimensions of QoS is. However, An analogy can be made with existing services for which consolidated data are available.

The main families of QoS parameters for conversational services are known from telephony. These are service accessibility, audio/video quality (part of service integrity including audio quality, video quality, and relations between simultaneous audio/video signals such as lip synchronization) and service continuity. A detailed list of the most relevant metrics pertaining to each QoS family is provided in clause 9.1.

Furthermore, ViLTE shares several characteristics with other services available over the same access technology, such as VoLTE [ITU-T G.1028] and video streaming over LTE (for video aspects). Concerning this last point:

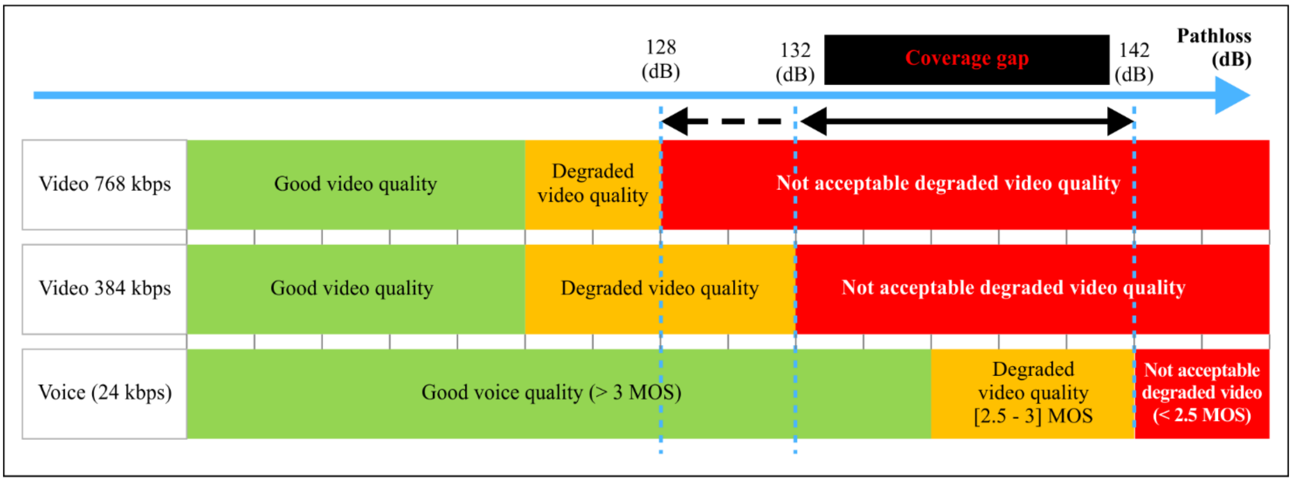
– The intrinsic quality of video rendering, highly correlated with video coding technology and bit rate, video size, resolution (and their adequation with screen size) and video frame rate;

– The occurrence of network (core or access) congestion, resulting in several visible artefacts (depending on decoding and buffering strategies at receiving side) like image freezing (similarly to re-buffering events in video streaming), pixelation, blocks, ghosting, etc.;

– A combination of the two last elements, bandwidth limitation or jitter buffer, that can be compensated by video coding bitrate adaptation, yielding potential visible quality degradations.

However, ViLTE is also characterized by the differences of media handling applied to voice and video, since the service profile for ViLTE, as defined in [b-GSMA IR.94], is based on QCI (see Table 1).

Thus, in case of network congestion or in the event a ViLTE terminal is at the edge of radio coverage, voice will be given a priority over video. A mechanism like transmission time interval (TTI) bundling, allowing retransmission of voice packets to ensure they are not lost, and thus limiting the bandwidth for other packets, amplifies this priority. TTI bundling actually defines the final threshold for ViLTE coverage beyond which only 64 kbit/s video can run with non-acceptable quality using ITU-T H.264 "Baseline". By reducing the video bit rate, coverage is improved but by only 4 dB gain with half the bit rate, as illustrated in Figure 2 below. In the most severe situations, depending on the strategy defined by service provider, end users will face either a communication reduced to its voice component or a call drop.



G.1028.1(19)_F02

Figure 2 – Video bit rate versus. coverage

Another element of consideration is how different overall ViLTE service quality, as experienced by users, can be when issues are related to audio only or to video only. It is known from user tests that customers are more sensitive to voice impairments than to video impairments during audio-video conversations. This generally results in better overall judgments when impairments affect video signal and to a lesser degree voice signal.

8.2 User equipment (codec design and implementation)

ITU-T H.264 constrained high profile level 1.2, as specified in clause 5.2.2 of [ETSI TS 126 114], is mandatory in UE. However, for backward compatibility, it is required that UE also supports constrained baseline profile level 3.1 of the same release. Support for ITU-T H.265 Main profile, Main tier, level 3.1 is also recommended.

Also, as part of procedures, in clause 2.2.2 of [b-GSMA IR.94], the UE and the network must be able to establish a video call directly during session establishment or by adding video to a voice session by sending SIP (re-) INVITE request with an SDP offer that contains both voice and video media descriptors. To ensure optimum QoS delivery it is imperative to adjust the maximum bit rate (MBR) of the video signal to levels far below the configuration settings of level 3.1 of [ITU-T H.264] and fine-tuned to the transmission capabilities of the network.

It is recommended to align codec implementations for ViLTE so that codecs can be used in use cases taken as assumptions for the development of relevant parametric models proposed in [ITU-T G.1070] and [ITU-T P.1202.1]. Suffice to indicate that ITU-T H.264/ITU-T H.265 codec resolution, frame rate and encoding bit rate constitute key dependencies as far as maximum user perceived quality of the ViLTE service is concerned. Manufacturers of terminal devices (mobile phones and personal digital assistants (PDAs)) that support video telephony over the LTE network can find interesting guidance in the Table 2 assumptions, whereas codec design requirements should consider the coefficient derivation functions cited in Appendix I of [ITU-T G.1070].

Table 2 – Assumptions about monitor characteristics

Monitor specifications	Nominal values
Diagonal length (Note)	2-10 inches
Dot pitch	< 0.30
Colour temperature	6500 K
Bit depth	8 bits/colour

Table 2 – Assumptions about monitor characteristics

Monitor specifications	Nominal values
Refresh rate	≥ 60 Hz
Brightness	100-300 cd/m2
NOTE – Diagonal length means the image size of monitor.	

The end-to-end delay that a ViLTE video packet experiences may fluctuate from packet to packet. This variation in end-to-end delay is referred to as delay jitter. Delay jitter is a crucial problem for ViLTE because the receiving terminal (UE) must receive/decode/display frames in realtime and at a constant rate, any late frames resulting from the delay jitter can produce annoying artefacts in the reconstructed video e.g., jerks in the video.

This problem is typically addressed by including a playout buffer at the receiver. While the playout buffer can compensate for the delay jitter, it can potentially introduce additional delay. Video jitter buffer management for guaranteed QoS in video channels requires placing a cap on the jitter buffer latency (delay threshold), probing the jitter buffer state and doing away with excess video packets from the jitter buffer. In the case of an overflow a latency exceeded message is sent to notify the application that there may be enough delay in the jitter buffer to affect media synchronization and this is addressed by purging the jitter buffer.

8.3 E-UTRAN (Radio resource management)

Within the evolved-UMTS terrestrial radio access network (E-UTRAN) segment of the ViLTE architecture model, it is the responsibility of the enhanced Node B (e-NodeB) to ensure the provisioning of the necessary QoS conditions for a dedicated (video) bearer over the radio interface, taking into consideration such key determinants as the QCI and the priority levels.

One very key requirement in QoS provisioning at the radio-interface level is the type of scheduling strategy that must be administered on the e-NodeB as part of the radio resource management functions for a multi-user orthogonal frequency-division multiple access (OFDMA)-based mobile system. A good and efficient scheduling algorithm is required to demonstrate the desired performance levels in accordance with tolerable limits specified in [ETSI TS 123 203] for video telephony traffic. The priority and packet delay budget, and to some extent the acceptable packet loss rate from the QCI label, is required to determine the radio link control (RLC) mode configuration and how the scheduler in the medium access control (MAC) handles packets sent over the bearer.

It is thus recommended to RAN equipment vendors and systems operators a scheduling strategy that overcomes some of the limits of traditional benchmark scheduling algorithms (e.g., packet loss fair (PLF), modified largest weighted delay first (M-LWDF) or proportionality fair (PF)) in terms of throughput, packet loss and fairness among others. An LTE network operating a radio cell coverage, reference signal received power (RSRP) level, of less than -105 dBm is required to guarantee basic rule of thumb on admission controls based on the appropriate QCI from the user equipment.

An IMS session request for a video call (originating or terminating) in E-UTRAN requires that one dedicated bearer resource for voice and another dedicated bearer resource for video as specified in [b-GSMA IR.94] is created by authorizing the flows utilizing dynamic policy and charging control (PCC). The network must initiate the creation of dedicated bearer resources to transport the voice and video media. The dedicated bearer for conversational video stream may be a GBR or a non-GBR bearer. If a GBR bearer is used it must utilize the standardized QCI value of two (2) and have the associated characteristics as provided in [ETSI TS 123 203]. In the case of IMS termination of a session using conversational media, dedicated bearer resources must be deleted by withdrawing the authorization of the flows. The network must initiate the deletion of the bearer resources.

8.4 Evolved packet core (QCI allocation and mobility management procedures)

Evolved packet core (EPC) provides support to the QoS classification (between the policy and charging enforcement function (PCEF) and the ViLTE client), as defined in clause 5 of [ETSI TS 122 105] and clause 6.1.7 of [ETSI TS 123 203]. The mobility management entity (MME) provides tracking area updates to mobile UEs.

When a UE attaches with the network, a mutual authentication of the UE and the network is performed between the UE and the MME/home subscriber server (HSS). This authentication function also establishes security keys that are used for encryption of the bearers. Signalling overhead due to excessive tracking area (TA) updates must be managed in such a way as to guarantee reduced delays during the video calling session setup.

The S-GW supports transport level QoS through marking IP packets with appropriate Diffserv code points based on the parameters associated with the corresponding bearer. The P-GW is the point of interconnect to external IP networks through the SGi interface. It also has a key role in supporting QoS for end-user IP services.

A good hierarchical design is required to provide for seamless coordination of control-plane signalling during mobility with the two (2) major QoS preconditions being minimization of interruption in QoS during handover as well as improved support for interoperability among mobility protocols (IP/IPv6).

8.5 IMS and IP transit core (call control and signalling)

The IMS core supports ViLTE client registration and authentication. Video over IP (VoIP) session setup and release is enabled by the IMS and requires SIP signalling operating at an assigned QCI-5 as well as a realtime transfer of voice and video RTP flow as at QCI-1 and QCI-2, respectively (see Table 1).

To fulfil these requirements, and in the context of a capacity-constrained network, the Diffserv (differentiated services code point (DSCP)) approach may be used to ensure efficient bandwidth allocation and scheduling among several traffic applications including video telephony.

An LTE operator providing triple-play service offerings (voice, video, data) can adapt to the varying traffic requirements on their network by creating a traffic class group for each of the service types.

9 Budget estimation and QoS parameterization

9.1 Relevant indicators

There are two categories of indicators to consider when assessing the quality of ViLTE services:

1) Session setup and continuity;

2) Integrity of the content.

In the first category, the target is to assess which level of quality a user can access and use the service over an entire ViLTE session. The recommended metrics are given in Table 3 below.

Table 3 – QoS parameters for session setup and continuity

Name	Definition
Video telephony (VT) service availability	End-to-end service availability in terms of capacity to establish a call, as well as its audio and video components, from, and to, a ViLTE customer. An attempted ViLTE call resulting in a voice-only session is considered as failed.

Table 3 – QoS parameters for session setup and continuity

Name	Definition
Video component availability	The availability of the video component if it is requested to be added to an existing VoLTE call.
VT setup time (post dialling delay (PDD))	Time interval (in seconds) between the end of dialling by a caller and the reception back of the appropriate ringing tone, in case of a successful ViLTE call.
Components setup time	Time interval (in seconds) between the reception of a ringing tone and the beginning of the corresponding audio and video sessions, in case of a successful ViLTE call, or the time it takes to add the video component after request from a VoLTE call. Narrowband (NB): this metric does not consider whether or not the relevant QCI has been assigned to each flow (QCI-1 for voice, QCI-2 for video).
VT service interruption time	Time interval (in seconds) during which the session is paused (at least one medium, audio or video, is missing) before the session starts again.
VT cut-off ratio	The possibility to use the service and/or its audio and video components until the user requests to release the call. A ViLTE call with an undesired release of the video component, but the audio component still working, is considered as dropped.

The second category concerns video quality (audio quality is considered in [ITU-T G.1028]), with two complementary points of view: global quality (expressed in terms of mean opinion score (MOS)) and detection and characterization of artefacts. The recommended metrics are given in Table 4 below.

Table 4 – QoS parameters video quality measurements

Name	Definition
Video quality (MOS)	Provides an objective view on the quality of the video signal as perceived by the VT customer
Detection of freezing	• number and rate of detections • cumulative duration of all detected events
Detection of blurriness	• number and rate of detections • cumulative duration of all detected events
Detection of pixelization	• number and rate of detections • cumulative duration of all detected events

Guidance is given in clause 10.3 on video quality measurement methods.

– **Freezing**: for reliable transmission freezing is the only distortion caused by transmission problems, in ViLTE it is just one in between others (and a minor one). In principle, it happens only if the (short) buffer runs empty. Based on current knowledge, the player decodes and plays-out what it gets, regardless of how destroyed the packets are. But it is a question of time that players apply other strategies as error concealment or freezing until the next I-frame is received for full synchronization.

– **Blurriness**: caused by low resolution along with compression. Depending on the markets, the native image resolution is usually limited at 240p or 360p (quite blurry on a high definition (HD) phone display). Even if standards allow higher resolutions and also adaptive bit rates. Detection of 'blocks' in case of a 240p I-frame is considered as blurriness.

– **Pixelization**: what can be seen in case of transmission errors is the full set of image distortions caused by wrong updates (erroneous intra-frames). These are false-colour macro-blocks appearing and moving around, wrongly moved macro-blocks in general, freezing part of the image, luminance information that does not fit to the chrominance and more. The effect of 'error propagation' also needs to be considered: one erroneous intra-frame destroys an image, so that, even if all following intra-frames are received without error, the update information is applied to a destroyed image.

9.2 Impact assessment of relevant operating conditions on QoS parameters

Below are feedbacks provided from laboratory or field tests concerning the influence of operating conditions on the various dimensions of QoS for ViLTE. This clause is to be completed in further revisions.

– Codec resolution vs. video quality

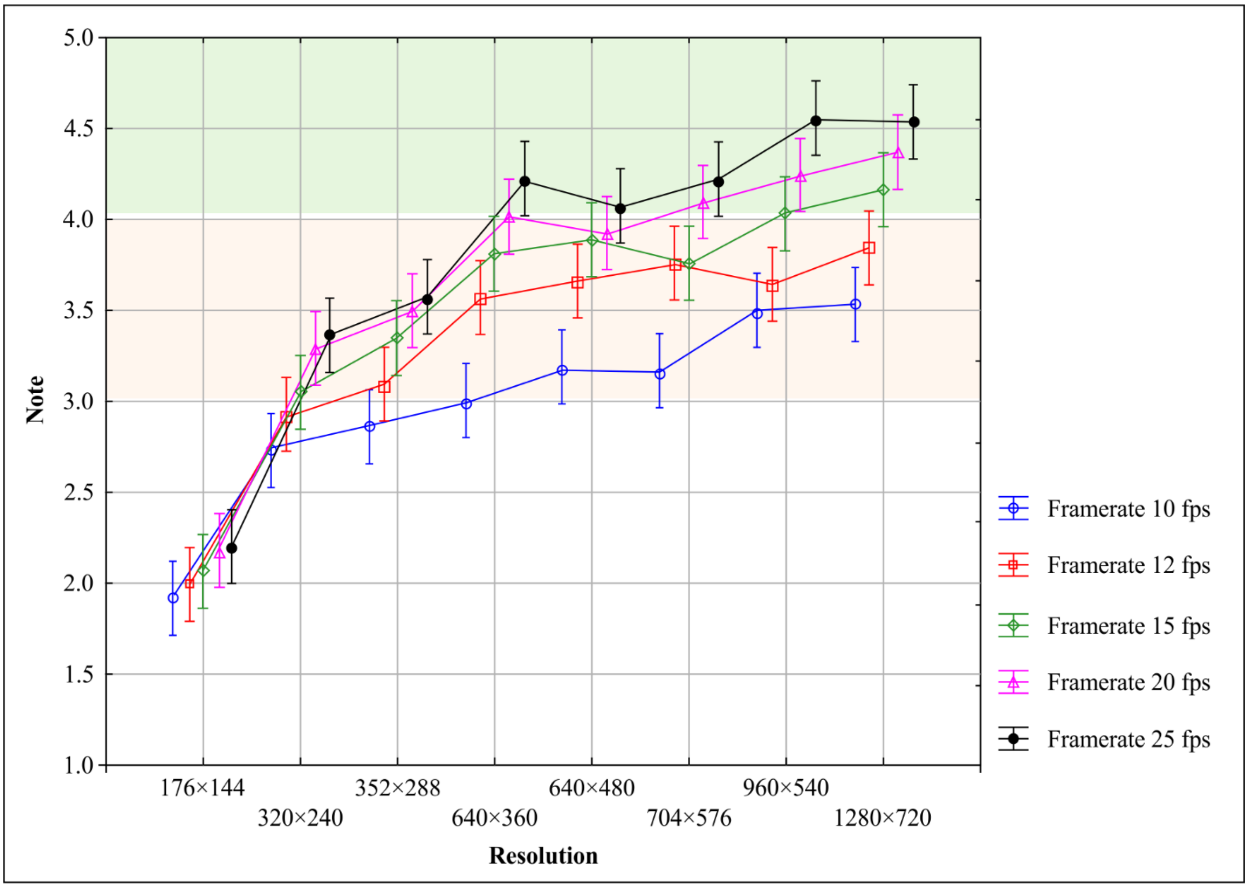


Figure 3 – Codec resolution versus video quality

Results from subjective tests show that video graphics array (VGA) (320×240) at 15 fps can only provide a medium quality user experience (MOS ≈ 3.0). A good quality (MOS ≈ 4) requires a minimum resolution of (640×360) at 15 fps. VGA (640×480) is however, the widely-supported resolution for achieving this quality level.

– Encoding bitrate vs. video quality

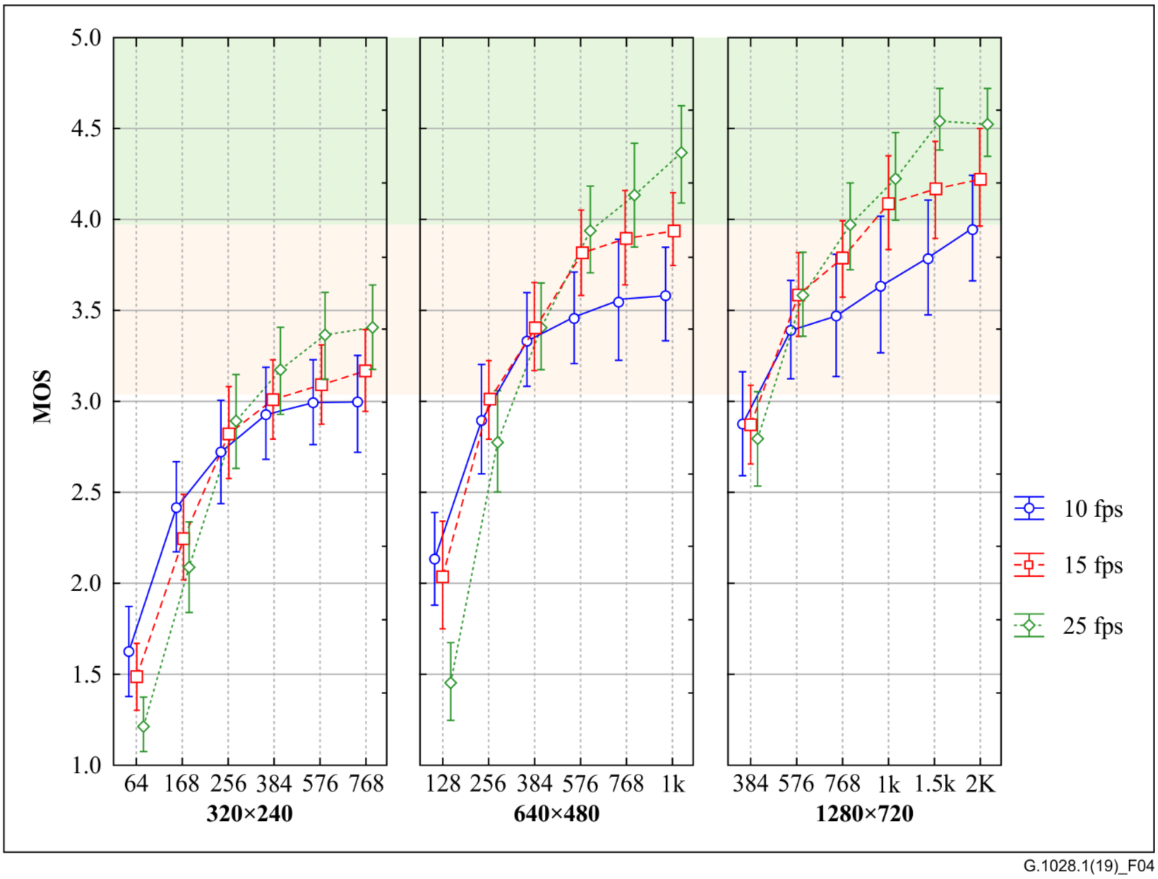


Figure 4 – Bit rate versus video quality

It can be inferred from subjective test results that the optimal operating range for good video MOS using ITU-T H.264 baseline level 3.1 is a VGA resolution with frame rate of 15 fps to 30 fps and a bit rate of 384 kbit/s to 768 kbit/s. Thus, 384 kbit/s is the minimum bit rate to ensure a quite good quality experience (approx. 3.5 MOS) whereas a very good video quality (≥ 4.0 MOS) requires a bit rate up to 768 kbit/s. A better codec will, however, not solve all capacity/coverage issues – rate adaptation is needed. Devices must be capable to detect transmission conditions (at receiver and sender side) and adapt bit rate/frame rate and resolution accordingly.

– Video bit rate vs. capacity

Dedicated bearer for ViLTE (with QCI-2) provides a GBR. The radio scheduler gives more radio resources to this bearer to ensure the GBR at cell edge. With a GBR at 768 kbit/s, a single ViLTE call consumes 20% of the radio resources in uplink (UL) (10 MHz bandwidth); thus, the overall data performance in the cell is impacted.

Table 5 – Video bit rate versus capacity

%rRadio resource allocation per ViLTE terminal		Cell-centre	Mid cell	ViLTE cell edge
10 MHz	384 kbit/s	2.0%	8.0%	12%
	768 kbit/s	3.2%	2.0%	20.6%
20 MHz	384 kbit/s	1.0%	4.0%	6.0%
	768 kbit/s	1.6%	5.5%	10.3%

To ensure ViLTE quality of experience (QoE) without affecting other user's throughput, practical best-fit measures would have to be taken. The use of GBR = MBR is actually not suitable for operation with rate adaptation. The possible options for maximum video quality are:

– Use of QCI-2 with GBR < MBR; or
– Use of a non-GBR QCI (6 or 7) with scheduling priority + cell based min bit rate + ViWifi whenever possible.

In summary, to ensure a more optimal ViLTE performance there is the need for efficient codecs to reduce the video bit rate, adapt the video bit rate to transmission condition with rate adaptation and possibly consider other QCI options than 2.

– Jitter buffer performance vs. video quality;
– RTP packet loss vs. setup time (PDD);
– Tracking area update vs. setup time;
– Coverage/interference vs. service availability and call cut-off ratio;
– Handover vs. service interruption time.

9.3 Quality targets

This clause is for further study. Table 6 below will be completed once feedback from field deployments are available.

Table 6 – Quality budget allocation

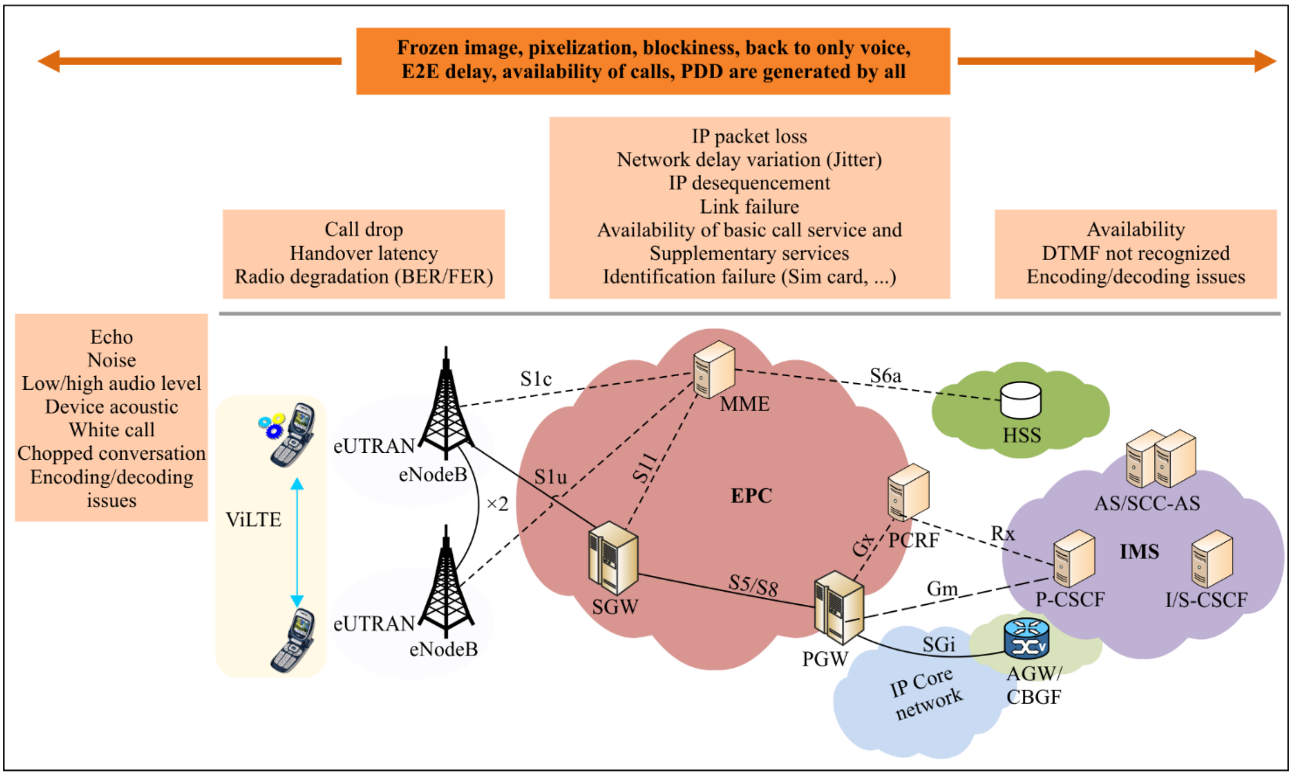
Network segment	**LTE-LTE (intra)**		**LTE-LTE (with interconnection)**		**LTE-LTE (with roaming)**	
	Indicator A	Indicator B	Indicator A	Indicator B	Indicator A	Indicator B
UE						
E-UTRAN						
EPC						
IMS/AS						
Total budget						
	Indicator C	Indicator D	Indicator C	Indicator D	Indicator C	Indicator D
UE						
E-UTRAN						
EPC						
IMS/AS						
Total budget						
	Indicator E	Indicator F	Indicator E	Indicator F	Indicator E	Indicator F
UE						
E-UTRAN						
EPC						

Table 6 – Quality budget allocation

Network segment	LTE-LTE (intra)		LTE-LTE (with interconnection)		LTE-LTE (with roaming)	
IMS/AS						
Total budget						
	Indicator G	Indicator H	Indicator G	Indicator H	Indicator G	Indicator H
UE						
E-UTRAN						
EPC						
IMS/AS						
Total budget						

10 Diagnostic strategy for QoS degradations

This clause explains the various video-centric degradations that can be encountered on a mobile LTE network. Main elements of the mobile network are depicted to show the signalling and media elements as well as the connections with public switched telephone network (PSTN) or mobile platforms.



G.1028.1(18)_F05

Figure 5 – Sources of potential audiovisual impairments in ViLTE

In order to have a point of comparison in terms of QoS delivered, a reference call is taken whose ideal characteristics are:

– Fourth generation (4G)-4G call with end-to-end codecs (adaptive multi-rate wideband (AMR-WB) for audio, ITU-T H.264/ITU-T H.265 for video) and associated video features (frame rate, ITU-T H.264/ITU-T H.265 profile, video orientation) correctly negotiated;
– No degradation on EPC (no IP loss, no load, etc.);
– No degradation on E-UTRAN (no radio degradation, no congestion, etc.);
– Wideband audio compliant devices, both on 4G, with excellent acoustic, voice quality enhancement algorithms (noise reduction (NR), acoustic echo control (AEC) and automatic gain control (AGC)), electronic;
– Quiet environment on both ends;
– All services are available (e.g., call transfer, dual-tone multi-frequency (DTMF)).

Below are the main possible technical reasons, which generate the encountered degradations. Separation is done according to the impact assessment due to the customer.

10.1 QoS problem source-linked to availability of service

Table 7 – Degradations related to availability of the service and their potential causes

Kind of degradation	Possible reasons	Location
UE identification failure	• problem with MME, HSS or policy and charging rule function (PCRF)	EPC
Unavailability of basic call	• error in scheduling • radio resource control (RRC) connection setup failure (reception of RRC connection reject, or expiry of timer T300, no RRC connection setup complete sent after reception of RRC connection setup)	EUTRAN
	• not available due to load (S-GW or P-GW) • failed negotiation (e.g., allocation of QCI, codec) • reception of several SIP error codes (e.g., 401 = Unauthorized, 405 = Method Not Allowed) • reception of SIP CANCEL from IMS • TD internal timer expired, causing a "SessionSetupFailureTimeout"	EPC
Unavailability of video component	• failed negotiation (e.g., allocation of QCI, codec, resolution)	EPC/Terminal
PDD	• load • interworking between systems • circuit switched (CS) fallback at call setup	All
Link failure	• bad negotiation between two pieces of equipment of the network during call establishment (bad codec management)	EUTRAN/ EPC
White call	• terminal is not able to code or decode speech while the signalling is OK for the communication	Terminal

10.2 QoS problem source-linked to network performance

In this sub-section, QoS degradations, which are linked to network performance, are depicted. In concrete terms, these degradations, specific to network mainly lead to video degradation from the customer's perspective.

Table 8 – Degradations related to network performance and their potential causes

Kind of degradation	Possible reasons	Location
Frozen image	• no reception of video frame • network congestion (several causes: traffic load, distance from cell center causing activation of TTI bundling, for instance) • jitter buffers not adapted to actual jitter amount	All
Blurriness	• no reception of infra video frame • recovery strategy of decoder in terminal	All
Blockiness / Pixelization	• no reception of infra video frame • recovery strategy of decoder in terminal	All
Encoding/Decoding issues		Terminal/ eUTRAN
E2E delay (latency)	• network load • media handling (packet construction, jitter buffer management) • speech processing in terminals • random access channel (RACH) upon receiving handover command • RACH/contention procedure • additional RACH attempts • dynamic scheduling, link adaptation • radio link failure/re-establishment during handover (possibly different cell)	All
Bad (lip) synchronization between voice and video	• network congestion associated with differentiated QoS (QCI) • different jitter buffer size and behavior • decoding time	All
RTCP/IP packet loss	• network congestion (several causes: traffic load, distance from cell center causing activation of TTI bundling, for instance) • jitter buffers not adapted to actual jitter amount or packet size (can depend on use of robust header compression (RoHC) or not)	EPC / Terminal
RTP/IP desequencing	• new route after a problem such as congestion	EPC
Network delay variation (Jitter)	• network congestion. • jitter buffers not adapted	EPC / Terminal
Radio degradations	• limit of the cell coverage • interference • area not well covered (e.g., obstacle) • bad radio optimization. • radio loss profile • bad radio scheduling • no or bad use of hybrid automatic-repeat-request (HARQ) • mechanisms etc.	eUTRAN

Table 8 – Degradations related to network performance and their potential causes

Kind of degradation	Possible reasons	Location
Handover latency	• latency due to new route after handover or single radio voice call continuity (SRVCC)	EPC / CS network
Call drop	• terminal bug, bad covered area, handover/SRVCC failures due to cells neighborhood problem, etc. RRC connection drop (at reception of RRC connection Re-establishment reject, or expiry of timer T301 or in case RRC connection release is received before new RRC connection setup attempt)	Terminal/ eUTRAN
	• link failure: system failure, bad re-negotiation between two equipments of the network during call • reception of SIP status code 500 (server internal rrror) • no RTP packet received during a period longer than "*SessionDropTimeout*" TD internal timer • no SIP 200 OK on BYE is received within the time measured by "*SessionHangupTimeout*" TD internal timer	EPC
Back to voice-only communication	• network congestion (several causes: traffic load, distance from cell center causing activation of TTI bundling) • strategy of service provided and/or device maker	

10.3 Tools and models for measurement and prediction of video quality

This clause is a complement for video to clause 10.3.2 of [ITU-T G.1028] where an overview of tools and models for voice quality is provided.

A global view of all standard quality assessment methods is given in Table 10.3 of [ITU-T G.1011] and shows a detailed application scope of each model in terms of supported resolutions and codecs.

Following the taxonomy provided there, the potential methods are:

– Media layer models: all models for video media streaming quality assessment:
  - full reference: [ITU-T J.144] (standard definition (SD)), [ITU-T J.247] (quarter common intermediate format (QCIF), common intermediate format (CIF), VGA), [ITU-T J.341] (HD);
  - reduced reference: [ITU-T J.249] (SD), [ITU-T J.246] (QCIF, CIF, VGA), [ITU-T J.342] (HD);
  - no reference: none.

– Packet layer models:
  - models for planning purposes: [ITU-T G.1070] (dedicated tool for video telephony, including also an audio quality module), [ITU-T G.1071] (for video streaming, SD, HD);
  - models for monitoring purposes (no reference) on UDP for video media streaming quality assessment: [ITU-T P.1201.1] (QCIF, quarter video graphics array (QVGA), half video graphics array (HVGA))), [ITU-T P.1201.2] (SD, HD), [ITU-T P.1201] Amd. 2, App. III (HVGA, HD (1080i50, 1080p24, 1080i60, 1080p30)).

– Bitstream layer models (no reference) on UDP for audiovisual media streaming quality assessment:
  - [ITU-T P.1202.1] (QCIF, QVGA, HVGA), [ITU-T P.1202.2] (SD, HD).

– Hybrid models: all models for video media streaming quality assessment:

- full reference: [ITU-T J.343.5] (HD, encrypted bitstream), [ITU-T J.343.6] (HD, not encrypted bitstream);
- reduced reference: [ITU-T J.343.3] (HD, encrypted bitstream), [ITU-T J.343.4] (HD, not encrypted bitstream);
- no reference: [ITU-T J.343.1] (HD, encrypted bitstream), [ITU-T J.343.2] (HD, not encrypted bitstream).

With the exception of [ITU-T G.1070], all these methods have been developed for an application on video or audiovisual streaming services, not for video telephony. Because of the relatively good similarity between contents of both types of services, their application for the evaluation of quality of video telephony services may be envisaged, though it need to be understood that it would require some hard validation work.

Indeed, implementation of these methods raises some important concerns:

– Rating the impairments of video quality based on (encrypted) bitstreams is rather complicated. A bitstream method can give some metrics, how it would look like in a general statistical view under assumption of medium players and encoding strategies, but the accuracy and the relevance of the measurement results must be considered with highest caution;

– Full reference methods require the possibility to inject a reference video or audiovisual content at the applicative level inside UEs instead of the content provided by the camera. This feature is currently not supported on almost all models of mobile devices;

– The video player strategy of mobile devices to deal with errors and to minimize their visibility is varying between devices. Models must be calibrated before they can be applied on a given model of device.

Bibliography

[b-GSMA IR.94] GSMA IR.94 v 11.0 (2016), *IMS Profile for Conversational Video Service*.

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, 2019