

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.1071

(11/2016)

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

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

Opinion model for network planning of video and audio streaming applications

Recommendation ITU-T G.1071

ITU-T

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.1071

Opinion model for network planning of video and audio streaming applications

Summary

Recommendation ITU-T G.1071 provides algorithmic models for network planning of IP-based video services. This Recommendation addresses two application areas:

- the higher resolution (HR) application area, including services such as IPTV;
- the lower resolution (LR) application area, including services such as mobile TV.

The algorithmic model addressing the HR application area is described in Annex A.

The algorithmic model addressing the LR application area is described in Annex B.

A further algorithmic model extends the HR application area to include ITU-T H.265 high efficiency video coding (HEVC), and is described in Annex C.

The application of the models is limited to quality of experience (QoE/ quality of service QoS planning. Other applications such as quality benchmarking and monitoring are outside the scope of this Recommendation.

As input, the models take network planning assumptions, for example, the video resolution, the audio and video codec types and profiles, the audio and video bitrates, the packet-loss-rate and the packet-loss distribution.

As output, the model algorithms provide individual estimates of audio, video and audiovisual quality in terms of the five-point absolute category rating (ACR) mean opinion score (MOS) scale. Further diagnostic information on causes of quality degradations can also be made available.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T G.1071	2015-06-29	12	11.1002/1000/12512
2.0	ITU-T G.1071	2016-11-29	12	11.1002/1000/13125

Keywords

Audio, audiovisual, IPTV, mean opinion score (MOS), mobile TV, network planning, quality, QoE, video.

* 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 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/>.

© ITU 2017

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

		Page
1	Scope.....	1
2	References.....	1
3	Definitions	2
3.1	Terms defined elsewhere	2
3.2	Terms defined in this Recommendation.....	2
4	Abbreviations and acronyms	2
5	Conventions	4
6	Areas of application.....	4
7	Model framework	7
8	Model output information and performance details	7
	Annex A – Description of the ITU-T G.1071 model algorithm for HR application area.....	9
A.1	Audio module	9
A.2	Video module	11
A.3	Audiovisual module.....	16
	Annex B – Description of the ITU-T G.1071 model algorithm for LR application area	17
B.1	Audio module	17
B.2	Video module	19
B.3	Audiovisual module.....	24
	Annex C – Description of the ITU-T G.1071 model algorithm for HR video applications using HEVC.....	28
C.1	Scope	28
C.2	Model framework	29
C.3	Model output information and performance details	30
C.4	Model algorithm	30
	Bibliography.....	35

Recommendation ITU-T G.1071

Opinion model for network planning of video and audio streaming applications

1 Scope

This Recommendation provides models which deliver estimates of the impact of typical IP network impairments on the quality experienced by the end user in multimedia mobile streaming and Internet protocol television (IPTV) applications over transport formats such as: real-time transport protocol (RTP) (over user datagram protocol (UDP)), motion picture experts group-2 transport stream (MPEG2-TS) (over UDP or RTP/UDP), 3rd generation partnership project packet-switched steaming service (3GPP-PSS) (over RTP).

The models are network planning tools. They are of help in selecting IP-network transmission settings such as the audio and video format, the audio and video codecs and the audio and video bitrates, under the assumption that the network is prone to packet loss.

This Recommendation targets the same services as [ITU-T P.1201] and [ITU-T P.1202]. In particular, this Recommendation covers the same coding technologies and unreliable network mechanisms such as UDP. Moreover, this Recommendation extends the higher resolution (HR) application area to include [ITU-T H.265] high efficiency video coding (HEVC). However, this Recommendation is limited to quality of experience (QoE)/ quality of service (QoS) planning, while [ITU-T P.1201] and [ITU-T P.1202] are dedicated to service monitoring and benchmarking.

The following Recommendations have also been developed for QoE/QoS planning, but for different applications:

- [ITU-T G.107]: speech (telephone band)
- [ITU-T G.107.1]: speech (wideband)
- [ITU-T G.1070]: videophone
- [ITU-T G.1030]: Appendix II: web browsing

For a summary of the services, encoding and network characteristics covered by the [ITU-T G.1071] model algorithms, refer to Table 1 of clause 6.

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.107] | Recommendation ITU-T G.107 (2015), <i>The E-model: a computational model for use in transmission planning</i> . |
| [ITU-T G.107.1] | Recommendation ITU-T G.107.1 (2015), <i>Wideband E-model</i> . |
| [ITU-T G.1030] | Recommendation ITU-T G.1030 (2014), <i>Estimating end-to-end performance in IP networks for data applications</i> . |
| [ITU-T G.1070] | Recommendation ITU-T G.1070 (2012), <i>Opinion model for video-telephony applications</i> . |

- [ITU-T H.264] Recommendation ITU-T H.264 (2014), *Advanced video coding for generic audiovisual services*.
- [ITU-T H.265] Recommendation ITU-T H.265 (2015), *High Efficiency Video Coding*.
- [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] Recommendation ITU-T P.1202 (2012), *Parametric non-intrusive bitstream assessment of video media streaming quality*.
- [ITU-T P.1401] Recommendation ITU-T P.1401 (2012), *Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following term defined elsewhere:

3.1.1 mean opinion score (MOS) [b-ITU-T P.800.1]: The mean of opinion scores, i.e., of the values on a predefined scale that subjects assign to their opinion of the performance of the telephone transmission system used either for conversation or for listening to spoken material.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 model, model algorithm: An algorithm used for estimating the subjective (perceived) quality of a media sequence.

3.2.2 sequence: A short decoded audio, video or audiovisual portion of a stream, typically shorter than 30 seconds.

3.2.3 compression artefacts: Artefacts that are introduced due to lossy compression of the encoding process.

3.2.4 slicing artefacts: Artefacts that are introduced when packet losses are concealed through use of a packet loss concealment (PLC) scheme to repair erroneous frames.

3.2.5 freezing artefacts: Artefacts that are introduced when the packet loss concealment (PLC) scheme of the receiver replaces the erroneous frames (either due to packet loss or error propagation) with the previous error-free frame until a decoded picture without errors has been received. Since the erroneous frames are not displayed, this type of artefact is also referred to as freezing with skipping.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

- AAC Advanced Audio Coding
- AAC-LC Advanced Audio Coding – Low Complexity

AC3	Audio Coding 3
ACR	Absolute Category Rating
AMR-NB	Adaptive Multi-Rate – Narrowband
AMR-WB	Adaptive Multi-Rate – Wideband
ARQ	Automatic Repeat Request
DASH	Dynamic Adaptive Streaming over HTTP
FB	Fullband
FEC	Forward Error Correction
GOP	Group of Pictures
HD	High Definition (television)
HE-AAC	High-Efficiency Advanced Audio Coding
HEVC	High Efficiency Video Coding
HR	Higher Resolution
HRC	Hypothetical Reference Circuit
HTTP	Hypertext Transfer Protocol
HVGA	Half Video Graphics Array
IPTV	Internet Protocol Television
LR	Lower Resolution
MBMS	Multimedia Broadcast/Multicast Service
MOS	Mean Opinion Score
MPEG	Motion Pictures Experts Group
NB	Narrowband
NTSC	National Television Standard Committee
PAL	Phase Alternating Line
PCC	Pearson Correlation Coefficient
PES	Packetized Elementary Stream
PLC	Packet Loss Concealment
PVS	Processed Video Sequence
QCIF	Quarter Common Intermediate Format
QoE	Quality of Experience
QoS	Quality of Service
QVGA	Quarter Video Graphics Array
RMSE	Root Mean Square Error
RTP	Real-time Transport Protocol
SD	Standard Definition
SRC	Source Reference Channel or Circuit
SWB	Superwideband

TS	Transport Stream
UDP	User Datagram Protocol
VSP	Visual Simple Profile
WB	Wideband

5 Conventions

None.

6 Areas of application

Table 1 below shows the application range of the models based on what the models have actually been developed for. The application range for the model applicable to HEVC is given in Annex C.

Table 1 – Factors and application ranges of the ITU-T G.1071 model algorithm

	ITU-T G.1071 – Lower resolution (LR)	ITU-T G.1071 – Higher resolution (HR)
Application information	Value range, unit	
Sequence duration (Ts)	It is expected that the model will give reliable prediction results for sequence durations within the range 8-24 seconds	It is expected that the model will give reliable prediction results for sequence durations of approximately 8-16 seconds
Packetization	3GPP MBMS, PSS or using RTSP directly (all three over RTP/UDP/IP)	MPEG2-TS/RTP/UDP/IP RTP/UDP/IP (Note 3) MPEG2-TS/UDP/IP (Note 3)
Video codec	MPEG4 visual simple profile (VSP) ITU-T H.264 baseline profile	ITU-T H.264 main profile, ITU-T H.264 high profile
Video size	QCIF, QVGA, HVGA	SD: PAL, NTSC HD: 720p, 1080p, 1080i (High profile: 1080; main profile: 720, SD)
Audio codec	AMR-NB, AMR-WB+, AAC-LC, HE-AAC (v1, v2)	MPEG-4 AAC-LC MPEG-4 HE-AAC (V1 and V2 = 3GPP enhanced AAC+) MPEG-1 Layer 2 AC3
Coded video bitrate	[ITU-T H.264] QCIF: 32-1000 kbit/s QVGA: 80-3000 kbit/s HVGA: 192-6000 kbit/s MPEG4 QCIF: 40-1500 kbit/s QVGA: 90-3500 kbit/s HVGA: 192-6000 kbit/s	HD [ITU-T H.264]: 0.5 up to 30 Mbit/s SD [ITU-T H.264]: 0.5 up to 9 Mbit/s

Table 1 – Factors and application ranges of the ITU-T G.1071 model algorithm

	ITU-T G.1071 – Lower resolution (LR)	ITU-T G.1071 – Higher resolution (HR)
Application information	Value range, unit	
Coded audio bitrate	AMR-NB: 4.75-12.2 kbit/s AMR-WB+: 10.4-48 kbit/s AAC-LC: 16-128+ kbit/s HE-AAC (v1, v2): 32-128 kbit/s	AAC-LC: 32-576 kbit/s HE AACv2: 16-96 kbit/s MPEG-1 Layer 2: 64-384 kbit/s AC3:64-384 kbit/s
Video decoder packet loss concealment	Two types of assumed decoder behaviour are covered: 1) freezing with skipping; 2) slicing with: MPEG4: 1 slices/frame [ITU-T H.264]: 1 slice/packet Both MPEG4 and [ITU-T H.264]: Fixed PLC (using fixed decoder, details and settings).	Types of decoder behaviour: two dimensions: slicing, PLC 1) freezing with skipping (duration(source)=duration(processed sequence)); 2) slicing with 1 slice/frame; 3) slicing with 1 slice per macroblock row, PLC with zero-motion copy (temporal from same region of previous good frame).
Audio decoder packet loss concealment	Decoder default modes: Codec-implementation specific loss concealment	Decoder default modes: Codec-implementation specific loss concealment
Retransmission mechanisms (ARQ); forward error correction (FEC); client jitter buffer behaviour	Developed models assume that the de jitter buffer, ARQ and FEC mechanisms have already corrected the stream.	(No rebuffering) Developed models assume that the de jitter buffer, ARQ and FEC mechanisms have already corrected the stream.
Encoder implementation	The model has been developed using the following encoders (Note 1): Video: - MPEG4 Part 2: ffmpeg [ITU-T H.264] (MPEG4 Part 10): x264 Audio: - AMR-NB/WB+: According to standard - AAC-LC, HE-AAC(v1, v2): Nero	The model has been developed using the following encoders (Note 1): Video: [ITU-T H.264] (MPEG-4 Part 10): x264 Audio: - AAC-LC, HE-AAC v2: Nero - MPEG1-LII and AC3: ffmpeg
Decoder implementation	Reference decoder was a proprietary decoder provided by one proponent, which also performed de-packetization and audio-video-demultiplexing. The [ITU-T H.264] -decoding is standard-conformant, with the PLC as described above (Note 2).	Reference decoder was a proprietary decoder provided by one proponent, which also performed de-packetization and audio-video-demultiplexing. The ITU-T H.264-decoding is standard-conformant, with the PLC as described above (Note 2).

Table 1 – Factors and application ranges of the ITU-T G.1071 model algorithm

	ITU-T G.1071 – Lower resolution (LR)	ITU-T G.1071 – Higher resolution (HR)
Application information	Value range, unit	
Group of pictures (GOP)	Typical GOP structure for which the model has been trained: M = 1, N = 40 (typically no B frames for mobile case) Length: fixed, variable, adaptive Structure (e.g., IPPP...PPPI)	Supporting default modes for typical GOP structures E.g., M = 3, N = 15 Length: fixed, variable, adaptive Structure (e.g., IBBPBB...PBBI) NOTE – GOP structure is explicitly estimated from stream.
Frame rate	5, 8.33, 12.5, 15, 20, 25, 30 fps	SD: 50i (PAL), 59.94i (NTSC) HD: 50p, 59.94p, 60p, 50i, 59.94i, 60i, 25p, 29.97p, 30p
Audio channel number	1 (diotic mono), 2 (stereo)	2 (stereo)
# of Audio frames per RTP packet	1 to 5 audio frames	Bitrate-specific (depending on both the audio and video bitrate)
Audio-video multiplexed?	Default: No, at RTP-level; no audio-video asynchrony	In MPEG-2 TS/RTP/UDP and MPEG-2 TS/UDP: Supported
Packet loss degradation, video	Uniform loss:0-10% Burst loss:0-10% (4-state Markov model)	Uniform loss:0-2% Burst loss:0-2% (4-state Markov model)
Packet loss degradation, audio	Uniform loss:0-10% Burst loss:0-10% (4-state Markov model)	Uniform loss:0-6% Burst loss:0-6% (4-state Markov model)
Symmetrical versus asymmetrical handling of audio and video in audiovisual case	Model application: Symmetrical, but can handle asymmetric cases due to specific model development process	Model application: Symmetrical, but can handle asymmetric cases due to specific model development process
NOTE 1 – It is assumed that the model can be used for estimating quality when other encoder implementations for the given codec have been used. However, if the encoder performance is significantly worse or better than for the encoder used, the model prediction accuracy will be reduced. NOTE 2 – One aspect not covered by decoder packet loss concealment is post-filtering. Guidance on how to adjust internal model parameters for specific other decoders, including set-top boxes, is for further study. NOTE 3 – The ITU-T P.1201.2 model has been trained on MPEG-2 TS/RTP/UDP. However, due to the design of the ITU-T P.1201.2 algorithm, it is also applicable to MPEG2-TS/UDP/IP. Further, to the model's design, it is assumed to also work for RTP/UDP/IP transport with similar, but so far unverified, accuracy as compared to MPEG2-TS/RTP/UDP/IP.		

Factors and applications not covered by the model are:

- evaluation of audiovisual quality including display/device properties;
- audiovisual streaming with significant rate adaptation (such as that used in dynamic adaptive streaming over HTTP (DASH) streaming);
- transcoding situations;

- the effects of audio level, noise and delay (and corresponding similar video factors);
- re-buffering degradation of audio, video and audiovisual;
- coding technologies the models are not intended for: [ITU-T H.261], MPEG-2, [ITU-T H.263], etc.

7 Model framework

As shown in Figure 1, the [ITU-T G.1071] models follow the same structure as the [ITU-T P.1201] models. Indeed, they are composed of three modules: the audio module, the video module and the audiovisual module. However, in contrast to the [ITU-T P.1201] models, the [ITU-T G.1071] models take as inputs network planning assumptions instead of the encrypted bitstream. Some of the [ITU-T P.1201] input parameters are available as planning assumptions as well. For the [ITU-T P.1201] input parameters which are not available as planning assumptions, the [ITU-T G.1071] models provide a set of rules to convert planning assumptions into these [ITU-T P.1201] input parameters (see block "conversion rules").

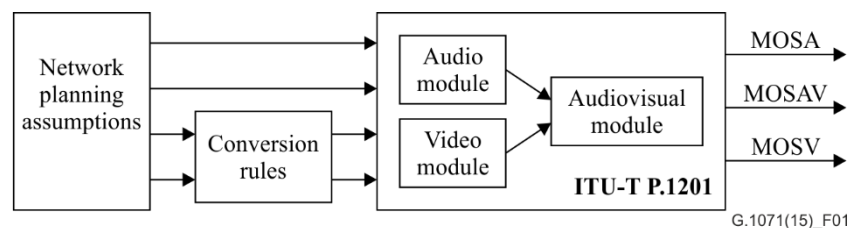


Figure 1 – Model framework

8 Model output information and performance details

ITU-T G.1071 has three output parameters:

- 1) estimated audiovisual MOS on the 1 to 5 scale, which is an estimation of the perceived audiovisual quality;
- 2) estimated video MOS on the 1 to 5 scale, which is an estimation of the perceived video quality (without audio present). The model is able to give both a video score for a stream without audio and a stream including audio;
- 3) estimated audio MOS on the 1 to 5 scale, which is an estimation of the perceived audio quality (without video present). The model is able to give both an audio score for a stream without video and a stream including video.

The performance information for the ITU-T G.1071 models can be found in Table 2 and Table 3 for the HR and LR application areas respectively. The statistical metrics root mean square error (RMSE) and the Pearson correlation are used to describe the performance, see [ITU-T P.1401]. (For these performance figures, the subjective ratings have been mapped to the model scores using a linear, i.e., 1st-order, mapping function at a per-database-level. This has been done in order to avoid misalignment due to bias in the different subjective tests, e.g., as a result of different test settings). Moreover, the performance information for the HR application area in ITU-T H.265 encoding can be found in Tables C.3 and C.4 of Annex C.