A New Parameterized Model for Determining Quality of Online Video Service Using Modern H.265/HEVC and VP9 Codecs

Tadeus Uhl\textsuperscript{1} and Christian Hoppe\textsuperscript{2}

\textsuperscript{1} Maritime University of Szczecin, Szczecin, Poland
\textsuperscript{2} Nextragen Solutions GmbH, Handewitt, Germany

Abstract—This paper describes a new measurement method (VS model) for determining the quality of online video services relying on modern H.265/HEVC and VP9 codecs. The said method has been developed on the basis of VQuad-HD curves (according to ITU-T J.341). This model does not refer to signal analysis, but protocol analysis instead. The parameters used are: type of video codec, encoding rate, transport technique, packet loss and burst size. The method may be implemented quickly and easily, which is one of the great advantages when using this method to measure QoS.

Keywords—H.265/HEVC, IP network, ITU-T, J.341, native RTP, QoS measurement techniques, video service, VP9.

1. Introduction

In current IP networks, new types of services continue to appear. Many of them are related to transporting video content that constitutes a major portion of today’s Internet traffic. According to CISCO reports [1]–[2], in 2017, video material accounted for 75% of global Internet traffic and for nearly 60% of mobile traffic. The volume of video traffic is growing rapidly, especially in mobile communications, and is estimated to reach nearly 80% of the world’s mobile traffic by 2022. It is expected that mobile video will increase nine-fold between 2017 and 2022. Therefore, it makes sense to take a closer look at video services.

Video services are currently rendered in two forms. The first category includes real-time services, e.g., television over IP (IPTV), video telephony over IP (VToIP). The other form encompasses services in which time does not play a principal role. These include video streaming services, e.g. video on demand (VoD). To guarantee a good-quality video service, effective transport techniques are needed. In real-time services, MPEG-2 transport streams based on ISO/IEC 13818-1 [3] and native RTP based on IETF RFC 3640 [4] are currently the most widely used techniques. In video streaming (non-real-time service), native RTP and MPEG-DASH based on ISO/IEC 23009-1:2012 [5] are currently the most widely used techniques. It is worth noting that the native RTP technique is used in both forms of video service. This is a sufficient reason to devote a complete paper to this particular transport technique that will serve as a basis of all investigations described herein.

In order to ensure high video service image quality on the one hand and to reduce the coding rate (and thus the required bandwidth in the network) on the other, improvements in image coding have been promoted in recent years. A new ITU-T standard H.265/HEVC (high-efficiency video coding) was published in 2013 [6]. A few years later, a new proprietary solution (Google), the VP9 Codec, was launched [7]. Image resolution is also increasing continuously, with full HD succeeding the HDTV format. New formats, 2K and 4K, are gaining in popularity as well. The impact that the developments mentioned here exert on the quality of video services must be measured in practice.

Quality of service (QoS) and quality of experience (QoE) should be continuously monitored in modern networks, preferably automatically. Dedicated measuring systems and methods are obviously indispensable for that purpose. When it comes to video communications, several methods are already available for measuring QoS and QoE. Most of them have been standardized, e.g. ITU-T Rec. J.144 [8], ITU-T J.147 [9], ITU-T J.247 [10], ITU-T J.341 [11], ITU-T P.1204.4 [12]. All these standards have evolved from signal-based QoE measuring methods with “full reference” (intrusive measurement necessary). That is what makes these measuring techniques so complicated, time-consuming, and expensive due to the licensing costs involved.

The second group of QoS measurement methods is based on parameterized models. They are simple, faster to execute and generally more accessible (without licenses). However, they do not provide as accurate results as the QoE measurement methods mentioned above. The most important parameterized QoS models for video services include: mean squared error (MSE) [13], new quality index (NQI) [14] and structural similarity index (SSIM) [15]. A comparison of performance of the above-mentioned QoS techniques was presented in papers [16], [17] and else-
where. Specialized QoS models exist as well, e.g. for VToIP [18] or for IPTV services [19]. The last two QoS models were developed for the “old” video codecs (e.g. H.264/AVC) and for lower picture resolutions (up to SDTV). There are currently no QoS models that do justice to the latest developments in the field of video codecs (such as H.265/HEVC, VP9) and picture formats (such as HDTV, Full HD). The work described in this paper aims to remedy that deficiency.

After the introduction presented above, a new video service (VS) model will be formulated and illustrated in the context of actual video codecs and picture formats. This model is then analyzed in various applications and its practicability is put to the test in a comparison study. The results obtained based on the analyses are presented graphically and are interpreted. The paper is concluded with a summary and an outlook on future work.

2. New Parameterized VS Model

It is generally known that packet losses are some of the worst impairment parameters in the IP environment. This is true in the case of both audio and video transmission. Other parameters that greatly affect QoS values include codec type, encoding rate and burst size. The size of jitter buffers in the terminal equipment has a significant influence on QoS as well. The effects of jitter and out-of-order packet delivery are converted into losses, bearing in mind that these errors may be smoothed out with the aid of a jitter buffer. Two types of packet losses and burst size, i.e. burst ratio, being the ratio of the measured burst size to the burst size expected in a burst-free environment, can be determined by means of practical measurements during an RTP session. Protocol analysis may also determine the video codec, the encoding rate and the type of transport technique used. This is one of the prerequisites enabling the parameterized QoS model to be used in practice. To investigate the effects of packet loss, burst size and encoding rate on the quality of video services, several extensive numerical calculations, which took several weeks, were carried out within the scope of this paper. The QoSCalc (IPTV) numerical tool [20] was used. The tool relies on the VQuad-HD algorithm [21] to determine the QoS (“full reference” technique). Two modern codecs were selected for the investigation scenarios, i.e. H.265/HEVC and VP9 (video sequence in full HD resolution of 1920 × 1080). The native RTP method was used here as the transport technique. This technique requires the UDP protocol within the transport layer. This combination of protocols is most frequently used for real-time web-based video services (e.g. VToIP, IPTV).

Figures 1 and 2 show examples of numerical results for the H.265/HEVC codec, i.e. the packet loss and burst size values assumed for the investigation. They show that all QoE curves progress exponentially. It is evident that as packet losses increase, the curves fall rapidly and then run asymptotically at a value between 1 and 1.5 MOS. It is also evident that burst size has an equally considerable impact on QoE values, with a burst size of “1” producing the worst quality of service in all three cases and increases in burst size leading to improvements in service quality. This is in keeping with the psycho-visual model of the human visual perception system, which asserts that viewers will accept a few isolated major disruptions in reception far more readily than numerous, regularly occurring disturbances of minor character. The QoE curves shown in Figs. 1 and 2, supplemented by curves for further encoding rates, form the basis for creating new parameterized QoS models.

![Fig. 1. VQuad-HD values as a function of packet loss and burst size at an encoding rate of 3875 kbps for the H.265/HEVC codec and native RTP.](image1)

![Fig. 2. VQuad-HD values as a function of packet loss and burst size at an encoding rate of 6750 kbps for the H.265/HEVC codec and native RTP.](image2)
In the study, for reasons of compatibility with paper [19], the encoding rate and the results obtained from the approximations made in Eqs. (2) to (5) in case of the encoding rate for the H.265/HEVC codec and the native RTP are:

Step 2. Approximate each of the SwissQual VQuad-HD curves as a function of packet losses for burst sizes of “1” to “5” and selected encoding rates using:

\[
VS\text{factor} = P \cdot e^{-a \cdot \text{packetloss}} + Q \cdot e^{-b \cdot \text{packetloss}}.
\]

(1)

Constants \(a\) and \(b\) are selected so that they display values equal to or lower than zero, with significantly smaller values being chosen for \(b\). The result of this is that the two summands are responsible for the steepness of the curves when packet losses are few (second summand) or many (first summand). All constants (\(P\), \(Q\), \(a\), and \(b\)) are now calculated iteratively as best possible values for each encoding rate.

Step 3. Determine the formulas the constants \(P\), \(Q\), \(a\), and \(b\). Once determined, the constants are recorded as functions of the encoding rate. The corresponding formulas can then be calculated by means of polynomial approximation. The degree of the polynomial is determined by the complexity of the curve. Equations (2) to (5) show the simple relationship:

\[
P = w_n \cdot \text{Bitrate}^n + w_{n-1} \cdot \text{Bitrate}^{n-1} + \cdots + w_1 \cdot \text{Bitrate} + w_0 ,
\]

(2)

\[
Q = z_n \cdot \text{Bitrate}^n + z_{n-1} \cdot \text{Bitrate}^{n-1} + \cdots + z_1 \cdot \text{Bitrate} + z_0 ,
\]

(3)

\[
a = x_n \cdot \text{Bitrate}^n + x_{n-1} \cdot \text{Bitrate}^{n-1} + \cdots + x_1 \cdot \text{Bitrate} + x_0 ,
\]

(4)

\[
b = y_n \cdot \text{Bitrate}^n + y_{n-1} \cdot \text{Bitrate}^{n-1} + \cdots + y_1 \cdot \text{Bitrate} + y_0 .
\]

(5)

In the study, for reasons of compatibility with paper [19], two terms were used in Eq. (1). The optimization steps carried out as part of the work described in this paper, e.g., three or four terms in Eq. (1) have not contributed to any improvement in the mode of operation of the model. Figures 3 to 5 show the parameters as a function of the encoding rate and the results obtained from the approximations made in Eqs. (2) to (5) in case of the H.265/HEVC codec.

Therefore, the actual equations for the H.265/HEVC codec and the native RTP are:

\[
P = 3.8 \cdot 10^{-14} \cdot \text{bitrate}^3 + 2.79 \cdot 10^{-10} \cdot \text{bitrate}^2 + 2.37 \cdot 10^{-5} \cdot \text{Bitrate} + 1.82 ,
\]

(6)

\[
Q = 1.71 \cdot 10^{-12} \cdot \text{bitrate}^3 - 5.81 \cdot 10^{-8} \cdot \text{bitrate}^2 + 6.34 \cdot 10^{-4} \cdot \text{Bitrate} + 1.14 ,
\]

(7)

\[
a = -6.36 \cdot 10^{-11} \cdot \text{bitrate}^2 + 1.42 \cdot 10^{-6} \cdot \text{Bitrate} - 2.53 \cdot 10^{-2} ,
\]

(8)

\[
b = -1.04 \cdot 10^{-13} \cdot \text{bitrate}^3 + 3.34 \cdot 10^{-10} \cdot \text{bitrate}^2 + 3.93 \cdot 10^{-5} \cdot \text{Bitrate} - 0.827 .
\]

(9)

For the clarity of presentations, the graphics with QoE values and with parameters \(P\), \(Q\), \(a\), and \(b\) are not presented here in the case of the VP9 codec.
The combination of Eqs. (1) and (6) to (9) describes the new, parameterized VS model for the H.265/HEVC codec. In contrast, the combination of Eqs. (1) and (10) to (13) forms the new, parameterized VS model for the VP9 codec. The two QoS models developed here complement the existing range of parameterized QoS models from papers [18] and [19] well, which is of great benefit to any practical utilization.

It should be noted that the new QoS model was determined for encoding bitrates up to 15,000 kbps. In practice, the new model is effective only in this range. The encoding rates chosen in this study correspond to the values that are used in the Internet (e.g. by the Netflix [22]) for video streaming. It should also be noted that in the new model, the following units are to be used for the variables used there: packet losses in the unit [%] and encoding bitrate in the unit [kbps].

Below is an example of the application of the newly created QoS model. The following assumptions are made: H.265/HEVC codec, encoding rate equal to 5000 kbps. The parameters determined according to Eqs. (6)–(9) are: $P = 1.71$, $Q = 3.07$, $a = -0.0197$, $b = -0.635$. The QoS values determined for packet losses of 0 to 5% are shown in Table 1.

| Packet loss [%] | 0  | 1  | 2  | 3  | 4  | 5  |
|----------------|----|----|----|----|----|----|
| QoS values [MOS] | 4.78 | 3.3 | 2.5 | 2.07 | 1.82 | 1.68 |

3. Comparison Study

The QoSCalc (IPTV) software tool from paper [20] was used in the following analyses as well. Figure 6 shows the block diagram of the numerical tool used in this paper. This tool operates in the following manner:

- a reference video file is loaded,
- the video is encoded in accordance with the selected codec by FFmpeg [23],
- the coded data is encapsulated according to the selected transport protocol (here native RTP [24]) by FFmpeg,
- the block “Error” represents the generation of a selected level of network impairments,
- the packed video is decoded to the same format as the reference (raw video, same resolution and encoding rate) by FFmpeg,
- finally, the decoded data and the video reference file loaded at the start are passed on to the evaluation algorithms (here VQuad-HD and VS model).

The “Error” block has been designed for nondeterministically distributed packet loss (binomial distribution with probability $P$) and non-deterministically distributed burst size (exponential distribution) with a selectable mean value.

The following parameters were assumed for the numerical comparison study:

- encoding: H.265/HEVC (medium); VP9 (FFmpeg default),
- packaging: native RTP with NAL size equal to 1200 bytes,
- full HD video sequence with image refresh frequency of 25 images/s [25],
- encoding rates: 3875, 6750 and 10,500 kbps,
- packet loss (binomial distributed): 0–12 (in increments of 1), 14, 16, 20%,
- burst size: 1, 2, 3 (practical values),
- SwissQual VQuad-HD and VS model as the QoE/QoS measuring techniques,
- 31 measurements for each determined performance value. This number of tests ensures a confidence interval of less than 10% of the estimated average (with a 5% error probability).

The results of the comparison study are presented in Figs. 7 to 12. Figures 7 to 10 refer to the H.265/HEVC codec. Here, two extremely different coding rates (3875 and 10,500 kbps) and burst sizes (1 and 3) were deliberately chosen to demonstrate the effectiveness of the new model in the entire range of its applicability.

Figures 7 to 10 show that QoS deteriorates exponentially as packet losses increase. This is the case for both QoE/QoS...
measuring techniques used. Furthermore, the curves fall less steeply as burst size increases. The reason for this is that synchronization of I/P/B images fails more frequently when small groups of packet losses occur regularly, rather than when large groups of packet losses occur infrequently. Moreover, the more numerous the breakdowns in synchronization are, the more frequently the images will freeze. This will naturally be reflected in a drop in QoE/QoS values.

Figures 7 to 10 also show that the curves of the VQuad-HD and the VS model progress very closely to each other. Only when burst size is “1” and large packet losses (16–20%) occur, relatively large deviations in the QoS values may be observed. However, they can be accepted from a practical point of view. In practice, such large packet losses indicate a network fault, which in turn leads to the connection being broken. In other words: the numerical comparison study has delivered strong arguments for using the VS model in everyday practice.

Figures 11 and 12 show the results of the investigation from this comparative study using the VP9 codec in two examples. In order to show the effectiveness of the new QoS model in comparison with the investigations carried out previously within the scope of this study (H.265/HEVC codec), the burst sizes of 1 and 2, and the bit rate of 6750 kbps were deliberately chosen. A considerable agreement between the numerical values and the values from the VS model has been observed. This once again confirms good practical functionality of the VS model.
In the course of the work described in this paper a new parameterized QoS measuring model for two new video codecs (H.265/HEVC and VP9) was developed to determine the quality of real-time video service over IP network, and its functionality was put to the test in a comparison study. The new VS model is based on the VQuad-HD curves. This is of special practical importance, since the VQuad-HD algorithm (according to ITU-T J.341) is the most objective QoE measuring method for video. The comparison study has proved, without a shadow of doubt, practical suitability of the new QoS model. The new, inexpensive, quick and easy-to-use VS model (offline) is far more than simply an adequate alternative to the laboriously slow (online) VQuad-HD method with its expensive licenses.

It would be extremely useful to verify the results presented in this paper in a subjective study (which is unfortunately, extremely time-consuming and resource-intensive). This is planned in the future.

Development in the field of video coding continues. First announcements of the new H.266/VVC standard may already be found [26]. Image resolution is also being increased on a continuous basis (2K and 4K). This should also be considered in further studies and developments. This would enable the currently available range of parameterized QoS models for live video streaming to be expanded and used in practice. In the new developments of such QoS models, one could very well apply the procedure presented in this paper. Further work is planned in this direction, too.

4. Summary and Outlook

References

[1] “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017–2022”, Cisco, 2019 [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-745460.html

[2] “Cisco Predicts More IP Traffic in the Next Five Years Than in the History of the Internet”, Cisco, 2018 [Online]. Available: https://www.cisco.com/c/dam/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-745460.pdf

[3] “Information technology—Generic coding of moving pictures and associated audio information: Systems”, ISO/IEC 13818-1, 2000 [Online]. Available: https://www.iso.org/standard/36048.html

[4] “RTP Payload Format for Transport of MPEG-4 Elementary Streams”, IETF RFC 3440, 2003 [Online]. Available: http://www.faqs.org/rfcs/rfc3440.html

[5] “Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats”, ISO/IEC 23009-1:2012, 2012 [Online]. Available: https://www.iso.org/standard/57623.html

[6] “H.265: High efficiency video coding”, ITU-T [Online]. Available: https://www.itu.int/rec/T-REC-H.265

[7] “VP9 Bitstream& Decoding Process Specification”, Google [Online]. Available: https://storage.googleapis.com/downloads.webmproject.org/docs/vp9/vp9-bitstream-specification-v0.6-20160331-draft.pdf

[8] “Objective perceptual video quality measurement techniques for digital-cable television in the presence of a full reference”, ITU-T [Online]. Available: https://www.itu.int/rect/T-REC-J.145-200403-I/en

[9] “Objective picture quality measurement method by use of in-service test signals”, ITU-T [Online]. Available: https://www.itu.int/rec/T-REC-J.147-200207-I/en

[10] “Objective perceptual multimedia video quality measurement in the presence of a full reference”, ITU-T [Online]. Available: https://www.itu.int/rec/T-REC-J.148-200808-I/en

[11] “Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference” ITU-T [Online]. Available: https://www.itu.int/rec/T-REC-J.341-201603-I/en

[12] “Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to full and reduced reference pixel information”, ITU-T [Online]. Available: https://www.itu.int/rec/T-REC-P.1204.4-202001-I/en

[13] Z. Wang and A. C. Bovik, “Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures”, in IEEE Signal Process. Mag., vol. 26, no. 1, pp. 98–117, 2009 (DOI: 10.1109/MSP.2008.930649).

[14] Z. Wang and A. C. Bovik, “A universal image quality index”, in IEEE Signal Process. Let., vol. 9, no. 3, pp. 81–84, 2002 (DOI: 10.1109/97.995823).

[15] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity”, IEEE Transac. on Image Process., vol. 13, no. 4, pp. 600–612, Split, Croatia, 2004 (DOI: 10.1109/TIP.2003.819861).

[16] U. Sarea, M. Akter, and M. Uddin, “Image quality assessment through FSIM, SSIM, MSE and PSNR – A comparative study”, J. of Comp. and Commun., vol. 7, no. 3, 2019 (DOI: 10.4236/jcc.2019.73002).

[17] J. Klink and T. Uhl, “Video quality assessment: some remarks on the selected objective metrics”, in Proc. of the 28th Int. Conf. on Software, Telecommun. and Comp. Networks, IEEE Transaction on Image Process., vol. 13, no. 4, pp. 600–612, Split, Croatia, 2004 (DOI: 10.1109/TIP.2003.819861).

[18] T. Uhl and S. Paulsen, “The new, parameterized VT Model for determining quality in the video telephone service”, Bull. of the Polish Acad. of Sci.: Techn. Sci., vol. 61, no. 3, pp. 431–437, 2014 (DOI: 10.2478/bpasts-2014-0045).

[19] T. Uhl and H. Juergensen, “The new, parameterized IPTV Model for determining the quality in the IPTV service”, Bull. of the Polish Acad. of Sci.: Techn. Sci., vol. 63, no. 2, pp. 495–500, 2015 (DOI: 10.1515/bpasts-2015-0057).

[20] T. Uhl and H. Juergensen, “New tool for examining QoS in the IPTV service”, in Proc. of the World Telecommun. Congress WTC’14, Berlin, Germany, 2014 (DOI: 10.13140/RG.2.1.3087.5608).

[21] “SwissQual VQuad-HD Selected as New ITU HD Video Testing Standard J.341” [Online]. Available: https://www.realwire.com/downloads.webmproject.org/docs/vp9/vp9-bitstream-specification-v0.6-20160331-draft.pdf

[22] “Internet Connection Speed Recommendations”, Netflix [Online]. Available: https://help.netflix.com/en/node/306
Christian Hoppe received his B.Eng. in Communications Technology from the Flensburg University of Applied Sciences, Germany, in 2010 and his M.Sc. in Information Technology from the Kiel University of Applied Science, Germany, in 2016. Today, he works as a Software Engineer for Nextragen Solutions GmbH in Handewitt, Germany. He focuses mainly on the following areas: quality assurance for triple play services and medical imaging solutions.

E-mail: christian.hoppe@nextragen-solutions.de

Nextragen Solutions GmbH Flensburg
Heideland 1
D-24976 Handewitt, Germany

Tadeus Uhl received his M.Sc. in Telecommunications from the Academy of Technology and Agriculture in Bydgoszcz, Poland in 1975, his Ph.D. from the Gdańsk University of Technology, Poland in 1982 and his D.Sc. from Dortmund University, Germany in 1990. Since 1992, he has been working as a Professor at the Institute of Communications Technology, Flensburg University of Applied Sciences, Germany and, in addition, since 2013, as a Professor at the Institute of Transport Engineering and Economics, Maritime University of Szczecin, Poland.

His activities focus on the following areas: traffic engineering, performance analysis of communications systems, measurement and evaluation of communications protocols, QoS and QoE by triple play services, Ethernet and IP technology. He is an author or co-author of five books and 130 papers dealing with LAN, WAN and NGN.

https://orcid.org/0000-0001-6849-9168

E-mail: t.uhl@am.szczecin.pl

Maritime University of Szczecin
Henryka Pobożnego 11
70-507 Szczecin, Poland