A Comparison Study of DASH Technique by Video Streaming over IP with the Use of RTP and HTTP Protocols

Christian Hoppe[1] and Tadeus Uhl[2]

[1] *Nextragen Solutions GmbH Flensburg, Handewitt, Germany*
[2] *Maritime University of Szczecin, Szczecin, Poland*

Abstract—Today's Internet knows no bounds. New applications are marketed every single day. Many of them incorporate video sequences. These must be transported over the Internet quickly (often in real time). However, the Internet has not been designed for live communications and, regrettably, this may become apparent all too quickly. Countermeasures are required in the form of new, efficient transport techniques facilitating online video services. MPEG-DASH is one of such modern techniques. But how good is this new technique really? This paper delves into the matter. The paper contains an analysis of the impact that the new technology exerts on the quality of video streaming over IP networks. It also describes a new numerical tool – QoSCalc (DASH-HTTP) which has been used to analyze MPEG-DASH under different use scenarios. The results are presented graphically and their interpretation is provided.

Keywords—*H.265, IP network technology, MPEG-DASH, network management, QoS measurements, video streaming.*

1. Introduction

Nowadays, a significant portion of Internet traffic is made up by video service applications, e.g. IP television (IPTV), Video Telephony over IP (VToIP) and non-real-time video streaming service. It is common knowledge that the Internet is a lossy transport platform causing deterioration of QoS. New, efficient transport techniques for video services are needed. MPEG-DASH [1] is one of such modern techniques. It is currently offered by several video service providers operating online, including Netflix [2] and YouTube [3]. The MPEG-DASH technique can be used with both Real-time Transport (RTP) and HTTP protocols. What influence do these transport protocols have on QoS? Paper [4] includes an analysis of the results from a study concerning the DASH technique in conjunction with the RTP Protocol. This article, examining the DASH technique used in conjunction with the HTTP protocol, is to be perceived as a study complementing [4].

In order to determine the QoS/QoE in a network, two models (dual-ended and single-ended) are generally used [5]. In the case of the dual-ended model, two signals are used: original and degraded. These two signals are available in their uncompressed form. Therefore, measurements may be carried out for both QoE (subjective evaluation) and QoS (objective evaluation). In the case of the single-ended model, only the impaired signal (often compressed) is available. This means that the objective QoS evaluation may be performed only.

The QoE measurement techniques used most widely for Video Streaming over IP (VSoIP) include, currently, perceptual evaluation of video quality PEVQ (J.247) [6] and VQuad-HD (J.341) [7]. These techniques are very accurate; they are, however, time-consuming and can often only be implemented based on a license. Both algorithms incorporate an electronic emulation of the human eye. Therefore, one can justifiably speak of QoE values here. The peak signal-to-noise ratio (PSNR) method [8] is a well-known measuring method for determining QoS in the video service domain. It will also be used in the course of the analyses described in this paper.

This paper begins with a brief explanation of the new numerical tool QoSCalc (DASH-HTTP). In the subsequent section, specific analysis scenarios are described in detail. The results obtained in this comparison study are presented graphically, and are interpreted. The paper concludes with a summary and an outlook on future works.

2. Upgraded Tool QoSCalc (DASH-HTTP)

The QoSCalc IPTV tool described in [9] is capable of taking a reference video and encoding it with a predefined codec and specific parameters. Afterwards, it may apply impairments in the form of packet loss and may compare the resulting video with the reference material. This tool has been upgraded to analyze the MPEG-DASH technique, as described in paper [4]. However, QoSCalc (DASH) from [4] still uses RTP and native RTP transport streams. Since MPEG-DASH is HTTP-based, a new version of the tool is necessary to support the HTTP transport stream. This can be realized by starting a local HTTP server and client services for each series of measurements. This complement has given rise to the new version of the

QoSCalc (DASH) tool which is referred to as QoSCalc (DASH-HTTP).

As with the previous QoSCalc (DASH) tool, there is still a need to encode the reference video in different formats and to make sure that the differently encoded videos may be quickly and easily switched by MPEG-DASH. That is why the configuration is split into multiple groups. When the measurement settings in all groups have been made, all sets of group options are combined. The settings for the reference video and the parameters have not changed and are presented in Fig. 1.
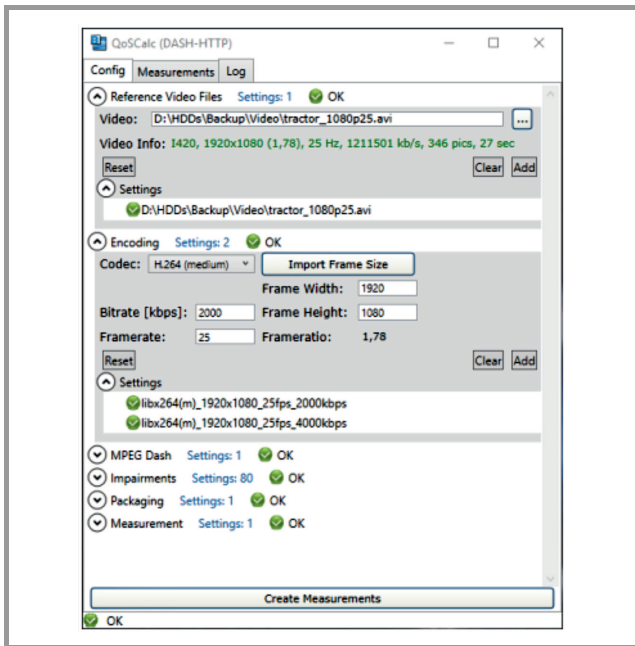

Fig. 1. Reference video and coding settings for the tool QoSCalc (DASH-HTTP).

The settings for MPEG-DASH are also unchanged and can be seen in Fig. 2. MPEG-DASH requires that multiple changes of the selected format be possible over time, which is difficult to achieve with a short reference signal (lasting, e.g. 8 s). That is why the reference signal is first looped a selectable number of times to increase its duration. Afterwards, the tool encodes the reference video with different sets of parameters available in the tool (as indicated in Fig. 1). Given the different qualities of the encoded videos, an additional option is required, which tells the tool how and when to change the selected video format during the measurement. Four options are available with which to pre-arrange the format selection procedure. To achieve that, the order in which the encoding options are added is important. The following format selection patterns are predefined:

- steps (deterministic) – encoding formats are selected in the sequence in which they were added, starting from the top again after the last has been reached,

- non-deterministic – the encoding format is selected at random,

- drop and rise – encoding formats are selected by using the first setting to the last, and then in the reverse order,

- rise and drop – encoding formats are selected in the reverse order, i.e. the last to the first and then back down again from the first to the last.
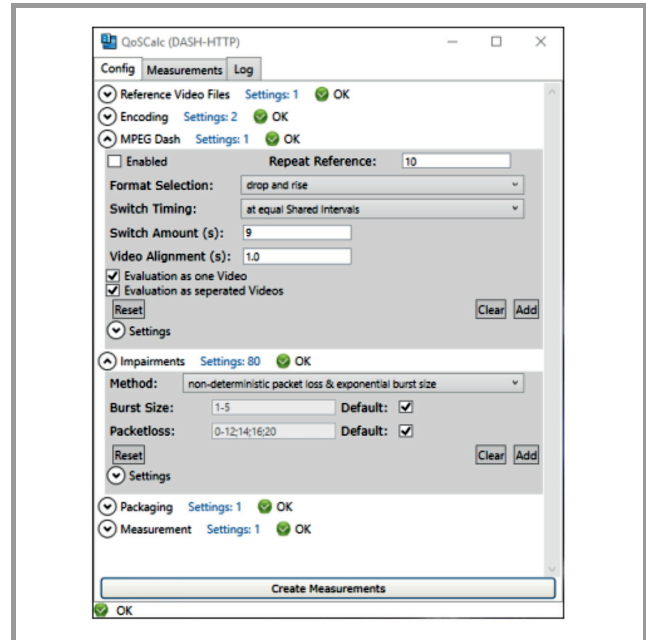

Fig. 2. MPEG-DASH and impairments settings for the tool QoSCalc (DASH-HTTP).

In addition to selecting which format is used, it is also possible to select when the changes are to occur. To that end, the following options are available:

- at reference boundary – this will change the format whenever the reference video is repeated (looped),

- at equally spaced intervals – the total length of the video is split into equally sized sequences determined by the number of changes set,

- at non-deterministic intervals – a selected number of changes are distributed at random over the total time of the measurement (equal distribution).

To accomplish a smooth switch between formats, I-frames have to be available in each format at the same moment in time (frame). Being the underlying codec library, FFmpeg [10] may be adjusted to force an I-frame at constant intervals, configurable in the new configuration window in the video alignment settings. Finally, there is an option in the MPEG-DASH settings with which to select whether the video should be evaluated as one single video or as each sequence individually, or as both.

The next group of settings is also unchanged in the new version. Different impairment methods can be selected which determine how packet loss is distributed. The distribution

is a function of the selected burst and packet loss values, and is applied to each packet of the video stream:

- deterministic packet loss and constant burst size,

- non-deterministic packet loss and constant burst size,

- non-deterministic packet loss and exponential burst size,

- 2-state Markov loss model/burst ratio.

The burst ratio corresponds to the ratio of the measured burst size to the burst size expected in a burst-free environment. Details may be found in [11].

The only visible change in the settings window of the new tool is found in the packaging section. HTTP is now selectable, and the size of the client-side buffer can be changed (see Fig. 3). This simple change, however, affects the process of the entire measurement run.
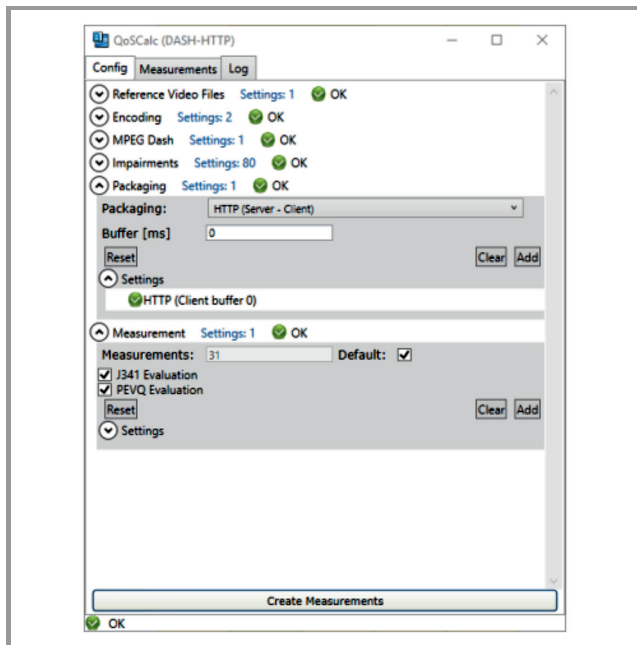

Fig. 3. Packaging and measurement settings for the tool QoSCalc (DASH-HTTP).

All other packaging methods (RTP, native RTP, MPEG2TS) are accomplished by using FFmpeg to split and package the encoded video. After applying the impairments (e.g. packet loss), FFmpeg is used as well to reassemble and save the video. This can be done because all channels are unidirectional. This has changed for HTTP, however, since the client can detect and request the resending of missing packets from the server. This will be taken into account in the new tool accordingly.

3. Measurement Scenarios and Measurement Results

The same measurement scenarios as in [4] were used to compare the results achieved with the use of different trans-

port protocols. Most parameters remain constant over all scenarios. First, the same reference video from [4] was used and repeated nine times to generate segments that were long enough to apply five different encoding formats in each scenario, using the drop and rise pattern. Drop and rise was selected because it simulates a temporary problem in a video stream in which the quality is first reduced and then increased again. The new measurements use HTTP on the transport layer instead of RTP. The HTTP client is set to a buffer size of 0 ms, which is an unusual value, but it allows the evaluation algorithms to show a degradation of QoE whenever packets are lost. Otherwise, the buffer would only add a starting delay on the client side, which is always the same. Missing packets would be requested within the buffer timeframe, and the degradation would not be visible in the resulting video. An overview of the five different scenarios analyzed is shown in Table 1. The H.265 codec is used in all measurement scenarios.

Table 1
Measurement scenarios

Scenario no.	Buffer size	Evaluation	Resolution	Bit rate
1	0 s	Constant reference	Constant	Variable
2	0 s	Variable reference	Variable	
3	0 s	Constant reference	Variable	
4	2 s	Constant reference	Variable	
5	2 s	Constant reference	Constant	

Scenarios 2-4 use 5 different resolutions and bit rate combinations which are common in popular web services (see Table 2). The format selection patterns here correspond to the "drop and rise" technique. Additionally, the reference video with which the degraded video was compared also had matching resolutions in scenario 2.

Table 2
Encoding formats for variable resolution scenarios

Format	Resolution [pixels]	Bit rate [kbps]
1	1080 × 1920	4500
2	720 × 1280	2500
3	480 × 853	1000
4	360 × 640	500
5	260 × 427	300

3.1. Scenario 1

The H.265 Codec and constant resolution with a variable bit rate were used. The reference videos used to evalu-

ate the different segments are constant at 1080p. The QoE measurement techniques used here were PEVQ (J.247) [6] and VQuad-HD (J.341) [7]. The QoS measurement technique used is the PSNR method [8]. The results obtained are shown in Figs. 4 and 5. The left-hand y axis represents QoE/QoS values on the MOS scale (for J.247 and J.341). The right-hand y axis, on the other hand, represents QoS values on the decibel scale (for PSNR).
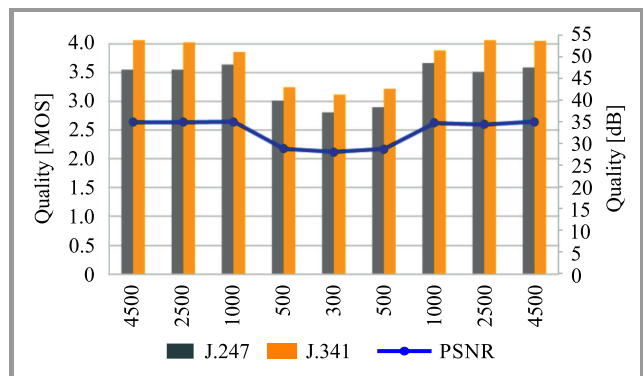

Fig. 4. QoE/QoS values as a function of coding rate for scenario 1 at 1% packet loss using RTP [4].
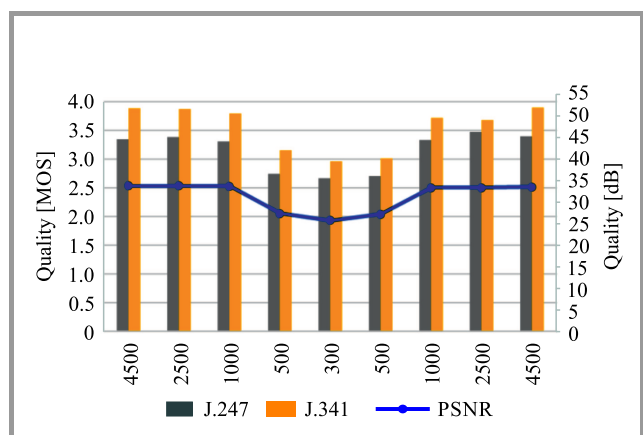
Fig. 5. QoE/QoS values as a function of coding rate for scenario 1 at 1% packet loss using HTTP.

The diagrams in Figs. 4–5 show that application of the RTP protocol offers slight advantages, especially at low resolution/coding rates, over the HTTP protocol. This can be explained by the varying sophistication of the protocols. The HTTP protocol has a larger overhead than RTP and requires more bandwidth in the network. Moreover, the difference in communication patterns between UDP and TCP slows down the last protocol.

3.2. Scenario 2

H.265 and variable resolution with variable bit rate were used. In this scenario the resolution is changed along with the bit rate. The format details are shown in Table 2. The reference video used to evaluate the different segments is

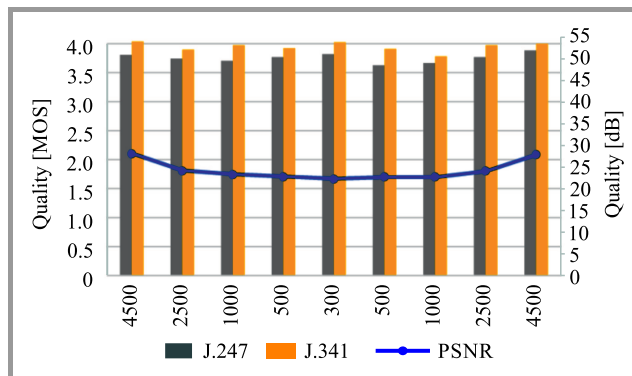also variable. The results obtained are shown in Figs. 6 and 7.

Fig. 6. QoE/QoS values as a function of coding rate and resolution for scenario 2 at 1% packet loss using RTP [4].
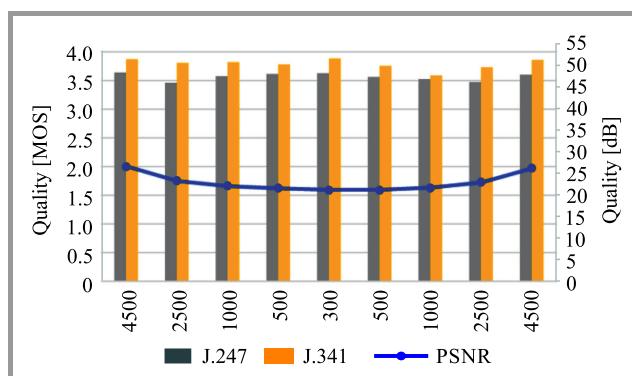

Fig. 7. QoE/QoS values as a function of coding rate and resolution for scenario 2 at 1% packet loss using HTTP.

The diagrams in Figs. 6–7 reveal better QoE and QoS values than those in scenario 1. This is because the reference signal corresponds to the resolution used in the network. The comparisons between the transport protocols show a similar improvement of QoE/QoS with RTP. Knowing this means, in practice, that any terminal equipment can have a low resolution and it will be fully adequate for applications involving transport streams with variable resolution and variable bit rates. This will reduce the bandwidth necessary within the network.

3.3. Scenario 3

The same setting as in scenario 2 with H.265 and variable resolution and variable bit rate were used, the difference being that the resolutions of the reference video used to evaluate the different segments are constant at 1080p. The results obtained are shown in Fig. 8.

The diagram in Fig. 8 shows the result which that could be expected. Whenever the received, low-resolution video is compared with the original high-resolution reference, the perceived QoE is reduced. This is relevant for devices with higher resolutions, where low resolution videos have a bigger impact on QoE/QoS.
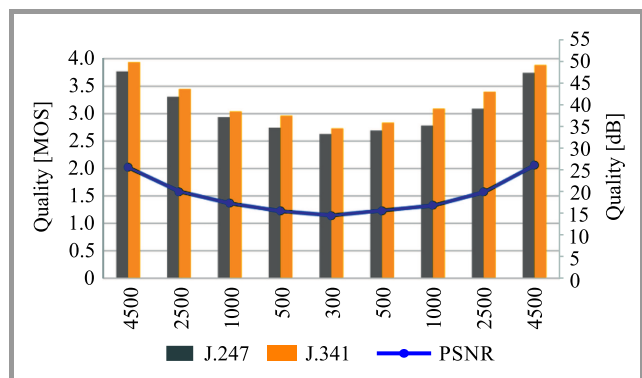
Fig. 8. QoE/QoS values as a function of coding rate and resolution for scenario 3 at 1% packet loss using HTTP.

3.4. Scenario 4

The same setting as in scenario 3 was used. Additionally, the jitter buffer size had a value of 2 s. The results obtained are shown in Fig. 9.
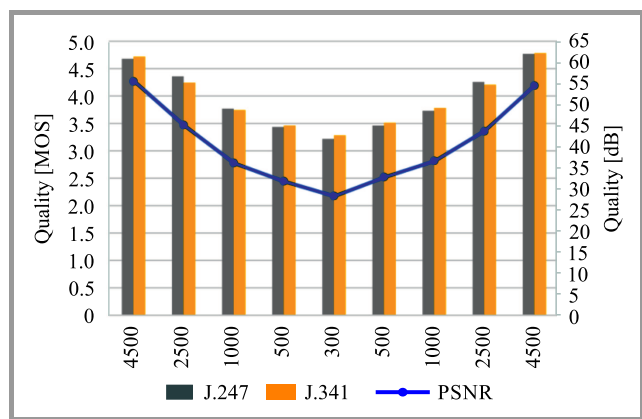

Fig. 9. QoE/QoS values as a function of coding rate and resolution for scenario 4 (buffer 2 s) at 1% packet loss using HTTP.

From Fig. 9 it is evident that using a jitter buffer will considerably improve quality in a lossy environment. It is just as evident that changes in resolution and bit rate also have a significant influence on video quality. The resending of the lost data packets is requested here by the client. The jitter buffer is able to compensate for the resulting network delay.

4. Scenario 5

H.265 and constant resolution with variable bit rate were used. Additionally, the jitter buffer size has a value of 2 s. The resolution of the reference video used to evaluate the different segments was at 1080p throughout. The results obtained are shown in Fig. 10.

The curves in Fig. 10 confirm, yet again, that the use of a sufficiently large jitter buffer will improve quality considerably. In fact, almost maximum video quality can be achieved because the system will demand lost data packets
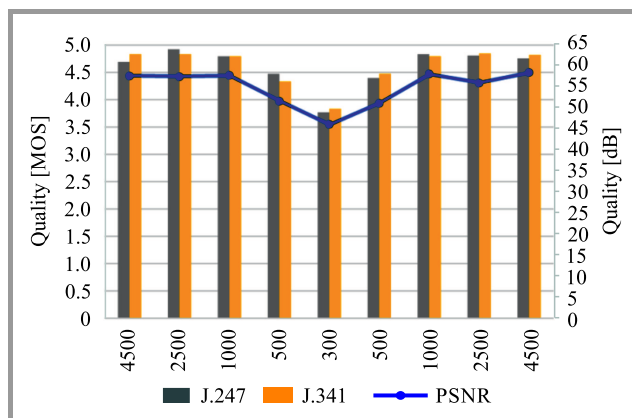

Fig. 10. QoE/QoS values as a function of coding rate and resolution for scenario 5 (buffer 2 s) at 1% packet loss using HTTP.

to be sent again and successfully delivered. This means that Fig. 10 shows only the influence of the decrease in bit rate on video quality. It is to be noted that the bandwidth used in the scenario 5 is greater than that used in scenario 4, which means that optimization is needed here in order to achieve an acceptable compromise between the bandwidth used in the network and video quality.

5. Summary and Outlook

This paper has focused on the assessment of the influence of the new MPEG-DASH technique on QoS VSoIP with the RTP and HTTP protocols applied. The new technique is flexible and it may be implemented to relieve congestion in networks. That can improve end-to-end QoS.

The analyses have shown that there are many benefits of using the MPEG-DASH technique in conjunction with the HTTP protocol. This is often done in practice despite the fact that it increases transmission times. For such applications as video streaming, for instance, which is not a real-time service, this can be tolerated. Analyses have also revealed that the size of the jitter buffer plays a decisive role in determining the quality of service. Optimization has now become critical because a compromise between QoS and necessary network bandwidth must be found.

The analyses conducted in the course of the work leading up to this paper have clearly demonstrated a significant correlation between QoE and PSNR values. PSNR values can very well be interpreted as QoS values. Consequently, the question will arise as to how faithfully PSNR values (which are easy to measure) reflect true QoE values (which are very difficult to determine). It would be worthwhile examining this point in future work.

New video codecs appear regularly on the telecommunications market, one of the latest being the VP9 Codec. Papers [12], [13] have shown that the new VP9 codec using the MPEG2 transport stream is more versatile and superior to the H.265 codec. It would be interesting to know how well the new VP9 codec works in conjunction

with the MPEG-DASH technique and a comparison of both the H.265 codec and the VP9 codec used in conjunction with the MPEG-DASH technique would be well worth conducting.

Further studies in these directions are already being planned.

References

[1] ISO/IEC DIS 23009-1.2, "Dynamic adaptive streaming over HTTP (DASH)" [Online]. Available: https://www.iso.org/standard/57623.html/2019.11.10

[2] "Netflix_Dash" [Online]. Available: https://bitmovin.com/status-mpeg-dash-today-youtube-netflix-use-html5-beyond/2019.11.10

[3] "Youtube_Dash" [Online]. Available: http://streamingcodecs.blogspot.com/2013/01/mpeg-dash-support-in-youtube.html/2019.11.10

[4] Ch. Hoppe and T. Uhl, "Video Streaming over IP using the DASH Technique", *J. of Telecommun. and Inform. Technol.*, no. 2, pp. 3–7, 2018 (doi: 10.26636/jtit.2018.125318).

[5] T. Uhl, "Quality of Service everywhere", in *Proc. of the 9th Int. Conf. on Adv. Service Comput.*, Athens, Greece, 2017, pp. 28–30 (ISBN 978-1-61208-528-9).

[6] Standard ITU-T J.247, "Objective perceptual multimedia video quality measurement in the presence of a full reference" [Online]. Available: http://www.itu.int/rec/T-REC-J.247-200808-I/2019.11.10

[7] Standard ITU-T J.341, "Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference" [Online]. Available: http://www.itu.int/rec/T-REC-J.341-201101-I/en/2019.11.10

[8] Q. Huynh-Thu and M. Ghanbari, "The accuracy of PSNR in predicting video quality for different video scenes and frame rates", *Telecommunication Syst.*, vol. 49, no. 1, pp. 35–48, 2012 (doi: 10.1007/s11235-010-9351-x).

[9] T. Uhl and H. Jürgensen, "New tool for examining QoS in the IPTV service", in Proc. of the World Telecommun. Congr. WTC'14, Berlin, Germany, 2014 (doi: 10.13140/RG.2.1.3087.5608).

[10] "FFmpeg (current Window builds)" [Online]. Available: http://FFmpeg.zeranoe.com/builds/2019.11.10

[11] S. Paulsen, "QoS/QoE-Modelle für den Dienst Voice over IP (VoIP)", Dissertation, University Hamburg, Department of Information Technology, 2016 [Online]. Available: https://ediss.sub.uni-hamburg.de/volltexte/2016/7908/pdf/Dissertation.pdf

[12] M. Řeřábek and T. Ebrahimi, "Comparison of compression efficiency between HEVC/H.265 and VP9 based on subjective assessments" [Online]. Available: https://infoscience.epfl.ch/record/200925/files/article-vp9-submited-v2.pdf/2019.11.15

[13] T. Uhl, K. Nowicki, J. Klink, and Ch. Hoppe, "Comparison study of H.264/AVC, H.265/HEVC and VP9-coded video streams for the service IPTV", in *Proc. of the 26th Int. Conf. on Softw., Telecommun. and Comp. Netw. SoftCOM 2018*, Split, Croatian, 2018 (doi: 10.23919/SOFTCOM.2018.8555840).

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 Handwitt, Germany. His main activities cover 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
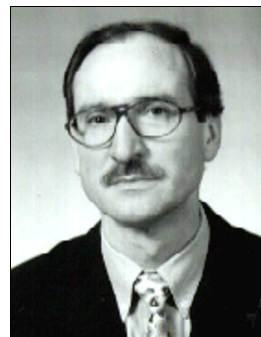
Tadeus Uhl received his M.Sc. in Telecommunications from the Academy of Technology and Agriculture in Bydgoszcz, Poland in 1975, his Ph.D. from the Gdańsk University of Technology, Poland in 1982 and his D.Sc. from Dortmund University, Germany in 1990. Since 1992, he has worked as a Professor at the Institute of Communications Technology, Flensburg University of Applied Sciences, Germany and, in addition, since 2013, as Professor at the Institute of Transport Engineering and Economics, Maritime University of Szczecin, Poland. His main activities cover the following areas: traffic engineering, performance analysis of communications systems, measurement and evaluation of communication 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