Paper

Quality of Variable Bitrate HD Video Transmission in New Generation Access Network

Piotr Makowski

Faculty of Electronics and Information Technology, Warsaw University of Technology, Warsaw, Poland

Abstract—Article presents influence of multiplying variable bitrate high definition (HD) video streams in an access network link on Quality of Service (QoS). The aim of a conducted study is to define key parameters influencing Quality of Experience (QoE). Numerous simulations were performed and indicators like packet loss, delay, jitter, frame loss and bandwidth utilization were observed. Moreover, two independent algorithms were used to indicate QoE values of video streams. These are SwissQual VQuadHD and Telchemy VQMon applications which provided credible full reference and no reference algorithms, respectively. In the article evaluation of accuracy of no reference algorithm is performed. In future work it is planned to build analytic model of VBR video transmission and to undertake more thorough research of transmitting multiplied HD video streams in an access network using various QoS policies and optimizing size of buffers.

Keywords*—*congestion in access network, Quality of Experience, streaming, variable bitrate.

1. Introduction

Nowadays, the popularity of HD and 3D television causes bandwidth demand to increase in the access network. Many publications [1], [2] prove dynamic increase of multimedia streaming in the Internet which supersedes peer-2-peer transmission model. Over-the-top television (OTT), i.e. Netflix or YouTube, becomes common and it swaps broadcast TV based on a carriers network. It is essential to deploy more efficient coding methods or guarantee more bandwidth to provide more content with better quality, particularly in the “last mile” of the network.

The European Union established a directive, which states that all members are obliged to provide connection to New Generation Access network (NGA) with bandwidth above 30 Mbit/s for 100% of households and above 100 Mbit/s for 50% [3]. Nevertheless, building NGA will proceed slowly due to the lowering prices of Internet access and the increasing costs of network investments and truck rolls. Telecommunication Service Providers (TSPs) are trying to build cheaper NGA systems using point-to-multipoint technologies like DOCSIS (Data Over Cable Service Interface Specification), PON (Passive Optical Network) or BPL (Broadband Power Line). Moreover, bandwidth demand constantly increases due to the multiplication of customer’s equipment that receives video content, i.e. smartphones, tablets. Additionally, more streams are transmitted to the customer due to the aforementioned OTT as well as multiroom and Picture-In-Picture (PIP) services. All these aspects lead to insufficient bandwidth in the “last-mile” of the network. Thus, providing satisfactory Quality of Experience with existing bandwidth boundaries is still up to date.

This research focuses on the effects of various bandwidth utilization level on QoE when providing HD video streaming in IP network. The bandwidth utilization can change due to sharing the same link with other video streams and because of Internet traffic. Simulations were made in which HD content was coded producing variable bitrate (VBR) streams. Comparing with constant bitrate (CBR), VBR has better coding efficiency [4]. It maintains constant quality for the whole sequence, which is important in high-motion content, i.e. sport. Moreover, VBR allows statistical multiplexing of video streams. This multiplexing allows not only to reduce the total bandwidth utilization but also to preserve quality at the same time. The main drawback of this solution is more complicated traffic engineering and network dimensioning.

The research concentrates on quality defined as Quality of Experience, which is more generic term compared to Quality of Service. The first term includes not only Packet Loss Rate (PLR), delay and jitter but video content quality parameters. VQuadHD [5] tool from SwissQual and VQMon tool from Telchemy [6] were used to estimate subjective Mean Opinion Score (MOS), which is commonly known QoE indicator described in ITU-T P.910 standard [7]. VQuadHD implements Full Reference (FR) algorithm, i.e., it compares source with processed video. Furthermore, VQuadHD is an objective perceptual video quality measurement method and is an implementation of ITU-T J.341 recommendation [8] which is one of the most accurate full reference algorithm according to VQEG findings described in [9]. VQMon implements No Reference (NR) algorithm, i.e., it calculates results based only on receiver’s data. Thus, it uses only information from RTP, MPEG2-TS, UDP and IP headers. Contrary to subjective MOS, VQMon is an objective algorithm similar to the one described in ITU-T P.1201 [10], where Pearson correlation with subjective MOS values is above 90%. In the article, the results of NR metric are compared with FR metric to validate this high correlation figure for video streams in presence of transmission errors.

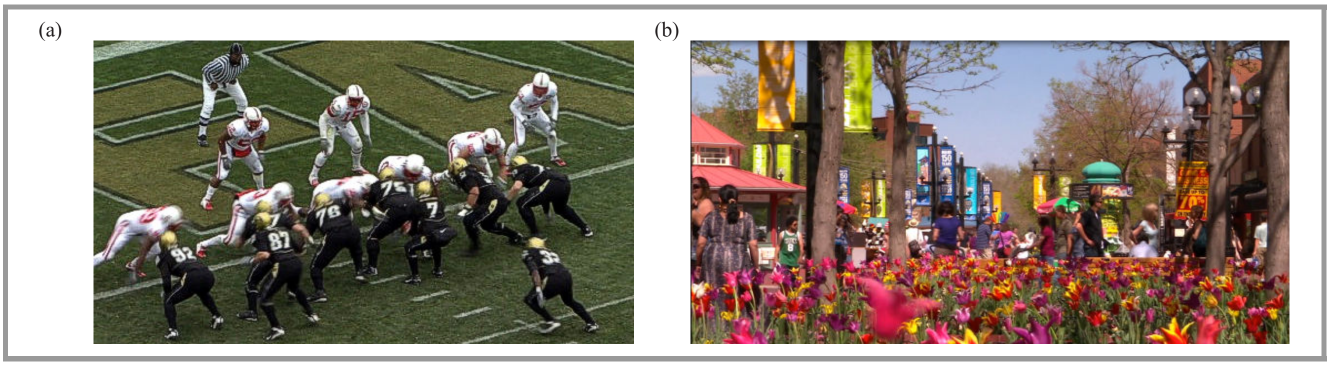

Fig. 1. Sample scenes from: (a) vqeghd1_src04 and (b) vqeghd1_csrc14 sequences.

Conducted tests are introduction to innovative methods aiming to improve video quality transmission in IP network. Second section describes testing environment and architecture, whereas third part concentrates on observations' analysis and conclusions.

2. Testing Environment

In this section all HD video sequences, measuring environment and conducted tests are described.
Testing sequences were chosen from VQEG Final Report of HDTV Validation Test [11]. Two sequences illustrated in Fig. 1 have different content and coding parameters:

- vqeghd1_src04 – football game with dynamic objects moving in different directions,
- vqeghd1_csrc14 – scene in a park with low dynamic objects, however colorful and with high contrast.

Consequently, tested streams have different traffic characteristics in IP network. Common source parameters are presented in Table 1.

Table 1
Source parameters of tested sequences

Resolution	Aspect ratio	Frames per second	Chroma subsampling
HDTV 1080p (1920×1080, progressive)	16:9	29.97	4:2:2

Source files were encoded according to ITU-T H.264/MPEG4 AVC recommendation [12] using x.264 open-source encoder [13]. Streams are VBR-coded with average bitrate of 10 Mbit/s achieved during 10 seconds. Moreover, RTP and UDP protocols were used due to high capabilities of RTP/UDP stack for connectionless transmission of TV channels in Future Internet [14], [15]. In the conducted tests, MPEG2-TS/RTP/UDP stack was used instead of native RTP. There are several reasons for this assumption. Firstly, it is recommended in Broadband Forum TR-126 [16] to use main profile with level 4 coding and MPEG2-TS container. Secondly, the quality assessment algorithms described in P.1201 standard have only been evaluated for MPEG-2 TS/RTP/UDP and not for native RTP. On top of that, the size of Network Adaptation Layer (NAL) used for tested streams is not configurable in x.264 encoder. Therefore, default settings were used in which average size of NAL of type "slice" was more than 20 Kbytes, which is much more than maximum transmission unit of Ethernet frame. In this case, the difference between native RTP and MPEG2-TS used for video transmission would be subtle. Nevertheless, author of the article is aware of advantages of defining the maximum NAL size and using native RTP stack presented in [11], [15]. Including these aspects in the test scenarios is planned in the future work.

Sequences were encoded with three maximum lengths of Group of Picture (GOP) to observe influence of changing number of I frames on QoE. Table 2 contains crucial coding parameters like number of I, P and B frames, average stream bandwidth for different maximum GOP lengths.

IP network simulation tool was used as a test environment, which is compliant with ITU-T G.1050 [17] recommendation. Figure 2 presents its architecture and main settings. Model consists of core network, access node, modem, firewall and router. Traffic is sent to the core network and received at the router behind the firewall. There are five routers in the core network connected with 1 Gbit/s links. They insert together 50 ms of delay. Link between core network and access node has also 1 Gbit/s bandwidth but 0.1 ms of delay. Access node stores settings of the link to the modem for both transmission directions. These are bandwidth, BER and delay. BER is set to 10^{-10} assuming optical link according to NGA standards. It is important to mention that probability of packet loss due to transmission error is very small because 100 Mbit/s link bandwidth and simulation time is shorter than 15 seconds. Main reason of possible packet losses is queue overflow on which this research concentrates. At the end of the transmission path, there are aforementioned modem, firewall and router which are connected in a sequence and have the same settings of link bandwidth and buffers. The total minimum delay of the system is 51.1 ms and the maximum delay depends from the buffer size and obviously from the variability of

Table 2
Coding parameters for various maximum GOP lengths

	Sequence name					
	qvegd1_src04			vqeghd1_csrc14		
Maximum GOP length [frames]	15	50	250	15	50	250
Number of frames [I/P/B]	21/14/138	7/148/145	3/148/149	21/189/89	7/198/94	3/199/97
Average stream bandwidth [Mbit/s]	9.35	9.13	9.06	10.0	9.7	9.62
Standard deviation of stream bandwidth [Mbit/s]	15.1	13.3	12.0	16.7	14.4	13.5
Min. stream bandwidth [Mbit/s]	0.89	0.94	0.94	0.02	0.02	0.02
Max. stream bandwidth [Mbit/s]	59.0	66.3	61.7	73.9	82.6	83.1

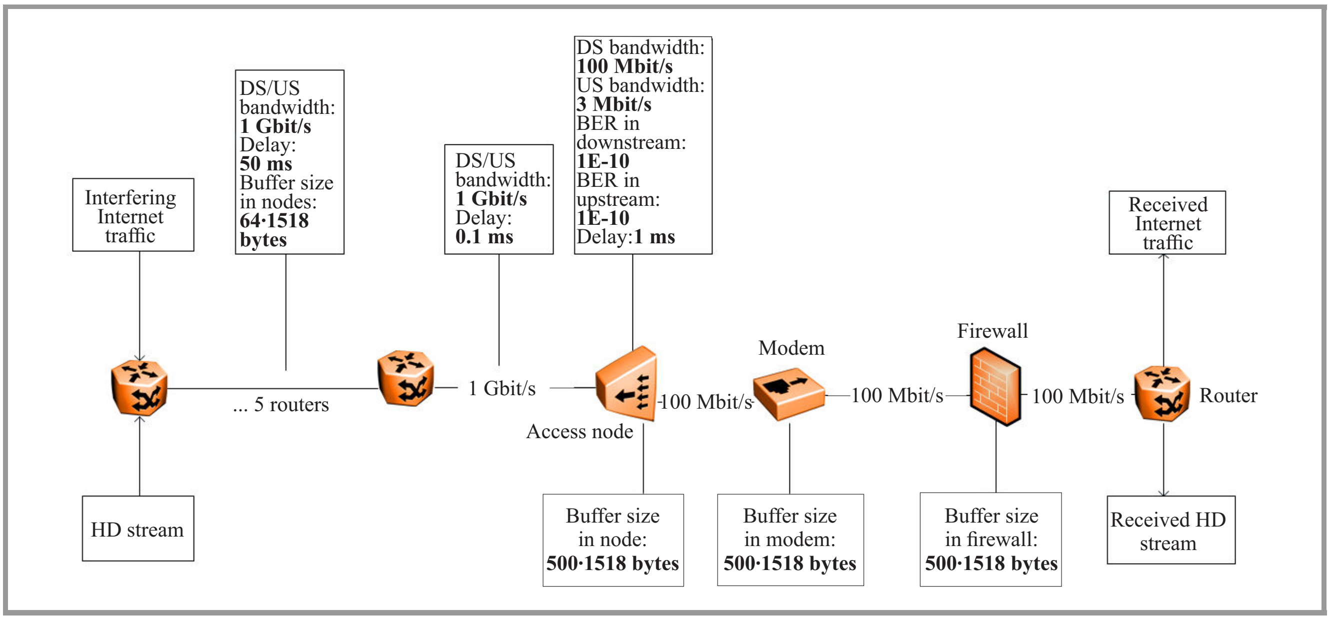

Fig. 2. Architecture and settings of G.1050 compliant network simulator.

the traffic. All elements apart from access node have buffer size limited to 97152 bytes. This value is equal to 64 Ethernet frames with size of 1518 bytes. Access node has a queue of 500 frames. This setting increases maximum delay of the whole system to 112 ms in each direction. It allows to analyse impact of the delay variations on the QoE of the video streams. 61 ms of additional delay results from the time to transmit 500 frames stored in the buffer with a speed of 100 Mbit/s.

For all six streams four tests were conducted. Firstly, bandwidth of interfering Internet traffic was increased from 0 Mbit/s to 150 Mbit/s with 15 Mbit/s step, whereas HD stream, transmitted simultaneously in the same link, was not prioritized. Therefore, buffer queues were shared between two classes of traffic (test name: same_prior). Second test differeds from first one as HD stream had higher priority compared to Internet traffic (test name: higher_prior). In the third test, interfering traffic had constant bandwidth of 15 Mbit/s, whereas quantity of HD streams was increased from 1 to 9. In the last test, all video streams were transmitted with higher priority. Therefore, only Internet traffic was not prioritized. In the tests number three and four names same_prior and higher_prior were maintained. Each simulation lasted 15 s and the total number of transmitted video streams was 683.

Testing environment allowed monitoring following parameters: packet loss rate, delay, jitter, I, P, B frame loss and bandwidth utilization. In the first and third test FIFO scheduling was applied, whereas Strict Priority Queuing (SPQ) was used in the test number two and four.

3. Observations and Conclusions

In this section, evaluation of NR algorithm by comparison with FR method is presented as well as key observations of dependencies between bandwidth utilization, number of HD streams and QoE.

Figure 3 illustrates scatter plot in which MOS-V [18] and J.341 Objective values are shown. They refer to Telchemy's VQMon and SwissQual VQuadHD metrics, respectively. From 683 test cases 521 (76%) matched to provide J.341 Objective values. Other 24% cases refer mainly to MOS-V values lower than 3 and with high percentage of impaired I-frames. The data did not matched, because it is more dif-

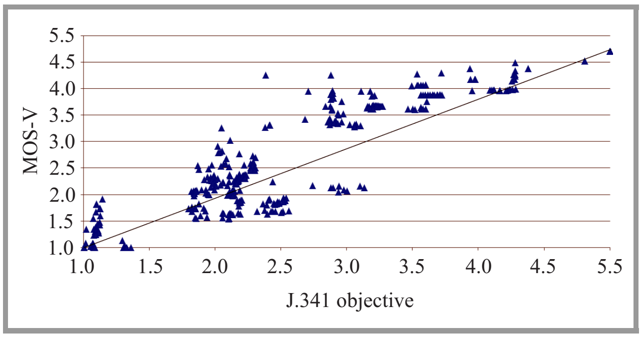

Fig. 3. Scatter plot presenting no reference MOS-V values as a function of full reference J.341 objective MOS values.

ficult to perform time and spatial alignment by VQuadHD tool for highly impaired video scenes. For 76% of data the Pearson correlation factor equals 95%. It proves that NR algorithm is sufficient for evaluation of transmission quality. Nevertheless, NR metric from Telchemy overestimates quality results in majority of cases. The reason for this is that reference video in FR algorithm provides substantial information to deeply analyze spatial artifacts, i.e. blurring, blockiness or temporal artifacts, jerkiness and freezing. These degradation types are not directly identifiable using only receiver's data. It leads to aforementioned overestimation of VQMon visible in Fig. 3.

Figure 4 illustrates results of first and second test. Providing higher priority for HD stream, i.e. higher QoS, causes packets entering the channel to be scheduled first, whereas Internet traffic as second. When stream bandwidth is not exceeding 100% utilization, queues for higher priority are not overwhelmed and packet loss rate does not occur, although Internet traffic encounters non-zero loss rate. Nevertheless, when priority for HD stream and interfering traffic is the same, higher utilization effects in decrease of MOS-V value.

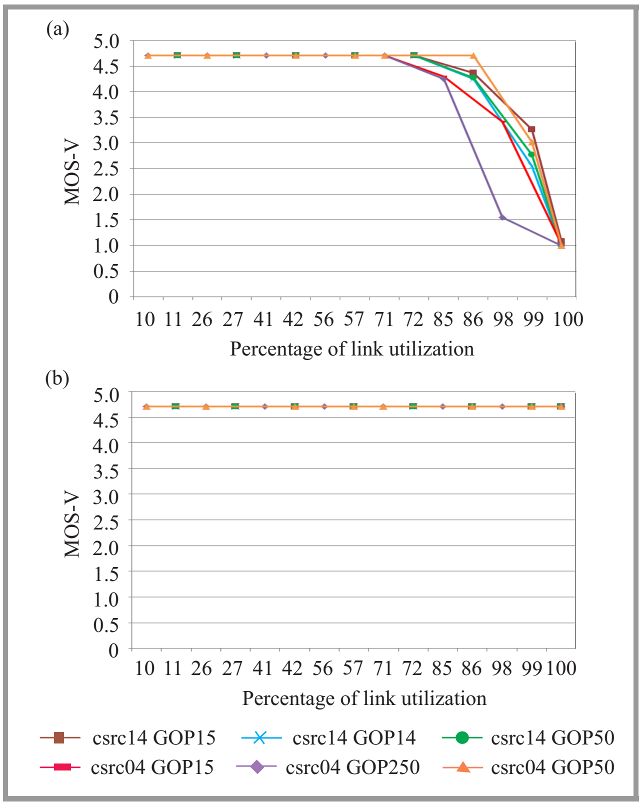

Fig. 4. Results of first and second test – MOS-V values as a function of link utilization for: (a) same_prior test and (b) higher_prior test.

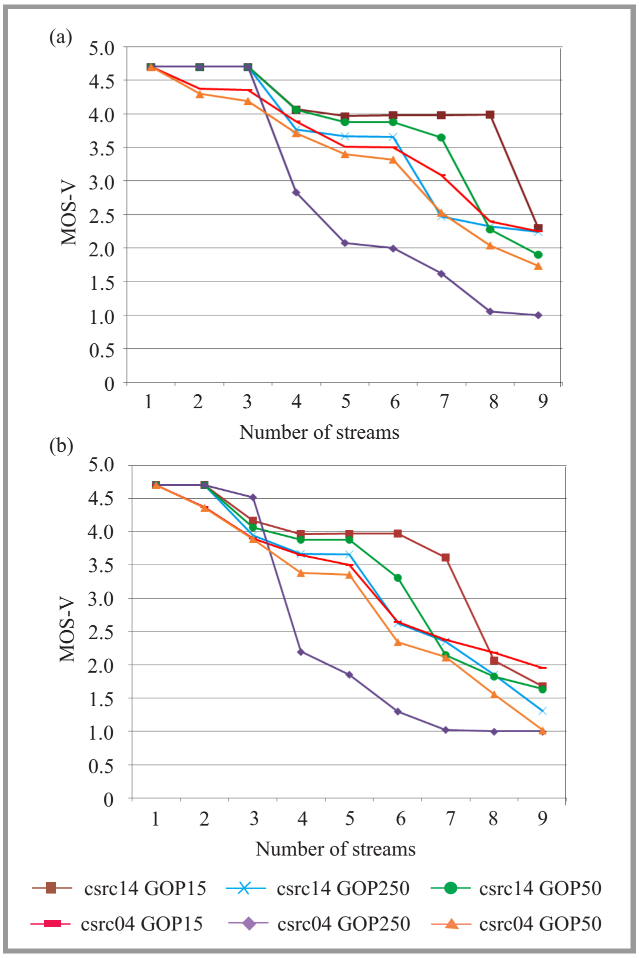

Fig. 5. Results of third and fourth test – MOS-V values as a function of number of HD streams in a link for: (a) same_prior test and (b) higher_prior test.

Packet loss rate occurs when utilization is higher than 75–80% which results in QoS drop. Reason for PLR at 75% of channel utilization and not more than 95% is VBR transmission characteristics. HD stream standard deviation of bandwidth varies from 12 to 17 Mbit/s, resulting in link congestion. Nevertheless, when streaming CBR video, the number of streams would not be higher than two assuming constant average bitrate that equals maximum bitrate. High channel utilization damages streams with long GOPs, i.e., when gap between I frames is very big. That causes substantial number of P and B frames to depend on I frames. Consequently, effects of damaged I frames are visible on numerous P and B frames, to which I frame is a reference. Therefore, increasing PLR and decreasing number of I frames, causes decay in QoE.

Figure 5 shows that increasing number of streams in a channel results in a decreased average quality. However, higher priority for video transmission is not a solution to maintain quality because the dominant share of channel utilization is video traffic and not the background Internet traffic.
As observed, in 94% of cases, for the same number of streams, videos encoded with GOP value 250 have worse quality than with GOP value 15. Therefore, profit from improved quality is not so costly due to small difference between bandwidth parameters of encoded streams (see Table 2). For tested sequences, the difference between stream with GOP value 250 and 15 is 400 kbit/s – 4% of average stream bandwidth.
From further observations, it can be concluded that serial loss of packets is more damaging for QoS than random one, which is presented in Fig. 6. Thus, using interleaving of single or multiple HD streams can result in higher average quality. However, substantial limitation of that solution is incrementing delay.

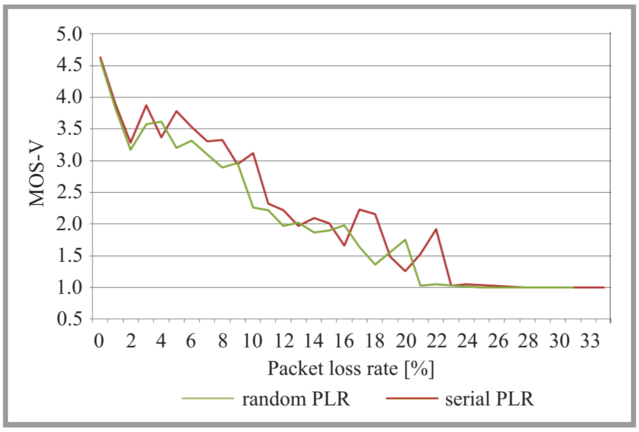

Fig. 6. Average MOS-V for random and serial packet loss rate.

Final remark refers to higher packet loss sensitivity of vqeghd1_src04 stream compared to vqeghd1_csrc14. Dynamic movement in the video causes that the traffic has lower deviation because motion vectors prediction done in the interframe domain by the encoder is not so efficient. It results in higher bandwidth consumed by P and B frames than it is in the static video sequences. It leads to less variable traffic. As a consequence, a video with lower deviation encounters more packet loss rate and is more fragile to network congestions what illustrates Fig. 7. In the presence of more than 90 Mbit/s of Internet traffic, buffers of access nodes are almost full. Congestion causes a delay of 112 ms to majority of video packets and probability of packet loss rises. Figure 7 presents that vqeghd1_src04 with lower deviation lost nearly 600 packets and vqeghd1_csrc14 lost less than 400. It is important to mention that the average bandwidth of vqeghd1_src04 is lower what proves that higher PLR is not an effect of more data sent in the channel (Table 2).
Furthermore, Fig. 7 shows that delay of 112 ms is a reason of packet loss, because it results in queue overflow. On one hand, appropriate buffer dimensioning can be a solution

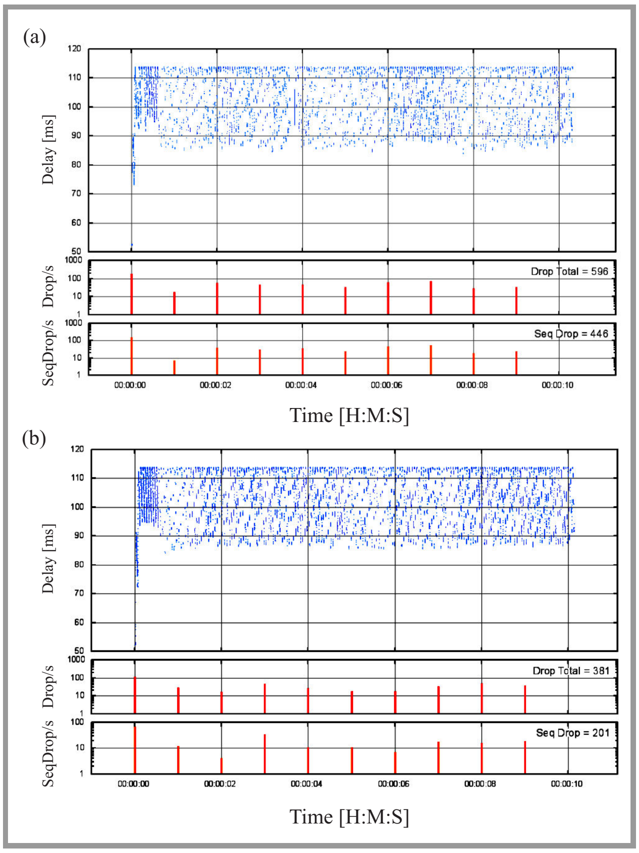

Fig. 7. Delay, packet drop and sequential packet drop values for: (a) vqeghd1_src04 and (b) vqeghd1_csrc14 streams in presence of more than 90 Mbit/s of concurrent Internet traffic.

for congestion avoidance in the presence of multiple HD streams. On the other hand, increasing buffer size causing delay of more than 3–5 seconds might be unsatisfactory solution, particularly for IPTV sport spectators.

4. Summary

This research presents influence of multiplying streams in an access network link on quality of HD video transmission. Consequences of an access network congestion evidently stunt development of new services, particularly when investments in very high broadband NGA networks progress moderately. VBR video streaming is proposed in order to facilitate transmission of multiple HD content to customer. Nevertheless, better coding efficiency causes higher deviation of bandwidth, thus traffic engineering and network dimensioning are getting more complicated.
Presented degradation of video quality is an obvious effect of network congestion. The research proved that increasing QoS for all video streams in a channel does not result in increasing QoE of HD content when video is dominant traffic in a channel. Moreover, it was observed that serial packet losses caused more severe effect on quality than random ones. In addition, sequences with lower deviation were more exposed on packet loss when congestion appears.

Furthermore, it was proved that VBR coding with high GOP values is not efficient. Short GOP values compared to long ones do not increase average bandwidth significantly, whereas quality gain is evident.

Last but not least, it was shown that NR algorithm is applicable to evaluate quality of video transmission.

In the future work, it is planned to take into consideration various NAL size settings and their influence on quality of service. It is also planned to analyze the influence of various QoS policies (i.e. shaping and scheduling) and network buffer sizes for multiple video streams in order to observe behavior and potential profit of statistical multiplexing.

References

[1] N. Anderson, "P2P traffic drops as streaming video grows in popularity", *Ars Technica* [Online]. Available http://arstechnica.com, retrieved 14.01.2013.

[2] C. Labovitz *et al.*, "ATLAS Internet Observatory 2009 Anual Report" [Online]. Available: http://www.nanog.org/meetings/nanog47/presentations/Monday/Labovitz_ObserveReport_N47_Mon.pdf, retrieved 14.01.2013.

[3] "Digital Agenda for Europe: Pillar IV: Fast and ultra-fast Internet access", European Commission [Online]. Available: http://ec.europa.eu/4/en/our-goals/pillar-iv-fast-and-ultra-fast-internet-access, retrieved 6.03.2013.

[4] T. Lakshman, A. Ortega, and A. Reibman, "VBR video: Tradeoffs and potentials", *Proc. of the IEEE*, vol. 86, no. 5, pp. 952–973, 1998.

[5] VQuadHD official Website [Online]. Available: http://www.vquad-hd.info, retrieved 10.10.2013.

[6] Telchemy [Online]. Available: http://www.telchemy.com/appnotes/Understanding%20IP%20Video%20Quality%20Metrics.pdf, retrieved 23.03.2013.

[7] "Subjective video quality assessment for multimedia applications", ITU-T Rec. P.910, 2008.

[8] "Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference", ITU-T Rec. J.341, 2011.

[9] "Report on the Validation of Video Quality Models for High Definition Video Content", VQEG, 2010.

[10] "Parametric non-intrusive assessment of audiovisual media streaming quality", ITU-T Rec. P.1201, 2012.

[11] T. Uhl, S. Paulsen, and K. Nowicki, "Transport possibility for MPEG-4/AVC and MPEG-2 encoded video data in IPTV: A comparison study", in *Proc. 10th Int. Conf. "Mulimedia w Biznesie i Zarządzaniu"* (*"Multimedia in Business and Management"*), Częstochowa, Poland, 2013, pp. 83–95.

[12] "Advanced video coding for generic audiovisual services", ITU-T Rec. H.264, 2010.

[13] VideoLan organisation [Online]. Available: http://www.videolan.org, retrieved 29.12.2012.

[14] J. Mongay Batalla and P. Krawiec, "Nowe scenariusze biznesowe wykorzystujące możliwosci protokołu IPv6", *Przegląd Telekomunikacyjny + Wiadomości Telekomunikacyjne*, nr 8–9, pp. 1323–1331, 2012 (in Polish).

[15] A. MacAulay, B. Felts, and Y. Fisher, "Whitepaper – IP Streaming of MPEG-4: Native RTP vs MPEG-2 Transport Stream", Envivio, 2005.

[16] Telchemy [Online]. Available: http://www.telchemy.com/vqmon.php, retrieved 23.03.2013.

[17] "Broadband Forum", Tech. Rep. TR-126 Triple-play Services Quality of Experience (QoE) Requirements, 2006.

[18] "Network model for evaluating multimedia transmission performance over Internet Protocol", ITU-T Rec. G.1050, 2011.

Piotr Makowski graduated from Telecommunication at Faculty of Electronics and Information Technology and received M.Sc. in 2009. In 2010 he started Ph.D. studies in video transmission quality domain. Since 2008 he is senior specialist in R&D at Orange Group responsible for researches and tests in xDSL access network domain and architecture of access network and development of new access technologies. His main activities are related to xDSL, video transmission, access network technology and video transmission quality domain. He is author of two articles regarding VDSL2 testing methodology and parameters analysis and a paper describing parametric model of video conference quality. In 2010 he applied for a patent to European Patent Office with invention crosstalk reduction by DPBO parameters determination.

E-mail: p.makowski@tele.pw.edu.pl
Institute of Telecommunications
Faculty of Electronics and Information Technology
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland