QoE based comparison of H.264/AVC and WebM/VP8 in an error-prone wireless network

Omer Nawaz, Tahir Nawaz Minhas, Markus Fiedler
Department of Technology and Aesthetics (DITE)
Blekinge Institute of Technology
Karlskrona, Sweden
{omer.nawaz, tahir.nawaz.minhas, markus.fiedler}@bth.se

***Abstract*—Quality of Experience (QoE) management is a prime topic of research community nowadays as video streaming, online gaming and security applications are completely reliant on the network service quality. Moreover, there are no standard models to map Quality of Service (QoS) into QoE. HTTP media streaming is primarily used for such applications due to its coherence with the Internet and simplified management. The most common video codecs used for video streaming are H.264/AVC and Google's VP8. In this paper, we have analyzed the performance of these two codecs from the perspective of QoE. The most common end-user medium for accessing video content is via home based wireless networks. We have emulated an error-prone wireless network with different scenarios involving packet loss, packet delay and delay variation. The focus of this paper is evaluating the end user satisfaction with the multimedia content by subjective assessment using Mean Opinion Score (MOS). We have observed the overall superiority of H.264/AVC but the VP8 codec strongly gains ground in highly error-prone networks in terms of user satisfaction.**

***Keywords*—Quality of Experience, H.264/AVC, WebM(VP8), Multimedia communication, QoE management**

I. Introduction

The emergence of affordable handheld devices with highly capable hardware/software for smooth execution of multimedia applications has brought a revolution towards Internet usage. According to the latest Cisco's report [1], the wireless (wifi) and mobile based access devices will generate the two third of the overall global IP traffic by 2020. The share of IP based video traffic is already 70%, which is forecasted to reach 82% by 2020. Thus the prime aim of internet service providers is to make sure that the end-users are satisfied with their services in terms of quality of video content. It's an established fact that typical Quality of Service (QoS) metrics, i.e. packet loss, delay, jitter, etc. used by service providers to measure the service quality does not correlate with the actual multimedia quality delivered to the consumers. The main reason is the different nature of multimedia traffic as some packets are more important than others. The loss of I frame would certainly impact huge loss in terms of video quality as compared to P or B frame types respectively. According to [2], the latest video encoding schemes have improved efficiency and error resilient features but the effect of losing high-priority frame in terms of video characteristics will incur greater damage to the overall video quality. The same notion is also applicable to delay variation. The delay of the sensitive frame on which the subsequent inter-frames are dependent will result in more quality loss as compared to lower priority frame. Hence, the traditional QoS metrics simply fails to analyze the network measurement's impact on the end-user service satisfaction.

The other approach to measure the user-satisfaction is by direct interaction via subjective assessment. But the downside is the time and cost associated with these qualitative subjective assessments and their inability to be applied in real-time networks. The objective measurement quality tools like Mean Squared Error (MSE), Peak signal-to-noise ratio (PSNR), Structural Similarity Index (SSIM), etc. are termed as indirect metrics to assess the video quality. However, these metrics may not correlate with human perceived video quality and are unable to assess the impact of play out interruptions [2], [3]. Therefore, its imperative to test the most common video streaming tools by replicating real network scenarios from the subjective point of view. The term Quality of Experience (QoE) is defined as *'the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user's personality and current state'* [4]. QoE can be considered as a framework that encompasses the established QoS metrics along with objective measurement tools and subjective evaluation of human satisfaction. So, the most common video encoding schemes, streaming methods, media access methods and data transmission protocols needs to be validated from the end-user perception. The user should be able to watch multimedia without noticing any delays, blurriness, etc.

The nature of video traffic, and especially of live streaming, makes it more resilient to packet losses as compared to traditional data, allowing an RTSP/UDP/IP based traditional approach for such applications. However, this selection rules out the possibility of bandwidth management at transport layer and best-effort service is provided by the IP [3]. The traditional streaming protocol like RTSP requires a connection and frequent communication between client and server. In the recent years, HTTP based streaming has gained popularity due to its scalability, reliability, proximity to the user's location and high availability via general purpose servers. There is a clear trend towards HTTP as main protocol for multimedia communication because underlying TCP/IP is widely implemented and its ability to use standard HTTP servers and caches to deliver the content. This means that the client is managing the whole session resulting in providing effortless streaming services via bypassing local firewall and NAT issues. Moreover, HTTP provides bandwidth management by automatically selecting initial content rate and modifying it without any negotiation with the streaming server [5]. HTTP based streaming is reliable due to TCP but the bandwidth fluctuations or packet loss will result in retransmissions. Hence, the high probability of Bit Error Rates (BER) and delay variations due to mobility, fading, handover, etc. would result in severe performance degradation and subsequent user perception.

In this paper, we have analyzed the user perception of two of the most common and widely implemented encoding standards, i.e. H.264/AVC and WebM/VP8. We have selected the HTTP based streaming method for both the encoding schemes. We have created an emulated setup with tightly controlled wireless connectivity at the end node to replicate the most common access mechanism nowadays. We have tested the different BER scenarios resulting in subsequent packet losses to monitor the actual effect on video quality. We have also applied the most common values of delay variation to replicate the received quality of the content. The focus of this paper is entirely on the human assessment about the performance degradation of these codecs in terms of video quality under identical network conditions. We have detailed the results along with regression coefficients for benchmarking with the objective metrics.

The paper is organized as follows. In Section II, we provide an overview of the related work. The Section III provides brief overview of H.264/AVC, WebM and QoE based subjective metrics. The detailed experimental setup information with every parameter is available in Section IV. The assessment results and necessary explanations are available in Section V. Finally, the conclusions are outlined in Section VI.

II. RELATED WORK

Although H.264/AVC and VP8 are the most common video encoding schemes, but we have failed to find a study that has performed a simple test of their performance over emulated test bed and subsequent qualitative subjective assessment. The authors in [6] have tested H.264/SVC performance over LTE networks. They have used both the full-reference and no-reference approaches. The study has concluded that no-reference metric like blurring can substitute the full-reference for real-time video adaptation. There is no comparison provided of any other encoding standard.

The performance of H.264 and VP8 in terms of traffic statistics and PSNR is discussed in [7]. However, the study does not include quality metrics.

The impact of QoE on HTTP video streaming with different frame rate and video resolutions is performed in [8]. The authors have used the perception evaluation of video quality (PEVQ) and temporal quality metric (TQM) algorithms to obtain MOS. The results of both PEVQ and TQM differ from the user ratings highlighting the importance of subjective assessment. The writers were focused on MOS correlation rather than comparison of different encoding standards.

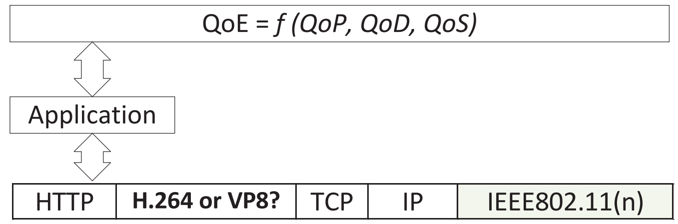


Fig. 1. QoE and layers involved in this paper.

Finally, the authors in [9] have carried out a simulative study using Open Evaluation Framework Multimedia Over Networks (OEFMON) to analyze the video quality of H.264/AVC, H.264/SVC and VP8 over wireless networks. The authors have used PSNR values to guess the MOS. As already mentioned in section I, this approach is faulty and could not be considered authentic.

III. OVERVIEW OF PROTOCOL STACK FOR HTTP STREAMING

In this section, we provide a brief overview of HTTP based video streaming using H.264/AVC and WebM/VP8 encoding standards.

The protocol stack for a typical HTTP streaming session is shown in Figure 1. The HTTP streaming is gaining popularity due to ease of its management and bandwidth management features of TCP. The stateless nature of HTTP means that the entire session is maintained by client and terminated after receiving data via HTTP GET methods. Thus, the majority of cloud based video streaming services like YouTube rely on HTTP to stream.

The Advanced Video Coding (AVC) used in our experiments is composed of two layers. The Video Coding Layer (VCL) is responsible for video presentation, coding, compression, decompression, etc. The important concept is isolation of VCL from underneath Network Abstraction Layer (NAL) which is responsible for transportation of compressed data. The NAL can automatically switch between steam or packet mode as per the requirement of the application. The H.264/AVC has three frame types, i.e. I (intra), P (predicted) and B (bidirectional predicted) frames. I frames are the independently coded block that holds the major chunk of data. P frame is dependent on either I or previous P frame. The B frame is also dependent on other frames in either direction for decoding. The P and B frames are almost 50% and 25% of the size of I frame respectively. Thus, the loss or delay of I frame can incur greater damage as compared to the other two types [3], [10].

VP8 is an open source codec that was released by Google. The advantage of using VP8 for content providers is its reduced cost as compared to H.264/AVC. The VP8 classifies frames into Intra-frame and Inter-frame. Intra-frame is the independent block just like I frame. Inter-frame prediction is done via 'golden frame', 'last frame', and 'alternate reference

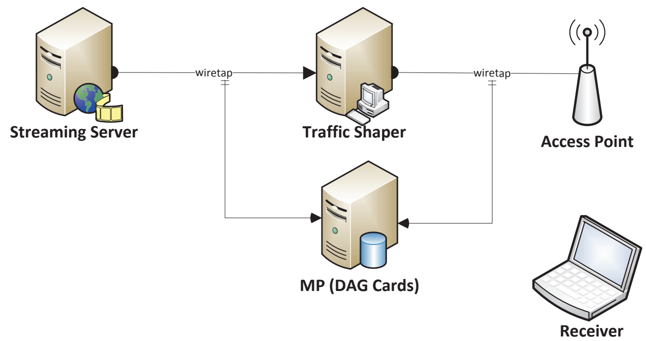


Fig. 2. Video Broadcast Setup: Emulated Approach

frame'. Golden frame is used as a copy of the background to improve coding efficiency [11].

A. A QoE Perspective

The QoE involves an end-to-end user satisfaction level. The methods used to evaluate user perception are generally classified as subjective and objective where:

- Subjective tests generally involve real humans (normal or expert) giving their opinion via Mean Opinion Score (MOS) with different scales. In this study, we have used a scale of 1 to 5. The score 5 is considered excellent whereas the score of 1 is considered bad. This type of ratings consumes time and money, so software alternatives based on different algorithms have been proposed.
- Objective tests generally rely on benchmarking the transmitted and received video in terms of MSE, PSNR, SSIM, PEVQ, etc.

Previous studies have exhibited the inefficiency of objective tests as compared to subjective measures due to priorities associated with video delivery [2]. In [12], [13], the authors have proposed a QoE hour glass model and termed QoE as a composite function of QoS metrics, content delivery and presentation to the end user as shown in Figure 1. The video presentation primarily depends on Quality of Delivery (QoD) which itself is a function of quality of data delivery of the application and the network. Thus, the important aspect is to make sure that the human subject's feedback is taken under well-defined repeatable conditions to allow for a valid evaluation of results from different studies.

IV. Experimental Setup

We have used emulated setup in our experiment to correctly measure the network artifacts on video delivery under controlled environment. The block diagram of experimental setup is shown in Figure 2. The streaming of videos was carried out on an HP Micro Tower machine powered by AMD Athlon x64@2.7GHz processor with 2048MB DDR3 RAM. The server was running Ubunto 12.04, and we have used an open source streaming server named Flumotion for streaming media. It supports all the leading audio and video codecs with good performance and stability. The traffic is shaped using Netem emulation software as a traffic shaper for dropping packets and managing variable delay due to its accuracy [14]. The traffic is captured using wiretaps on a Measurement Point (MP). The MP is a Linux system equipped with two Endace Digital Acquisition and Generation (DAG) 3.6E cards. This setup allowed passive traffic capturing by avoiding clock synchronization issues. The flushing command is used to empty the buffers of the DAG cards after every session. The wireless connectivity was obtained using a DLink DAP-1522 access point. It's important to mention that to reduce the external interferences; the access point was placed using a Ramsey STE3000FAV Radio Frequency (RF) shield test enclosure. The client machine has the same specification as of the streaming server, but the received videos were captured and saved on network storage using VLC media player version 2.0.6. In order to accurately emulate the network impairments effects, the jitter buffer value of VLC player was set to zero.

A. Video Selection

Four common videos were selected from xiph.org test media to perform the experiment. The technical specification of these videos is provided in the Table I. Four common videos were selected from xiph.org test media to perform the experiment. The technical specification of these videos is provided in the Table. We have followed the ITU-T recommendations to make sure that the selected videos should have different spatial and temporal parameters [15]. The Football and News are considered fast in the temporal domain while the Forman and Hall Monitor could be rated as medium. The common profiles (baseline for H.264) with default settings are used for video encoding and decoding. According to the recent study, the lowest resolution used in low end smart phone is QVGA [16]. Due to this reason, we have used this resolution as it's common and more suitable for error-prone networks.

B. Network Impairments

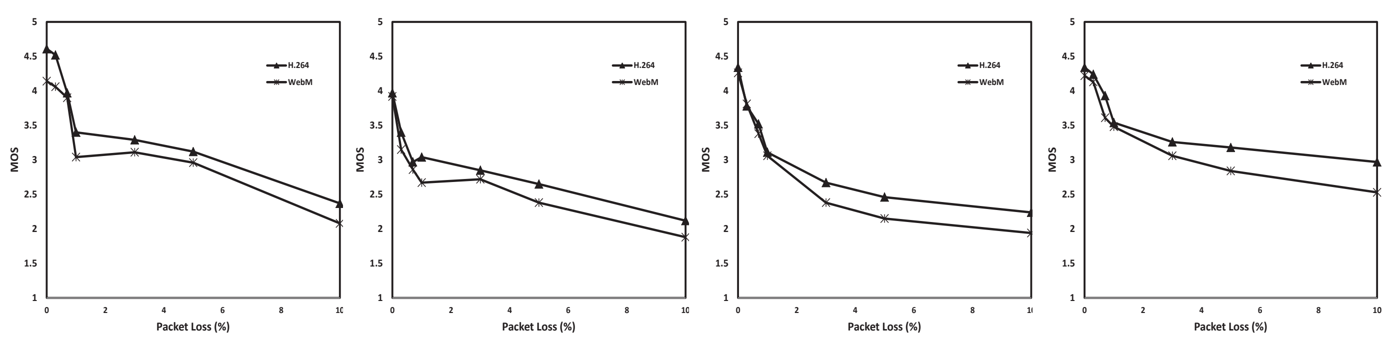
In this paper, we have used the most common QoS metrics like packet loss, delay and delay variation to evaluate the performance of the two codecs. As already explained in section I, the effect of packet loss could be catastrophic in video streaming as all packets are not equal. We have tested 0%, 0.3%, 0.7%, 1%, 3%, 5% and 10% packet loss ratios as these correspond to the most common loss scenarios in real time networks. Moreover, any video traffic having more than

TABLE I
Video Specification for Streaming and Subjective Evaluation

Parameters	H.264/AVC & WebM/VP8
Streaming Server	Flumotion
Client	VLC Player
Aspect Ratio	4:3
Frame Rate	30 fps
Duration (Football)	8s
Duration (Foreman, News, Hall Monitor)	10s
Bit rate	768kbps
Resolution	320 × 240
Encoder	libx264 & libvpx

TABLE II
MOS VALUES ALONG WITH POWER AND LOGARITHMIC REGRESSION EQUATIONS

Video Name	L [%]	H.264 Packet Loss MOS	H.264 Packet Loss MOS Trend	WebM Packet Loss MOS	WebM Packet Loss MOS Trend	D [ms]	H.264 Packet Delay Variation MOS	H.264 Packet Delay Variation MOS Trend	WebM Packet Delay Variation MOS	WebM Packet Delay Variation MOS Trend
Foreman	0.0	4.61 ± 0.18	N.A	4.14 ± 0.21	N.A					
	0.3	4.52 ± 0.15	$3.70L^{-0.159}$	4.06 ± 0.23	$3.43L^{-0.163}$	150 ± 0	4.61 ± 0.18	$4.52D^{-0.29}$	4.14 ± 0.21	$4.03D^{-0.22}$
	0.7	3.97 ± 0.33	$R^2 = 0.900$	3.90 ± 0.18	$R^2 = 0.8061$	150 ± 50	3.56 ± 0.25	$R^2 = 0.9858$	3.34 ± 0.2	$R^2 = 0.9494$
	1.0	3.40 ± 0.26	- - - - - - - - - - - -	3.04 ± 0.21	- - - - - - - - - - - -	150 ± 110	3.27 ± 0.20	- - - - - - - - - - - -	3.06 ± 0.19	- - - - - - - - - - - -
	3.0	3.29 ± 0.20	$-0.54ln(L) + 3.75$	3.11 ± 0.20	$-0.50ln(L) + 3.48$	150 ± 130	3.02 ± 0.19	$-1.07ln(D) + 4.49$	3.05 ± 0.2	$-0.76ln(D) + 4.02$
	5.0	3.12 ± 0.23	$R^2 = 0.9151$	2.96 ± 0.25	$R^2 = 0.834$	150 ± 150	2.87 ± 0.20	$R^2 = 0.9673$	2.88 ± 0.19	$R^2 = 0.9328$
	10.0	2.37 ± 0.29		2.08 ± 0.20						
Football	0.0	3.97 ± 0.18	N.A	3.92 ± 0.27	N.A					
	0.3	3.40 ± 0.22	$3.00L^{-0.112}$	3.15 ± 0.31	$2.77L^{-0.123}$	150 ± 0	3.97 ± 0.18	$3.96D^{-0.30}$	3.92 ± 0.27	$3.91D^{-0.32}$
	0.7	2.97 ± 0.26	$R^2 = 0.8519$	2.86 ± 0.17	$R^2 = 0.8228$	150 ± 50	3.16 ± 0.21	$R^2 = 0.9935$	3.18 ± 0.25	$R^2 = 0.9927$
	1.0	3.04 ± 0.23	- - - - - - - - - - - -	2.67 ± 0.24	- - - - - - - - - - - -	150 ± 110	2.88 ± 0.19	- - - - - - - - - - - -	2.68 ± 0.19	- - - - - - - - - - - -
	3.0	2.85 ± 0.19	$-0.31ln(L) + 3.02$	2.72 ± 0.26	$-0.31ln(L) + 2.79$	150 ± 130	2.64 ± 0.20	$-0.95ln(D) + 3.92$	2.54 ± 0.2	$-0.98ln(D) + 3.88$
	5.0	2.65 ± 0.26	$R^2 = 0.8849$	2.38 ± 0.27	$R^2 = 0.858$	150 ± 150	2.39 ± 0.20	$R^2 = 0.9905$	2.37 ± 0.2	$R^2 = 0.986$
	10.0	2.12 ± 0.19		1.88 ± 0.22						
News	0.0	4.34 ± 0.18	N.A	4.26 ± 0.18	N.A					
	0.3	3.78 ± 0.27	$3.18L^{-0.156}$	3.81 ± 0.26	$3.04L^{-0.203}$	150 ± 0	4.34 ± 0.18	$4.42D^{-0.36}$	4.26 ± 0.18	$4.16D^{-0.35}$
	0.7	3.52 ± 0.26	$R^2 = 0.9862$	3.38 ± 0.24	$R^2 = 0.9925$	150 ± 50	3.58 ± 0.19	$R^2 = 0.9865$	3.18 ± 0.25	$R^2 = 0.9859$
	1.0	3.11 ± 0.27	- - - - - - - - - - - -	3.06 ± 0.22	- - - - - - - - - - - -	150 ± 110	2.91 ± 0.25	- - - - - - - - - - - -	2.74 ± 0.24	- - - - - - - - - - - -
	3.0	2.67 ± 0.27	$-0.46ln(L) + 3.23$	2.38 ± 0.26	$-0.56ln(L) + 3.11$	150 ± 130	2.73 ± 0.20	$-1.20ln(D) + 4.35$	2.6 ± 0.2	$-1.14ln(D) + 4.13$
	5.0	2.46 ± 0.28	$R^2 = 0.9791$	2.15 ± 0.24	$R^2 = 0.9859$	150 ± 150	2.44 ± 0.19	$R^2 = 0.991$	2.41 ± 0.2	$R^2 = 0.9648$
	10.0	2.24 ± 0.35		1.94 ± 0.24						
Hall monitor	0.0	4.34 ± 0.19	N.A	4.22 ± 0.19	N.A					
	0.3	4.24 ± 0.18	$3.73L^{-0.106}$	4.13 ± 0.18	$3.49L^{-0.134}$	150 ± 0	4.24 ± 0.19	$4.23D^{-0.28}$	4.22 ± 0.19	$4.11D^{-0.29}$
	0.7	3.93 ± 0.20	$R^2 = 0.958$	3.61 ± 0.20	$R^2 = 0.9952$	150 ± 50	3.46 ± 0.18	$R^2 = 0.9951$	3.24 ± 0.19	$R^2 = 0.9739$
	1.0	3.54 ± 0.17	- - - - - - - - - - - -	3.48 ± 0.25	- - - - - - - - - - - -	150 ± 110	3.13 ± 0.20	- - - - - - - - - - - -	2.88 ± 0.18	- - - - - - - - - - - -
	3.0	3.26 ± 0.22	$-0.38ln(L) + 3.75$	3.06 ± 0.17	$-0.44ln(L) + 3.53$	150 ± 130	2.94 ± 0.20	$-0.94ln(D) + 4.19$	2.78 ± 0.19	$-0.99ln(D) + 4.09$
	5.0	3.18 ± 0.26	$R^2 = 0.94$	2.84 ± 0.21	$R^2 = 0.9917$	150 ± 150	2.68 ± 0.20	$R^2 = 0.9916$	2.62 ± 0.2	$R^2 = 0.9528$
	10.0	2.97 ± 0.22		2.53 ± 0.23						



(a) Foreman (b) Football (c) News (d) Hall Monitor

Fig. 3. MOS for Packet Loss: (a) Foreman (b) Football (c) News (d) Hall Monitor

10% packet loss is not considered adequate from the end-user perspective. The fixed delay and delay variations also reflect a big part in traffic characteristics and have a direct impact on the video quality. It's practically impossible to eliminate the delay and delay variation in real networks so the ITU-T [17] recommends that the one way delay value of 150ms is acceptable for video transmission. The delay and delay variation could be defined as $D \pm \Delta D$ where D corresponds to the delay and ΔD as delay variation. So, we have tested 150 ± 0ms, 150 ± 50ms, 150 ± 110ms, 150 ± 130ms and

150 ± 150ms values to check the performance of the codecs under normal to extreme network conditions.

C. Subjective Assessment

The final set of videos was used for subjective assessment as per the guidelines of ITU-T BT.500 standard [18]. The viewing distance, lighting and display characteristics were carefully monitored as per the specification. Thus, the MOS ratings are obtained using specifically designed setup, and a training session was conducted prior to every assessment. Every subject was verbally briefed about the procedure along with the written instructions. The users were shown the original videos without any network impairments for the reference point. The encoded videos were shown randomly, and the user response was collected manually using the Single-Stimulus Method [18]. Each video was ranked before moving to the next one with an average delay of 8 to 10 seconds between them. A total number of 88 videos were shown to the subjects resulting in approximated total time of 30-35 minutes for every subject as we wanted to make sure that the number of evaluated videos should remain under a reasonable time limit.

V. Results and Discussion

All results of the subjective assessment are shown in Table II. A total of 38 subjects took part in the study, and five outliers were detected and subsequently removed due to inconsistency of MOS compared to the other subjects. The average MOS with 95% confidence interval for both packet loss and delay variation is calculated. The student's t distribution is used for confidence intervals due to the low number of subjects whom data is used in this paper. There are some discrepancies in MOS values e.g football video streamed via H.264 has slightly greater value of MOS at a loss ratio of 1% as compared to the loss ratio at 0.7%. One can notice a little pause (freeze) in low loss video, which may be the result of high-priority packet loss. Another, possible reason could be the sheer length of the qualitative assessment (almost 35 minutes), as the subjects started to lose attention. One interesting feedback is the fact that those who paid close attention throughout the experiment complained about stress on the eyes. More than 75% of the subjects who took part in the study classified their frequency of watching online videos as low and didn't seem to bother by most of the impairment issues in general. The fairly consistent values of confidence intervals for average MOS as shown in Table II reflect the accuracy of the subjective assessment.

All the received videos with 10% packet loss involved pronounced smearing, blurriness, freezing and jerkiness effects regardless of the encoding scheme used. The effect of packet loss on average MOS is shown in the Figure 3. It is important to mention that both H.264/AVC and VP8 have performed neck to neck until 0.7% packet loss ratio, which is fairly common in real-time networks. However, the H.264 has outperformed the VP8 codec at the loss ratio of 1% and higher due to its macro-blocking and error resilient features. One interesting finding is the low ranking of high motion videos like News and Football in the same packet loss ratios. This illustrates the fact that subjects are more critical about the quality of content with high temporal characteristics. This effect could not be replicated by any QoS or objective metric.

On the other hand, the impact of delay variation as shown in Figure 4 has interesting findings. The WebM/VP8 performs much better as compared to packet loss scenarios and shows its flexibility to the effect of delay variation. Although it's unable to surpass the H.264 but the performance of Google's codec dramatically remains constant in high delay variation test cases. This shows the ability of VP8 to cope with high delay variation via effective bandwidth management. The performance of H.264 quickly deteriorates at 150 ± 110ms levels, and its videos show the freezing effect. Again, the videos with fast motion got low average MOS as compared to the slow videos. Finally, we have calculated the logarithmic and power regression models to benchmark the user trend as shown in Table II. As the power regression does not allow for $x = 0$, so this value is not considered. The logarithmic model is most fit with user-ratings for the packet loss scenarios. In delay variation scenarios, it is clear that both power and logarithmic models are almost identical in terms of user-ratings. But, the power regression is the better fit in most cases, as the Coefficient of Determination (R^2) is closer to 1.

VI. Conclusion

In this paper, we have benchmarked the performance of two of the most widely deployed encoding schemes, i.e. H.264/AVC and WebM/VP8 by qualitative subjective assessment. The emulated setup was created to replicate the precise requirement of network impairments using HTTP streaming. Both encoding schemes demonstrated satisfactory trust level from end-users in the error-prone networks. H.264 has a superiority edge over VP8 in almost every situation, but its relative quality degrades when exposed to high delay variation. On the other hand, the VP8 holds its ground during such environment. We have also noticed a clear user trend to be more critical of video quality in fast motion clips. Finally, the results motivate for an extended study to evaluate the video qualitative database and network traces via objective measures.

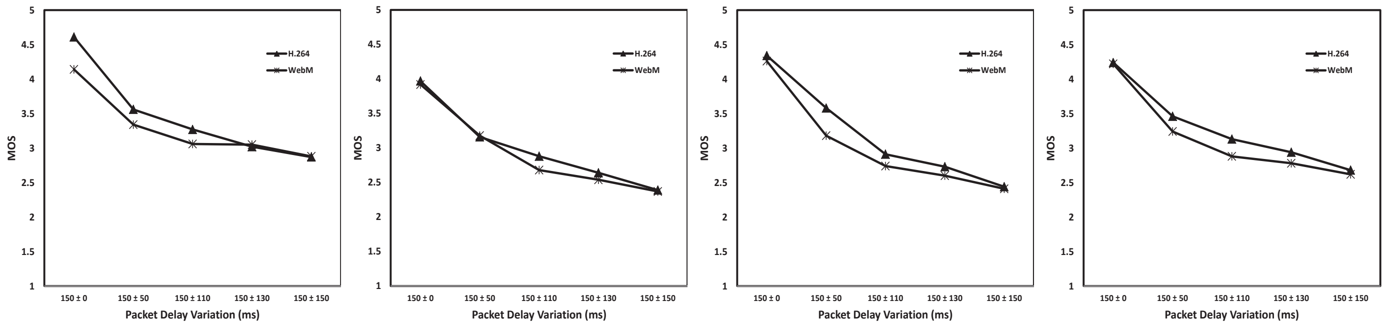
Acknowledgment

The authors would like to thank Raj Kiran Addu and Vinod Kumar Potuvardanam for their contribution to this paper via help during measurement setup and constructing video qualitative database.

References

[1] "Cisco vni forecast and methodology, 2015-2020." [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html [Accessed: 2016-12-15]

[2] J. Greengrass, J. Evans, and A. C. Begen, "Not all packets are equal, part 2: The impact of network packet loss on video quality," *Internet Computing, IEEE*, vol. 13, no. 2, p. 74–82, 2009. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4797940 [Accessed: 2014-09-17]



(a) Foreman (b) Football (c) News (d) Hall Moniter

Fig. 4. MOS for Packet Delay Variation: (a) Foreman (b) Football (c) News (d) Hall Monitor

[3] O. Nawaz, T. N. Minnas, and M. Fiedler, "Optimal MTU for realtime video broadcast with packet loss #x2014; A QoE perspective," in *The 9th International Conference for Internet Technology and Secured Transactions (ICITST-2014)*, Dec. 2014, pp. 396–401.

[4] "Qualinet white paper." [Online]. Available: http://www.qualinet.eu/index.php?option=com_content&view=article&id=45&Itemid=52 [Accessed: 2014-09-30]

[5] T. Stockhammer, "Dynamic Adaptive Streaming over HTTP –: Standards and Design Principles," in *Proceedings of the Second Annual ACM Conference on Multimedia Systems*, ser. MMSys '11. New York, NY, USA: ACM, 2011, pp. 133–144. [Online]. Available: http://doi.acm.org/10.1145/1943552.1943572 [Accessed: 2017-01-02]

[6] P. McDonagh, C. Vallati, A. Pande, P. Mohapatra, P. Perry, and E. Mingozzi, "Quality-oriented scalable video delivery using H. 264 SVC on an LTE network," in *Proc. WPMC*, 2011. [Online]. Available: https://www.researchgate.net/profile/P_Perry/publication/267957474_Quality-Oriented_Scalable_Video_Delivery_using_H.264_SVC_on_an_LTE_Network/links/54d4da100cf24647580675e3.pdf [Accessed: 2017-01-05]

[7] P. Seeling, F. H. P. Fitzek, G. Ertli, A. Pulipaka, and M. Reisslein, "Video Network Traffic and Quality Comparison of VP8 and H.264 SVC," in *Proceedings of the 3rd Workshop on Mobile Video Delivery*, ser. MoViD '10. New York, NY, USA: ACM, 2010, pp. 33–38. [Online]. Available: http://doi.acm.org/10.1145/1878022.1878031 [Accessed: 2017-01-05]

[8] T. N. Minhas, M. Shahid, B. Lövström, A. Rossholm, H.-J. Zepernick, and M. Fiedler, "QoE rating performance evaluation of ITU-T recommended video quality metrics in the context of video freezes," *Australian Journal of Electrical and Electronics Engineering*, vol. 13, no. 2, pp. 122–131, Apr. 2016. [Online]. Available: http://dx.doi.org/10.1080/1448837X.2015.1094855 [Accessed: 2017-01-05]

[9] Y. Yoon, M. Kim, S. Lee, B. Lee, S. J. Hyun, and K. Lee, "Performance analysis of H.264/AVC, H.264/SVC, and VP8 over IEEE 802.11 wireless networks," in *2012 IEEE Symposium on Computers and Communications (ISCC)*, Jul. 2012, pp. 000 151–000 156.

[10] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, p. 560–576, 2003. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1218189 [Accessed: 2014-09-17]

[11] J. Bankoski, P. Wilkins, and Y. Xu, "Technical Overview of VP8, an Open Source Video Codec for the Web," in *Proceedings of the 2011 IEEE International Conference on Multimedia and Expo*, ser. ICME '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/ICME.2011.6012227 [Accessed: 2017-01-02]

[12] T. N. Minhas and M. Fiedler, "Quality of experience hourglass model," in *Computing, Management and Telecommunications (ComManTel), 2013 International Conference on*. IEEE, 2013, p. 87–92. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6482371 [Accessed: 2014-09-17]

[13] T. N. Minhas, *Network Impact on Quality of Experience of Mobile Video*. Blekinge Institute of Technology, 2012. [Online]. Available: http://www.diva-portal.org/smash/record.jsf?pid=diva2:834917 [Accessed: 2017-02-15]

[14] "netem | the linux foundation." [Online]. Available: http://www.linuxfoundation.org/collaborate/workgroups/networking/netem [Accessed: 2014-09-30]

[15] "ITU-T P.910 subjective video quality assessment methods for multimedia applications." [Online]. Available: https://www.itu.int/rec/T-REC-P.910/en [Accessed: 2014-09-30]

[16] "The most used smartphone screen resolutions in 2016." [Online]. Available: https://deviceatlas.com/blog/most-used-smartphone-screen-resolutions-in-2016 [Accessed: 2017-01-01]

[17] "G.114 : One-way transmission time." [Online]. Available: https://www.itu.int/rec/T-REC-G.114-200305-I/en [Accessed: 2016-12-15]

[18] "ITU-T BT.500-10 methodology for the subjective assessment of the quality of television pictures." [Online]. Available: https://www.itu.int/rec/R-REC-BT.500-10-200003-S/en [Accessed: 2014-09-30]