Video Streaming Based On Scalable Video Coding Over JXTA Peer-to-Peer Network

Lahbabi Youssef^a, Ahmed Hammouch^b

^aENSET, Mohammed V Souissi University, Rabat, Morocco

^bENSET, Mohammed V Souissi University, Rabat, Morocco

lahbabi_youssef@yahoo.fr, hammouch a@yahoo.com

Abstract

In this paper, we aim to improve Peer-to-Peer (P2P) video streaming by using advanced video coding techniques, mainly layered video and to design appropriate video streaming techniques based on Hypertext Transfer Protocol (HTTP) and the Scalable Video Coding (SVC) over JXTA P2P Network. On the other hand, we investigate the behavior of scaling a video stream by using the objective Quality of Experience (oQoE) with PSNR and the subjective Quality of Experience (sQoE), typically described by the Mean Opinion Score (MOS). Finally, we believe that our results show that the use of multi-layer video coding, such as SVC, is an important step towards the future generation of P2P Video on Demand (VoD) systems.

Keywords: H.264/SVC, Video Streaming, VoD, HTTP, P2P, JXTA, QoE, sQoE, oQoE.

1. Introduction

Live video streaming over networks, especially over the Internet, is more and more used in large applications, such as VoD and Internet TV. The broadcast is usually insured by deploying several streaming servers. These services are accessible to a large number of receivers, which is increasing and can result in a server link bottleneck. Furthermore, these clients are connected to the network via various network accesses (LAN, ADSL, Wi-Fi...) and using diverse access terminals such as PC, PDA, and cell phone. These heterogeneous conditions make the stream adaptation a hard task to achieve in order to guarantee a proper Quality of Service (QoS) to the client. Stream parameter negotiation can be an effective solution to the heterogeneous conditions and be achieved by using the SVC. Such can provide several sub-streams, each one providing certain fidelity of the original video, so the client can choose the one adapted

to these capacities. On the other hand, P2P applications have become popular in recent years and offer an effective solution to overcome the limitations of centralized servers.

In the audiovisual streaming area, the popularity of P2P real-time and VoD streaming applications such as PPLive [1], PPStream [2], UUSee [3], Pando [4], Zattoo [5] has been demonstrated. As an example, PPLive has registered over 110 million users; 2 million users concurrently connected, offer more than 600 channels and has users in more than 200 countries. In addition, Youtube [6], a worldwide well-known web 2.0 streaming applications, is preparing to use P2P computing in order to improve its download rate while reducing transmission costs (especially interesting in current financial crisis time). Kawashima and Ma [7] developed a collaborative tool called TOMSCOP, which is based on the elementary peer group services offered by the JXTA general framework. TOMSCOP provides four types of services: synchronous message transportation, peer room administration, peer communication support and application space management. However, TOMSCOP suffers from inefficient pipe advertisement problems and it takes a long time. On the other, STARCast [8], which is a JXTA based platform that offers streaming services. STARCast uses an ALM (Application Layer Multicast) structure for streaming tasks. But, opposite to platform CIMSLive (Collaborative and Interactive Media Streaming Platform) which is a hybrid solution based on the P2P Java API of JXTA, DONET (Data-driven Overlay Network) [9] and ALM structure, it treats all the streams in the same manner, that is, it does not distinguish between low and high bitrate streams. In addition, there is no adjusted the level of SVC layer and no optimized by QoE. Nowadays, P2P supposes the major part of current Internet traffic (more than 60% of Internet traffic in 2006 was consumed by P2P file-sharing applications [10]). In addition, regarding to high-consuming P2P video

applications, statistics in one of the biggest Chinese Internet Service Providers (ISP) show that PPLive accounts 10% of the total Internet backbone traffic, even more than file-sharing (Bittorrent [11] represents the 8% of the traffic). Some studies [12] manifest that streaming is taking over P2P users for video content. So, video and P2P contents are both rapidly increasing Internet bandwidth demands. Recent reports predict an “exaflood” [13] from advances in video over the Internet, rich media content, and User Generated Content (UGC). Moreover, it is expected that by 2013, the sum of all forms of video (TV, VoD, Internet video, and P2P) will exceed 90 percent of global consumer IP traffic. In that sense, new systems and studies to optimize future P2P and video traffic may have a very high impact on the future of the Internet.

In this paper we propose to design and implement such solution by using a JXTA P2P architecture to share and stream scalable videos. The SVC extension of the H.264/AVC standard strives to overcome the problem of coding the underlying videos. For this purpose we conducted an extensive measurement by scaling method of the oQoE with PSNR and sQoE with MOS. The communication between a client and the sender peers is ensured by establishing several Hypertext Transfer Protocol (HTTP) sessions.

The rest of this paper is organized as follows. In Section 2, we propose architecture for scalable P2P video streaming. In Section 3 and 4, we present our simulative analysis and then conclude the paper in Section 5.

2. PROPOSED ARCHITECTURE

Consider distributed video peers for storing and streaming videos where single video processing machines are connected to a network of distributed peers connected via switches and routers. This peer interconnection provides a virtual video server to the player application. The management for the virtual server can be implemented by a JXTA-based middleware which ensures the QoS for video streaming and manages on the other hand the storage and distribution of recorded videos. Each peer offers video content and may play videos from other peers by streaming simultaneously. The basic architecture of P2P video streaming SVC is presented in Figure 1. Based on the combination of SVC and P2P networks. This system is composed of several modules as

follows:

Fig. 1. Architecture of scalable P2P video streaming.

2.1 Graphical user interface

The client needs the GUI module to display the GUI software and to facilitate communication between users of peers.

2.2 Mplayer

For the actual playback, we use a modified version of the MPlayer [14], which supports SVC Encoding [15].

2.3 Quality of Experience (QoE)

Quality of Experience is defined as the subjectively perceived acceptability of a service [16]. The perceived quality can be investigated in subjective tests, where presented stimuli such as impaired video sequences are rated by subjects under controlled conditions. The obtained rating expresses the sQoE, typically described by the MOS. However, subjective tests are time-consuming, expensive and have to be undertaken manually, which does not allow for automatic quality ratings by software. This aspect motivates objective metrics, which are designed to correlate with human perception, and, thus avoid cost and time intensive empirical evaluations. Estimates for the quality obtained by metrics are called oQoE. A more comprehensive discussion on this subject can be found in [17][18].

2.4 Scalable Video Coding (SVC)

The MPEG-4 SVC standard is an extension of the H.264/AVC standard and provides a number of different layers within one encoded bitstream. While the H.264/AVC compliant base layer of a scalable bitstream provides the minimum quality, the enhancement layers are used to further increase the quality, resolution or frame rate of the bitstream [19][20]. Thus, a client only needs to receive a small part of the scalable bitstream to consume the data in low quality, while it has to receive and decode the complete

scalable bitstream to consume the data in best quality. The usage of scalable codecs simplifies the adaptation of bitstream significantly, as an adaptation of such a bitstream can be performed by simply skipping some or all of the data related with enhancement layers. The SVC base layer may be enhanced in three dimensions: the temporal dimension (frame rate), the spatial dimension (resolution) and the quality dimension (SNR). When considering networks with fluctuating bandwidth, especially temporal and SNR scalabilities enable powerful adaptation by diminishing the video bitrate. However, when several terminals with unique device capabilities exist also spatial scalability is a considerable alternative for saving the encoding time of various different types of sequences. An essential feature of the design of the SVC extension is that the majority of the components of the H.264/AVC standard were adopted. This implies that transform coding, entropy coding, motion compensation, intra-prediction, the deblocking filter or the structure of the NAL units (NALU) are used as intended for the H.264/AVC standard. One advantage of this approach is that the base layer of an SVC encoded bitstream can generally be processed by a H.264/AVC compatible decoder, as the extensions of the H.264/AVC standard are only used to support spatial and signal-to-noise ratio (SNR) scalability. The levels are the same as the H.264/AVC levels. Figure 2 gives an example of different possible scalabilities for a video file.

Fig. 2. SVC cube, illustrating the possible scalability dimensions for a video file.

2.5 Hypertext Transfer Protocol

In spite of the general tendency of using the UDP protocol [21] for video and audio real-time delivery due to its lower latency, HTTP streaming has raised the interest of many researchers and service providers in the last years. Relying on HTTP [22]/TCP [23] allows for reusing existing network infrastructures such as the widely deployed network caches, which reduce the amount of outbound traffic

that servers have to support and consequently prevent scalability issues with respect to the system size. Although it is technically not difficult to build a similar infrastructure for RTP/UDP, it is much more costly to build it from scratch than to reuse the HTTP infrastructure that already exists. Furthermore, the use of HTTP/TCP resolves the common traversal issues with firewalls and NATs, which typically arise when the data is transmitted over RTP [24]/UDP. Moreover, implementation of HTTP streaming systems is simple, where servers are typical web servers agnostic of the actual data they provide and therefore do not need any special functionality to deal with the media files.

2.6 JXTA Peer-to-Peer Network

JXTA [25], pronounced 'juxta', comes from the word juxtapose, that reflects the operations by which peers establish temporary associations to form a P2P network (See Figure 2). Reference [26] shows a method which sends a stream media file via JXTA pipe, it denotes that we can also receive real time video package on JXTA overlap network. It is quite noticeable for JXTA beginners that the URI which is offered by official containing Relay list and Rendezvous list have not worked any more after version 2.5[25]. We should establish our own Rendezvous seed peer and Relay seed peer (if necessary) on a known IP address, so that our JXTA peer can contact each other via the services provided by them. There are three types of nodes in JXTA framework, edge peer, relay peer and rendezvous peer, they play different roles respectively when JXTA is running.

Fig. 3. JXTA Technology Virtual Network.

2.6.1 Edge Peer

An Edge peer is nearly the most frequently used peer in JXTA. As a basic JXTA peer, it tries to connect and to remain the connection to one and only one Rendezvous peer. An Edge peer can use both unicasting and multicasting to contact with other JXTA peers. Each communicate mode can be implemented as TCP or HTTP by modifying the configuration file before Edge peer starts. If multicasting is enabled, an Edge peer can communicate with other Edge peers directly on LAN.

2.6.2 Edge Relay Peer

A Relay peer is a special Edge peer who has routing ability. We can compare a Relay peer to a router on network. It is needed when a peer with private IP attempts to contact peers from WAN. Typically, a remote peer wants to establish a connection to NATed peer or peer behind firewall, it should try to find a route which can go "through" NAT server or firewall device under the help of Relay peer. It provides means for peers having private IP on a LAN to become reachable. A Relay peer must have a public IP, namely, a private IP is useless for Relay peer.

2.6.3 Rendezvous Peer

Rendezvous peers are Edge peers those offer services of JXTA rendezvous for a peer group. A peer group can assimilate to a JXTA LAN, and Rendezvous peer can assimilate to a gateway of this JXTA LAN. So, a peer group should have at least one Rendezvous peer. Edge peers can know each other via Rendezvous peer who performances as the leader of the group. Rendezvous peers accept lease requests from Edge peers and propagate these messages through the peer group to Edge peers.

3. Evaluation

3.1 Simulation setup

The video traffic trace used for the experiment is "Paris" video that uses H.264/SVC video coding with Joint Scalable Video Model (JSVM 9.16) codec [27]. This video is collected from the standard test media [28]. The encoding parameters can be seen in Table 1. In order to have a variation to PSNRs and bitrates we encoded the test sequences without enabling the rate control. From Table 2 the quality scalable sequence takes on three values namely spatial, temporal and quality identifiers denoted by (D,T,Q) respectively. In this case, the extraction is done on bitrate

allocation for each scalable layer, where as the layer extraction was done based on the D,T,Q identifiers in combined scalability. For this video to extract layer 17 with the (D,T,Q) values (1,4,1) from the Table 2, the extraction was based on the bitrate that is equal to 419.1 kb/s. The comparisons were done between the layers (0,x,0) to (0,x,1) where x is the temporal level ranging from 0 to 3 for base layer and enhancement layer 1 and also layers (1,y,0) to (1,y,1) where y is the temporal level ranging from 0 to 4 for enhancement layers 2 and 3.

Table 1. SVC encoding parameters.

Number of Layers	4
Base Layer & Enhancement layer resolution	176×144;352×288
Frames to be encoded	100
GoP size	16
Frame Rate	30
Quantization parameters	28 ;30 ; 34;36
Base Layer bitrate	53.9 ; 75.3 ; 97.8;119.5 kb/s
Enhancement layer 1 bitrate	54.6 ;76.4 ;99.6 ;122.6 kb/s
Enhancement layer 2 bitrate	163.8 ;224.9 ;221.9 ;361.3;409.5 kb/s
Enhancement layer 3 bitrate	165.6;227.2;295.3;366.8;419.1 kb/s

Table 2. Layer structure for quality scalable sequence.

Base Layer	(0,0,0) L ₀	(0,1,0) L ₁	(0,2,0) L ₂	(0,3,0) L ₃	-----
Enhancement layer 1	(0,0,1) L ₄	(0,1,1) L ₅	(0,2,1) L ₆	(0,3,1) L ₇	-----
Enhancement layer 2	(1,0,0) L ₈	(1,1,0) L ₉	(1,2,0) L ₁₀	(1,3,0) L ₁₁	(1,4,0) L ₁₂
Enhancement layer 3	(1,0,1) L ₁₃	(1,1,1) L ₁₄	(1,2,1) L ₁₅	(1,3,1) L ₁₆	(1,4,1) L ₁₇

3.2 Demonstration scenario

The demonstration will make use of the Scalable Video for P2P Streaming. It will start by presenting how the streaming system is built over this JXTA Virtual P2P Networks. After a stream has been selected, data transmission starts and, after some startup time, MPlayer is automatically launched and playback starts. The GUI of the

Scalable Video for P2P Streaming application is presented in Figure 4.

Fig. 4. A screen shot of the Scalable Coding over JXTA Peer-to-Peer Streaming.

4. RESULTS and ANALYSIS

MOS and PSNR Analysis. There are basically two approaches to measure the digital video quality, namely subjective quality measure and objective quality measure. Subjective quality metrics always grasp crucial factors which are the impression of the user when watching the video. The human quality impression is usually given in a scale from 5 (best) to 1 (worst). This scale is called MOS (Table 3) [29],[30]. While the objective quality is measured by taking the average PSNR (Peak-Signal-to-Noise-Ratio) over all the decoded frames. PSNR computes the maximum possible signal energy to noise energy. PSNR measures the difference between the reconstructed video file and the original video file. As long as the video content and the codec type are not changed, PSNR is a valid quality measure [31],[32]. Figure 5 represents a real decoded of the video sequence paris with 17 layers by using the JSVM simulator and Figure 6 indicates the plot of PSNR versus layers for this sequence. The x-axis denotes number of layers and the y-axis denotes PSNR obtained. From Figure 7 we can see the MOS scaling method corresponds to index layer. This simulation show that the fidelity values for PSNR have high correlation with the MOS scores observed. The MOS score drops as the quality or temporal identifiers are reduced and are increased when expanding their identifiers. This observation is due to the lower or higher quality and temporal identifiers, which affect the video quality.

Table 3. Possible PSNR to MOS conversion.

PSNR	MOS
>37	5 (Excellent)
31-37	4(Good)
25-31	3 (Fair)
20-25	2 (Poor)
<20	1 (bad)

Fig. 5. Frame Decoder for each layer.

Fig. 6. PSNR of a test video versus the number of layers.

Fig. 7. MOS versus layers.

5. Conclusion

In this paper, we introduced a new design for P2P streaming systems based on Scalable Video Coding and HTTP on JXTA P2P Network. Taking into consideration the heterogeneity of the peers and the network fluctuations. For that we used SVC, where a video

is divided in quality layers and each user can take more or less layers according to its capabilities. Furthermore, we aim to throw light on the differences between user experience and the objective measurements using a video quality assessment tool. Finally, we believe that our results shows that the necessity of scalability to support robustness in a network scenario and could provide research insight towards development of newer and efficient peering strategies in P2P media streaming systems.

References

- (1) PPLive official site, www.pplive.com (last visit July 2014).
- (2) PPStream official site, www.ppstream.com (last visit July 2014).
- (3) UUSee official site, www.uusee.com (last visit July 2014).
- (4) Pando official site, www.pando.com (last visit July 2014).
- (5) Zattoo official site, www.zattoo.com (last visit July 2014).
- (6) Youtube official site, www.youtube.com (last visit July 2014).
- (7) Kawashima T, Ma J (2004) TOMSCOP—A synchronous P2P collaboration platform over JXTA. In: Proceeding of the international workshop on multimedia network systems and applications (MNSA'2004), In: Conjunction with the 24th international conference on distributed computing systems. Tokyo, Japan, March 2004, pp 85–90.
- (8) T. Tsuchiya, H. Yoshinaga, K. Koyanagi, A. Honda and A. Minami, "STARCAS: Streaming Collaboration Platform using the Overlay Technology", In Proc. SAINTW'06, IEEE Computer Society, January 2006.
- (9) X. Zhang, J. Liuy, B. Liz, and T. P. Yum, "CoolStreaming/DONet: A Data-Driven Overlay Network for Efficient Live Media Streaming". In Proc. IEEE INFOCOM, March 2005.
- (10) Cachelogic official site, www.cachelogic.com/ (last visit July 2014).
- (11) Bittorrent official site, www.bittorrent.com/ (last visit July 2014).
- (12) Internet Study 2008-2009 by Ipoque, <http://www.ipoque.com/resources/internet-studies>, 2014.
- (13) CISCO IP traffic forecast, http://newsroom.cisco.com/dlls/2009/prod_060909.html, 2014.
- (14) MPlayer, "The Movie Player. Online: <http://www.mplayerhq.hu/>", 2014.
- (15) Mederic Blestel and Mickael Raulet. Open SVC Decoder. available at <http://sourceforge.net/projects/opensvcdecoder/>.
- (16) ITU-T Delayed Contribution D.197, "Definition of Quality of Experience," 2004.
- (17) B. Staehle, A. Binzenhofer, D. Schlosser, and B. Boder, "Quantifying the influence of network conditions on the service quality experienced by a thin client user," in 14. GI/ITG MMB, Dortmund, Germany, apr 2008.
- (18) Thomas Zinne, Osama Abboudy, Oliver Hohlfeldz, Tobias Hossfeld, Phuoc Tran-Gia, "Towards QoE Management for Scalable Video Streaming". In: 21st ITC Specialist Seminar on Multimedia Applications -Traffic, Performance and QoE, March 2010.
- (19) H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard", IEEE Trans. on CSVT, vol. 17, no. 9, pp. 1103-1120, September 2007.
- (20) A. Kouadio, M. Clare, L. Noblet and V. Bottreau "SVC a highly-scalable version of H.264/AVC", EBU TECHNICAL REVIEW, vol. 22, 2008.
- (21) J. Postel, User Datagram Protocol (UDP) IETF RFC 768, August 1980, <<http://tools.ietf.org/html/rfc768>>.
- (22) R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, T. Berners-Lee, Hypertext Transfer Protocol—HTTP/1.1 IETF RFC 2616, June 1999, <<http://www.ietf.org/rfc/rfc2616.txt>>.
- (23) Transmission Control Protocol (TCP) IETF RFC 793, September 1981, <<http://tools.ietf.org/html/rfc793>>.
- (24) H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, RTP: A Transport Protocol for Real-Time Applications July 2003, <<http://www.ietf.org/rfc/rfc3550.txt>>.
- (25) JXTA, <https://jxse.kenai.com/>, 2014.
- (26) Sébastien Vénot, Lu Yan, "On-demand mobile peer-to-peer streaming over the JXTA overlay," Proceedings - International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, UBICOMM. Turku, Finland, vol. 16, pp. 131–136, 2007.
- (27) SVC Reference Software (JSVM software). [Online]. Available: http://ip.lhi.de/imagecom_G1/savce/downloads/SVCReference-Software.htm, accessed on 2014.
- (28) <http://trace.eas.asu.edu/yuv/>, 2014.
- (29) J. Klaue, B. Tathke, and A. Wolisz, "Evalvid – A framework for video transmission and quality evaluation", In Proc. Of the 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation, Urbana, Illinois, USA, 2003, pp. 255-272.
- (30) J. Korhonen, J. You, "improving objective video quality assessment with content analysis", In Fifth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM'10), Scottsdale, AZ, USA, Jan 2010.
- (31) ITU-T Rec. J.247, "Objective perceptual multimedia video quality measurement in the presence of a full reference", International Telecommunication Union, Geneva, 2008.
- (32) Q. Hyunh-Thu, and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," Electronics Letters, 44(13), pp.800-801, June 2008.