Time-constrained Aggregated-ARQ Video Streaming in Wireless LAN

Bo-Yao Huang*, Chang-Ming Lee*, Hui-Kai Su†, and Cheng-Shong Wu*
*Department of Communications Engineering, National Chung Cheng University, Chia-Yi, Taiwan.
†Department of Electrical Engineering, National Formosa University, Yunlin, Taiwan
E-mail: cmlee@ccu.edu.tw

Abstract— **In this paper, TA-ARQ (Time-constrained Aggregated Automatic Repeat-reQuest) with an adaptive rate control for SVC streaming based on UDP transport protocol is proposed. To meet the requirement of terminal equipment in the heterogeneous network environments, we consider SVC for different applications. Selected NALUs (Network Abstraction Layer Units) are sent according to the quality estimation of scalable video coding in the server. Moreover, a retransmission mechanism is adopted to reduce loading of the aggregated ARQ retransmissions in multimedia networks and enhance the quality of video streaming. Based on the evaluation of NALU quantity and display order, the priority of each unit is determined. In transmission of SVC sequences, the server can extract different transmission packets, especially for retransmitting lost data. The experimental results show that the bandwidth utilization ratio of the proposed scheme can be up to 91.25% compared with decoding-order-based decision when the theoretical bandwidth is 5 Mbps. In advance, the proposed scheme with TA-ARQ can have quality gain about 0.4 dB in the unicast receiver.**

I. INTRODUCTION

Multimedia streaming services become more popular with the development of video compression and network techniques. Especially, the wide bandwidth of network transmission is necessary to satisfy popularity of high-quality video streaming service. However, the techniques to fit this requirement are complicated because of the various resources and devices. Scalable video coding (SVC) [1] is one of key techniques to satisfy these heterogeneous and dynamic demands. However, the bandwidth is really variant in the wireless environments. For example, the traffic congestion in the network or buffer overflow in the receiver usually incurs data packet loss. Consequently, the media playback may be unsuccessful.

Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) are frequently adopted to transmit video data in network. The main concern of TCP is about the transmission rate variation. According to the network transmission quality, the transmission rate and request of retransmission of lost packets can be adjusted by the AIMD [2] which realizes the congestion control. However, this rate adaptation changes dramatically to result in low perception quality in multimedia streaming service. Moreover, TCP usually treats the transmission with dense miss caused by the channel variety in WLAN as network congestion, and the slow-start mechanism always decreases the transmission rate. Thus the reflection of rate adaptation would wait for a certain interval to fit the actual available bandwidth (ABW). This management in the slow-start phase may cause serious delay and unsuccessful multimedia streaming. On the other hand, UDP is unreliable because of lacks of congestion control mechanism. The fast growing real-time application may lead Internet into congestion collapse, especially with the unresponsive UDP protocol over the error-prone wireless network. To solve these issues, TCP-friendly rate control [3] and Real-time TCP [4] are famous to improve the bandwidth utilization for multimedia communication. In [5], two congestion control schemes (CCID2 and CCID3) are proposed for TCP-like and TCP-Friendly Rate Control (TFRC) in DCCP (Datagram Congestion Control Protocol). The latter can predicts the available bandwidth by the received and lost pack instead of sliding-window and retransmission mechanism. In general WLAN transmission, the lost packets are not only caused by network congestion, but also by the variety of channel quality. This feature is considered in our proposed TA-ARQ.

II. RELATED WORKS

A. *SVC extractor*

The concepts of SVC can be realized by retaining the base layer and removing several enhancement layers according to the channel bandwidth and the layer dependency. Extractor following the SVC coding can provide adjustment of SVC-coded bit-stream with selecting appropriate enhancement layers according to the instrument properties and transmission conditions. In JSVM, extractor (called Basic) can access the SEI (Supplemental enhancement information) in NALU (Network Abstraction Layer Unit) to realize the video scalability. In [6], the layer effect value (*LE*) of each NALU in an Intra period is determined. Based on the Lagrange cost function, the unit-level optimization of the video extraction can be achieved in the video streaming. In addition, the priority ID (PID) is also considered in this paper. This value can be obtained as the ratio of *LE* and the size of a NALU.

B. *Aggregated ACK*

Since per-packet ACK can cause numbers of ACK signals and time-out issue of packet loss, Wang et al [7] utilize the reliable transmission protocol, Service Specific Connection Oriented Protocol (SSCOP), in the asynchronous transfer mode (ATM) to report the receive error in WLAN transmission, instead of MAC layer retransmission including the time-constrained aggregated ARQ (TA-ARQ) and adaptive rate control mechanism. SSCOP can provide the reliability in the network transmission by bi-directionally sending the control packets, including POLL, STAT (Solicited Status), and USTAT. The improvement of Bitmap in STAT is proposed in [8].

III. RATE-ADAPTIVE STREAMING SERVICE BASED ON AGGREGATED ARQ

In this paper, the proposed scheme is based on UDP. The dynamic control scheme to adapt the SVC transmission rate according to the network status in the application layer is capable of fairness and TCP-friendliness. The network status can be observed by the TA-ARQ which can provide a reliable transportation mechanism. The main idea is the adjustment of video transmission rate (for the data sent to the network) in according to the packet loss ratio. The high packet loss ratio seen as network congestion is a sufficient condition to decrease the transmission rate, and inversely.

A. Adaptive rate control over WLAN

The algorithm of proposed adaptive rate control is demonstrated in Fig.1. In the initialization, *initial_rate* (the default of transmitter sending rate) is assumed after analyzing several network conditions to avoid the network congestion incurred by the high transmitting rate. Generally, only the content in the base layer can be sent. Therefore, the number of corresponding packets ($N_{expected}$) can be expected. Following the initialization, the transmitter can make the decision (to increase or decrease the transmission rate) based on the Bitmap information (defined in [8]) recorded in the latest STAT. The packet loss rate is seen as the disturbance in the wireless channel instead of the network congestion. The rate increase process is considered for the case that the packet loss rate ($RecentPacket_{loss}$) is lower than the value α. Therefore, more packets can be sent and video quality can be improved in the SVC decoder. In the second case (the loss rate is in the interval [α, β]), the transmission rate remains unchanged for the satisfactory of the current network status. In the last case (the loss rate is higher than β), less packets are transmitted by the decrease process, because the network congestion is resulted by the high-rate transmission which is over the substantial network bandwidth. According to SVC streaming structure, the transmission of base layer in the SVC streaming is considered for the minimum sending bitrate (*minimum_rate*) in the transmitter. The maximum sending rate (*maximum_rate*) is noted for transmitting all layers. The *increase* (sending more packets) and *decrease* (sending less packets) rates are 0.8 Mbps and 0.7 Mbps, respectively. The transmission rate (T_{rate}) can be determined by the adaptive rate control algorithm. Based on T_{Rate}, the number of enhancement layers ($EL_{approprriate}$) can be calculated by the extraction/decision process.

B. Time-constrained aggregated ARQ

In the multimedia network, the video streaming quality improved by SVC can properly meet the application requirements of the various devices in an extremely heterogeneous network environment. TCP sliding window mechanism may adjust sending rate dramatically, and its retransmission scheme for real-time video packet recovery may suffer a long round trip time delay. Unfortunately, the lack of inbuilt congestion and flow control mechanisms within UDP may frequently cause the network congestion. Besides, the retransmission mechanism to ensure the integrity of the video data is invalid in UDP.

To cope with these transmission issues over UDP, the sender can pass the POLL packet to ask the receiver to report the packet status of the base and enhancement layers dynamically or in the end of time-constraint interval for each GOP. In Fig. 2, a Bitmap is considered in the modified STAT packet, where the transmission status of each packet can be indicated bit-by-bit. The retransmission would consider the dependency between base and enhancement layer in SVC streaming along with the time constraint.

Initial:
  T_{rate} = *initial_rate*;
  $N_{expected}$ = the number of packets in the first EL;
New STAT:
  if ($RecentPacket_{loss}$ <= α)
    $T_{rate} = T_{rate} + increase$;
  else if ($RecentPacket_{loss}$ >= α && $RecentPacket_{loss}$ <= β)
    $T_{rate} = T_{rate}$;
  else
    T_{rate}=max(min(T_{rate}**decrease*,T_{rate}*(1−$RecentPacket_{loss}$)), *minimum_rate*);

  if (T_{rate} <= *maximum_rate*)
    $EL_{approprriate}$ = appropriate number of ELs to match T_{rate};
    $N_{expected}$ = the number of packets from first EL to $EL_{approprriate}$;
  else
    $N_{expected}$ = the number of packets in all ELs;

Fig.1 Adaptive rate control algorithm

0 15 31

Control Message Type (00 00 00 02)	
POLL Sequence Number	First Sequence Number
Last Sequence Number	Bitmap
Bitmap . .	

Fig.2 STAT packet.

C. Video streaming Framework

The SVC streaming system are proposed in two parts (streaming sender and unicast receiver). In Fig. 3, following the SVC encoder, the impact of each NALU is normalized to obtain PID as dividing the factor (*LE*) by the data size. Instead of the PID-based decision, we also consider the optimized extraction (called Opt) in [6] while the transmission rate could be predicted by the STAT analysis or adjusted by the adaptive rate control. Afterward, the information would be stored in the sending buffer and accessed by the sending process. According to the transmission time constraint of a GOP, the sending buffer can temporally save the selected SVC data. Besides, according to the received STAT information, the

RTP (or NALU) selected for resisting the transmission error could be sent again. In the sending process, RTP and POLL packets are sent via the network interface for the video content and packet status, respectively. Based on the report information in STAT from the receiver, the sender can adjust the SVC streaming by the extraction/decision process.

In Fig. 4, the receiver should report the transmission error (loss or timeout of packet) and reconstruct the remaining SVC bitstream. Based on the sequence number addressed in RTP packet, the out-of-order state can be detected in the receiving process. In the STAT building, report packet can be built according to the receive status of packet transmission and passed to the sender while the report request noted in POLL arrives. Therefore, the sender can analyze the network character (e.g., packet loss ratio) and predict the current bandwidth to realize the adaptive rate control (through the retransmission process). The lost packets can be selected for the retransmission by the TA-ARQ. Accordingly, two kinds of video packets (transmitted and retransmitted) are re-ordered in the receiver. The video can be reconstructed by the video decoder and the error concealment (for the lost frame caused by the transmission errors).

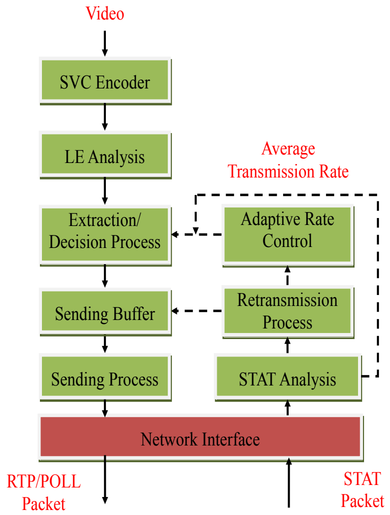


Fig. 3 Streaming sender

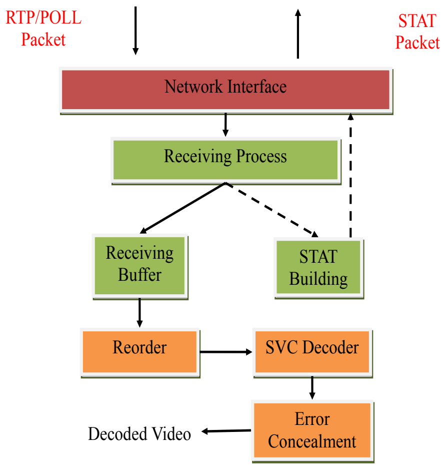


Fig. 4 Unicast receiver

After transmitting the NALUs of one GOP in the base layer, the streaming server would send the request packet POLL to ask the unicast receiver to send the status packet STAT. The Bitmap in STAT can record the information of unsuccessful transmission. Accordingly, the retransmission and sending process in the server can retransmit lost RTP (NALU) packets among the transmission of EL NALUs (in the same GOP). An example of one error (for transmitting P1) in the base layer is shown in Fig. 5. After the communication of POLL and STAT packets, the lost RTP packet can be retransmitted before transmitting the following GOP.

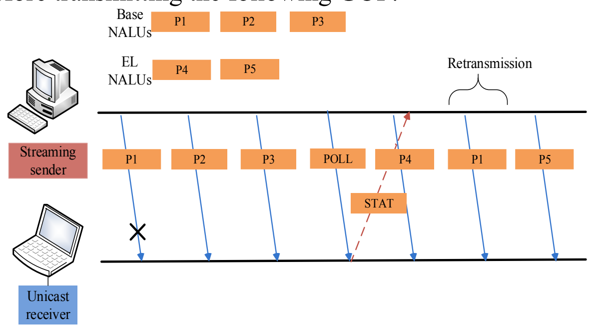


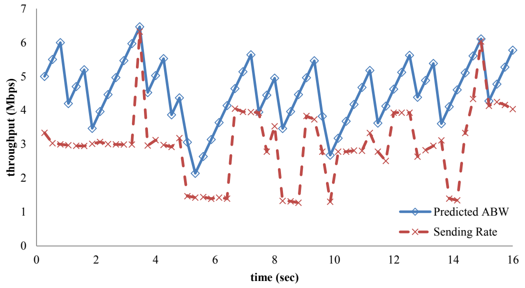
Fig. 5 Transmissions of NALU, POLL and STAT between streaming sender and unicast receiver.

IV. IMPLEMENTATION AND EVALUATION

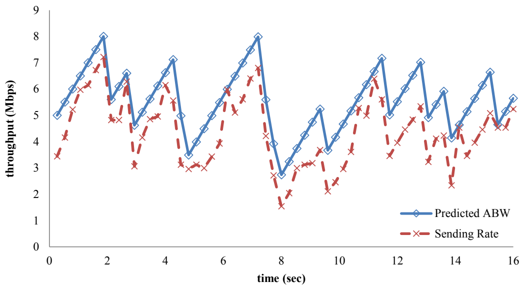
In this paper, the applications of video codec and extraction are based on the popular reference software JSVM9.18. To fit the environment with TA-ARQ and the adaptive rate control in the WLAN, the UDP transmission rate would be determined in the application layer (Enhance RTP). The total WLAN bandwidth is 27 Mbps. The WLAN background traffic (22 Mbps) is generated from SmartBits (Spirent 600B) to a background traffic receiver. The theoretical bandwidth (5 Mbps) is given from the difference between the total bandwidth and the background traffic.

In this experiment, the time-constrained aggregated rate control follows one of four extractions (including basic extraction, decoding dependency, PID, and Opt). The Dec is compatible with PID but with NALU ID depending on the decoding order. In the WLAN with available bandwidth 5 Mbps controlled by SmartBits, the channel environment remains dynamic in this four tests. After video coding by SVC encoder, the adaptive rate control can predict the transmission rate. Then, NALUs in a GOP are extracted by four different methods. The extracted units are packetized into RTP packets and data rate can be obtained. In Fig. 6, the bandwidth utilizations of Dec, PID and Opt schemes are better than that of Basic, while the transmission rate is dynamic in the actual environment. In Table I, the comparisons of bandwidth utilization exhibit the benefit of the proposed schemes, especially for the Opt scheme.

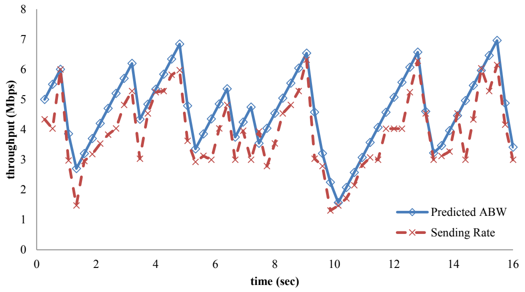
In the TA-ARQ scheme, the sender can transmit the RTP packet according to the TA-ARQ for the lost packets. All statistic results are calculated with repeating three times in the same condition. In Table II, the end-to-end system (one streaming sender and one unicast receiver) is evaluated with the available bandwidth 5 Mbps for big-buck-bunny and elephants_dream sequences. The TA-ARQ schemes can improve the streaming quality by transmitting more RTP packets to recover the transmission error. Comparing to the system without ARQ, the TA-ARQ scheme always perform more transmission efficiency. The quality improvement in both sequences is 0.4 dB.



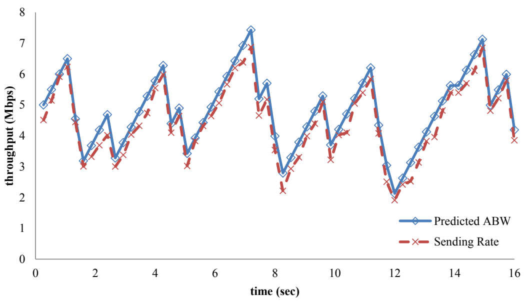
(a)



(b)



(c)



(d)

Fig. 6 Bandwidth usage with theoretical bandwidth 5 Mbps based on (a) Basic extraction, (b) Dec, and (c) PID, and (d) proposed (Opt).

Table I Comparisons of bandwidth utilization

Theoretical available bandwidth	Basic	Dec	PID	Opt
5 Mbps	77.12%	83.67%	88.25%	91.25%

Table II Performance of transmission and quality in unicast receiver with theoretic ABW 5 Mbps.

	big_buck_bunny		elephants_dream	
	No ARQ	TA-ARQ	No ARQ	TA-ARQ
Number of transmission	32785	33972	37541	38842
Number of retransmission	0	524	0	917
Retransmission ratio (%)	0	1.54%	0	2.36%
Ave. PSNR(dB)	35.9	36.3	35.3	35.7

V. CONCLUSIONS

With the development of network transmission and video compression technologies, multimedia streaming services become part of everyday life. TCP sliding window mechanism usually adjusts sending rate dramatically, and its retransmission mechanism for real-time video packet recovery may suffer a long round trip time delay and make the protocol impractical for streaming applications. In order to preserve the end-to-end semantics of the transport layer protocol and the layered structure of the Internet, TA-ARQ with an adaptive rate control for SVC streaming based on UDP transport protocol is proposed in this paper. Moreover, a retransmission mechanism is adopted to reduce loading of the aggregated ARQ retransmissions in multimedia networks and enhance the quality of video streaming. Finally, the experimental results show that the bandwidth utilization ratio of the proposed scheme can be up to 91.25% compared with decoding-order-based decision when the theoretical bandwidth is 5 Mbps. In advance, the proposed scheme with TA-ARQ can have quality gain about 0.4 dB in the unicast receiver.

ACKNOWLEDGMENT

This paper was supported by National Science Council of Taiwan (NSC102-2221-E-150 -001 and NSC102-2221-E-194 -003).

REFERENCES

[1] H. Schwarz, D. Marpe, and T. Wiegand, “Overview of the scalable video coding extension of the H.264/AVC standard,” *IEEE Trans. Circuits Syst. Video Technol.*, v(17)9, pp. 1103–1120, Sep. 2007.

[2] Y. R. Yang and S. S. Lam, “General AIMD congestion control,” ICNP, pp. 187–198, November 2000.

[3] L. Cai, X.S. Shen, J.W. Mark, and J. Pan, “QoS Support in Wireless/Wired Networks Using the TCP-Friendly AIMD Protocol”, *IEEE Trans. Wireless Communications*, v6(2), pp. 469–480, Feb. 2006.

[4] G. Lusilao-Zodi, M. Dlodlo, G. Jager, K. Ferguson, “RRB-SIMD: RTP Rate-Based SIMD Protocol for Media Streaming Applications over the Internet,” Communication Networks and Services Research Conference (CNSR), pp. 69–76, May 2011.

[5] Yuan-Cheng Lai, “DCCP: Transport Protocol with Congestion Control and Unreliability,” *IEEE Internet Computing*, v(12)5, pp. 78–83, Sept./Oct. 2008.

[6] C.-M. Lee, C.-Y. Lee, B.-Y. Huang, and K.-C. Chang, “Unit-level Optimization for SVC Extractor,” European Conference of Computer Science, pp. 73–78, Dec. 2012.

[7] H.-L. Wang, A. Velayutham, “An SSCOP-based Link Layer Protocol for Wireless LANs,” GLOBCOM, pp. 453–457, Dec. 2013.

[8] Jia-Shi Lin, Kai-Ten Feng, Yu-Tzu Huang, and L.C. Wang, “Novel Design and Analysis of Aggregated ARQ Protocols for IEEE 802.11n Networks,” *IEEE Trans. Mobile Computing*, v(12)3, pp. 556–570, March 2013.