Active Packetization and Priority Description for Scalable Video over IPv6 Based Wireless Networks

Chu-Chuan Lee
Department of Electrical Engineering
National Central University, R.O.C
Chunghwa Telecom Co., Ltd., Taiwan, R.O.C

Shao-Wei Chen
Department of Electrical Engineering
National Central University
Taiwan, R.O.C

Pao-Chi Chang
Department of Electrical Engineering
National Central University
Taiwan, R.O.C

Abstract

To cope with vastly increased unique IP addresses demand, an IPv6 specification with enhanced QoS capabilities was standardized by IETF. However, an effective interface for the priority mapping between the video application and RTP protocol is still required for the QoS provision of IPv6 networks. In addition, involving IPv6 based wireless networks, how to decide a suitable video payload length is still a trade-off problem. Therefore, this paper first proposes a Payload Priority field into the video packet header of application layer for providing sufficient priority information to the lower OSI layers. Moreover, this work develops an adaptive packetization mechanism integrated with scalable video coding for providing unequal error protection to the layered video data. Furthermore, a simplified IPv6 multimedia platform that comprises of a video streaming server, video clients, and an AAA server is implemented. Through simulations with different video sources, the proposed mechanism shows great efficiency and strength.

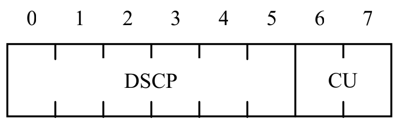
1. Introduction

With improved capabilities of wireless terminals and network infrastructures, multimedia applications such as real time news, streaming movie, video phone, and so on, are dramatically boosted to wireless networks [1][2]. However, unlike wireline links that can be models as lossless pipes, wireless links are unreliable due to the variable Bit Error Rate (BER). Therefore, video applications over wireless networks face many technical challenges that are significantly different from the problems typically encountered in the wired line desktop environment [3]. Besides, due to the great number of wireless terminals, current Internet Protocol version 4 (IPv4) cannot provide a sufficient number of unique IP addresses for all elements connected to the Internet. To cope with vastly increased demand from a wide range of users, the Internet Engineering Task Force (IETF) standardizes an IPv6 specification, RFC 2460 [4]. The IPv6 protocol can provide numerous strengths, such as huge number of addresses with hierarchical addressing structure, streamlined header format, address auto-configuration, flow label capability, enhanced IP Security (IPsec), IP mobility, and so on. Extending the Type of Service (ToS) field of IPv4 protocol, IPv6 protocol simultaneously provides a Class Field (CF) and Flow Label fields (FL) in the IPv6 header for enhancing the QoS provision capability. Therefore, this paper utilizes IPv6 network as the network infrastructure.

On the other hand, for adapting to the network environment with high bandwidth variation and limited QoS guarantee, video applications require a compression solution having powerful error resilience capabilities. MPEG-4 is an ISO/IEC standard developed by MPEG (Moving Picture Expert Group) [5][6]. It can integrate and synchronize the data associated with multiple objects so that they can be transported over network channels. MPEG-4 provides powerful error resilience capabilities, such as resynchronization, reversible variable length coding, header extension code, and data partitioning, to deal with the highly variable network environment. In addition, MPEG-4 provides scalable coding and object-based coding representation of audio-visual information that are suitable for the transmission of limited available bandwidth of networks. Therefore, this paper selects MPEG-4 as the video coding technology.

However, although the IETF RFC 3016 [7] provides rich MPEG-4 payload formats for RTP packetization, the RTP protocol still cannot obtain sufficient priority information of a layered video data from the application layer, and thus cannot reply any effectively unequal error/loss protection mechanism for the part of layered video data having higher significance [8][9]. Additionally, involving the wireless network environment, it is intuitive that a small payload length has smaller Frame Error Rate (FER) than a long payload length with the same BER. Therefore, one of possible solutions for suiting the wireless network having high BER is to reduce the payload length directly. However, a small payload length also may cause high header overhead, especially for the IPv6 header. How to seek the most suitable payload length in wireless environment is a trade-off problem.

Proceedings of the 2004 International Symposium on Applications and the Internet Workshops (SAINTW'04)
0-7695-2050-2/04 $20.00 © 2004 **IEEE**



DSCP : Differentiated Services CodePoint
CU : Currently Unused

Figure 1. The DS field structure of Diffserv model

Therefore, this paper proposes an adaptive packetization mechanism integrated with scalable video coding for suiting video applications to the variable WLAN environment. The proposed mechanism can adaptively adjust the payload length based on the current BER and the priority of current packetized video data. Moreover, this paper presents a new interface for the priority mapping between the video application and RTP protocol. Finally, a simplified IPv6 multimedia platform that comprises of an IPv6 video streaming server, IPv6 clients and an AAA server is implemented for exploring the difference between IPv4 and IPv6 video applications and the integration with the AAA server.

The rest of this paper is structured as follows. In Section II, the priority description criterion for the layered video data is first presented. In Section III, the proposed adaptive packetization mechanism integrated with scalable video coding is described in detail. In Section IV, an IPv6 multimedia platform is presented. In Section V, simulation results are discussed. Finally, section VI concludes this paper.

2. Priority description for layered video data

From the viewpoint of scalable video coding, there are several concerns to the QoS provision capability of networks, which may affect the received picture quality. The first concern to networks is how to discriminate two services, such as a voice service and video service, which have similar delay and loss requirements. The second concern is how to distinguish the parts with different importance in the layered video traffic. For example, a layered video traffic may comprise of two parts, one is the based layer data with higher significance and the other is the enhanced layer data with lower significance. Currently, the IPv4 protocol uses the Type of Service (ToS) field in the IPv4 header to distinguish the priority of IP packets. Moreover, considering the Differentiated Services (Diffserv) model that is suitable for IPv4 and IPv6 networks [10], the ToS field of IPv4 header is replaced by the Differentiated Services field (DS field) that consists of an octet IP header, as shown in Fig. 1.

However, even related QoS fields in the IPv6 header are defined well and the scalable capability of video coding are robust, an effective interface for the priority mapping between the video application and RTP protocol is still needed. Therefore, this work adds a Payload Priority (PP) field with one byte into the header of video packets, as shown in Fig. 2.

According to the given significance of a video packet, the PP field is actively filled in by the application layer. Meanwhile, the X bit in the RTP header is also set and a Header Extension Field (HE) is appended to the RTP header, following the Contributing Source (CSRC) list if present. While RTP receives a video packet coming from the upper application layer, RTP directly copies the PP value into the HE field of RTP header. Note that the content of HE field of RTP header is fully decided by the upper applications layer. After getting the priority information, RTP can then stripe the progressive video encoding layers of a hierarchically represented data across multiple RTP sessions.

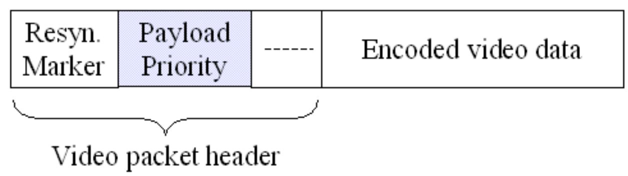


Figure 2. An improved video packet header

3. Adaptive packetization integrated with scalable video coding

As mentioned early, it is a trade-off problem to decide the most suitable payload length in wireless environment. The relationship among the BER, payload length and bandwidth utilization is formulated by [11]

$$\rho \cong \frac{l}{l^* + l}(1-(l^* + l)p) \tag{1}$$

where

ρ : bandwidth utilization
l : payload length
l^* : header overhead
p : bit error rate

Moreover, the relationship between the encoded video output rate and practical network bandwidth requirement also can be expressed by

$$\begin{aligned} R_N &= \left(L_{IPv6} + L_{UDP} + L_{RTP}\right) \times 8 \times \frac{R_v}{L_v \times 8} + R_v \\ &= R_v \times \left(\frac{L_{IPv6} + L_{UDP} + L_{RTP}}{L_v} + 1\right) \end{aligned} \tag{2}$$

where

IEEE COMPUTER SOCIETY

L_{RTP} : header of RTP Layer
L_{UDP} : header of UDP Layer
L_{IPv6} : header of IPv6 Layer
L_V : payload length
R_V : encoded video output rate (bps)
R_N : practical network bandwidth requirement (bps)

From (2), it is obvious that the ratio of R_N to R_v is expanded if the payload length decreases or the header overhead increases.

To solve the mentioned trade off problem and provide unequal error protection to the layered video data, this paper proposes an Adaptive Packetization mechanism integrated with Scalable Video coding (AP-SV), which is implemented at the application layer. Figure 3 shows the detailed algorithm of AP-SV, where ε denotes the weighting factor of BER. The α_I, α_P, and α_B represent the significance factor of I frame, P-frame, and B frame, respectively. From the temporally scalable video characteristic, the I-frame always has the highest priority but the B-frame usually has a lower priority than the P-frame. Therefore, the relationship among α_I, α_P, and α_B is $0 \leqq \alpha_I \leqq \alpha_P \leqq \alpha_B \leqq 1$. Moreover, for reacting to the impact of BER, the value ε in the case of high BER is smaller than that in the case of low BER. In summary, the resulted payload length γ_n is not only the function of scalable video characteristics, but also the function of current BER.

If T_n = I-frame **Then**
$\gamma_n = \alpha_I \times \varepsilon \times (\gamma_{mtu} - \gamma_{overhead})$
Else If T_n = P-frame **Then**
$\gamma_n = \alpha_P \times \varepsilon \times (\gamma_{mtu} - \gamma_{overhead})$
Else If T_n = B-frame **Then**
$\gamma_n = \alpha_B \times \varepsilon \times (\gamma_{mtu} - \gamma_{overhead})$
End If

T_n : n-th frame type
γ_n: packet size of the n-th frame
α_I, α_P and α_B :I, P, B frame weighting factor
$\gamma_{overhead}$: network header
γ_{mtu}: maximum transmission unit
ε : weighting factor of BER

Figure 3. The algorithm of AP-SV

4. Implementation of IPv6 video server integrated with AAA server

To evaluate the difference between IPv4 and IPv6 video applications and to integrate with the AAA server, an IPv6 multimedia platform is implemented, which consists of an IPv6 video streaming server, IPv6 clients and an AAA server, as shown in Fig. 4. When a client has accomplished authentication and authorization procedures with the AAA server, the AAA server assigns a dynamic IPv6 address to the client. Meanwhile, the AAA server also sends the corresponding user profile to the IPv6 video server. After that, the client may select a video program through the portal and the IPv6 video server then delivers the selected streaming video service to the client over IPv6 network. During the playback period, the IPv6 video server also traces and records the status of video traffic and forwards it to the AAA server for accounting purposes. The maximum served user number is set to be five at present. In the near future, we shall implement the proposed criterion for cross layer QoS mapping and the adaptive packetization mechanism to the platform.

5. Simulation results

In the following simulation scenarios, an IPv6 based IEEE 802.11b WLAN with PCF mode is constructed. The ARQ mechanism is enabled in which the max permitted retransmitted number is fixed to two for all video frame types. This work uses the Gilbert Model [12] to generate various error patterns where the burst error length is set to 10 bits. Three BER cases, including 2×10^{-4}, 4×10^{-4}, and MIX, is used in this paper. The MIX simulation case comprises of two different BER, 2×10^{-4} and 4×10^{-4}. Moreover, This study uses the MS-FDIS V1.0 codec to

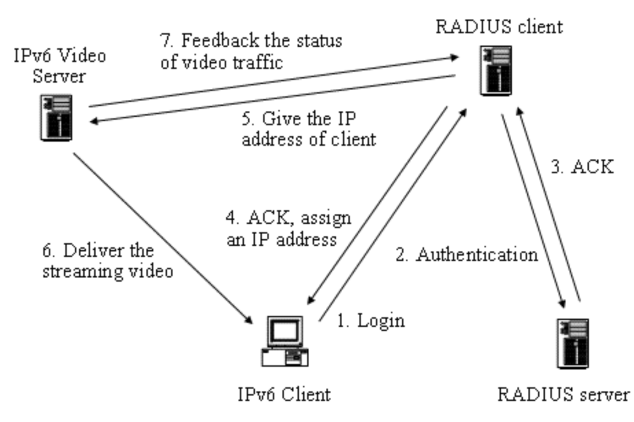


Figure 4. The structure of a simplified IPv6 multimedia platform.

Proceedings of the 2004 International Symposium on Applications and the Internet Workshops (SAINTW'04)
0-7695-2050-2/04 $20.00 © 2004 **IEEE**

generate three compressed test video sequences, "*Foreman*", "*Mobile*", and "*Coastguard*". All test video sequences have following properties: 1) All video sequences are CIF format with the frame rate of 30 fps. 2) A GOP consists of 30 frames and its pattern is set to I-B-B-P format. 3) The targeted encoding rate is 1Mbps and the TM5 rate control mechanism is enabled. Besides, this work assumes all admissible video connections with identical priority to be homogeneous. Table 1 summarizes the parameters used in the simulated 802.11b WLAN.

Figure 5 first displays the relationship among the IPv6 header overhead, system throughput and BER. From Fig.5, it is obvious that the packetization using long payload length is not suitable for the network condition with high BER. However, considering the packetization case using short payload length, it is noted that the impact of header overhead to the throughput is serious even in the network condition with light BER.

In Table 2 and Table 3, performances of the proposed AP-SV mechanism are evaluated. In both simulation scenarios, the default MTU, γ_{mtu}, is set to 1400bytes and the header overhead, $\gamma_{overhead}$, is fixed to 60 bytes which includes headers of RTP, UDP, and IPv6. A traditional Statistic-ARQ (S-ARQ) mechanism with no adaptive packetization is used here in which the max permitted retransmitted number is fixed to two for all video frame types. From Table 2, the S-ARQ integrated with AP-SV

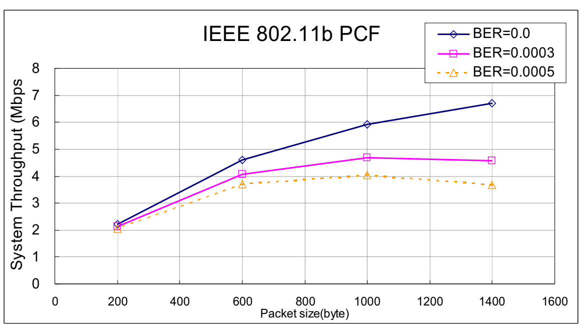


Figure 5. The relationship among the through-put, header overhead and BER.

approach can provide better PSNR performance than the original S-ARQ scheme in the case of six simultaneous served users, especially for the high BER situation. Moreover, using the *Foreman* video sequence and testing on various user numbers, the AP-SV+S-ARQ mechanism still has better PSNR than the compared scheme. However, the effect of adaptive packetization decreases while the simultaneous served user number increases. The main reason is that the large scheduling delay may cause several video packets to be dropped because of time out

Table 1. Parameters used in 802.11b

Beacon interval	100 ms	T_{H_PHY}	48 μs
CFP	90 ms	T_{DIFS}	50 μs
R_{DATA}	11 Mbps	$T_{Preamble}$	144 μs
R_{ACK}	11 Mbps	CF-End	20 bytes
T_{SIFS}	10 μs	L_{MAC}	28 bytes
T_{PIFS}	30 μs	L_{LLC}	8 bytes

Table 2. AP-SV performance in the case of fixed served user number

Video Source	Served Users	Scheme	2×10^{-4} PSNR (dB)	4×10^{-4} PSNR (dB)	MIX PSNR (dB)
Coastguard	6	S-ARQ	31.77	18.03	20.74
		AP-SV+S-ARQ	31.82	18.82	21.47
Mobile	6	S-ARQ	26.37	14.68	15.77
		AP-SV+S-ARQ	25.81	15.61	16.25

Table 3. AP-SV performance in cases of variable served user number

Video Source	Served Users	Scheme	2×10^{-4} PSNR (dB)	4×10^{-4} PSNR (dB)	MIX PSNR (dB)
Foreman	6	S-ARQ	34.13	18.38	20.22
		AP-SV+S-ARQ	35.06	18.80	20.53
	12	S-ARQ	31.43	18.46	19.37
		AP-SV+S-ARQ	31.04	18.23	19.66

Table 4. Performances of video frame loss rate with different video frame type

Video Source	Served Users	Scheme	2×10^{-4}			4×10^{-4}		
			I-Frame Loss rate	P-Frame Loss rate	B-Frame Loss rate	I-Frame Loss rate	P-Frame Loss rate	B-Frame Loss rate
Coastguard	6	S-ARQ	1.2 %	1.0 %	0.5 %	5.8 %	5.3 %	2.2 %
		AP-SV+S-ARQ	0.5 %	0.6 %	0.6 %	3.9 %	4.1 %	2.8 %
Mobile	6	S-ARQ	1.0 %	0.9 %	0.4 %	5.3 %	5.1 %	2.7 %
		AP-SV+S-ARQ	0.9 %	0.7 %	0.5 %	4.6 %	4.4 %	2.8 %

of delay tolerance.

Finally, Table 4 presents the details of video frame loss rate for each video frame type. While the AP-SV mechanism is utilized, the effect of protecting important parts of layered video data such as I-frame is significant, particularly in the case of high BER.

6. Conclusions

Based on existing RTP header and video packet header formats, this paper first adds the PP field into the video packet header to solve the cross layer QoS mapping problem. With the priority information captured from the video packet header, RTP can then stripe the progressive video encoding layers by means of multiple RTP sessions. Moreover, this paper proposes an adaptive packetization mechanism integrated with the layered video encoding to adjust dynamically the payload length. Through simulations involving three different video sources and various served user numbers, the proposed AP-SV mechanism can effectively provide unequal error protection to the layered encoded video data, especially for high BER cases. However, the effect of adaptive packetization may decrease because of large scheduling delay in the case of large served user number. In the near future, the proposed criterion of cross layer QoS mapping and adaptive packetization mechanism will be implemented on the IPv6 multimedia platform that is constructed in this paper.

Acknowledgement

The authors would like to thank Prof. Yen-Wen Chen for his assistance on the design and implementation of AAA server.

References

[1] A. Dutta, J. C. Chen, S. Das, M. Elaoud, D. Famolari, S. Madhani, A. Mcauley, M. Tauil, S. Maeda, N. Nakajima, Y. Ohba, and H. Schulzrinne, "Implementing a testbed for mobile multimedia," in *proc. IEEE GLOBECOM'01*, 2001, pp. 1944-1949.

[2] C. Kodikara, S. N. Fabri, and A.M. Kondoz, "Performance Improvement for Real-Time Video Communications by Link Adaptation in E-GPRS Networks," in *Proc. Int. Conf. on 3G Mobile Commun. Technool*, 2002, pp. 489-494.

[3] S. Hartwig, M. Luck, J. Aaltonen, R. Serafat, and W. Theimer, "Mobile multimedia-challenges and opportunities," *IEEE Trans. Consumer Electron.*, vol. 46, pp. 1167-1178, Nov. 2000.

[4] S. Deering and R. hinden, "Internet Protocol, Version 6 (IPv6) Specification," IETF RFC 2460, Dec. 1998.

[5] R. Koenen, "MPEG-4 – Multimedia for our time," IEEE Spectrum, vol. 36, pp. 26-33, Feb. 1999.

[6] D. Wu, Y.T. Hou, W. Zhu, H.J. Lee, T.H. Chiang, Y.Q. Zhang, and H.J. Chao, "On End-to-End Architecture for Transporting MPEG-4 Video over the Internet," IEEE Trans. on Circuits and Systems for Video Technology, vol.10, no.6, pp.923-941, Sept. 2000.

[7] Y. Kikuchi, T. Nomura, S. Fukunaga, Y. Matsui, and H. Kimata, "RTP Payload Format for MPEG-4 Audio/Visual Streams," IETF RFC 3016, Nov. 2000.

[8] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," IETF RFC 3550, July 2003.

[9] ISO/IEC JTC1/SC29/WG11, "MPEG-4 Video Verification Model version 18.0," N3908, Jan. 2001.

[10] K. Nichols, S. Blake, F. Baker, and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," IETF RFC 2474, Dec. 1998.

[11] M. Schwartz, *Telecommunication Net- works : Protocols, Modeling and Analysis*. Reading, Massachusetts: Addison-Wesley, 1987, pp. 119-156.

[12] E.N. Gilbert, "Capacity of a Burst-Noise Channel," *Bell System Technical Journal*, pp. 1253-1266, Sep. 1960.

IEEE COMPUTER SOCIETY