Journal of Zhejiang University SCIENCE A
ISSN 1009-3095 (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

An adaptive joint source and channel coding scheme for H.264 video multicasting over wireless LAN

LIU Zheng-ye[1], LIU Hang[2], WANG Yao[1]
([1]Six MetroTech Center, Polytechnic University, Brooklyn, NY 11201, USA)
([2]Corporate Research Lab, Thomson Inc, Princeton, NJ 08540, USA)
E-mail: zhengye@vision.poly.edu; hang.liu@thomson.net; yao@poly.edu
Received Dec. 15, 2005; revision accepted Feb.20, 2006

Abstract: This paper proposes an adaptive joint source and channel coding scheme for H.264 video multicast over wireless LAN which takes into account the user topology changes and varying channel conditions of multiple users, and dynamically allocates available bandwidth between source coding and channel coding, with the goal to optimize the overall system performance. In particular, source resilience and error correction are considered jointly in the scheme to achieve the optimal performance. And a channel estimation algorithm based on the average packet loss rate and the variance of packet loss rate is proposed also. Two overall performance criteria for video multicast are investigated and experimental results are presented to show the improvement obtained by the scheme.

Key words: Multicast over 802.11 wireless local area networks (WLAN), Adaptive joint source and channel coding, H.264
doi:10.1631/jzus.2006.A0837 **Document code:** A **CLC number:** TN919.8

INTRODUCTION

Video multicasting over wireless local area network (WLAN) (IEEE Standard 802.11, 1997) enables the distribution of live or pre-recorded programs to many receivers efficiently. An example application is to redistribute TV programs or location-specific information in hot spots such as airport. Users can watch their favorite TV programs on mobile devices while browsing the Internet. For enterprise applications, an example is multicasting video of a lecture or training session over WLANs. Other examples include movie previews outside cinemas, replay of the most important scenes in a football match, etc.

A challenging problem for video multicast over WLAN is that the underlying wireless channel is error prone due to fading and channel interference. For multicast, the 802.11 link layer does not retransmit lost packets. A data frame is discarded at the receiving MAC in the event of an uncorrectable error. Hence users with poor channel conditions may experience very high packet loss rates. Therefore appropriate error protection mechanisms are required to guarantee satisfactory video quality for all multicast receivers. The necessary error protection level depends on the user topology (i.e., distribution of users with different channel conditions). Therefore it is desirable to adapt the error protection level based on the user topology in a service area.

To overcome packet losses in WLAN video transmission, solutions targeted at different network layers have been proposed (van der Schaar *et al*., 2003) including the selection of appropriate physical layer mode, MAC layer retransmission, packet size optimization, etc. Among these, application-layer FEC (Majumdar *et al*., 2002; Wang and Zhu, 1998) can effectively reduce the error rates seen by a data sink, and error resilience provided by a video codec can maintain reasonable video service quality when corrupted data is consumed. To optimize the overall performance, these two techniques can be designed jointly.

In video multicast, each user may have a different channel condition and users may join or leave the service during a session so that the user topology can change in time. The key issue is therefore to design a system to optimize an appropriate multicast performance metric that measures the overall user satisfaction. For a chosen performance metric, this can be achieved by appropriately configuring the source codec and the FEC subject to a constraint on the available bandwidth at the application layer.

In this work, we investigate multicasting H.264 video over WLAN. Given the user topology, we jointly configure the source coder (in terms of both the quantization parameter and error resilience features) and the FEC codec (in terms of code rate) to optimize a chosen multicast performance metric. Furthermore, we propose a channel estimation algorithm that is based on the average packet loss rate and the variance of packet loss rate. Based on the channel estimation algorithm, we propose to collect feedback regarding estimated channel conditions from multicast receivers and dynamically update the source and channel coder configurations based on these feedbacks. Two overall performance criteria for video multicast and their effects on the video quality at individual receivers are investigated. We present simulation and experiment results to show that the proposed scheme improves the overall video quality of all the served users.

The remainder of this paper is organized as follows. Section 2 presents an overview of the video multicast system. In Section 3, we propose the adaptive joint source and channel coding scheme in the single user scenario. In Section 4, we investigate two overall performance criteria for video multicast. In Section 5, we extend the adaptive joint source and channel coding scheme to the multicast scenario. Section 6 concludes the paper.

SYSTEM ARCHITECTURE

The video multicast system under consideration is shown in Fig.1. The video servers are connected to the wireless access points (APs) through a high-speed Ethernet LAN which is also connected to Internet through a router or other broadband access. Stored video contents are transcoded, traffic-shaped and multicasted to a number of clients through WLAN by the video server. Some video servers also equip video capture and encoding cards with which live video contents, fed from cable/satellite set-top boxes or video cameras, are real-time encoded into H.264 format, traffic shaped and multicasted to a number of clients. The users can view one or more video programs and simultaneously access the Internet with an 802.11 WLAN card.

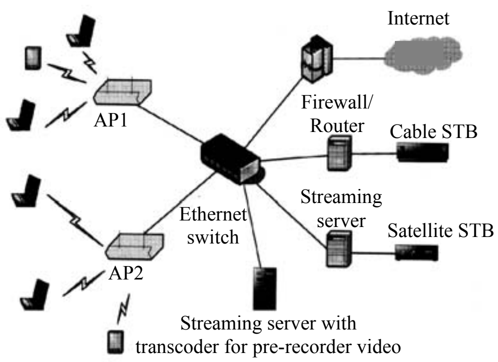

Fig.1 End-to-end architecture for video multicast over WLAN

For multicast, the 802.11 link layer does not support retransmission of lost packets. Thus additional error correction and error resilience mechanisms are required to provide satisfactory services for users within the serving area. One of the effective approaches for WLAN multicast operation is to jointly use FEC codes at the application layer and error resilience redundancy in video coding. We consider the use of Reed-Solomon (RS) codes for application-layer cross-packet FEC because the RS code is a maximum distance code with excellent error correction capability. The RS coding is applied across the RTP packets.

The optimal configuration of the source coder (or transcoder) and FEC depends on the user topology. In a typical multicast session, users may join and leave a session at any time. Also, channel conditions of mobile receivers can change dramatically as they move around. To adapt to changes in user topology, we further propose collecting feedback regarding channel conditions from multicast receivers and dynamically update the source and channel coder configurations based on these feedbacks.

JOINT ADAPTATION OF SOURCE AND CHANNEL CODING IN SINGLE-USER CASE

In this section, we focus on how to jointly consider the video source coding and channel coding to achieve an optimal performance for a single user with known channel conditions. Given an appropriate multicast performance metric and user channel conditions, the results obtained in single user case can be easily extended to the multi-user case.

When video is streamed over a lossy packet network, such as WLAN, the distortion D of the decoded video at a receiver depends both on the quantization incurred at the encoder and the channel errors that occurred during transmission and consequent error propagation in the decoded sequence. We will call the former the source-induced distortion, denoted by D_s, and the latter the channel-induced distortion, D_c. The total distortion D depends on D_s and D_c in an unnecessarily additive form.

Typically, there are multiple operating parameters that the source encoder can choose from, including the quantization parameter (QP) and the intra frame rate (the frequency that a frame is coded using the intra-mode, without prediction from a previous frame, noted as β), etc., to be denoted collectively as A. The encoding parameters A determine D_s as well as the source coding rate R_s. QP regulates how much spatial detail is saved. Smaller QP introduces lower D_s with higher R_s. Intra frame rate affects the error resilience of the video stream. More periodically inserted intra-coded frames can limit transmission error propagation and hence reduce D_c, but it will also lead to higher source rate for almost the same source distortion.

The channel distortion D_c depends both on A as well as channel error characteristics. In a simplified version, we can characterize the channel error statistics by the residual packet loss rate P, which depends on the raw packet loss rate P_e, and the FEC rate r. For (n,k) RS coding, r is defined as k/n.

For a given target bit rate R_{tot}, higher R_s will reduce the channel rate R_c allocated to FEC coding, hence P and D_c will increase. For a particular user with a given channel condition P_e, there is an optimal operation point $S^*=(A^*, r^*)$ at which D is minimal. The specific relation between D_c and A and P is very complex and accurate modelling of this relation is still an active research area. The challenge in modelling this distortion lies partly in the fact that the error in a frame can propagate to future frames because of the use of inter-frame prediction. In our current work, we generate packet loss traces based on a chosen channel loss model, and determine the total distortion based on decoded (including concealed) frames.

To achieve the optimal operation point S^*, we can formulate an optimization problem as follows:

$$D_{opt} = \min D(S, P_e) \quad \text{subject to } R_s + R_c \le R_{tot}. \tag{1}$$

Furthermore, we focus on β and r to investigate which parameter is more powerful for error resilience. Since more intra-coded frames result in higher R_s without changing D_s greatly, we separate the bit rate R_{si} which is induced by inserting more intra coded frame from R_s, and define the minimum bit rate R_{sb}, where R_{sb} is the source rate with only one intra-coded frame per GOF (Group of Frames). The bit rate used for error resilience and error correction R_r depends on R_c and R_{si}.

$$R_r = R_{si} + R_c,\ R_{tot} = R_{sb} + R_r. \tag{2}$$

Given QP, we can formulate an optimization problem to minimize the channel distortion D_c.

$$D_{c,opt} = \min D_c(\beta, r) \quad \text{subject to } R_{si} + R_c \le R_r. \tag{3}$$

In this work, we consider S as a triple-set (QP, β, r) and obtain the optimal operation point S^*. Also, we investigate when QP and corresponding R_{sb} are given, how to jointly design β and r in order to minimize D_c.

Exhaustive searching for the optimal operation point

Given a video sequence and R_{tot}, the optimal operation point S^* could be obtained by exhaustive searching from all feasible S that satisfy the constraint in Eq.(1).

In our simulation, we code the "Kungfu" video sequence in SD (720×480) resolution using the latest JM9.6 H.264 codec. Each GOF has duration T=2 s and comprises 48 frames. We encode the first 240 frames and loop the encoded video sequence 30 times to generate a 5-min long video sequence. QP is changed from 34 to 39 and intra frame rate is changed from 4 frames per GOF to 1 frame per GOF. The corresponding source coding rate ranges from 599 kbps to 366 kbps. The target bandwidth R_{tot} is set to be 600 kbps. Given R_{tot}, QP and β, all the left-over bandwidth besides source coding is allocated to R_c, hence r is determined. Herein, we use "slice mode", which means one frame could be encoded into several slices. The slice size is smaller than 1450 bytes.

To simulate the burst packet loss in WLAN, two-state Markov model characterized by the average packet loss rate (PLR) and the average packet loss burst length (ABL) is used in our experiments. To simulate the fluctuation of channel conditions, 4 different channel conditions are modelled using Markov model with different parameters (PLR, ABL): A (0.01,1.1), B (0.05,1.2), C (0.1,1.5), D (0.2,2.0). On the receiver side, the "motion copy" method available in the JM9.6 H.264 decoder is used for error concealment.

Simulation result

Fig.2 shows the effect of operation points on the received video quality. Here we use the average of PSNRs for all reconstructed 7200 frames as the video quality measure. We vary QP from 34 to 39 on the x-axis, and plot the corresponding video quality in terms of $PSNR$ on the y-axis. Different curves represent different β. The corresponding FEC rate for the given QP and β is also indicated in the figure showing that smaller QP, smaller β, and correspondingly lower r leads to higher $PSNR$, but when QP is large, video qualities for different β are similar.

These observations can be explained as follows: when QP is relatively small and hence R_{sb} the source coding rate is high, the available redundancy rate R_r is insufficient for FEC to recover all packet losses. In this case, it is more efficient to minimize packet loss rather than to stop error propagation when packet loss occurs. On the other hand, when QP is high so that R_{sb} is low and R_r is high, no matter what β is, the remaining bandwidth for FEC enables FEC to correct all lost packets, hence D_c=0 and the overall video quality only depends on the source coding.

Thus, it can be concluded that in this multicast video application, FEC rate r is the more dominant factor for error resilience than intra frames rate β. It is more efficient to allocate redundancy bits to the FEC coder than to the source coder.

Fig.3 compares the achievable video quality under different channel conditions at different operation points. We plot QP on the x-axis, and plot video quality in terms of $PSNR$ on the y-axis. We also indicate for each channel condition, the optimal operation points (QP, β, r) that leads to the highest $PSNR$. Notice that when the channel condition is poor, optimal QP is large, so that more bits can be allocated to FEC coding for channel protection. When channel condition is good, it is a waste to add a lot of redundancy through FEC. Thus we use a smaller QP to reduce source coding distortion. Also it may be noticed that for a particular channel condition, the video quality is quite different using different QP. The video quality degradation between a particular QP chosen arbitrarily and optimal QP is noticeable, especially when channel condition is poor. It is crucial to adapt the operation point to the channel condition.

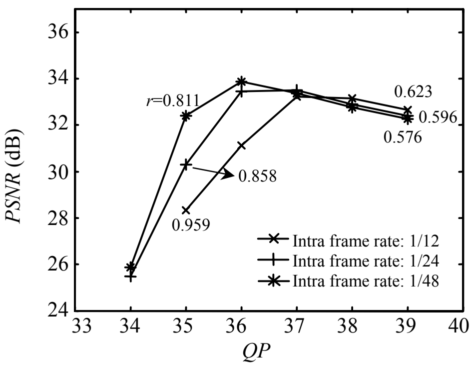

Fig.2 Achievable video quality under different operation points for a given channel condition (packet loss rate=10%)

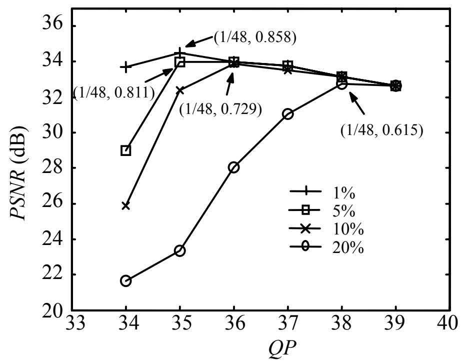

Fig.3 Achievable video quality under different operation points for different channel conditions

COMPARISON OF MULTICAST PERFORMANCE CRITERIA

In the multicast scenario, the same video signal is transmitted to multiple users by the AP. Each user has a different channel condition and receiving channel quality. The optimal operation point of source and channel coding for one user may not be optimal for other users. It is desirable to optimize some composite performance criteria for all the users of the same multicast session under the total rate constraint. However, the optimal operation point is dependant on the overall performance criterion. We will investigate the effect of two different criteria in determining the optimal operation point in multi-user case, and also evaluate the video quality variation among users (or for the same user under different channel conditions) under these multicast criteria.

In this paper, we employ and compare two performance criteria, i.e.,

(1) Weighted average criterion (Wang *et al*., 2004)

With this criterion, we maximize the weighted average of the video quality (in terms of PSNR) in all users in a multicast group. Mathematically, this can be written as:

$$Q_{\mathrm{opt}} = \max\left[\sum_{k=1}^{N} W(k) Q_k(S, P_{\mathrm{e},k})\right], \tag{4}$$

where N is the number of users in the multicast group, $Q_k(S, P_{\mathrm{e},k})$ is the individual video quality of each user, $W(k)$ is the weight function for user k, satisfying $\sum_{k=1}^{N} W(k) = 1$. The weight function $W(k)$ depends on the channel conditions of user k. One simple but practical form of $W(k)$ is

$$W(k)=\begin{cases} \dfrac{1}{N_{\mathrm{g}}}, & P_{\mathrm{e},k} \le P_{\mathrm{th}}, \\ 0, & P_{\mathrm{e},k} > P_{\mathrm{th}}, \end{cases} \tag{5}$$

where $P_{\mathrm{e},k}$ is the packet loss rate of user k, P_{th} is the threshold of packet loss rate, and N_{g} is the number of users with $P_{\mathrm{e},k} < P_{\mathrm{th}}$. This criterion averages the individual performance over the users with reasonable channel conditions and ignores the users with very bad channel conditions.

(2) Minimax degradation criterion

In this case we minimize the maximum performance degradation due to multicast among multiple users, following the minimax criterion proposed in (Qian and Jones, 2001). Different from that in (Qian and Jones, 2001), our criterion requires that a user must meet a minimum requirement for receiving channel condition if it is to be served. This prevents a user with a very bad channel condition to cause dramatic quality degradation at other users. Similar to the weight average criterion, we can use a weight $W(k)$ to achieve it. Our minimax degradation criterion is defined as follows:

$$Q_{\mathrm{opt}} = \min\{\max\{W(k)[Q_{\mathrm{opt},k}(P_{\mathrm{e},k}) - Q_k(S, P_{\mathrm{e},k})]\}\}, \tag{6}$$

where k=1,2,3,…,N, N is the total number of users in a multicast group; $Q_{\mathrm{opt},k}(P_{\mathrm{e},k})$ is the maximum video quality in terms of PSNR of the kth user obtainable with an operation point that is optimized for this user; and $Q_k(S,P_{\mathrm{e},k})$ is the actual received video quality for a chosen operation point for the entire multicast group. The weight for a user depends on its channel condition. Similar to $W(k)$ in Eq.(5), we can define:

$$W(k)=\begin{cases} 1, & P_{\mathrm{e},k} \le P_{\mathrm{th}}, \\ 0, & P_{\mathrm{e},k} > P_{\mathrm{th}}. \end{cases} \tag{7}$$

Given each user's individual channel condition, this criterion equalizes the degradation of video quality among all users from their individual optimal operation points.

In order to see the effects of these two criteria on optimal operation point selection, we use the similar simulation platform described in Section 3.1. In this experiment, a video stream is multicasted to 100 users, which every user experiencing one of the four different channel conditions A, B, C, D given in Section 3.1 in each 30 s period. For a new 30 s period, each user will be assigned to a new channel condition with probabilities P_a, P_b, P_c, P_d. The entire test time is 10 min. For this set of simulations, we assume the streaming server has perfect knowledge of the channel condition distributions during different time slots, and determines the optimal operation point at each time slot based on a chosen multicast performance criterion.

Here, we choose P_a, P_b, P_c and P_d to be 70%, 20%, 5% and 5% respectively.

To compare two criteria fairly, the threshold P_{th} in Eqs.(5) and (7) should be the same for both criteria. Without loss of generality, in our simulation we set threshold P_{th}=0.3, so that all users are to be considered.

Fig.4 plots the received video quality at a chosen time slot for all users using the selected optimal operation points under different criteria. x-axis is the index of users, and y-axis is the video *PSNR* for each user. The solid line represents the minimax degradation criterion and dotted line represents the weighted average criterion. From Fig.4, we can see that based on the minimax degradation criterion, the individual video qualities of different users intend to be consistent with each other no matter which channel conditions the users are experiencing. However, using the weighted average criterion, there is a much larger variance of video quality between different users. In the meantime, the average video *PSNR* among all users based on the weighted average criterion is higher than that based on minimax degradation criterion.

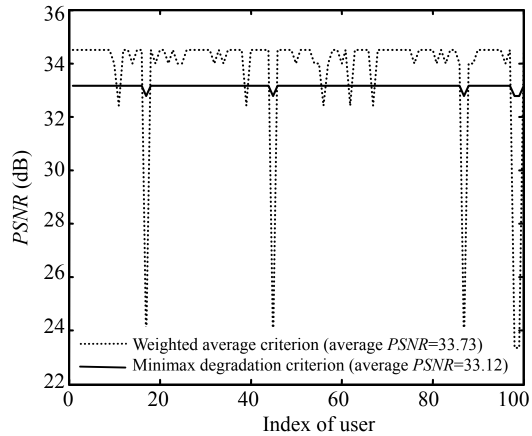

Fig.4 Effect of different criteria on video quality for different users in a particular time slot

As we mentioned, besides the heterogeneity of channel conditions among different users, the channel condition of an individual user is unstable. It is also interesting to see the variation of video quality for a particular user using different criteria.

Fig.5 plots the video quality for one particular user in different time slots. The individual video quality is more stable using the minimax degradation criterion than using the weighted average criterion. The stability of video quality is appealing for subjective video quality.

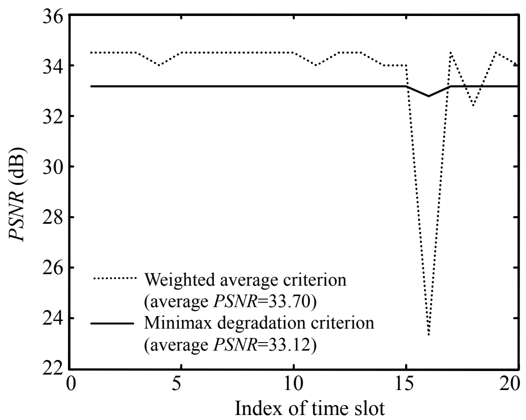

Fig.5 Effect of different criteria on video quality variation for a particular user

Thus, it can be concluded that the minimax degradation criterion can maintain a stable video quality for the same user as he/she moves around and can yield similar video quality among users with different channel conditions. However, this stability is achieved at the expense of a lower average quality. If the user topology is such that most users have very good channel conditions most of the time, then the weighted average criterion will make most users seeing better video most of the time than the minimax criterion. But occasionally, some users may see very bad video.

ADAPTATION OF SOURCE AND CHANNEL CODING CONFIGURATION BASED ON FEEDBACK OF ALL MULTICAST RECEIVERS

In wireless environment, channel condition for each user is not always stable. In order to dynamically make adaptation decisions at the transmission time, the packet loss rate at the receiver side should be known. This can be achieved by means of periodic feedbacks from receivers. The receivers predict their packet loss rates in next time slots based on their previous packet loss rates and send the feedbacks to the video streaming server. Based on the estimated channel conditions, the video streaming server determines the operation point for the next set of frames.

Finally, we propose an adaptive joint source and channel coding algorithm for video multicast over WLAN.

Prediction of channel condition based on feedback

Due to the heterogeneity and instability of channel conditions of receivers in the multicast group, it is desirable to dynamically make adaptation decisions at transmission time according to the most recent estimation of packet loss rate for each receiver. The receivers estimate their future packet loss rates based on the observed loss rates in the past, and send their estimates to the video streaming server by means of periodic feedbacks.

For prediction of the future loss rate at any receiver, we propose the following method, which keeps a running estimate of the average loss rate and the variance of packet loss rate. It can be formulated as follows:

$$\begin{aligned} D &= P_{\mathrm{m}}(t) - P_{\mathrm{a}}(t), \\ P_{\mathrm{a}}(t+1) &= P_{\mathrm{a}}(t) + aD, \\ P_{\mathrm{v}}(t+1) &= P_{\mathrm{v}}(t) + b(|D| - P_{\mathrm{v}}(t)), \\ P(t+1) &= P_{\mathrm{a}}(t+1) + cP_{\mathrm{v}}(t+1), \end{aligned} \tag{8}$$

where $P_{\mathrm{m}}(t)$ is the real packet loss rate in last time slot t, $P(t+1)$ is the estimated packet loss rate in time slot $t+1$, $P_{\mathrm{a}}(t)$ and $P_{\mathrm{v}}(t)$ are estimated average packet loss rate and variance of packet loss rate respectively, based on the observed loss rates up to time t, D is the difference between the real packet loss rate in time slot t and the estimated average packet loss rate, a, b are two numbers between 0 and 1, c is a non-negative number.

Parameters a, b are selected based on the dependence of channel conditions between two consecutive time slots. If the channel condition changes slowly, larger a and b are chosen, and vice versa. Note that $P_{\mathrm{a}}(t+1)$ is merely the expectation of packet loss rate in the next time slot. The actual loss rate may be higher or lower than this expected value. However, from the discussion in Section 3.2, the video quality is very sensitive to the residual packet loss. Underestimating the packet loss rate will lead to higher residual loss rate and induce dramatic degradation of video quality. Thus, we correct $P_{\mathrm{a}}(t+1)$ by adding a scaled version of $P_{\mathrm{v}}(t+1)$ to avoid underestimation. Obviously, this may overestimate the actual packet loss rate, and accordingly a larger *QP* rather than the actual optimal *QP* will be chosen. Since the packet loss has more impact than *QP* on the video quality, we would rather overestimate packet loss. Parameter c controls how conservatively the estimation is done to avoid underestimate.

Fig.6 plots the (a) real packet loss rate, (b) estimated packet loss rate, (c) estimated average packet loss rate and (d) estimated variance of packet loss rate in each time slot when the feedback time slot is 4 s and the parameters are set to be a=1/4, b=1/8, c=1. It can be seen that the estimated packet loss rate follows the same trend as the real packet loss rate and responds rapidly when real packet loss changes dramatically. Moreover, in most cases, the real packet loss rate is smaller than the corresponding estimated packet loss rate.

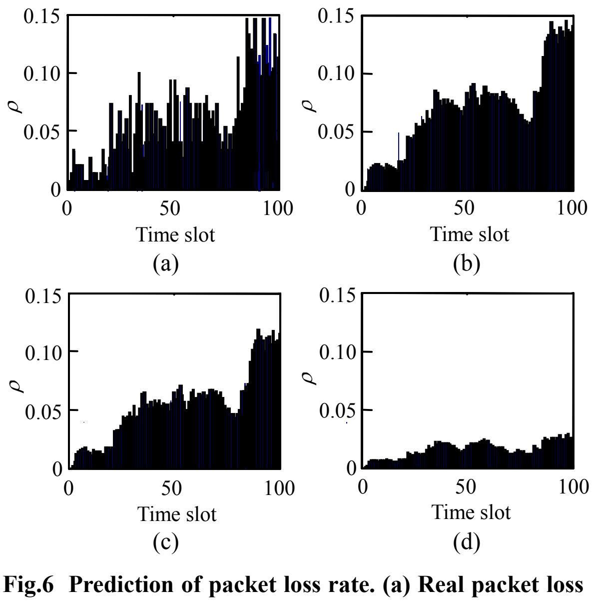

Fig.6 Prediction of packet loss rate. (a) Real packet loss rate; (b) Packet loss rate predicted; (c) Estimated average of packet loss rate; (d) Estimated variance of packet loss rate

It should be noticed that the proposed prediction approach is a more general case of the averaging packet loss prediction (Yajnik *et al*., 1999), which only considers the average of packet loss rate. The proposed prediction algorithm combines both the mean and the variance to do a conservative prediction. When c=0, the proposed prediction algorithm reduces to the averaging packet loss prediction (Fig.6c).

Adaptive joint source and channel coding for multicast video

Based on what has been concluded above, in this subsection we describe in detail the joint adaptation of source and channel coding scheme for multicasting video over WLAN.

In this scheme, based on the results obtained in Section 3, it is not necessary to add more intra-coded frames, so the minimum β is always used where there is only one intra-coded frame per GOF. Given QP and β, source rate R_s is determined, and all the surplus bandwidth $R_{tot}-R_s$ is allocated to FEC coding. Hence QP in the video source encoder is the only parameter to be tuned in order to achieve optimal overall performance. We assume the video source encoder or the video transcoder are capable of changing QP in real time based on the estimated packet loss rates of all receivers. We assume that the video quality curves corresponding to different source/channel coder operations points (similar to those in Fig.3) for a particular video under different channel conditions can be calculated beforehand based on simulations. Receivers need to sense their channel conditions, predict their future channel conditions and send feedbacks of their packet loss rates estimated to AP.

The joint adaptation of source and channel coding scheme could be achieved in the following procedure:

(1) Choose an appropriate intra frame rate and packet loss threshold P_{th} based on prior knowledge of typical user topology in the serving area under consideration.

(2) Based on varying user topology and channel conditions fed back by multiple receivers (each receiver estimates its packet loss rate in next time slot and sends feedback to the video streaming server), the system dynamically adapts QP and FEC rate to optimize a chosen multicast performance metric, based on the video quality curves achievable with different operations points for different possible channel conditions (which we assume can be estimated based on simulations in advance). The users with packet loss rate larger than P_{th} are not considered.

Simulation setting and result

In this subsection, we show the simulation results of the proposed adaptive joint source and channel coding scheme for video multicast over WLAN. Similar to the experiment set up in Section 4, the simulation is also based on the simulation platform illustrated in Subsection 3.1. A video stream is multicasted to 10 users, each of which experiences one of the four different channel conditions A, B, C, D illustrated in Subsection 3.1 in a 30 s period. For a new 30 s period, each user will be assigned to a new channel condition with probabilities P_a, P_b, P_c, P_d. The entire test time is 5 min. We consider 5 sets of P_a, P_b, P_c, P_d (Table 1), which represent 5 different overall channel conditions of the whole multicast group. We will investigate the performance of the proposed scheme in different overall channel conditions.

Table 1 User percentage of different channel conditions

Set	P_a	P_b	P_c	P_d
1	50%	30%	15%	5%
2	40%	30%	20%	10%
3	30%	30%	30%	10%
4	20%	30%	30%	20%
5	10%	20%	40%	30%

We compare the proposed scheme with two simpler schemes. Both of them use application layer FEC coding for error correction.

Scheme 1 QP is chosen arbitrarily and fixed during the entire test time.

Scheme 2 In this scheme, no feedback is employed. We assume the video streaming server can pre-estimate the channel condition distribution among all receivers over the entire streaming session duration and chooses the QP and FEC rate to optimize a given multicast performance. This scheme adapts to to a certain degree the channel conditions of receivers, but ignores the instability of channel conditions and changes of network topologies.

Fig.7 shows the performance comparison between different schemes. AJSC 1 is the proposed adaptive joint source and channel coding scheme based on the weighted average criterion. AJSC2 is that based on the minimax degradation criterion. AJSC3 is Scheme 2 which fixes the operation points according to the overall packet loss rate distribution of receivers without employing feedback. For this scheme, the weighted average criterion is used to derive the optimal operation points. We plot two more curves for fixed QP case (Scheme 1), QP=35 and QP=39. x-axis is the overall average packet loss rate of all users in the multicast group during the entire test time, and was obtained with different combinations of P_a, P_b, P_c, P_d (Table 1). *y*-axis is the overall average video quality of all users in the multicast group in terms of *PSNR*.

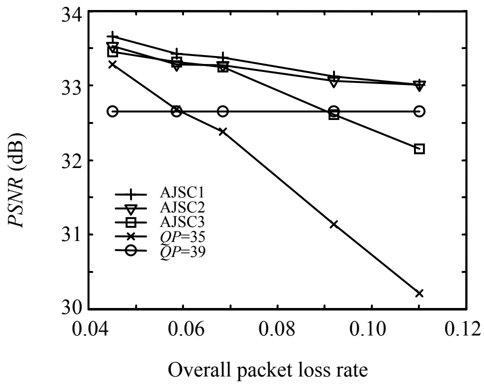

Fig.7 Comparison of different schemes in terms of average video quality

From Fig.7, we can see that AJSC1 provides the best overall performance in terms of average video quality. AJSC2 also outperforms other schemes except AJSC1. When overall channel condition is poor, the curves of AJSC1 and AJSC2 converge. This can be explained as: AJSC2 tends to assign more channel protection bits than AJSC1 when determining the optimal operation point, and the decision is less sensitive to channel conditions in order to generate more stable individual video quality. As overall channel condition worsens, for weighted average criterion, it also needs to assign more channel protection bits to combat the packet loss. Thus, when overall channel condition is bad, both of them choose similar optimal operation points, which result in similar performances. It is somewhat unexpected that AJSC3 performs as well as AJSC2 when the overall packet loss rate is up to 7%, and then becomes increasingly worse when the loss rate becomes higher. This result suggests that adaptation is more important when there are more users with poor channel conditions. Another interesting observation is that when *QP* is 39, the curve is constant. At this high *QP*, the source coding rate is low and the remaining bandwidth allocated to FEC can recover all packet loss no matter what channel condition a receiver experiences. Thus the overall video quality is the source coding video quality.

Fig.8 plots the deviation of video quality in the multicast group during the entire test time. The deviation considers both difference of video qualities between different users in the same time slot and the instability of video quality of individual user in different time slots. As we expected, the video quality is more stable when the minimax degradation criterion is adopted. Considering that the average PSNRs obtained with the two criteria are fairly close, the minimax degradation criterion seems to be a better choice.

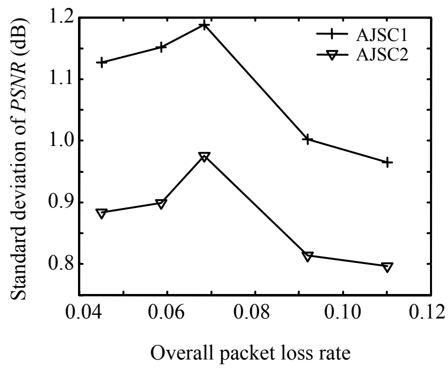

Fig.8 Comparison of different multicast performance criteria in terms of standard deviation of video quality

CONCLUSION

In this work, we investigate multicast of H.264 video over WLAN. We propose a joint source channel coding scheme that dynamically allocates the available bandwidth to the source coding and the FEC coding to optimize the overall system performance, by taking into account the user topology changes and varying channel conditions of multiple users. We jointly consider the error resilience in the source coder and error correction of FEC coding in channel coder, and investigate how to achieve the best performance in terms of a chosen multicast performance criterion. Two multicast performance criteria are proposed and compared. We found that the minimax degradation criterion can yield more constant video quality among all users and for the same user at different times. We also proposed a scheme for estimating the packet loss rate based on past observed packet loss rates and an adaptation scheme based on the feedbacks of the estimated packet loss rates at all receivers. We present simulation results to show that the joint optimization of the source and channel coder parameters for the overall channel condition distribution can provide substantial gains, and that adapting the operation points based on the instantaneous channel condition distribution can provide further improvements.

ACKNOWLEDGEMENT

We sincerely thank Zhenyu Wu of Thomson U.S. for fruitful discussions and careful reviews for this paper.

References

IEEE Standard 802.11, 1997. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.
Majumdar, A., Sachs, D.G., Kozintsev, I.V., Ramchandran, K., Yeung, M.M., 2002. Multicast and unicast real-time video streaming over wireless LANs. *IEEE Transactions on Circuits and Systems for Video Technology*, **12**(6): 524-534. [doi:10.1109/TCSVT.2002.800315]
Qian, L., Jones, D.L., 2001. Minimax disappointment criterion for video broadcasting. *Proc. Int. Conf. Image Processing*, **1**:449-452.
van der Schaar, M., Krishnamachari, S., Choi, S., Xu, X., 2003. Adaptive cross-layer protection strategies for robust scalable video transmission over 802.11 WLANs. *IEEE Journal on Selected Areas in Communications*, **21**(10): 1752-1763. [doi:10.1109/JSAC.2003.815231]
Wang, Y., Zhu, Q.F., 1998. Error control and concealment for video communications: A review. *Proc. IEEE*, **86**: 974-997.
Wang, Y., Wu, Z., Boyce, J.M., 2004. A Performance Measure for Video Multicast. 3GPP TSG System Aspects WG4, S4-AHVIC034.
Yajnik, M., Moon, S.B., Kurose, J.F., Towsley, D.F., 1999. Measurement and Modeling of the Temporal Dependence in Packet Loss. INFORECOM, **1**:345-352.