

# CROSS LAYER ADAPTATION FOR H.264 VIDEO MULTICASTING OVER WIRELESS LAN

*Zhengye Liu\**, *Hang Liu†*, *Yao Wang\**

\*Polytechnic University, Brooklyn, NY 11210, USA. Email: {zhengye,yao}@vision.poly.edu

†Corporate Research, Thomson Inc., Princeton, NJ 08540, USA. Email:hang.liu@thomson.net

## ABSTRACT

This paper describes cross-layer optimization strategies and simulation results for H.264 video multicast over wireless LAN. The proposed scheme takes into account the varying channel conditions of multiple users, and dynamically allocates available bandwidth between source coding and channel coding. In particular, source coding parameters (intra update and quantization) and application-layer FEC code rate are chosen jointly to optimize a multicast performance criterion, based on feedbacks from all multicast receivers. Two performance criteria for video multicast are investigated and compared.

## 1. INTRODUCTION

Video multicasting over wireless local area network (WLAN) enables the distribution of live or pre-recorded programs to many receivers efficiently. However, a challenging problem for video multicast over WLAN is that the underlying wireless channel is error prone due to fading and channel interference. Appropriate error protection mechanisms are required to guarantee satisfactory video quality for all multicast receivers. In video multicast, each user may have a different channel condition at any time and users may join or leave the service during a session so that the user topology also changes in time. The key issue is therefore to design a system to optimize an appropriate multicast performance metric that measures the overall user satisfaction over an entire multicast session.

In this paper, we investigate multicasting H.264 video over WLAN. The video servers are connected to the APs through a high-speed Ethernet LAN. Stored video contents (resp., live video contents) are transcoded (resp., real-time encoded), traffic-shaped by the video server and multicasted to a number of clients through WLAN.

Given the channel condition of each user, 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. Joint source and channel coding for video transmission has been extensively studied and reported in the literature, e.g. [1][2]. Herein, we focus on how to apply this general framework to the specific application of video multicast over WLAN. In order to allow joint adaptation of source

and channel codecs, we further propose to collect feedbacks regarding channel conditions from multicast receivers and dynamically update the source coder and channel coder configurations based on these feedbacks. Towards this goal, we have developed a channel estimation algorithm that is based on the average packet loss rate and the variance of packet loss rate at each receiver. Two overall performance criteria for video multicast and their effects on the video quality at individual receivers are investigated. We present simulation results to show that the proposed scheme improves the overall system performance.

The remainder of this paper is organized as follows. In section 2, we propose the adaptive joint source and channel coding scheme for video multicast over WLAN. In section 3, we show the simulation results. Section 4 concludes the paper.

## 2. ADAPTIVE JOINT SOURCE AND CHANNEL CODING SCHEME FOR VIDEO MULTICAST

### 2.1. Adaptive Joint Source and Channel Coding Scheme

When video is streamed over a lossy packet network, such as WLAN, the distortion of the decoded video at a receiver depends both on the quantization incurred at the encoder and the channel errors occurred during transmission and consequent error propagation in the decoded sequence. Smaller QP introduces lower source distortion with higher source rate, and vice versa. More periodically inserted intra-coded frames can limit transmission error propagation and hence reduce channel distortion, but it will also lead to higher source rate for almost the same source distortion. For channel coding, we consider the use of Reed-Solomon (RS) codes for application-layer cross-packet FEC. For a given target bit rate, higher source rate will reduce the rate allocated to FEC coding, hence channel distortion will increase. For a particular user with a given channel condition, there is an optimal operation point at which total distortion is minimal.

In the multicast scenario, 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. Furthermore, 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 estimations to the video streaming server by means of periodic feedbacks.

In our proposed system, each receiver estimates its packet loss rate in next time slot and sends this information to the video streaming server. Based on varying channel conditions fed back by multiple receivers, the system dynamically adapts QP, intra frame rate 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).

## 2.2. Multicast Performance Criteria

### 2.2.1. Weighted average criterion

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:

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

where  $N$  is the number of users in the multicast group,  $S$  is the triple-set of QP, intra frame rate and FEC rate,  $P_{e,k}$  is the packet loss rate of user  $k$ ,  $Q_k(S, P_{e,k})$  is the individual video quality of user  $k$ ,  $W(k)$  is the weight function for user  $k$ , satisfying

$$\sum_{k=1}^N W(k) = 1 \quad (2)$$

The weight  $W(k)$  depends on the channel condition of user  $k$ . One simple but practical form of  $W(k)$  is

$$W(k) = \begin{cases} 1/N_g & P_{e,k} \leq P_{th} \\ 0 & P_{e,k} > P_{th} \end{cases} \quad (3)$$

where  $P_{th}$  is the preset threshold, and  $N_g$  is the number of users with  $P_{e,k} \leq P_{th}$ . This criterion averages the individual performance over the users with reasonable channel conditions and ignores the users with very bad channel conditions.

### 2.2.2. Minimax degradation criterion

In this case we minimize the maximum performance degradation due to multicast among multiple users, following the minimax criterion proposed in [3]. Different from that in [3], 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 from causing dramatic quality degradation at other users. Similar

to the weighted average criterion, we can use a weight function  $W(k)$  to achieve it. Our minimax degradation criterion is defined as follows:

$$Q_{opt} = \min \{ \max [ W(k) (Q_{opt,k}(P_{e,k}) - Q_k(S, P_{e,k})) ] \} \quad (4)$$

where  $Q_{opt,k}(P_{e,k})$  is the maximum video quality in terms of PSNR of the  $k^{th}$  user obtainable with an operation point that is optimized for this user, and  $Q_k(S, P_{e,k})$  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 equation (3), we can define:

$$W(k) = \begin{cases} 1 & P_{e,k} \leq P_{th} \\ 0 & P_{e,k} > P_{th} \end{cases} \quad (5)$$

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

## 2.3. Prediction of Channel Condition Based on Feedback

For prediction of the future packet 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:

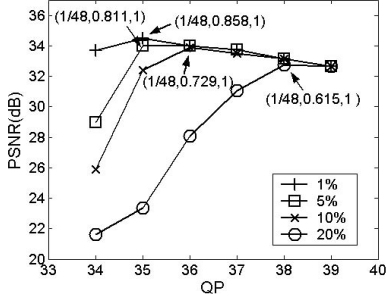
$$\begin{aligned} D &= P_m(t) - P_a(t) \\ P_a(t+1) &= P_a(t) + a * D \\ P_v(t+1) &= P_v(t) + b * (|D| - P_v(t)) \\ P(t+1) &= P_a(t+1) + c * P_v(t+1) \end{aligned} \quad (6)$$

where  $P_m(t)$  is the measured packet loss rate in last time slot  $t$ ,  $P(t+1)$  is the estimated packet loss rate in time slot  $t+1$ ,  $P_a(t)$  and  $P_v(t)$  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  $P_a(t+1)$  is merely the expectation of packet loss rate in next time slot. Since the video quality is very sensitive to the residual packet loss rate, we correct  $P_a(t+1)$  by adding a scaled version of  $P_v(t+1)$  to avoid underestimation. Parameter  $c$  controls how conservatively the estimation is done to avoid underestimation. Using this algorithm, the estimated packet loss rate follows the same trend as the real packet loss rate and has a fast response when real packet loss rate changes dramatically.

## 3. SIMULATION SETTINGS AND RESULTS

In our simulation, we code the "Kungfu" video sequence in SD (720X480) resolution using the latest JM9.6 H.264 codec. Each Group of Frames (GOF) has the duration of  $T=2$  seconds



**Fig. 1.** Video quality under different operation points for different channel conditions

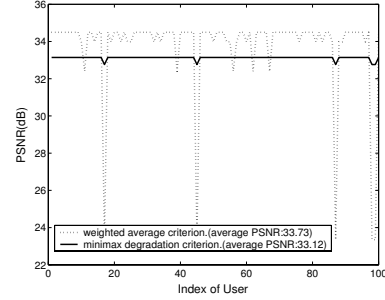
and comprises 48 frames. We encode the first 240 frames and loop the encoded video sequence 30 times to generate a 5-minute long video sequence. QP is changed from 34 to 39 and intra frame rate is changed from 4 I-frames within each GOF to 1 I-frame within each GOF. The corresponding source coding rate ranges from 599kbps to 366kbps. The target bandwidth is set to 600kbps. Given target bandwidth, QP and intra frame rate, all the left-over bandwidth besides source coding is allocated to FEC. The FEC block length is set to 2 seconds. Thus the delay introduced by FEC coding is 2 seconds. On the receiver side, the “motion copy” method available in the JM 9.6 H.264 decoder is used for error concealment. To simulate packet losses in WLAN which are usually bursty, a two-state Markov model characterized by the packet loss rate (PLR) and the average burst length (ABL) is used. To simulate the fluctuation of channel conditions, 4 different channel conditions are simulated 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).

### 3.1. Optimal Operation Point in Single-User Case

Figure 1 shows the maximum PSNR obtained at each QP value for different packet loss rates with the corresponding optimal intra frame rates, FEC rates and FEC decoding success rates. From the simulation, the correspondingly maximum PSNR is always achieved when intra frame rate is set to 1/48 for any tested QP and packet loss rate. Also, the QP should be chosen high enough so that there is enough FEC coding bits to guarantee all lost packets can be recovered. We see that, as expected, the optimal QP moves towards higher values when channel condition becomes worse.

### 3.2. Comparison of Multicast Performance Criteria Assuming Known Channel Conditions

In this experiment, a video stream is multicasted to 100 users, and every user experiences one of the four different channel conditions A, B, C, D in each 30 seconds period. For a new 30 seconds period, each user will be assigned to a new channel condition with probabilities  $P_a, P_b, P_c$  and  $P_d$ . For this set of



**Fig. 2.** Effect of different criteria on video quality for different users in a particular time slot.

simulations, we assume the streaming server has the perfect knowledge of the channel condition distributions at different time slots, and determines the optimal operation point at each time slot based on a chosen multicast performance criterion. Here, we choose  $P_a = 0.7$ ,  $P_b = 0.2$ ,  $P_c = 0.05$  and  $P_d = 0.05$ .  $P_{th}$  is set to 0.3 for both criteria.

Figure 2 plots the received video quality at a chosen time slot for all users under different criteria. We can see that based on the minimax degradation criterion, the individual video qualities of different users tend 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. At the meantime, the average video PSNR among all users based on the weighted average criterion is higher than that based on minimax degradation criterion.

Similar to the above result, the individual video quality is also more stable over the entire streaming time using the minimax degradation criterion than using the weighted average criterion.

### 3.3. Comparison of Different Adaptation Schemes

In this simulation, we consider 5 sets of  $P_a, P_b, P_c, P_d$ , which represent 5 different overall channel conditions of the whole multicast group. The feedback period of each user is set to 4 seconds. We will investigate the performance of the proposed feedback-based adaptation scheme under different performance criteria in different overall channel conditions. We compare the proposed scheme with some other simpler schemes.

Figure 3 shows the performance comparison between different schemes. AJSC1 is the proposed feedback-based adaptive joint source and channel coding scheme based on the weighted average criterion. AJSC2 is that based on the minimax degradation criterion. AJSC3 also uses the weighted average criterion, but no feedback is employed. We assume the video streaming server can pre-estimate the overall channel condition for each receiver over the entire streaming session duration and choose the QP, intra-coded frame rate and

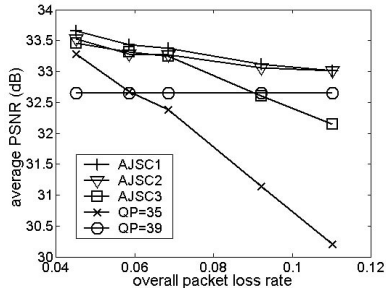


Fig. 3. Comparison of different schemes in terms of average video quality.

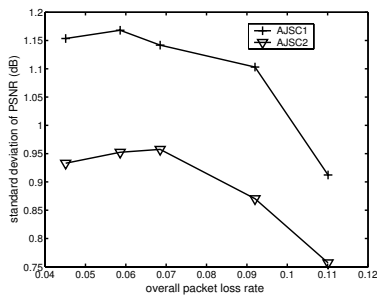


Fig. 4. Comparison of different multicast performance criteria in terms of standard deviation of video quality.

FEC rate to optimize a given multicast performance criterion. QP=35 and QP=39 are two simple non-adaptive schemes, where the QP and intra-frame rate are chosen arbitrarily and fixed during the entire test time. The intra frame rate are set to 1/48 in both cases. X-axis is the overall average packet loss rate of all users during the entire test time, obtained with different combinations of  $P_a$ ,  $P_b$ ,  $P_c$ ,  $P_d$ . Y-axis is the overall average video quality of all users in the multicast group in terms of PSNR. From Fig.3, we can see 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 become convergent. This can be explained as follows: AJSC2 tends to assign more channel protection bits when determining the optimal operation point than AJSC1, and the decision is less sensitive to channel conditions in order to generate more stable individual video quality. As overall channel condition becoming worse, 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. AJSC1 outperforms AJSC3 especially when the loss rate becomes higher. This result suggests that adaptation is more important when there are more users with poor channel conditions. We also see that the two

non-adaptive schemes perform substantially worse.

Figure 4 plots the standard 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 each user in different time slots. As we expected, the video quality is always more stable when the minimax degradation criterion is adopted. One reason that the points corresponding to the intermediate overall loss rates of around 0.06 and 0.07 have a higher variation in received quality is that the channel conditions among all the receivers corresponding to these points are more heterogeneous than other points. Considering the fact that the average PSNRs obtained with the two criteria are fairly close (with 0.25dB), the minimax degradation criterion seems to be a better choice.

#### 4. CONCLUSION

In this paper, 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 of varying channel conditions of multiple users. In this specific application, our simulation results show that FEC protection is more effective in reducing the channel distortion. Inserting intra frame on the other hand is not as beneficial. 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 mobile user at different times at only slight reduction of the average video quality (in terms of PSNR). 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.

#### 5. REFERENCES

- [1] R. E. Van Dyck, D. J. Miller, "Transport of wireless video using separate, concatenated, and joint source-channel coding," *Proceedings of the IEEE*, vol. 87, no. 10, pp. 1734–1750, October 1999.
- [2] Z. He, J. Cai, C-W. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate," *IEEE Transactions On Circuits And Systems For Video Technology*, vol. 12, no. 6, pp. 511–523, June 2002.
- [3] L. Qian, D. L. Jones, "Minimax disappointment criterion for video broadcasting," in *Proceedings of International Conference on Image Processing*, October 2001, pp. 449–452.