International Journal of Hybrid Information Technology
Vol. 9, No.11 (2016), pp. 1-12
http://dx.doi.org/10.14257/ijhit.2016.9.11.01

Content and Network Adaptive Cross-layer Optimization for Video Transmission

Zheng Wan

School of Information Technology,
Jiangxi University of Finance and Economics,
Nanchang, China, 330013
wanzheng97@163.com

Abstract

A cross-layer optimization video transmission scheme, which includes two mechanisms, is proposed in this paper. The first is the encoding adjustment mechanism to handle buffer insufficiency. Large frames are divided into several sub-frames and the maximum number of packets in these sub-frames depends on the minimum buffer size in forwarding path. In addition, sending time of adjacent frames or sub-frames should be separated. And time intervals between adjacent frames or sub-frames are determined by the numbers of packets those frames or sub-frames have. The second is the priority based transmission mechanism to handle bandwidth insufficiency. Different dropping probabilities or scheduling opportunities are assigned to different frame types. Less important frames always have a high dropping probability or a small scheduling opportunity. In performance evaluation, various video sequences, buffer sizes and bandwidth reservations are employed. Simulation results verify that: (1) Cross-layer optimization is adaptive to different video sequences, buffer sizes and bandwidth reservations. That is to say, it is a content and network adaptive scheme. (2) Proposed scheme outperforms existing schemes significantly, especially when buffer size is small or bandwidth reservation is large. (3) A further advantage is that priority based transmission mechanism enhances video decoding quality while receives less packets. Thus it is friendly to other streams.

Keywords*: Video Transmission; Cross-layer; Content Adaptive; Network Adaptive*

1. Introduction

With the development of network and video compressing technologies, more and more video are streamed in all kinds of networks. To ensure transmission quality of video streams, massive enhancements were proposed in the last decade. Firstly, studies focused on a single layer in the network. For example, many application layer [1-3] and link layer [4, 5] schemes were proposed. Since video streams often have a great deal of data to be sent and the network conditions are not steady, researchers found that improvement of single layer schemes are limited. Thus they paid more attentions to cross-layer schemes, which can be classified into several categories:

(1) Network adaptive video encoding strategies [6,7,8], which will adjust encoding based on the network condition parameters returned by network layer or link layer.

(2) MAC centric optimization [9,10,11], including “application layer + MAC” and “application layer + MAC + physical layer”. Key issue is adaptive frame scheduling in MAC.

(3) “Application layer + physical layer” schemes [12,13,14], perform cross-layer optimization according to wireless channels.

(4) Non-centric optimizations [15,16], which often include more than two layers.

ISSN: 1738-9968 IJHIT

Although most existed schemes have acceptable performance, they are complicated too. In this paper we design a simple cross-layer optimization scheme which is adaptable to various video sequences and network conditions.

The rest of the paper is organized as follows. Motivation for proposing this cross-lay optimization scheme is described in Section 2. Details of our proposed scheme are presented in Section 3. Section 4 discusses simulation settings, including simulation environments, video sequences and comparative schemes. Section 5 gives simulation results and corresponding discussions. Finally, Section 6 concludes the paper.

2. Motivation

In the past, we performed lots of experiments to evaluate and improve video transmission in DiffServ (Differentiated Services) domain. Several video sequences were streamed in various network conditions and under different enhanced transmission schemes. Results of one experiment drew our attentions. The experiment and its results are described as follows.

Figure 1 presents the simulation topology. In the network, S1 generates a video stream and S2 produces a CBR data stream with a data rate of 1.5Mbps (R_d). Packet size of data steam is 1500 bytes. Bandwidth of each link is marked in the figure. Buffer sizes of RED queues in all routers are 50 packets.

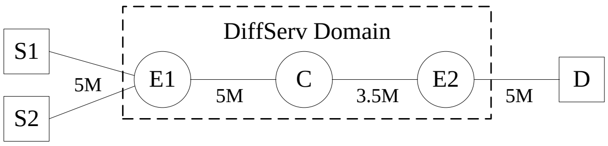


Figure 1. Simulation Topology

Foreman with CIF resolution is adopted as streaming video. Table 1 gives data rates of all half-second cycles. The maximum size of video packet is 1024 bytes. Consequently, foreman sequence has 3077 packets and a maximum average PSNR of 41.014316.

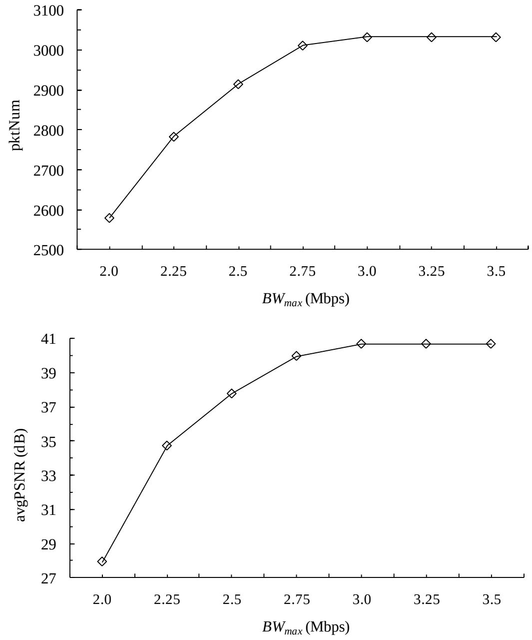
Table 1. Data Rate at Each Cycle/Half-Second of Foreman (Bps)

cycle	rate	cycle	rate	cycle	rate	cycle	rate
1	1947504	6	1932688	11	2494064	16	2948336
2	2262688	7	2266352	12	1851184	17	3352688
3	1669312	8	1755728	13	2685648	18	2953632
4	2134224	9	1767152	14	2254640	19	2943776
5	2071664	10	2473040	15	2480128	20	3721792

Receiving number of packets (pktNum) and resulting average PSNR (avgPSNR) under various bandwidth reservations (BW_{max}) for video stream at link(C,E2) are calculated and shown in Figure 2.

From the figure we find that receiving video quality gets better as BW_{max} increases. However, no enhancement emerges after BW_{max} exceeds 3.0Mbps, even if BW_{max} reaches 3.5Mbps which should ensure perfect receiving of foreman sequence (buffer in the RED queue could hold exceeded packets in the last cycle).

To find the reason, we monitored the numbers of lost packets in all links of the forwarding path and found that packet loss occurs in both link(S1,E1) and link(C,E2). Figure 3 shows pktNum of both links. Results tell that the number of lost packets in link(C,E2) decreases as more bandwidth is reserved. However, there are always 44 packets dropped in the first link -- link(S1,E1). Recall that the bandwidth of this link is 5Mbps. How does this happen?



(a) pktNum result (b) avgPSNR result

Figure 2. PKTNUM and AVGPSNR Results of Foreman

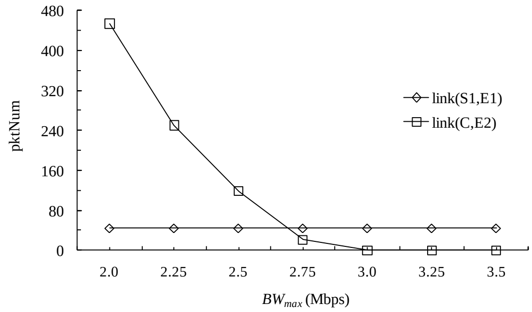


Figure 3. The Numbers of Lost Packets In Link (S1,E1) and Link(C,E2)

Look inside the video encoding which is partly shown in Table 2, we can find the exact reason. Taking frames 233, 234 and 235 as example, these frames are sent synchronously so that buffer of the sender cannot hold all packets, no matter how much bandwidth the link has. Similar situation occurs in the other continuous frames listed in Table 2.

Table 2. Packet Numbers and Sending Time of Certain Frames

frame	num	time	frame	num	time	frame	num	time
233	43	7.834	260	45	8734	287	45	9.635
234	6	7.834	261	4	8.734	288	7	9.635
235	6	7.834	262	8	8.734	289	6	9.635
242	44	8.135	269	45	9.045	296	45	9.936
243	6	8.135	270	5	9.045	297	8	9.936
244	6	8.135	271	3	9.045	298	7	9.936
251	46	8.436	278	45	9.334			
252	5	8.436	279	3	9.334			
253	3	8.436	280	3	9.334			

3. Cross-Layer Optimization

From the above analysis we know that buffer size and bandwidth reservation are key factors which influence video streaming. Packet losses are always caused by insufficient buffer and/or bandwidth. To guarantee the performance of video streaming, we must find ways to adapt to the absence of these resources. In this section, our proposed cross-layer optimization scheme is discussed in details.

3.1. Mechanism Answering to Insufficient Buffer

In order to handle buffer insufficiency, two encoding adjustments could be considered:

(1) The first task is to divide those “big” frames into several sub-frames. Let’s take frame 233 in Table 2 as example. Since there are 43 packets in this frame, we’d better divide it into several parts. pktNum of sub-frames should be smaller than buffer size. For example, if the buffer can only hold 15 packets, we must divide frame 233 into at least 3 sub-frames: the first two sub-frames have 15 packets each and the last sub-frame has 13 packets.

(1) The second task is to avoid synchronously sending of multiple frames. Taking the above example again, we must assign different sending time for frame 234/235 and three sub-frames of frame 233. Such assignment should consider pktNum of adjacent frames and available time intervals among these frames. The more packets a frame or sub-frame has, the larger time interval should be given.

Table 3 presents the sending time assignment of frame 233/234/235/236.

Table 3. Sending Time Assignment Example

frame	pktNum	time	frame	pktNum	time	frame	pktNum	time
233-1	15	7.834	234	6	7.912	236-2	7	7.963
233-2	15	7.861	235	6	7.924	…	…	…
233-3	13	7.888	236-1	15	7.936			

3.2. Mechanism Answering to Insufficient Bandwidth

When encountering bandwidth insufficiency, packet loss is inevitable. Under this situation, the best way is to drop the least important packets. For example, MPGE4 encoder has three frame types, i.e. I/P/B frames. Among these frame types, I frames are the most important and B frames are the least important. We should first recognize the frame type of a video packet, and then assign appropriate dropping probability or scheduling opportunity to it.

For example, if the video is streamed in DiffServ domain, we can use different traffic classes (physical queues) to hold I/P/B frame packets separately, or establish several virtual queues within one physical queue to hold them.

4. Simulation Settings

Simulations are based on the integrated platform of ns-2 [17] and Evalvid [18], implemented by C. H. Ke [19]. Experiments are performed in NS-2 DiffServ module.

4.1. Simulation Environments

NS-2 DiffServ module supports four traffic classes, each of which holds three dropping priorities (virtual queues). Six policy models are defined and Null policy (uses only one dropping priority) is adopted for traditional video streaming mechanism. In addition, we construct a new policy model to implement the mechanism mentioned in section 3.2. In this policy, frame type of a video packet is recognized and corresponding virtual queue is assigned. Different virtual RED queues have distinct parameters, including the lower and the higher queue length thresholds, and the dropping probability.

Simulation topology is similar to that described in Figure 1 except for link bandwidths. To avoid packet loss caused by bandwidth limitation in other links except for link(C,E2), bandwidths of other links are all set to 10Mbps. Bandwidth of link(C,E2) depends on the streaming video.

4.2. Video Sequences

To perform a content adaptive evaluation, three video sequences are chosen: foreman, coastguard and container with CIF resolution. These sequences have different data rates and bursts. Data rates at each half-second, average data rates (avg) and standard deviations of these sequences are shown in Table 4. Total bytes of different frame types are given in Table 5. The frame rates of all sequences are 30/s.

Table 4. Data Rates at Each Half-Second of Three Sequences (Bps)

No.	foreman	coastguard	container	No.	foreman	coastguard	container
1	1947504	4028720	1601952	11	2494064	3516048	1646672
2	2262688	4093744	1650208	12	1851184	3255968	1282000
3	1669312	4201248	1300832	13	2685648	3708336	1621344
4	2134224	4966032	1637904	14	2254640	3133648	1634000
5	2071664	4273392	1664592	15	2480128	3006784	1284384
6	1932688	3996656	1320480	16	2948336	3081296	1599232
7	2266352	4780848	1681616	17	3352688	3176208	1642544
8	1755728	3980688	1663024	18	2953632	2901616	1338304
9	1767152	3565296	1311392	19	2943776	3048400	1623008
10	2473040	3869488	1634752	20	3721792	3444256	1712912
avg	2398312	3701434	1542558				
stdev	560539	591374	161084				

Table 5. Total Bytes of I/P/B Frames of Three Sequences (Bytes)

type	foreman	coastguard	container
I	1094951	1270042	1067748
P	1318951	2521507	686307
B	583988	835243	174142

4.3. Comparative Schemes

To verify the performance of proposed cross-layer optimization scheme, we introduce three comparative schemes as follows:

(1) Traditional scheme (denoted as “Neither”): no encoding or transmission enhancement mechanism in this scheme.

(2) Priority based video transmission scheme (denoted as “PriTrans”): employing transmission enhancement mechanism described in section 3.2.

(3) Encoding adjustment scheme (denoted as “ReCode”): using the mechanism introduced in section 3.1.

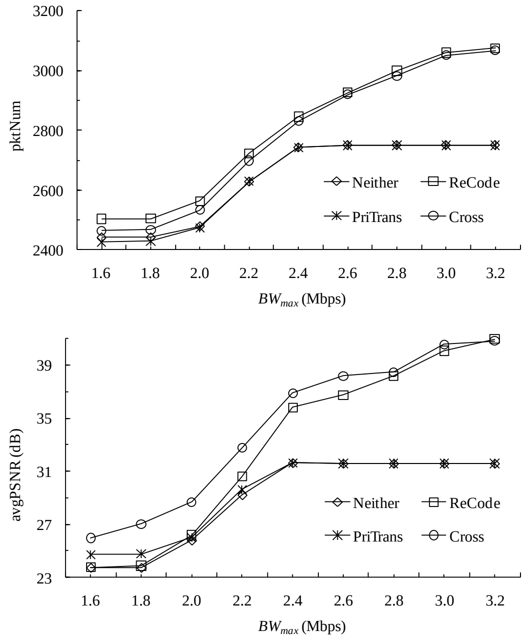
In the next section, the proposed scheme is denoted as “Cross”. Parameters values are the same as those in “PriTrans” and “ReCode”.

5. Performance Evaluation

5.1. Results under Various Bandwidth Reservations

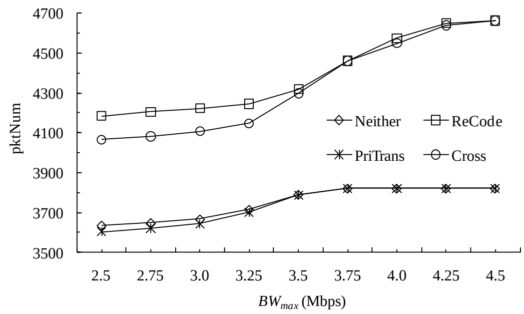
Buffer size in the following experiments is 30 packets. Parameters of RED queue for I frames in "PriTrans" and "Cross" schemes are "30-30-0", in which the first and the second "30" mean the lower and the higher queue length threshold values respectively, and "0" means the dropping probability value. Since the buffer size is 30, "30-30-0" means that do not drop I frame packets actively. Parameters of RED queues for P and B frames are "15-20-0.25" and "10-15-0.5" respectively. In "ReCode", the maximum pktNum of a sub-frame is 15.

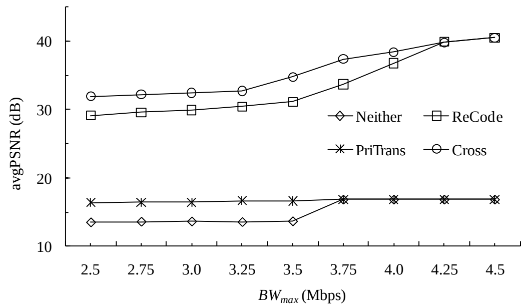
Data rate of CBR data stream in each experiment is set to 1.5Mbps. And $BW_{C,E2}$ is set to an appropriate value based on the streaming video to produce various congestion conditions. Figure 4 to Figure 6 present the simulation results.



(a) pktNum result (b) avgPSNR result

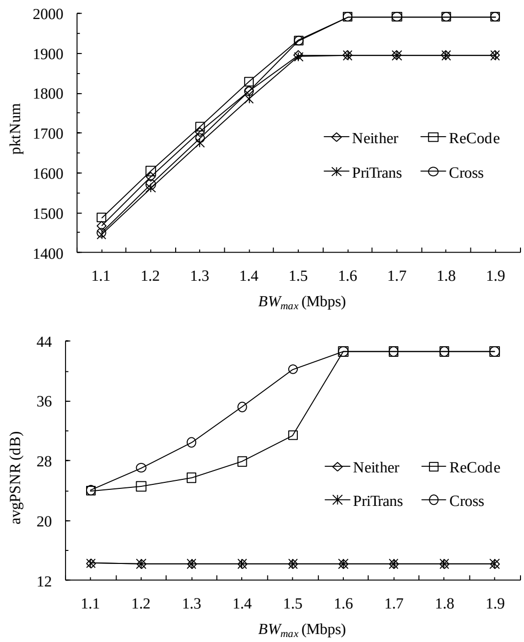
Figure 4. Results of Foreman Sequence, $BW_{C,E2}$=3.5Mbps





(a) pktNum result (b) avgPSNR result

Figure 5. Results of Coastguard Sequence, $BW_{C,E2}$=5Mbps



(a) pktNum result (b) avgPSNR result

Figure 6. Results of Container Sequence, $BW_{C,E2}$=2.5Mbps

From these figures we can draw the following conclusions.

(1) Proposed cross-layer optimization scheme always shows the best performance and adaptability to various videos and different bandwidth reservations.

(2) Encoding adjustment mechanism improves video receiving quality significantly. We notice that pktNum and avgPSNR of “ReCode” and “Cross” are much better than those of “PriTrans” and “Neither”. Furthermore, if encoding adjustment mechanism is not employed, certain number of packets will be dropped for insufficient buffer, no matter how many bandwidths are reserved. Consequently, pktNum and avgPSNR both reach their bottlenecks as BW_{max} increases continuously.

(3) Comparing “PriTrans”/“Neither” and “Cross”/“ReCode”, we find that less packets are received when priority based transmission mechanism is adopted. However, avgPSNR increases on the contrary. Although more packets are dropped by RED actively, transmission of important packets is guaranteed. Employing this mechanism, more bandwidth could be reserved for other video or data streams. Consequently, the integral performance of all streams in the network is improved.

(4) Finally, let’s focus on container sequence. We find that avgPSNR is fairly low if encoding does not adjust. From Table 5 we know that I frame have much more bytes than

the other two frame types. However, each I frame in container sequence has more than 30 packets. That is to say, all I frames cannot be decoded. With regard to the other two sequences, P and B frames could improve decoding quality.

5.2. Results under Various Buffer Sizes

In the following experiments, fixed bandwidth reservation for video stream at link(C,E2) and different buffer sizes are employed. Figure 7 and Figure 9 show the results.

From these figures we can draw the conclusions:

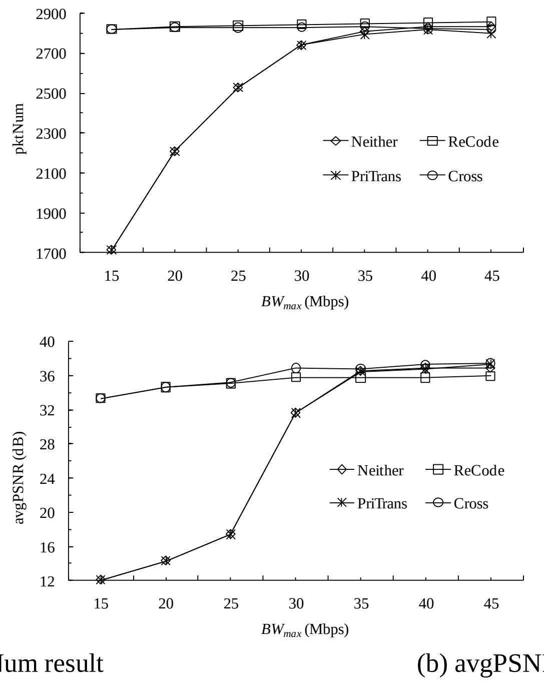
(1) Encoding adjustment mechanism (“ReCode” and “Cross”) leads to remarkable performance improvement, especially when buffer size is relatively small. The reason is straight. Dividing into sub-frames which have at least 15 packets and separating their sending time, help the buffer hold more packets.

(2) Performance of Encoding adjustment mechanism (“ReCode” and “Cross”) shows its adaptability to various buffer sizes. From the figures we find that pktNum and avgPSRN are fairly steady as buffer size varies, especially when streaming videos are foreman and coastguard sequences.

(3) As for the performance of priority based transmission mechanism, the results are similar to those in section 5.1. That is, fewer packets received but higher avgPSNR achieved.

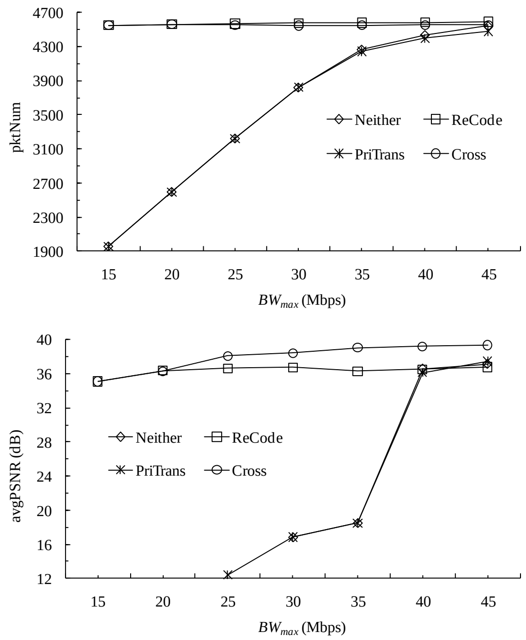
(4) In some cases (buffer size is 15 or 20 in “Neither” and “PriTrans” schemes when coastguard sequence is streamed), pktNum is too small to complete decoding.

(5) There is an “accident” when container sequence is streamed. Proposed “Cross” does not outperform “PriTrans” when buffer size is equal to or larger than 35. We’ll find the reason in the future work.



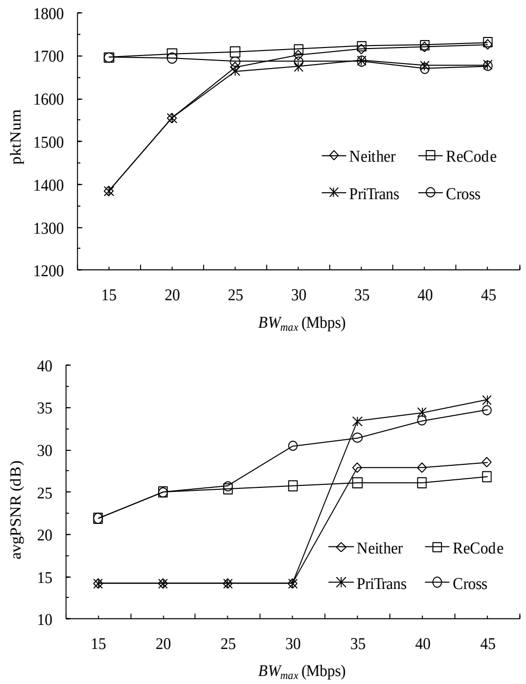
(a) pktNum result (b) avgPSNR result

Figure 7. Results of Foreman Sequence, BW_{max} =2.4Mbps



(a) pktNum result (b) avgPSNR result

Figure 8. Results of Coastguard Sequence, *BW_{max}* =4Mbps



(a) pktNum result (b) avgPSNR result

Figure 9. Results of Container Sequence, *BW_{max}*=1.3Mbps

6. Conclusions

In this paper we propose a cross-layer optimization scheme for video transmission, including two separate mechanisms. One is encoding adjustment mechanism, answering to buffer size insufficiency. The other one is priority based transmission mechanism, answering to bandwidth insufficiency. Simulation results show that:

(1) Proposed scheme shows its adaptability to content (different video sequences) and network (various buffer sizes and bandwidth reservations). That is to say, the performance of proposed scheme keeps steady when streaming different videos or employing different network parameters.

(2) Proposed scheme outperforms existing schemes greatly. First, encoding adjustment mechanism improve receiving video quality significant when buffer size is small or bandwidth reservation is large. Second, priority based transmission mechanism enhances video decoding quality by receiving less packets, leaving more bandwidth for other streams.

Acknowledgements

This work was supported by National Natural Science Foundation of China (No. 61162009), Natural Science Foundation of Jiangxi Province (No. 20142BAB217004), Science and Technology Project of Jiangxi Education Department (No. GJJ12273), and Visiting Scholar Special Fund of Young and Middle-Aged Teacher Developing Plan of Jiangxi Ordinary Universities.

References

[1] N. Kamnoonwatana, D. Agrafiotis and C. N. Canagarajah, "Flexible Adaptive Multiple Description Coding for Video Transmission", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 1, **(2012)**, pp. 1-11.

[2] G. J. Sullivan, J. Ohm, W.-J. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, **(2012)**, pp. 1649-1668.

[3] L. Zhang, Q. Peng, X. Wu and Q. Wang, "SSIM-Based error resilient video coding over packet-switched networks", Journal of Signal Processing Systems, vol. 74, no. 1, **(2014)**, pp. 103-113.

[4] J. Pokhrel, I. Paudel, B. Wehbi and A. Cavalli, "Performance Evaluation of Video transmission over 802.11 n wireless network: A MAC layer perspective", 2014 International Conference on Smart Communications in Network Technologies (SaCoNeT), **(2014)**, pp. 1-6.

[5] H. Abouzeid and H. S. Hassanein, "Efficient lookahead resource allocation for stored video delivery in multi-cell networks", 2014 IEEE Wireless Communications and Networking Conference (WCNC), **(2014)**, pp. 1909-1914.

[6] W.-S. Lim, D.-W. Kim and Y.-J. Suh, "Design of Efficient Multicast Protocol for IEEE 802.11n WLANs and Cross-Layer Optimization for Scalable Video Streaming", IEEE Transactions on Mobile Computing, vol. 11, no. 5, **(2012)**, pp. 780-792.

[7] M. G. Martini, M. Mazzotti, C. Lamy-Bergot, J. Huusko and P. Amon, "Content Adaptive Network Aware Joint Optimization of Wireless Video Transmission", IEEE Communications Magazine, vol. 45, no. 1, (**2007)**, pp. 84-90.

[8] Yongju Cho, S. Karande, K. Misra, H. Radha, Jeongju Yoo, Jinwoo Hong. Channel Capacity Estimation and Prediction for Rate-Adaptive Wireless Video. IEEE Transactions on Multimedia, **2008**, 10(7):1419-1426.

[9] Adlen Ksentini, Mohamed Naimi, Abdelhak Guéroui. Toward an Improvement of H.264 Video Transmission over IEEE 802.11e through a Cross-layer Architecture. IEEE Communications Magazine, **2006**, 44(1): 107-114.

[10] Chuan Heng Foh, Yu Zhang, Zefeng Ni, Jianfei Cai, King Ngi Ngan. Optimized Cross-Layer Design for Scalable Video Transmission over the IEEE 802.11e Networks. IEEE Transactions on Circuits and Systems for Video Technology, **2007**, 17(12): 1665-1678.

[11] Zheng Wan, Naixue Xiong, Nasir Ghani, Athanasios V. Vasilakos, Liang Zhou. Adaptive Unequal Protection for Wireless Video Transmission over IEEE 802.11e Networks. Multimedia Tools and Applications, **2014**, 72(1):541-571.

[12] R. Hormis, E. Linzer and X. Wang, "Adaptive Mode- and Diversity-Control for Video Transmission on MIMO Wireless Channels", IEEE Transactions on Signal Processing, vol. 57, no. 9, **(2009)**, pp. 3624-3637.

[13] M. M. Ghandi, B. Barmada, E. V. Jones and M. Ghanbari, "Wireless Video Transmission Using Feedback-Controlled Adaptive H.264 Source and Channel Coding", IET Communications, vol. 3, no. 2, (**2009)**, pp. 172-184.

[14] A. Alinejad, N. Y. Philip and R. S. H. Istepanian, "Cross-Layer Ultrasound Video Streaming over Mobile WiMAX and HSUPA Networks", IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 1, **(2012)**, pp. 31-39.

[15] Y. Ju, Z. Lu, D. Ling, Xi. Wen, W. Zheng and W. Ma, “QoE-based cross-layer design for video applications over LTE”, Multimedia Tools and Applications, vol. 72, no. 2, **(2014)**, pp. 1093-1113.
[16] R. Radhakrishnan and A. Nayak, “Cross layer design for efficient video streaming over LTE using scalable video coding”, 2012 IEEE International Conference on Communications (ICC) , **(2012)**, pp. 6509-6513.
[17] The Network Simulator - NS (version 2), http://www.isi.edu/nsnam/ns/.
[18] J. Klaue, B. Rathke and A. Wolisz, “EvalVid-A Framework for Video Transmission and Quality Evaluation”, In Proceedings of 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation, **(2003)**, pp. 255-272.
[19] C. H. Ke, C. K. Shieh, W. S. Hwang and A. Ziviani, “An Evaluation Framework for More Realistic Simulations of MPEG Video Transmission”, Journal of Information Science and Engineering, vol. 24, no. 2, **(2008)**, pp. 425-440.

Author

Zheng Wan, he received the B.E. and Ph.D. degree in computer science in 2001 and 2006 from Zhejiang University of China. He is now an associated professor of Jiangxi University of Finance and Economics, Nanchang, China. He is currently working in the areas of wireless video communications, wireless networks, Quality of Service and assessment of video transmission quality.