IPB-frame Adaptive Mapping Mechanism for Video Transmission over IEEE 802.11e WLANs

Xin-Wei Yao
College of Computer Science & Technology
Zhejiang University of Technology
Hangzhou, P.R.China
xwyao@zjut.edu.cn

Wan-Liang Wang
College of Computer Science & Technology
Zhejiang University of Technology
Hangzhou, P.R.China
wwl@zjut.edu.cn

Shuang-Hua Yang
Department of Computer Science
Loughborough University
Leicestershire, UK
S.H.Yang@lboro.ac.uk

Yue-Feng Cen
College of Computer Science & Technology
Zhejiang University of Technology
Hangzhou, P.R.China
cyf.zjut@hotmail.com

Xiao-Min Yao
College of Computer Science & Technology
Zhejiang University of Technology
Hangzhou, P.R.China
yaoxm@zjut.edu.cn

Tie-Qiang Pan
College of Electrical and Information Engineering
Quzhou University
Quzhou, P.R.China
ptq@qzu.edu.cn

ABSTRACT
This paper proposed an IPB-frame Adaptive Mapping Mechanism (AMM) to improve the video transmission quality over IEEE 802.11e Wireless Local Area Networks (WLANs). Based on the frame structure of hierarchical coding technology, the probability of each frame allocated to the most appropriate Access Category (AC) was dynamically updated according to its importance and traffic load of each AC. Simulation results showed the superior performance of the proposed AMM by comparing with three other existing mechanisms in terms of three objective metrics.

Categories and Subject Descriptors
C.2.1 [Network Architecture and Design]: Wireless communications; C.2.2 [Multimedia Information Systems]: Video

General Terms
Algorithms

Keywords
WLANs, 802.11e, Video Transmission, Mapping Mechanism, Hierarchical Video Coding

1. INTRODUCTION
With the development of wireless technologies (WiFi, 3G, 4G, WiMAX and Bluetooth), wireless multimedia transmission over WLANs has been gained more attention in the recent years, especially for video transmission [1], by means of video conference, video surveillance, on-line video streams and games, etc. To support the Quality of Service (QoS) requirement of video transmission, IEEE 802.11e Enhanced Distributed Channel Access (EDCA) [2] mechanism was designed by introducing four ACs, and classifying all data flows to different ACs according to their priorities. Four ACs were defined as AC(3), AC(2), AC(1) and AC(0) from the highest priority to the lowest priority. Each AC had its own value set of EDCA parameters, the higher the priority, the more the opportunity to transmit data. However, most researches just focused on the adjustments of EDCA parameters to improve the video transmission without considering the video contents and coding technologies [3][4][5].

For example, all video frames were allocated to the second priority queue AC(2) according to the original EDCA mechanism and its extensions. When massive video frames were instantly allocated into AC(2), the limited queue buffer space would be used up quickly, leading to queue-level congestion or queue overflow, which would degrade the quality of the delivered video streams. Furthermore, network resource was not fully utilized for other queues which might be idle or light traffic load, especially in an unsaturated case. In order to solve these deficiencies, some researches had been focused on the mapping mechanisms for hierarchical encoded video frames over a EDCA-based network [6][7][8], such as Static Mapping Mechanism (SMM) [9][10][11], Dynamic Mapping Mechanism (DMM) [12][13], Dynamic Frame Assignment Algorithm (DFAA)[14]. SMM and DMM were downward mapping mechanisms (i.e. video frames were mapped to the queues with a lower priority), which introduced unnecessary transmission delay and high packet loss if low priority queues were almost full of packets at the same time. DFAA took the priority of video frame and queue length of each AC as inputs to differentiate frames, and used fuzzy logic controller to produce adjustment of DFAA parameter, but it ignored the coding structures of different video streams.

Therefore, in order to improve the quality of transmitted video streams and the utilization of network bandwidth, it is desirable to design an effective mapping mechanism to apportion video frames to the most appropriate queue based on the integrated information of video coding structure, frame importance and traffic load of each queue together. In this paper, a novel simple and efficient IPB-frame adaptive mapping mechanism is proposed to support video transmission with high quality as shown in Fig.1, which has the function-
Fig. 1: The structure: GOP(12, 3)

Fig. 2: IPB-frame Adaptive Mapping Mechanism

Figures 1 and 2 illustrate the hierarchical video coding technology and existing related works. Section 3 describes the proposed IPB-frame AMM. Its performance is evaluated by comparing with EDCA, SMM and DMM in Section 4. Finally, we conclude this paper in Section 5.

2. BACKGROUND

2.1 Hierarchical video coding technology

The development of hardware and video coding technologies [15] has made video transmission over wireless network possible and efficient, and dramatically increasing video streams are becoming the main traffic in wireless networks. For example, the latest widely-used international hierarchical video coding technologies H.264 and MPEG-4 provide an excellent video quality even over wireless networks with low bit rate and do not increase the complexity of coding algorithm.

According to the hierarchical coding technology, there are three major frame types designed for compressing video streams: I-frame, P-frame and B-frame. They are different in the following characteristics: 1) I-frames are the least compressible and don’t require the information of other video frames to be decoded; 2) P-frames use the information of previous frames (I-frame and P-frame) to decompress and are more compressible than I-frames; 3) B-frames use both previous and forward frames as reference data to get the highest amount of data compression. In the process of MPEG-4 and H.264 coding, the whole video sequence is decomposed into a set of smaller units, known as Group Of Pictures (GOP). The GOP is a group of consecutive pictures, and begins with an I-frame. The structure of GOP $G(N, M)$ is referred to two parameters, where $N$ indicates the distance between two consecutive I-frames and $M$ indicates the distance between I-frame and P-frame. For example, the structure of G (12,3) is BBBBBBBBBBB as shown in Fig.2.

In detail, from Fig.2, it is observed that I-frame is encoded independently and decoded by itself, P-frame is encoded or decoded by using the information from preceding I-frame or P-frame in the same GOP, and B-frame needs the information of preceding and succeeding I-frames or P-frames to be encoded or decoded correctly. If the I-frame in a GOP cannot be decoded due to packet loss or time delay, then all the video frames in this GOP will be undecodable and useless, which will dramatically deteriorate the video transmission quality. Therefore, the importance of these three

frames from the highest to the least is: $I > P > B$. However, the original EDCA mechanism unfortunately neglects the significance of different video frames, which distributes all video frames to AC(2) with the same priority, and the coding structure is also neglected by other existing mapping mechanisms.

2.2 Existing Related Works

To support the service differentiation for various data flows, the original EDCA mechanism specified that all video frames were allocated to AC(2). Due to neglecting the significance of different video frames, this mapping mechanism could not guarantee the transmission of the most important video frames. On the other hand, without considering the traffic load of each AC, the queue congestion might occur under massive video frames and the network resource was not fully utilized under varying traffic load.

By taking the importance of video frames into consideration, static mapping mechanisms (SMM)[9][10] were proposed to allocate video frames into different priority queues as shown in Fig.3 (top), thereby reducing the impact of traffic congestion on video quality [11]. In detail, they were designed by introducing the following rules: all I-frames were allocated to AC(2), while all P-frames were allocated to AC(1) and B-frames were allocated to the lowest priority AC(0). However, these kind of SMM could not be adaptive to time-varying network traffic load, especially when the traffic load of lower priority AC(1) and AC(0) were very heavy. Moreover, when the network load was light, the video frames mapped to lower priority ACs might result in unnecessary transmission delay and packet loss.

According to the deficiencies of SMM, dynamic mapping mechanisms (DMM)[15][14] for improving the video transmission were proposed as shown in Fig.3 (bottom). The packets from different video frames were dynamically mapped to the lower priority ACs by introducing different mapping probabilities, which were updated based on the video frame importance and the network traffic load. Similar to SMM, these DMM were also downward mapping mechanisms [16], i.e., all video frames were mapped to the lower priority ACs. When the traffic load of the highest priority AC(3) was light or even empty, extra transmission delay and packet loss might be introduced by these downward mapping mechanisms, while the lower priority queues were congested. The queue resource could not be fully utilized as well. Moreover, the mapping parameters of each frame were predefined without considering the video coding structure.
3. IPB-FRAME ADAPTIVE MAPPING MECHANISM

3.1 Description of IPB-frame AMM

Contest-based EDCA mechanism provides smaller values of competitive parameters to higher priority AC to guarantee its QoS requirement. Therefore, AC(3) with the highest priority has more opportunities and less delay to access the channel than other three ACs. Real-time video streams also have the requirements of high bandwidth and low end-to-end delay. Thus, to fully utilize the network resources, IPB-frame AMM aims to build the relationship between video frames and AC(3), named forward mapping process, but only focuses on the two most important frame types (I-frame and P-frame) in order to simplify the mapping mechanism and reduce the computation load. Furthermore, due to the limited buffer size of each queue, congestion should be prevented when any new I-frames and P-frames are injected into AC(3) and should not interfere with the transmission of its original traffic. According to the above analysis, the probability that a I-frame is mapped to AC(3) in AMM is defined as:

\[ P_{I \rightarrow AC(3)} = \max\left\{ \frac{\text{qlen}(AC(2)) \times \text{threshold} - \text{qlen}(AC(3))}{\text{max}\ AC(2)} , 0 \right\} \]  

(1)

where \( \text{max}\ AC(2) \) indicates the maximum queue length (i.e. buffer size) of AC(2), \( \text{qlen}(AC(2)) \) is the function which can calculate the current queue length of AC(2) at any moment, and \( \text{threshold} \) is a constant representing queue congestion level. The value of parameter \( \text{threshold} \) is the same for all ACs, and is independent on the video coding structure. When the queue length of a AC exceeds the value of \( \text{threshold} \), this AC is going to be congested. When the current queue length of AC(3), \( \text{qlen}(AC(3)) \), is bigger than the constant \( \text{threshold} \), the probability \( P_{I \rightarrow AC(3)} \) is equal to zero, i.e. no more I-frame will be assigned to AC(3) to alleviate the congestion. In this paper, the values of \( \text{max}\ AC(2) \) and \( \text{threshold} \) are 50 and 40, respectively. Without loss of generality, when the current queue length is bigger than 80% of the maximum queue length, this queue is going to be congested, thus the value of \( \text{threshold} \) is set as 40, but which can be changed according to the particular requirements of different applications.

From the coding structure of video transmission unit GOP (N, M) shown in Fig.2, each GOP includes only one I-frame, \( \frac{N-M}{M} \) P-frames and \( N \cdot \frac{M-1}{M} \) B-frames. We assume that I-frame and all P-frames have equal opportunities to be assigned to AC(3), then the probability of a P-frame being mapped into AC(3) can be given as:

\[ P_{P \rightarrow AC(3)} = \frac{M}{N-M} \times P_{I \rightarrow AC(3)} \]  

(2)

On the other hand, when the highest priority queue AC(3) is extremely busy with transmitting data, no more video frame should be mapped into AC(3). At the same time, AC(2) is not capable to host all the video frames due to its limited queue buffer. Some frames will be discarded randomly due to queue congestion or buffer overflow. Under this condition, some of B-frames and P-frames have to be assigned to AC(1), named downward mapping process. We assume that all B-frames and all P-frames have equal opportunities to be assigned to AC(1), then the probabilities of B-frame, P-frame being assigned to AC(1) can be calculated through the following equations:

\[ P_{P \rightarrow AC(1)} = \max\left\{ \frac{\text{qlen}(AC(2)) \times \text{threshold} - \text{qlen}(AC(1))}{\text{max}\ AC(2)} , 0 \right\} \]  

(3)

\[ P_{B \rightarrow AC(1)} = \frac{N-M}{M} \times \frac{M}{N \cdot (M-1)} \times P_{P \rightarrow AC(1)} \]  

(4)

3.2 Realization of IPB-frame AMM

Based on the above obtained mapping probabilities, IPB-frame AMM can be realized as shown in Fig.4. When a node receives a request to transmit a GOP, the “Frame Recognition” module first checks the priority of each frame (i.e. frame type), because each frame type has its own mapping rules:

1) If the arrival frame belongs to the group of I-frames, due to its importance for the whole GOP encoding and decoding, it will be priority assigned to AC(3) or AC(2). The final mapping result is determined by comparison of the mapping probability \( P_{I \rightarrow AC(3)} \) and a random value generated from the function Rand (0, 1).

2) If the arrival frame belongs to the group of P-frames, its mapping rules are much more complicated than the rules

}\]
of I-frame. First, it will calculate the mapping probability $P_{ac} \rightarrow AC(3)$ to check whether it has the opportunity to be assigned to $AC(3)$. If not, it will be assigned to $AC(2)$ until the queue length of $AC(2)$ is bigger than the constant threshold. Then it will be assigned to lower priority queue $AC(1)$ when the probability $P_{ac} \rightarrow AC(1)$ is bigger than a random value as shown in Fig.3. Otherwise, in order to guarantee the transmission of other more important frames (e.g. I-frame) and avoid congestion, it has to be discarded.

3) If the arrival frame belongs to the group of B-frames, for its least importance in GOP, it has the opportunity to be assigned to $AC(2)$ only when the current queue length of $AC(2)$ is less than the constant threshold, i.e. the traffic load is much lighter. When a large amount of frames congregate in $AC(2)$, B-frame tries to access $AC(1)$, and it will succeed if the probability $P_{ac} \rightarrow AC(1)$ is bigger than a random value. Otherwise, it has to be discarded.

According to the proposed AMM, each video frame will be mapped to the most appropriate $AC$, and each $AC$ has its own transmission probability and collision probability. Based on the Markov Chain model, the packet loss probability and transmission delay of each video frame can be calculated according to the work [17]. Then the enhancement of video quality with the proposed AMM is obtained.

4. PERFORMANCE ANALYSIS

A simulation study based on NS-2.34 [18] was conducted to evaluate the performance of the proposed IPB-frame AMM. We first introduced the tested video source and three objective criteria. The proposed mechanism was tested by using publicly available video sequences [19], such as “Foreman” sequence, “News” sequence and “Stefan” sequence. The selected video sequences (“Foreman” and “News”) were in YUV QCIF (176*144) format, and the sequence (“Stefan”) has higher resolution in YUV CIF (352*288) format. Three objective criteria, DFR (Decoded Frame Rate) [17], PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural SIMilarity index) [20] were employed to provide quantitative and objective evaluations. Each video frame was fragmented into packets before transmission, and the maximum transmission packet size over the simulated network was 1024 bytes. Moreover, the transmission was in unicast mode. There were ten wireless nodes in the network topology where one was the video transmitter and another was the video receiver, other nodes were transmitting background flows. The data rate of voice flows was 64 kbps, the data rate of UDP flows was 25 kbps and the data rate of the wireless link was 2Mbps. The video transmission unit was GOP(12,3). All experiment results in this paper were obtained under the assumption of the value of play-out buffer at the receiver is 0.008s (it was very small and approximate to a real-time application). The parameters of IEEE 802.11e MAC-layer and PHY-layer used in simulations were listed in Table 1.

In order to evaluate the quality of video transmission with the proposed AMM, light and heavy traffic conditions were adopted in Test 1 and Test 2 respectively. In Test 1, only video streams were transmitted, there were no background data flows, i.e. light traffic load condition case, the observed results of four simultaneously transmitted video streams were shown in Figs. 5, 6 and 7. In order to distinguish different video sequences, video index in Figs. 5, 6 and 7 indicated the corresponding video ID, e.g. video index 1 meant the first transmitted video sequence, video index 2 meant the second transmitted video sequence, the similar to the video index 3 and 4. In Test 2, mixed traffic load was adopted, we introduced three voice flows (allocated to $AC(3)$), three video flows, three TCP flows (allocated to $AC(1)$) and three UDP flows (allocated to $AC(0)$). In detail, three video streams were currently transmitted between the video transmitter and video receiver, i.e. heavy video traffic condition, and its results were listed in Table 2. All video sequences were transmitted from the video sender to the same video receiver.

Fig.5 showed the DFR of each delivered video by using four different mapping mechanisms. It was observed that IPB-frame AMM had better video transmission performance than others. Due to the limited network bandwidth, the DFR of each video decreased with the increment of its video index. In detail, IPB-frame AMM got an advantage of average 6% over DMM, average 9% over SMM, and average 6% over EDCA.

Table 1: IEEE 802.11e parameters used in simulations

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHY Header</td>
<td>Physical Layer Header</td>
<td>192 bits</td>
</tr>
<tr>
<td>MAC Header</td>
<td>MAC Layer Header</td>
<td>272 bits</td>
</tr>
<tr>
<td>ACK Frame</td>
<td>Acknowledgement Frame</td>
<td>304 bits</td>
</tr>
<tr>
<td>RTS Frame</td>
<td>Request To Sent</td>
<td>352 bits</td>
</tr>
<tr>
<td>CTS Frame</td>
<td>Clear To Sent</td>
<td>304 bits</td>
</tr>
<tr>
<td>Payload</td>
<td>data Payload</td>
<td>8000 bits</td>
</tr>
<tr>
<td>DataRate</td>
<td>sending Data Rate</td>
<td>54 Mbps</td>
</tr>
<tr>
<td>Time slot</td>
<td>Time Slot</td>
<td>20 μs</td>
</tr>
<tr>
<td>SIFS</td>
<td>Short Inter-Frame Space</td>
<td>{2, 2, 3, 7}</td>
</tr>
<tr>
<td>AIFS</td>
<td>Arbitrary Inter Frame Space</td>
<td></td>
</tr>
<tr>
<td>CW&lt;sub&gt;min&lt;/sub&gt;</td>
<td>Minimum Contention Window size</td>
<td>{7, 15, 31, 31}</td>
</tr>
<tr>
<td>CW&lt;sub&gt;max&lt;/sub&gt;</td>
<td>Maximum Contention Window size</td>
<td>{15, 31, 1023, 1023}</td>
</tr>
</tbody>
</table>

Fig.6 showed the average PSNR of each video transmitted over WLANs with different mapping mechanisms. The results shown in Fig.6 demonstrated that the IPB-frame AMM achieved the superior performance, which got average 1dB better than DMM, 0.5dB better than SMM, and 1.2dB better than EDCA.

Fig.7 showed the average SSIM of each video transmitted...
with four mapping mechanisms. Since SSIM was the structure similarity index of video, therefore the higher value of SSIM, the better video quality. The SSIM of IPB-frame AMM was about 4% bigger than DMM, 1.3% bigger than SMM, 4% bigger than EDCA.

In Test 1, only four video streams were transmitted simultaneously in the wireless network. According to the observation of above three figures, it was found that: 1) For EDCA mechanism, video transmission quality declined quickly with the increase of video index due to all video frames were allocated to AC(2), which led to disastrous queue congestion when more video frames were mapped to the almost congested queue AC(2), such as video index 4; 2) For SMM mechanisms, they had a poor performance for light video traffic due to all P-frames and B-frames were mapped to low priority queues, especially for small number of video streams, such as video index 1 and 2. However, for big index value of video streams, they had a relative good performance as a result of all I-frames were mapped to AC(2), such as video index 4; 2) For SMM mechanisms, they had a poor performance for light video traffic due to all P-frames and B-frames were mapped to low priority queues, especially for small number of video streams, such as video index 1 and 2. However, for big index value of video streams, they had a relative good performance as a result of all I-frames were mapped to AC(2), such as video index 4; 3) By comparing with DMM and AMM, although mapping probabilities of video frames were both adopted in these two mechanisms, AMM achieved better performance than DMM due to only downward mapping mechanism used by DMM.

In Test 2, under the heavy traffic condition, Table 2 showed the detailed frame/packet loss, PSNR and DFR under four mapping mechanisms. In detail, the results listed in Table 2 were obtained by using the video sequence “foreman”, the adopted video codec was the widely-used “ffmpeg” codec. Through video coding, the foreman sequence had 480 frames, so the DFR in Table 2 was computed as n/400. The proposed AMM had the superior performance through the comparison with three other existing mechanisms on each criteria. In detail, queue-level congestion or overflow was decreased or even avoided effectively as a result of considering the traffic load of each queue in the proposed adaptive mapping mechanism, and more video frames, especially the most important I-frames, were guaranteed to be transmitted successfully with adaptive mapping mechanism. Therefore, the packet loss ratio was reduced more than 50% compared with other three mechanisms, the value of DFR was increased almost 10% than others.

In order to reconstruct the transmitted video at the receiver, an error concealment strategy was required. The main task of error concealment was to replace missing parts of the video content by previously decoded parts of the video sequence in order to eliminate or reduce the visual effects of errors caused by corrupted areas in the decoded frame. Error concealment strategy in the H.264 video decoder exploited the spatial and temporal correlations between the neighboring image parts within the same frame or from the past and future frames. Figs. 8, 9 and 10 illustrated the snapshot of reconstructed video of “Foreman” sequence, “News” sequence and “Stefan” sequence at the receiver, demonstrating that the perceived video quality of the proposed adaptive mapping algorithm was better than that obtained with the dynamic mapping scheme, where the part of frame in red rectangle was distorted during transmission. In detail, video sequence “Stefan” presents a highly irregular motion, the proposed AMM could guarantee the video frame transmission by considering the video coding structure, frame importance and queue traffic level, and then improved its delivered video quality.

5. CONCLUSION

In this paper, an IPB-frame adaptive mapping mecha-
Table 2: The performance of video transmission under different mapping mechanisms (Test 2)

6. ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their insightful comments and constructive suggestions that improved the quality of this paper significantly. This work was supported by the National Natural Science Foundation of China (Grant No. 61070043, 61371923) and National Science & Technology Pillar Program during the Twelfth Five-year Plan Period (2012BAD10B01).

7. REFERENCES


