Congestion Resiliency of Data Partitioned H.264/AVC Video over Wireless Networks

Ismail Ali[1], Sandro Moiron[1,2], Martin Fleury[1], and Mohammed Ghanbari[1]

[1]University of Essex, Colchester CO4 3SQ, United Kingdom
[2]Instituto de Telecomunicações, Portugal
{iaali,smoiro,fleum,ghan}@essex.ac.uk

Abstract. Data partitioning is an error resiliency technique that provides unequal error protection for transmission over lossy channels. Including a per-picture, cyclic intra-refresh macroblock (MB) line guards against temporal error propagation. This paper considers the impact of this form of error resilience on access control and proposes a selective dropping scheme based on dropping the partition bearing intra-coded MBs. The paper shows that by this scheme, when congestion occurs, it is possible to gain up to 3dB in video quality prior to transmission over assigning a stream to a single IEEE 802.11e access category. The scheme is shown to be consistently advantageous over other ways of assigning the partitioned data packets to different access categories. This counter-intuitive scheme for access control purposes reverses the priority usually given to partition B data packets over partition C data packets.

Key words: data-partitioning, H.264/AVC, IEEE 802.11e, intra-refresh macroblocks.

1 Introduction

The proven attraction of video services such as the BBC's iPlayer and YouTube [1] means that video streaming to the IEEE 802.11 LANs available at wireless hotspots will continue to grow. Currently, YouTube uses progressive download, a form of pseudo-streaming associated with Adobe's Flash Player but the stop-go nature of this form of streaming implies that true streaming will emerge in future services. As the H.264/AVC (Advanced Video Coding) encoder can achieve up to 50% increase in compression ratio [2] compared to an MPEG-2 encoder, it is particularly appropriate for bandwidth-limited WLANs, especially if they are operating in multi-hop ad hoc mode, which is why H.264/AVC has already been adopted for 3GPP's Multimedia Broadcast Multicast Service [3].

As part of its wireless-friendly approach [4], H.264/AVC introduced a set of error resilience tools [5] to improve robustness to packet loss in transmission over error-prone networks. Among these tools were data partitioning and distributed intra-refresh macroblocks (MBs). Data partitioning is a way of separating out data from a compressed bit-stream according to its importance in reconstructing a video stream. H.264/AVC Network Abstraction Layer units (NALUs) are

per-slice containers for packet transmission. Data partitioning of NALUs into partition A, B, and C types [6] of decreasing importance for decoder reconstruction purposes allows the more important data to be given preferential treatment. To protect the stream against temporal error propagation it is becoming common to distribute intra-refresh MBs across the video pictures [7] rather than provide periodic intra refresh, as periodic refresh causes a sudden spike in the data rate. As a consequence, partition B will become enlarged as it is this partition that normally contains naturally intra coded MBs. With or without an enlarged partition B, protecting partitions A, B, C in an unequal error protection scheme seems sensible. For example, in [8] a combination of hierarchical modulation and forward error protection was applied.

However, when accessing the network in the face of traffic congestion either self-congestion or congestion from other sources can change the priorities. In this paper, we show that when cyclic intra-refresh lines are applied [7] then partition B packets become a better candidate to drop at network access time than partition C. That is, despite their higher reconstruction priority, the order of priority is changed when congestion occurs compared to the degree of protection applied during transmission. IEEE 802.11e [9] governing priority-controlled access to a WLAN is applied to illustrate the proposed dropping scheme for data-partitioned, cyclic intra-refresh provisioned video streaming.

In [10], the authors proposed a cross-layer architecture to improve H264/AVC video delivery over IEEE 802.11e. Each data partition of the compressed bitstream was mapped to a different IEEE 802.11e access category (AC), allowing the prioritization of packets bearing the more important partition A and B data. However, their scheme is aimed at transmission issues, especially reduction of end-to-end delay and minimization of packet loss rate, rather than congestion during access to the network. Furthermore, unlike our scheme partition B data is favored over partition C. In [11], a similar scheme was compared to wireless transmission with a multiple description coding (MDC) scheme. If partition A data could be well protected then data-partitioned data transmission results in over 10 dB improvement in quality over the MDC scheme but otherwise MDC is preferable. However, again this paper's primary concern was transmission issues. In a similar vein, in [12] measurements of an IEEE 802.11e data-partitioned scheme indicated that using a single IEEE 802.11e AC was preferable to splitting the H.264 data-partitioned layers across the IEEE 802.11e ACs. However, in the experiments partition B and C packets were assigned to the same IEEE 802.11e AC. Conventional periodic intra-refresh was applied, with the Instantaneous Decode Refresh (IDR)-pictures being given the same higher priority as partition A packets.

This paper's contribution is a congestion-resilient scheme for data-partitioned video streams protected by intra-refresh MBs. In these circumstances, partition B is generally larger than would otherwise be the case (see Section 2.2). However, the data in partition B that potentially helps to mitigate channel errors, intra-coded MBs, is given lower priority at network access time in the proposed scheme. Because the area occupied by this data corresponds to a comparatively limited

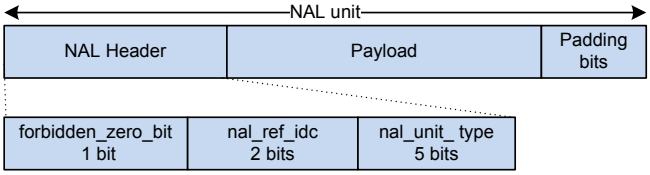

Fig. 1. NAL unit format.

part of a video frame, the impact of dropping packets formed from this partition is reduced. The paper supplies generalized confirmation of this counter-intuitive dropping policy as well as an IEEE 802.11e case study.

The remainder of this paper is organized as follows: Section II introduces data partitioning and intra refresh (IR) line error resilience techniques. The impact of introducing IR line on the relative size of data partition B is also analyzed, then a practical implementation of the proposed scheme is presented. Section III presents the framework used to assess the video quality under uniform and network drop conditions. The Section goes on to present results for a random packet drop model, where drops occur according to a uniform distribution, and a generalized multi-hop wireless network congestion scenario. Finally in Section IV, some concluding remarks are made and the future development of the proposed scheme is outlined.

2 Methodology

2.1 Data Partitioning

The H.264/AVC codec standard conceptually separates the Video Coding Layer (VCL) from the NAL. The VCL specifies the core compression features, while the NAL supports delivery over various types of network. This feature of the standard facilitates easier packetization and improved video delivery. The NAL facilitates the delivery of the H.264 VCL data to the underlying transport layers such as RTP/IP, H.32X and MPEG-2 transport systems. Each NAL unit could be considered as a packet that contains a header and a payload. The 8-bit header, Figure 1, specifies the NALU payload type (nal_unit_type) and the relative importance of NALU (nal_ref_idc) while the payload contains the related data. The payload padding bits are added to make the payload a multiple of bytes. Table 1 is a summarized list of different NALU types. NALUs 6 to 12 are non-VCL units containing additional information such as parameter sets and supplemental information.

In H.264/AVC, when data-partitioning is enabled, every slice is divided into three separate partitions and each partition is located in one of the type 2 to type 4 NALUs, as listed in Table 1. A NALU of type 2, also known as partition A, comprises the most important information of the compressed video bit-stream of predictively-coded P and optional bi-predictively-coded B-picture, including

Table 1. NAL unit types.

NAL unit type	Class	Content of NALU
0	-	Unspecified
1	VLC	Coded slice
2	VLC	Coded slice partition A
3	VLC	Coded slice partition B
4	VLC	Coded slice partition C
5	VLC	Coded slice of an IDR picture
6-12	Non-VLC	Suppl. info., Parameter sets, etc.
13-23	-	Reserved
24-31	-	Unspecified

Fig. 2. Cyclic intra-refresh MB line technique for the *Paris* test sequence, showing successive MB lines in lighter shading, with some slice boundaries also shown.

the MB addresses, motion vectors and essential headers. If any MBs in these pictures are intra-coded, their Discrete Cosine Transform (DCT) coefficients are packed into the type-3 NALU, also known as partition B. Type 4 NAL, also known as partition C, carries the DCT coefficients of the motion-compensated inter-picture coded MBs.

2.2 Intra Refresh MBs

The scheme exploits the periodic insertion of intra-refresh MB lines in a cyclic pattern within successive temporally predicted video pictures. The objective of inserting intra-refresh MB lines [7] is to mitigate error propagation at the cost of lower coding efficiency than purely predictive inter coding. Using a vertical (or horizontal) sliding intra-refresh line, Figure 2, reduces the error drift arising from packet loss without the instantaneous rise in the data rate when periodic intra-refresh pictures are sent. Provided a cyclic pattern of lines are transmitted the sequence is completely refreshed after each cycle, just as it is after the insertion of an IDR or I-picture at the start of each Group-of-Pictures. However, it should be carefully noted that an intra-coded MB line within a temporally predicted frame represents a significant percentage of the bits devoted to compressing the whole frame. Nonetheless, a packet containing data from a line of intra-coded MBs represents a small portion of the image area. Therefore, only a small potential quality penalty arises from the loss of packets containing intra refresh

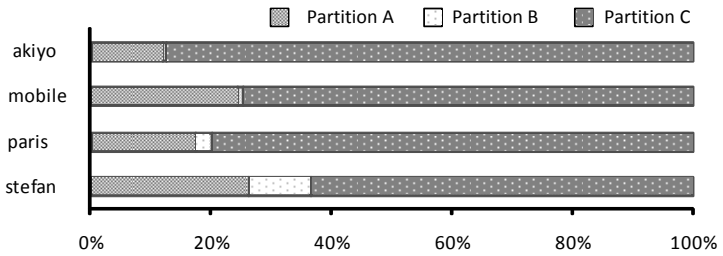

Fig. 3. Percentage of data for partitions A, B and C in a variety of sequences when no intra-refresh line is used.

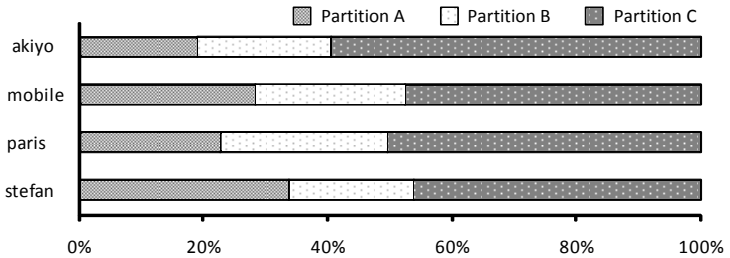

Fig. 4. Percentage of data for partitions A, B and C in a variety of sequences when an intra-refresh line is used.

MBs due to the small image area affected. The scheme described herein exploits this feature and proposes to selectively discard the intra-refresh MB line data contained in the B-partition in preference to dropping data belonging to partition C packets. This is of course the residual information contained in the quantized DCT coefficients.

Although each picture contains three data partitions, A, B and C, each data partition may be split into several packets due to network packetization constraints, which impose a maximum number of bytes per packet. The result is a variable number of packets per frame, the number of which depends on the content complexity. As a slice contains resynchronization markers to enable entropy decoder to be re-initialized, it is advisable to configure the codec itself to restrict the packet size rather than allow segmentation at the network-level.

Figure 3 shows the percentage of data occupied by each type of partition for a selection of well-known test videos in Common Intermediate Format (CIF) (352×288 pixel/picture), with a CBR target bitrate of 1Mbps. The data from partition C dominates, while in some sequences partition B is hardly used. When intra-coded MBs are included in partition B, it is because they have been naturally encoded, that is if the encoder cannot form a suitable prediction that reduces the difference signal sufficiently. This is an implementation dependent decision as only the decoder is standardized in the standard codecs. In Figure 4, an intra-refresh MB line has been included. The growth in the size of partition B data is the main feature of this Figure. Figure 5, shows the resulting packet

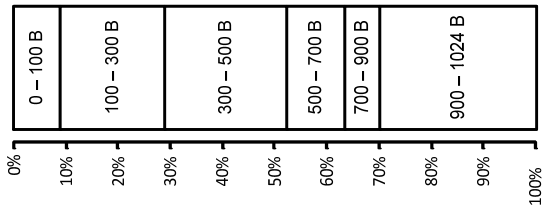

Fig. 5. Percentage packet size distribution when using data partitioning and a packet size limit of 1 kB for the *Paris* sequence.

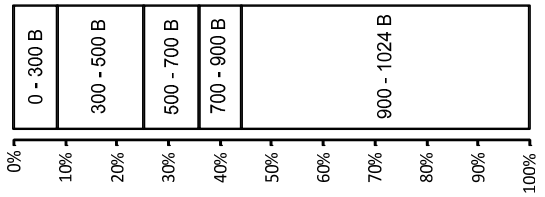

Fig. 6. Percentage data contribution for packet size ranges when using data partitioning and a packet size limit of 1 kB for the *Paris* sequence.

size distribution (in terms of numbers of packets) for the *Paris* video sequence with the same characteristics as employed in our evaluation of the proposed scheme (see Section 3). A 1 kB packet size constraint was imposed, representing a typical maximum transport unit size. Though the sizes of packets appear reasonably well distributed, as Figure 6 shows, the larger packets dominate the data transported.

2.3 IEEE 802.11e EDCA and Access Category Mapping

IEEE 802.11e Enhanced Distributed Channel Access (EDCA) adds quality-of-service (QoS) support to legacy IEEE 802.11 wireless networks by introducing four Access Categories (ACs), i.e. AC0, AC1, AC2, and AC3 for best-effort (BE), background (BK), Video (VI) and Voice (VO) traffic respectively in order of increasing priority. Each AC has different Distributed Coordination Function (DCF) parameters for the Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) back-off mechanism. The parameters include Contention Window (CW) minimum and maximum sizes and arbitrary inter-frame space (AIFS).

Figure 7 represents a practical implementation of the proposed prioritization scheme in which video packets are selectively mapped to IEEE 802.11e's Enhanced Distributed Channel Access (EDCA) AC queues. In this scheme, Data Partition A and C packets where mapped to AC2 as this is default AC for video traffic while partition B was mapped to the lower priority AC1. The effect is to delay access to the channel for numerically lower indexed AC queues, thus increasing the probability that packets from these queues will be dropped through

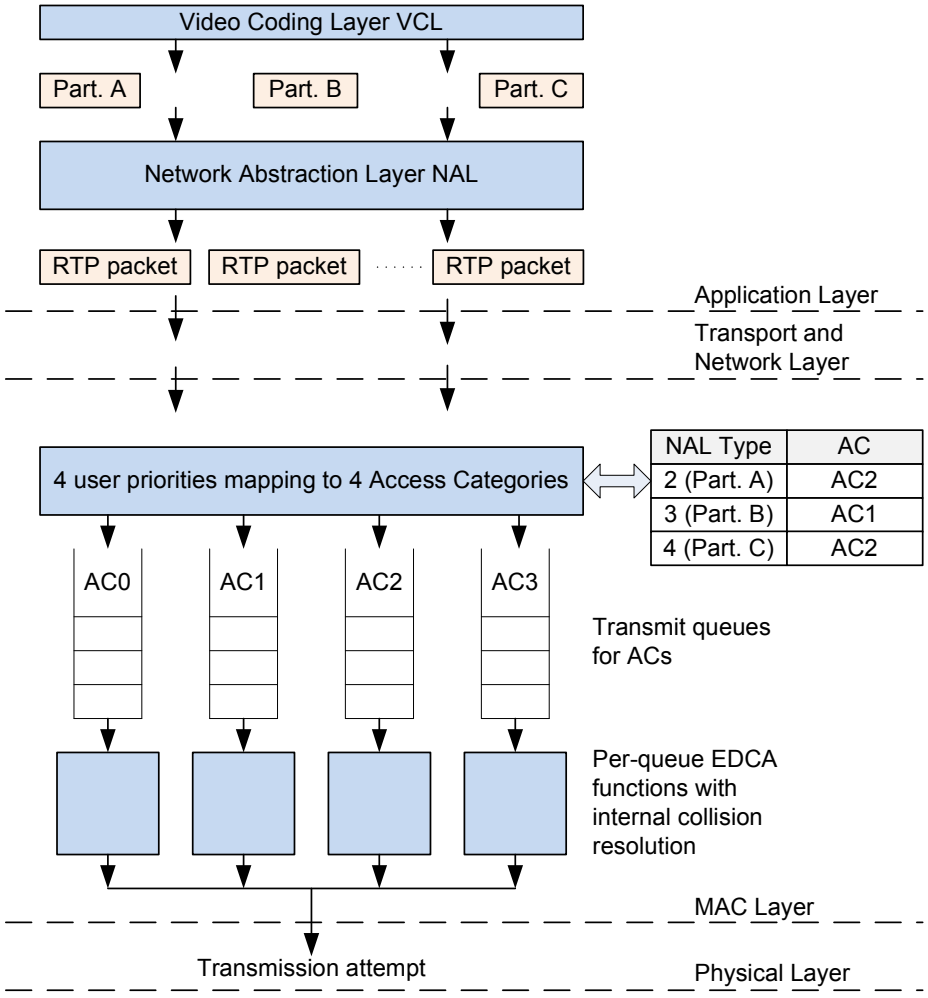

Fig. 7. IEEE 802.11e access category mapping for the data-partitioned packets.

buffer overflow as additional packets arrive. This mapping scheme is further examined in the next Section.

3 Evaluation

3.1 Video Assessment Framework

In order to evaluate the performance of the proposed scheme, four CIF test sequences (*Akiyo*, *Paris*, *Mobile* and *Stefan*), with 299 frames, were coded using

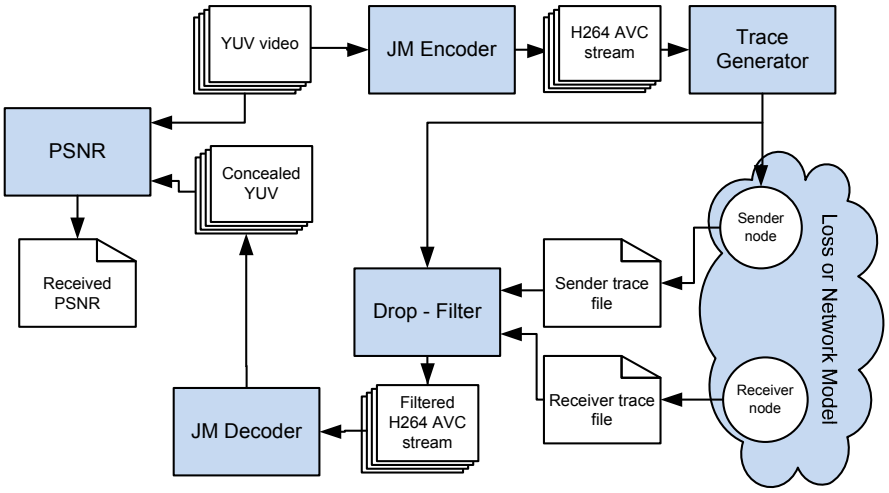

Fig. 8. Video assessment framework.

the H.264/AVC reference software JM16.1 [13] at 30 frame/s. An IPPP coding scheme and a target Constant Bit Rate (CBR) of 1Mbps was used. RTP packetization, data partitioning and one intra-refresh MB line per frame were also enabled to evaluate the error resiliency under lossy networks.

Figure 8 shows the evaluation framework for assessing the impact on video quality. The uncompressed test sequences are encoded according to the previously described settings to generate H.264/AVC compressed streams. After compression, these streams were parsed using an RTP parser to produce trace files containing the size and transmission schedule for each packet. These traces were fed to both a loss model an ns-2 simulated congestion scenario. Comparing the send trace file with the receiver's trace of packets received, the dropped packets were found. The dropped packets were removed from the original compressed stream to generate the damaged received stream. The received stream was then decoded using the H.264/AVC decoder with error concealment enabled to reconstruct the uncompressed video and regenerate the lost slices. The resulting PSNR of the received video was then obtained by comparing the original with the decoded uncompressed video. When decoding the distorted video, frame copy error concealment was used. A total of 2000 runs were used for each set of simulations and the standard absolute Peak Signal-to-Noise (PSNR) deviations were calculated over intervals of 0.5 % packet drop. To make the results more accurate, the percentage of video data drop has been used instead of the packet drop rate. Since packet size depends heavily on the packet content (headers or transform coefficients), a high packet drop might not represent a high data drop.

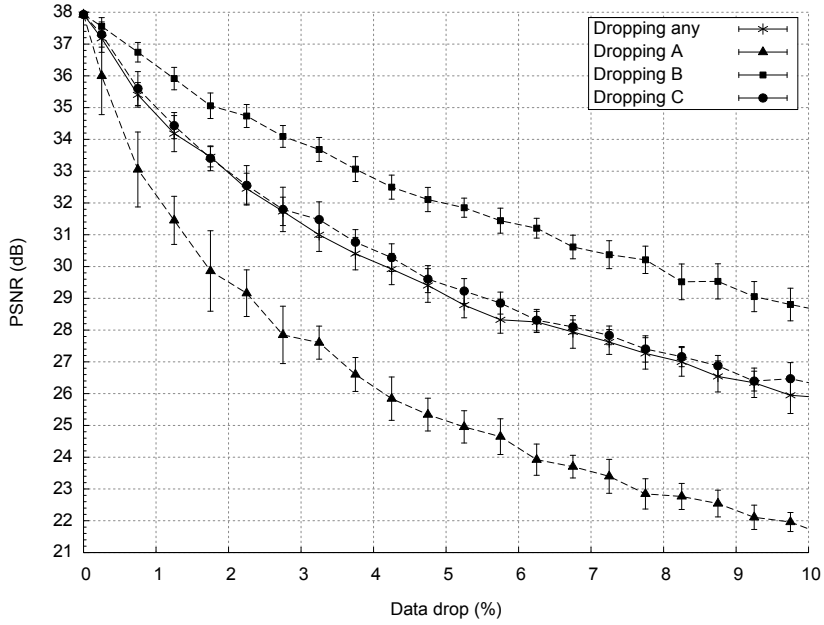

Fig. 9. PSNR vs. percentage video data drop for *Paris*, with absolute standard error bars included.

3.2 Results for uniform packet drops

In this Section, network congestion was assumed to cause uniformly distributed packet drops. Simple analysis of the NAL headers allows an easy identification of the data partition of each packet. A percentage of video packets were dropped according to the following criteria. The packets were simply separated into three different categories according to the data partition type (A, B or C) and a number were randomly selected to be dropped up to a desired percentage. Figure 9 shows the video quality obtained after: (i) dropping packets from partition A packets only (Dropping A); (ii) dropping packets from partition B packets only (Dropping B); (iii) dropping packets from partition C packets only (Dropping C); and (iv) dropping data from any partition packets (Dropping any).

Figure 9 shows that dropping packets from data partition A introduces a significant PSNR penalty (3 dB) when compared with dropping randomly from any partition type. Contrary to the previous case, dropping packets exclusively from data partition B results in a quality gain (3 dB). Finally, dropping only from data partition C produces results close to dropping randomly from any data partition type. Figure 3 and Figure 4 show that an IR MB line can contribute to a significant percentage of the video data as is noticeable from the increased percentage of partition B data occurring when an IR MB line is introduced. All the same, an IR line only represents a single row of MBs per frame. Although a significant portion of the whole bitrate is expended encoding intra MB slices, only

a small area of the image is represented by such a high number of bits. Therefore, the penalty in picture quality is low because error concealment is more efficient when small image areas are lost. The above procedures were repeated for many other sequences and the results are shown in Table 2. The results show the quality gain when using the proposed scheme for different data drop percentages. The overall conclusion is that the proposed scheme can give quality gains for all the tested sequences.

Table 2. Video Quality gain from proposed scheme across a range of difference sequences.

% Data Drop	Akiyo	Mobile	Paris	Stefan
0.25	0.6	0.4	0.1	0.4
0.75	3.2	1.4	0.3	1.3
1.25	3.2	1.9	0.6	2.2
1.75	5.3	2.4	0.8	2.7
2.25	3.9	2.5	1.2	3.5
2.75	4.9	3.2	1.2	3.4
3.25	5.4	3.2	1.4	3.8
3.75	4.9	2.8	1.5	4
4.25	5.6	3.7	1.8	4.3
4.75	5.6	3.2	2.1	4.1
5.25	5.1	3.1	2.3	4.4
5.75	4.7	3.3	2.5	4.6
6.25	5.2	3.3	2.4	4.7
6.75	4.6	3.3	2.8	4.8
7.25	5.1	2.9	3.1	5
7.75	4.5	3.2	3.1	5.2
8.25	5.3	3.1	3	4.9
8.75	4	2.8	3.1	5.5
9.25	4.1	2.6	2.9	5
9.75	4.1	2.6	3.1	5.2

3.3 Effect of congestion in a wireless network

In this Section, we simulate the behavior of the scheme in the presence of congestion within a multi-hop wireless network in which one of the sources is streaming video. To isolate the effect of the scheme upon congestion resilience, the wireless channel is assumed to be error-free.

To analyze the performance, we applied our mapping of Figure 7 to the static scenario shown in Figure 10, with all adjacent nodes 200m apart. IEEE 802.11e access control was applied on all nodes. The IEEE 802.11b physical layer was modeled with a transmission range of 250 m and a data-rate of 11 Mbps and a basic rate of 1 Mbps. As IEEE 802.11b was in ad hoc mode, the well-known Ad-

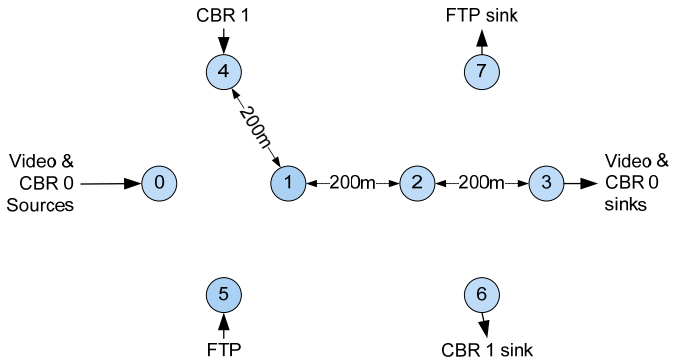

Fig. 10. Network topology used in our simulations.

hoc On-Demand Distance Vector (AODV) routing protocol was deployed with control packets carried at higher-priority AC3.

The video sequence was encoded with a target bitrate of 1 Mbps and packet size was limited to a maximum of 1 kB. The video data packets belonging to data partitions A and C were assigned to AC2 while those belonging to partition B were assigned to AC1 or alternatively no prioritization was used and all the video stream packets were simply assigned to IEEE 802.11e AC2. In these tests, previous frame copy was used as a simple form of error concealment.

We employed the IEEE 802.11e EDCA Media Access Control (MAC) simulation model developed by the Technical University of Berlin [14] for the well-known ns-2 simulator . The EDCA parameter values in Table 3 for the IEEE 802.11b radio were employed.

Table 3. IEEE 802.11e MAC parameter values for the IEEE 802.11b radio

Access Category	AIFSN	CW-min	CW-max	TXOP Limit (ms)
AC0 (Best Effort)	7	31	1023	0
AC1 (Background)	3	31	1023	0
AC2 (Video: VI)	2	15	31	6.016
AC3 (Voice: VO)	2	7	15	3.264

By repeating simulations runs, this arrangement enabled a wide range of total packet drop rates to be arrived at. Packet drops were totaled across the network path taken by the video from node 0 to node 3. Several CBR sources contributed to the congestion. A CBR traffic source from node 4 to node 6 was assigned to AC1 (to emulate background traffic), using UDP transport with a

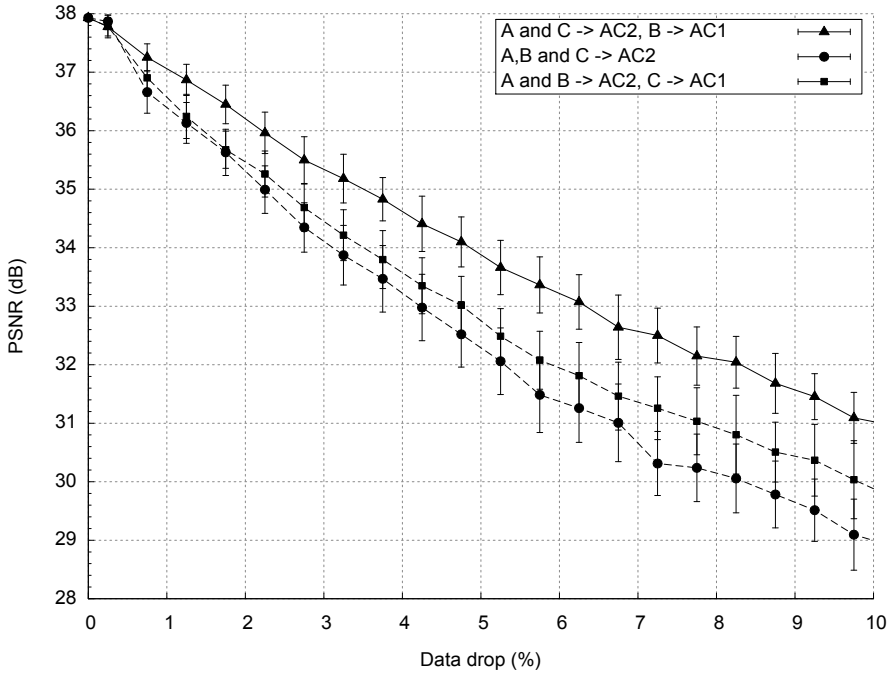

Fig. 11. PSNR vs. percentage video data drop for *Paris*, with absolute standard error bars included.

rate of 128 kbps and packet size of 500 B. Another CBR source from node 0 to node 3 was assigned to AC3 (to emulate voice traffic), also using UDP transport but with a variable rate (48 224 kbps) and a packet size of 160 B. The intention of varying the rate was to cause a range of packet drop percentages (up to 10%) for the video as it crossed the network path. Additionally, an FTP stream, assigned to AC0 (to emulate best effort traffic), was sent from node 5 to node 7 with TCP transport, thus contributing to the congestion. 1400 simulation runs were conducted and the resulting packet drops were totaled across the video's network path for each run.

Figure 11 shows the video quality versus mean percentage video data drop. It is clear that applying the proposed mapping scheme will give a PSNR gain (up to 2dB) than when the video stream is entirely assigned to AC2. It can also be seen that giving priority to partition B over partition C results in a smaller PSNR gain (up to 1dB) when compared with unprioritized transmission (sending all partition types over AC2). Figure 12 shows the end-to-end delay. As expected, the penalty of the scheme is that some video packets are assigned to the lower priority AC1 queue. As a result their queuing time is increased. The increase is greater when there is more congestion, which is also when there are more packet drops.

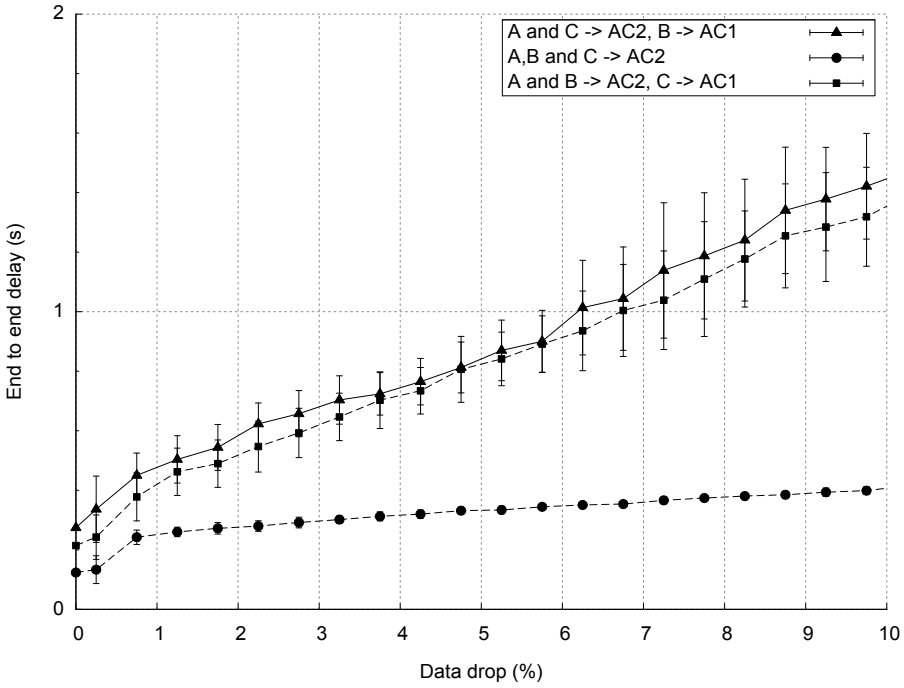

Fig. 12. Mean end-to-end delay vs. percentage video data drop for *Paris*, with absolute standard error bars included.

4 Conclusions

In this paper, the contrasting goals of congestion resilience and transmission error resilience have been highlighted. Because the presence of intra-refresh MBs assist the process of error resilience it is easy to assume that partition B data should be given higher priority access to a WLAN in preference to the transform coefficients collected into partition C packets. When the size of the B-partition is increased due to the presence of intra-refresh MBs then this is not so. In fact, all the results in this paper dropping partition B data was preferable. However, including an intra-refresh MB line is a normal procedure to avoid the presence of sudden increases in bandwidth when periodic intra-refresh is used, which is a significant problem for bandwidth-limited wireless networks. The scheme in this paper proposes assigning low priority to data partition B packets when intra refresh lines are used. The proposed scheme shows that, for congested networks, a PSNR gain of up to 2 dB is obtained in comparison to the case when no prioritization is used at all. Compared with other data partition prioritization schemes, the proposed scheme is still able to achieve up to 1 dB PSNR gain. Therefore, this paper has shown that previous schemes, mainly intended for prioritized transmission of data-partitioned video should also take into account the presence of intra refresh in the stream. Future work will consider whether the

scheme should distinguish between inserted intra-refresh line MBs and naturally encoded intra MBs.

References

1. P. Gill, M. Arlitt, Z. Li, and A. Mahanti, *'Youtube traffic characterization: a view from the edge'*, in 7th ACM SIGCOMM Conf. on Internet Measurement, Oct. 2007, pp. 15-28.
2. T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, *'Overview of the H.264/AVC video coding standard'*, IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560-576, July 2003.
3. J. Afzal, T. Stockhammer, T. Gasiba, and W. Xu, *'Video streaming over MBMS: A system design approach'*, J. of Multimedia, vol. 1, no. 5, pp. 25-35, 2006.
4. L. Liu, X.-J. Ye, S.-Y. Zhang, and Y. Zhang, *'H.264/AVC error resilience tools suitable for 3G mobile video services'*, J. of Zhejiang University SCIENCE, vol. 6, no. 4, pp. 1-46, 2005.
5. T. Stockhammer, M. M. Hannuksela, and T. Wiegand, *'H.264/AVC in wireless environments'*, IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 657-673, July 2003.
6. S. Wenger, *'H.264/AVC over IP'*, IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 645-656, July 2003.
7. R.M. Schreier, and A. Rothermel, *'Motion adaptive intra refresh for the H.264 video coding standard'*, IEEE Trans. Consumer Electronics, vol. 52, no. 1, pp. 249- 253, 2006.
8. B. Barma, M. M. Ghandi, E. V. Jones, and M. Ghanbari, *'Prioritized transmission of data-partitioned H.264 video with hierarchical modulation'*, IEEE Signal Proc. Letters, vol. 12, no.8, pp. 577-580, Aug. 2005.
9. IEEE 802.11e, *'Wireless LAN Medium Access Control and Physical Layer specifications Amendment 8: Medium Access Quality of Service Enhancements'*, 2005
10. A. Ksentini, M. Naimi, and A. Guroui, *'Toward an improvement of H.264 video transmission over IEEE 802.11e through a cross-layer architecture'*, IEEE Communs. Mag., vol. 44, no. 1, pp. 107114, 2006.
11. R. Bernardini, M. Durigon, R. Rinaldo, P. Zontone, and A. Vitali, *'Real-time multiple description video streaming over QoS-based wireless networks'*, in IEEE Int'l Conf. on Image Processing, Oct. 2007, pp. 245248.
12. R. Haywood, S. Mukherjee and X.-H. Peng, *'Investigation of H.264 video streaming over an IEEE 802.11e EDCA wireless testbed'*, in IEEE Int'l Conf. on Communications, Jun. 2009, pp. 1-5.
13. JVT Reference Software Version JM16.1 Online *http://iphome.hhi.de/suehring/tml/download/old_jm/jm16.1.zip*
14. An IEEE 802.11e EDCA and CFB Simulation Model for ns-2. Online *http://www.tkn.tu-berlin.de/research/802.11e_ns2*