This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2010 proceedings.

Traffic-Aware Video Streaming in Broadband Wireless Networks

Ehsan Haghani and Nirwan Ansari
Advanced Networking Lab
Department of Electrical and Computer Engineering
New Jersey Institute of Technology
Newark, NJ 07102, USA

Shyam Parekh and Doru Colin
Bell Labs
Alcatel-Lucent
Murray Hill, NJ 07974, USA

***Abstract*—With increasing implementation of broadband wireless networks and extensive deployment of multimedia services such as Video on Demand (VoD) or IPTV, the demand for video streaming applications will increase and more people will use their wireless devices to reach numerous video contents available in the Internet. Streaming real-time video in wireless networks is a challenging problem due to the stringent service requirements of video traffic and impairments of wireless channels. Providing the required Quality of Service (QoS) through efficient resource allocation is a complicated problem that service providers are confronting. In this paper, we propose a traffic-aware, cross-layer solution for enhancing the perceived video quality at the end user in wireless networks. Our solution incorporates the characteristics of the MPEG traffic to give more priority to the more important frames and to protect them against dropping when available resources of the network are not sufficient for providing the desired QoS to the traffic flow. It is shown that the proposed solution will improve the perceived video quality over the broadband wireless networks.**

I. Introduction

The growing interest in deployment of broadband wireless networks is providing the mobile users with a broadband access through which they can run various applications. It is expected that multimedia applications, and in particular video streaming applications will grow owing to the increasing population of online video servers. Therefore, an excessive number of people will use their wireless devices to access to the numerous video streaming contents on the Internet.

The increasing demand for broadband wireless access has called for the design and implementation of different wireless technologies such as WiMAX and LTE. These technologies are supposed to provide the users with network connectivity to run different applications with various QoS requirements. Hence, the network designers of the broadband wireless technologies should provision adequate arrangements to support different QoS requirements. The stringent QoS requirements of video streaming applications including high bitrate and low latency are some of the challenges the service providers are confronting. Moreover, supporting QoS requires proper resource allocation which makes the video streaming a complicated problem. Efficient allocation of scarce resources in wireless networks is a crucial and challenging problem. The importance of this problem has attracted a keen interest in the research community.

In this section, we review some of the research work devoted to the problem of video streaming in wireless networks. Rate adaptive video streaming is one of the solutions highly discussed in the literature [1]–[3]. Most of the rate adaptive solutions assume that the video server receives some feedback information such as the end-to-end delay or the loss rate from the end client. By using the feedback information, the video server can choose the optimum coding option. The rate adaptive solutions may not be feasible in wireless networks where the adaptive methods may not be able to track the fast changes in the channel. Moreover, this solution increases the computational complexity at the video server which may result in overloading the server in large networks. Furthermore, sending feedback information may not be possible in some video streaming applications such as IPTV in which the video server has to multicast the same content to different clients.

One of the issues for streaming real-time video over wireless networks is sustaining the satisfactory video quality even when congestion happens or the wireless channels become less reliable. Different techniques have been proposed to achieve an acceptable video quality with respect to the limited network resources; nevertheless, in some cases, it is inevitable to prevent the video packets from being lost due to transmission errors over wireless channels or dropped due to overflow of the queues at the Base Station (BS). The effect of packet loss on the video quality has been studied by many researchers. Liang *et al.* [4] analyzed the effect of bursty errors which occur frequently in wireless networks. Meanwhile, there has been extensive effort to find solutions to mitigate the loss effects. Examples of error moderating schemes include, rate adaptive coding [1], forward error correction schemes [5], and scalable video coding [6].

In this paper, we propose a cross-layer solution for streaming video over wireless networks. Unlike the rate adaptive solution, we do not increase the computational complexity of the video servers. Our solution is, therefore, appropriate for video multicasting applications. We consider MPEG video streaming, and use the characteristics of the MPEG traffic to enhance the video quality perceived by the end user. In our traffic aware video streaming mechanism, the video server includes the necessary information, required by the BS to handle the video traffic, in the IP header of the video packets.

The BS can determine the importance of each packet and its effectiveness on the perceived video quality at the end user. The BS can, therefore, protect the more important packets against dropping in the wireless medium, and thus enhances the video quality. Our solution may increase the processing load of the BS, but, as will be discussed later, the availability of the required information in the IP header of each packet will minimize the extra processing load because the IP header of all packets are usually processed by the BS for other networking purposes.

We examine the validity of our solutions by simulating a broadband wireless network which has implemented our traffic aware video streaming. Without loss of generality, we consider the WiMAX technology as the benchmark for evaluating the performance of our proposed mechanism. WiMAX is a promising technology expected to deliver broadband wireless connectivity to mobile users in near future. We must note that although we will present the impact of our solution on the quality of the video stream transmitted over a WiMAX network, similar quality enhancement is expected on other emerging wireless technologies comprised of base stations and subscriber stations. Video streaming over WiMAX networks will be a popular application for users and an attractive market for service providers. There are many research works trying to tailor the general solutions for WiMAX networks. Kim *et al.* [7] introduced a channel adaptive scheme while Hillestad *et al.* [8] proposed the scalable video coding to transmit video traffic over WiMAX. Unlike most of the proposed solutions that either increase the complexity of video server or are not applicable to multicasting applications, our cross-layer solution is scalable and suitable for both unicasting and multicasting applications of real-time video streaming while it does not increase the complexity of the video servers.

The outline of the rest of the paper is as follows. Section II explains the video streaming traffic model used in our research. In Section III, we explain our proposed cross-layer solution for video streaming in broadband wireless networks. We will present our simulation results in Section IV. Concluding remarks are given in Section V.

II. Video Streaming Traffic Model

The availability of various video applications requiring different services, from low bitrate videoconferencing to high bitrate movie streaming, has called for development of different video codecs. Understanding and modeling the main characteristics of the video traffic is needed to create efficient methods for transporting video from different applications with different codecs over various networks.

Moving Picture Expert Group (MPEG) has produced several standard codecs, of which MPEG-2 and MPEG-4 are most widely used. An MPEG encoder converts the input video into a sequence of frames called Group of Pictures (GOP). Generally, there is a high correlation between consecutive frames, and thus the encoder can compress the video by coding the small differences between consecutive frames. The MPEG coded videos have three types of frames. I (Intra coded) frames are

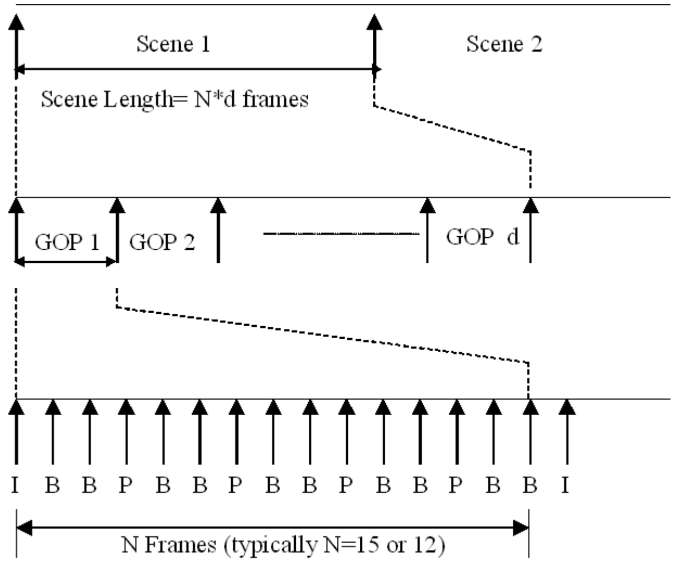

Fig. 1. Source Model [9]

single still images carrying video pictures. The I frames are encoded without reference to other frames, and they are used as the reference frame of each GOP. P (Predicted) frames are built by encoding the differences from a preceding I or P frame. Owing to the difference prediction, P frames provide increased compression, and the size of P frames is typically 20 to 70 percent of the corresponding I frame. B (Bi-directional) frames use previous and next I or P frames located close to them in their GOP as their reference point to predict the changes and compensate for the motion. B frames provide both backward and forward prediction, and thus support further compression and their typical size is about 10 to 40 percent of the size of their corresponding I frame. By considering the characteristics of MPEG frames, we understand that frames of each GOP are interrelated, and thus, the loss of an I or P frame affects some other P or B frames in their GOP.

MPEG encoders usually use a fixed pattern of frames for all GOPs of the video stream. The GOP pattern specifies the number of frames and their permutation in all GOPs of the video stream. Fig. 1 shows the schematic of a general GOP that begins with an I frame. The number of B frames between I and P frames or between two consecutive frames is constant. In Fig. 1, the I-to-I distance is $N = 15$ and the I-to-P or P-to-P distance is $M = 3$.

In this paper, we incorporate the traffic model suggested by the WiMAX forum [9]. In this model, the number of frames in each GOP is constant. The MPEG encoder divides each scene of the video into a number of consecutive GOPs. A scene is part of the video stream that does not contain sudden changes in the picture while it can exhibit some zooming or object movement. The number of GOPs comprising a scene, $'d'$, is a function of the scene complexity, and it is modeled as an exponentially distributed random variable.

In the incorporated model derived from real video traces, the size of B and P frames are modeled by log-normal distribution

TABLE I
MPEG MODEL PARAMETERS [9]

Model Parameter	Value
Display size	320x240
GOP pattern	$N = 15$, $M = 3$
Frame rate	30 frames per second
Scene size parameter (d)	Exponential with p=0.1
I frame size (bytes)	Log-normal ($\mu = 17068$, $\sigma = 7965$)
P frame size (bytes)	Log-normal ($\mu = 9190$, $\sigma = 7005$)
B frame size (bytes)	Log-normal ($\mu = 6839$, $\sigma = 5323$)
AR coefficients	$a_1 = 0.39$, $a_2 = 0.15$, $\sigma_\varepsilon = 4.36$

with parameters (μ_P, σ_P) and (μ_B, σ_B). The correlation between P frames and (similarly B frames) is neglected, and thus the model assumes them as independent random variables; however, it considers and models the correlation between I frames belonging to the GOPs of the same scene. The variation of the size of I frames is attributed to two scales: 1) the small variations within a scene period; 2) the large variations among different scenes. The model thus uses two independent random variables to model these variations. The size of the n^{th} I frame of the video stream, $X_I(n)$, located at the k^{th} scene is

$$X_I(n) = \bar{X}_I(k) + \Delta_I(n) \tag{1}$$

where $\bar{X}_I(k)$ is the mean activity of scene k and represents the large variations of I frames. It is constant for all I frames belonging to the k^{th} scene, and is modeled by a log-normal random variable with parameters (μ_I, σ_I). $\Delta_I(n)$ represents the small variations of I frames around the mean value of the scene. The $\Delta_I(n)$ is modeled by a second order auto regressive process, AR(2).

$$\Delta_I(n) = a_1\Delta_I(n-1) + a_2\Delta_I(n-2) + \varepsilon(n) \tag{2}$$

The values of a_1 and a_2 are assumed constant for each video stream, and $\varepsilon(n)$ is a normal random variable with zero mean and constant variance. Although the parameters defining the model depend on the content of the video, we will use a constant set of parameters suggested by [9] in our simulations. The parameters are presented in Table I.

Table 1 shows that there are two GOPs per second, and therefore the model generates two I frames per second, eight P frames per second, and twenty B frames per second. Hence, the average bitrate for each frame type is as follows: $\bar{R}_I = 273Kbps$, $\bar{R}_P = 588Kbps$, and $\bar{R}_B = 1094Kbps$, and the overall average bitrate for each video stream is $\bar{R}_{tot} = 1955Kbps$. We note that the average bitrate of B frames is higher than those of I and P frames as there are more B frames in each GOP.

III. TRAFFIC-AWARE VIDEO STREAMING

In this section, we introduce our novel solution for enhancing the performance of video streaming in broadband wireless networks. We will incorporate the characteristics of the video traffic in our video streaming solution. As discussed in Section II, MPEG video is a variable bitrate traffic which is delay sensitive. In WiMAX, this kind of traffic is mapped into real-time Polling Service (rtPS) class based on the QoS requirements. As a result, the BS has to guarantee a minimum reserved bitrate for the rtPS traffic flow at the call admission process. The BS periodically polls the rtPS queue and assigns resources based on the bandwidth request. We note that increasing the guaranteed minimum bitrate will decrease the queuing delay and the probability of packet loss happened due to queue overflow; however, it also decreases the network utility since the BS can admit fewer users into the network. Choosing the optimum bitrate which provides the end users with at least the minimum satisfactory video quality is an open problem. In this paper, we assume that the BS has already chosen a minimum bitrate for the video stream, and we will show that the video quality will be enhanced by incorporating our novel solution.

As discussed in Section II, the loss of the video frames will degrade the video quality, and this is more severe if an I frame is lost. Owing to the hierarchical structure of GOPs and the inter-dependency of frames, we understand that the network should protect the I frames and then the P frames from loss to prohibit the *propagation of errors* effect. We note that the probability of a frame loss is increased when the network becomes more congested. The main characteristics of the MPEG video traffic has inspired us to propose a cross-layer, traffic-aware queuing solution called *Intelligent Queue Management*, to be discussed next.

A. Intelligent Queue Management

In our solution, the video server indicates the type of each frame in the Type of Service (ToS) field of the IP header, and thus the BS can determine the frame type of each packet. The BS can also understand the GOP number of each packet and its frame number. We recall that in our model, the video encoder generates a constant number of frames per second, and thus the BS can distinguish the frame number of each packet and its GOP number by inspecting the sequence number and ToS fields of the IP header. As mentioned earlier, our solution protects the more important frames against dropping when congestion happens at the BS.

In WiMAX, the MPEG video streaming traffic is enqueued as an rtPS class flow at the BS. As mentioned in Section II, the traffic bitrate is about 1955 kbps, but the BS may not be able to guarantee this bitrate for the traffic flow. We note that most of the bitrate is made by the B frames while they have least impact on the video quality. Hence, the BS may guarantee a lower bitrate for this traffic flow to increase the number of admitted users in the network. If congestion happens, the BS would drop the least important frames while it tries to adhere to the minimum video quality by transmitting the more

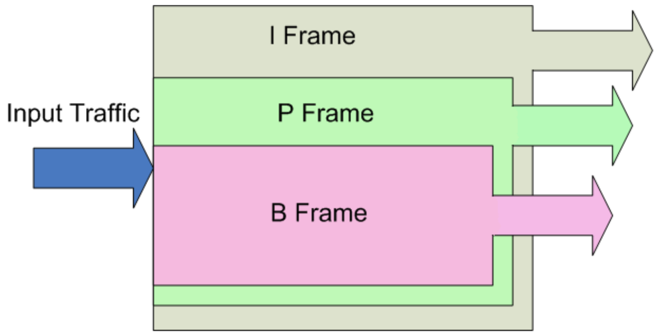

Fig. 2. Intelligent Queue Schematic

important frames on time. We note that during a congestion period, the BS provides the user with only the minimum bitrate which is guaranteed at the call admission process. In this case, the queue size of the corresponding video stream and the queuing delay will be increased. The BS considers a maximum possible size for queuing each traffic flow; the excessive incoming traffic beyond the maximum queue size will result in frame loss. Long queuing delay may also result in frame dropping because of the maximum caching time at the end user. The video traffic encountering delay time more than the maximum caching time cannot be used by the end user, and thus the BS should keep track of queuing delay of each packet. In WiMAX, the maximum tolerable delay is also determined at the call admission process. Therefore, we understand that the overflow of the queue or long queuing delay may result in frame loss which degrades the video quality at the end user.

In order to sustain the video quality, it is crucial to send as many frames as possible to the end user. Moreover, it is necessary to protect the more important frames, i.e., the I frames and then the P frames, against dropping. In our solution, we consider priority queuing at the BS. To better understand the queue structure, we assume a virtual queuing structure depicted in Fig. 2. The arriving packets will be enqueued at the BS based on their frame type. The packets are ordered in each queue according to their GOP and frame number. At the arrival of each packet, the BS determines its frame type, GOP and frame number by checking its IP header. If the queue has enough space, the BS will put the packet in its corresponding sub-queue, but if the maximum queue size is reached, the BS checks the frame type of the incoming packet. If there is any lower priority packet in the queue, the BS will drop as many lower priority frames as needed from the queue to enqueue the arriving packet and protect it against dropping. We note that loss of a fragmented packet will result in the loss of the whole frame, and thus the BS will prefer to drop the frames whose fragments have not still been sent in part to the end user. The BS will drop the incoming packet if there is not enough lower priority packets available in the queue suitable for dropping. The BS will also drop all the dependent packets that will arrive subsequently in the future and require the dropped packet for decoding. It is because the end user will not be able to decode them without receiving the dropped frame. The procedure executed upon arrival of a new packet by the BS is shown in Fig. 3. The priorities of the frames in each GOP are as follows:

1) I frame.
2) P frames in descending order. In each GOP, the higher ordered P frames have less priority because less number of frames is derived from them.
3) B frames.

As explained earlier, the BS should drop the packets awaiting in the queue for more than the maximum tolerable delay which is equal to the caching time at the end user. With that said, the BS regularly monitors the delay time of the first packet of each sub-queue, and if the maximum delay is reached, it will drop that packet from the queue. We note that the BS has to drop all the lower priority packets that are dependent on an already dropped packet for decoding. For example, the drop of the third P frame, which is the tenth frame of the GOP, will result in dropping of all B and P frames located from the 8^{th} to 15^{th} frame of that GOP, as depicted in Fig. 1. The delay monitoring procedure regularly performed by the BS is shown in Fig. 4.

It is worth noting that to avoid long delays, the queues should be served properly in time; however, this may not be possible during the congestion period when the BS can provide each queue with only the minimum reserved bitrate. Inspired by this fact, we incorporate the *multi-level service classification* solution [10]. In that method, the available bitrate is divided among three sub-queues, and more minimum reserved bitrate is conservatively allocated to the I frame sub-queue. Therefore, we can avoid the long delays for more important sub-frames. We also note that if the I frame traffic cannot consume all the allocated bandwidth, the BS will spare the remaining part to other traffic flows. We also choose higher bitrate for the P frame sub-queue as compared to that of B frames, and thus we reduce the probability of dropping the P frames. We will elaborate on the chosen bitrates for each of the sub-queues in Section IV where we will discuss the simulation results.

In order to allocate resources to the video stream, the BS polls each of the sub-queues separately. The polling period depends on the minimum reserved bitrate of each sub-queue and the higher the bitrate, the more frequent polling. During the congestion time the BS can only provide the user with the minimum reserved bitrate. In this case, the minimum amount of traffic associated to the minimum reserved bitrate is transmitted to the end user upon polling the sub-queue. The BS will not allocate any resources to the sub-queue if there is no traffic available in that sub-queue. Frames are sorted according to their frame number and GOP number in each sub-queue. Therefore, the departure permutation of the frames from each sub-queue is regulated. In the intelligent queue management, the loss of the higher priority frames will result in the loss of the dependent, and lower-order frames of their GOPs. Moreover, the sub-queues with higher priority are guaranteed to receive higher bitrates while they have smaller

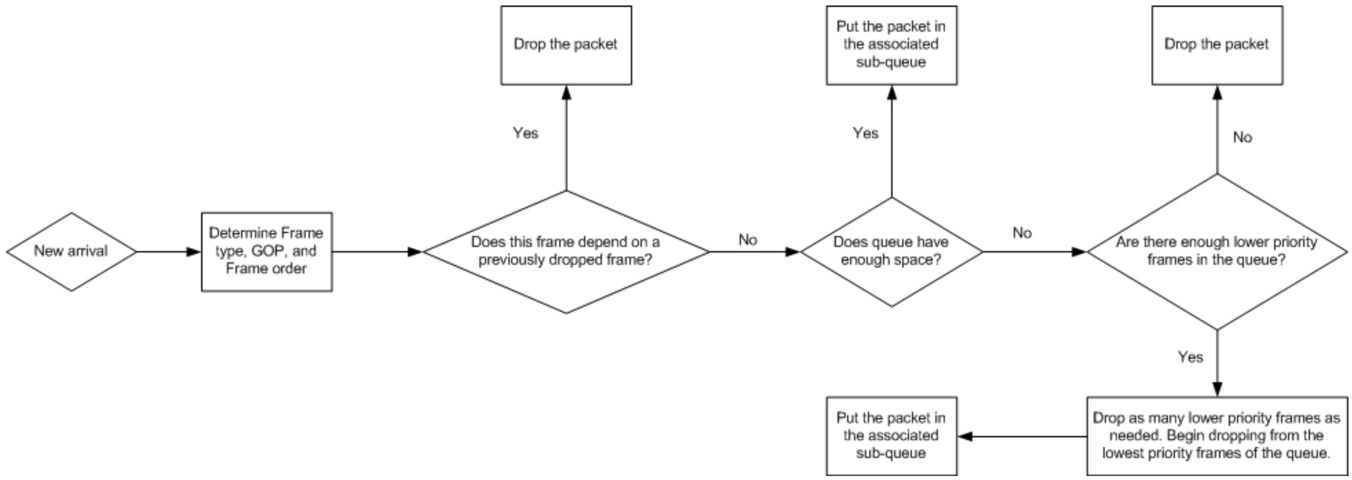

Fig. 3. Frame Arrival Management

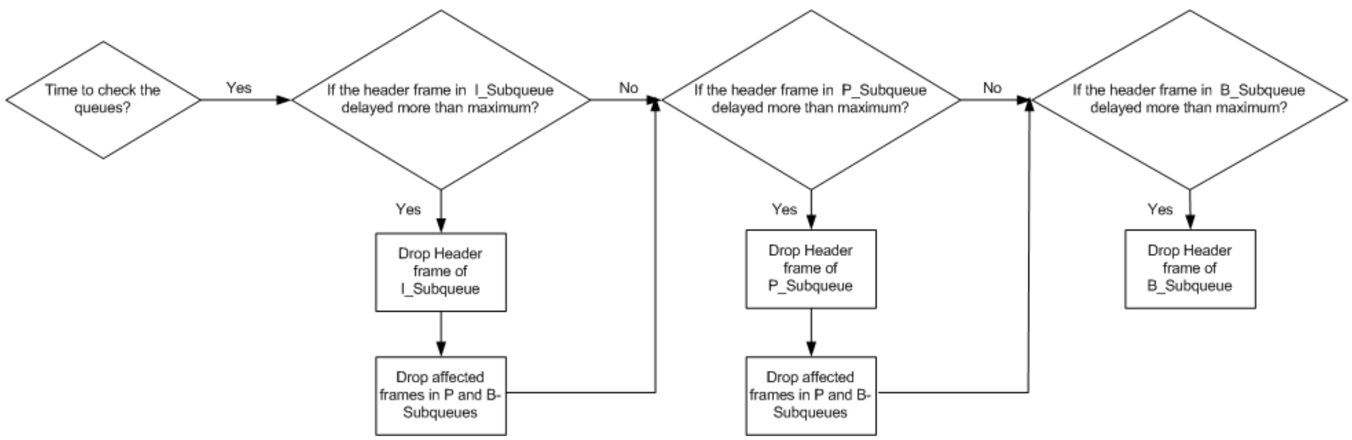

Fig. 4. Delay Management Procedure

average bitrate. Therefore, it is not necessary to synchronize the permutation of departing frames among different sub-queues. We note that the end user caches the arrival frames, and thus out of sequence arrivals will be adjusted accordingly.

IV. Simulation Results

In this section, we evaluate the performance of the proposed intelligent queue management scheme discussed in Section III. We study the number of frames received by clients through some comprehensive system level simulations. We generate the traffic based on the model explained in Section II. In our simulations, we only examine the performance of the MAC layer of WiMAX, and we do not involve the PHY layer effects. We simulated our network under congestion circumstances in which the BS can only provide each traffic flow with the minimum reserved bandwidth. We also simulate the conventional WiMAX network in which the video stream is classified as one rtPS service class and is enqueued in a regular FIFO queue at the BS. In conventional scheme, packets are dropped regardless of their frame types when the queue becomes full or the delay becomes long. We assume that the maximum possible queue size is 6MB in both schemes and the maximum tolerable delay is 10 seconds. Therefore, the frames which are delayed for more than 10 seconds in the queue, will be dropped.

We assume that the BS has reserved 900 Kbps for the traffic flow which is less than 1955 Kbps required for streaming the video. Since the departure rate of the queue is less than the arrival rate, the queue will overflow, and we are expecting packet loss in both queuing schemes. In the intelligent queue management scheme, we assume that the BS divides the reserved bitrate among the sub-queues and allocates 360 Kbps for the I frames, 340 Kbps for the P frames and 200 Kbps for the B frames sub-queues, respectively. We note that the BS has assigned a value more than the average rate of the I frames to them. This protects I frames against dropping. As mentioned earlier, the BS can spare the bandwidth to other traffic flows if the I frames cannot consume all the allocated bandwidth. We expect to observe the loss of the P fames and B frames due to shortage of the pre-allocated bandwidth to these flows, and it is expected to observe a higher number of dropped B frames since they receive the least share of the bandwidth.

We conducted simulations for both of the intelligent queue

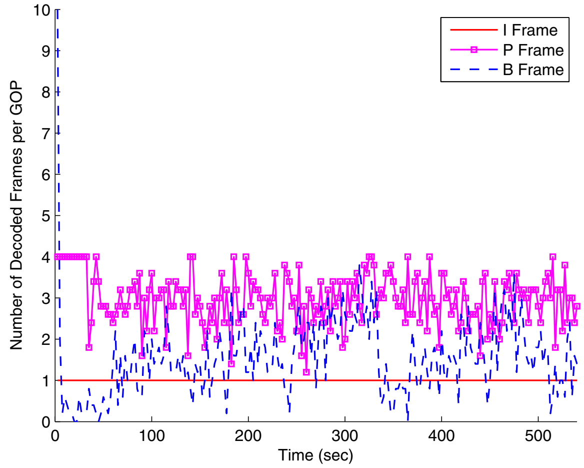

Fig. 5. Intelligent Queue Management Performance

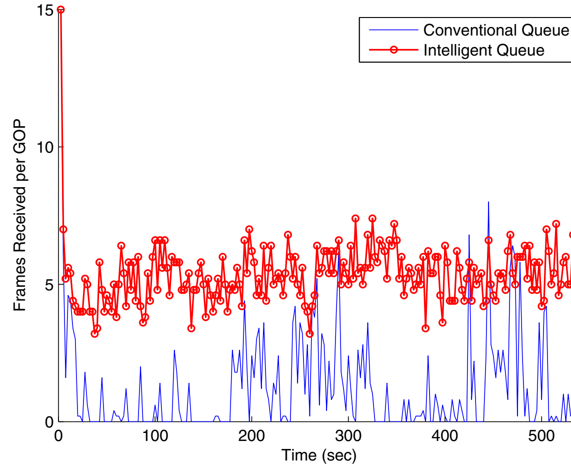

Fig. 6. Intelligent Queue and Conventional Queue Comparison

management and conventional queuing schemes, and compared their performances with respect to the number of MPEG frames received and decoded correctly at the end user. We know that receiving a higher number of frames will result in better perceived video quality. This can be a useful criterion for comparing the performances of different networking schemes while considering the importance (or priority) of the different frame types. Fig. 5 shows the number of frames of each GOP decoded by the end user in the intelligent queue management scheme. We observe that all I frames have been delivered, and on average, three P frames and two B frames were also delivered. In the intelligent queuing scheme, by receiving all of the I frames and most of the P frames, the end user can adhere to the minimum perceived video quality during the congestion periods. Fig. 6, shows the comparison on the number of frames of each GOP received and decoded by the end user when applying the intelligent queuing and conventional queuing schemes. We observe that the intelligent queuing scheme delivers more frames to the end user, and thus it provides better video quality. We also note that at some moments, the end user will not be able to decode any frames of some GOPs due to the loss of the higher priority frames. Owing to the sever disruptions in the video streaming, the conventional queuing scheme is not a feasible solution for the networks with the possibility of congestion or shortage of resources.

V. Conclusion

In this paper, we have proposed a novel solution for enhancing the quality of MPEG streaming over broadband wireless networks. We have briefly explained the characteristics of the MPEG traffic and introduced the traffic model suggested by the WiMAX Forum. We have discussed the structure of the MPEG GOPs which is comprised of I, P and B frames. We have also elaborated the effect of frame loss on the video quality of the end user. Inspired by the traffic model, we have designed a cross-layer scheme that provides the BS with the frame information of each incoming video streaming packet. In our solution, the BS can deduce the frame type and frame order of each incoming packet by examining its IP header. We have proposed a queue management strategy in which the BS can intelligently drop less effective frames when congestion happens. We have shown through some simulations that by incorporating our cross-layer solution and intelligent queuing scheme, we can deliver more video frames to the end users. Moreover, our design protects the most important frames, i.e., I frames and P frames, against dropping, and thus provides better perceived video quality.

References

[1] C.-Y. Hsu, A. Ortega, and M. Khansari, "Rate control for robust video transmission over burst-error wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 5, pp. 756 – 773, May 1999.

[2] M. Hassan and M. Krunz, "Video streaming over wireless packet networks: An occupancy-based rate adaptation perspective," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 8, pp. 1017–1027, Aug. 2007.

[3] X. Zhu and B. Girod, "Video streaming over wireless networks," in *Proc. European Signal Processing Conference (EUSIPCO)*, 2007, pp. 1462–1466.

[4] Y. J. Liang, J. G. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: Effect of burst losses and correlation between error frames," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 7, pp. 861–874, Jul. 2008.

[5] W. T. Tan and A. Zakhor, "Video multicast using layered FEC and scalable compression," *IEEE Trans. Circuits Syst. Video Technol*, vol. 11, pp. 373–386, 2001.

[6] D. Wu, Y. Hou, and Y.-Q. Zhang, "Scalable video coding and transport over broadband wireless networks," *Proceedings of IEEE*, vol. 89, no. 1, pp. 6–20, Jan. 2001.

[7] H.-S. Kim, H.-M. Nam, J.-Y. Jeong, S.-H. Kim, and S.-J. Ko, "Measurement based channel-adaptive video streaming for mobile devices over mobile WiMAX," *IEEE Transactions on Consumer Electronics*, vol. 54, pp. 171 – 178, Feb. 2008.

[8] O. I. Hillestad, A. Perkis, V. Genc, S. Murphy, and J. Murphy, "Adaptive H.264/MPEG-4 SVC video over IEEE 802.16 broadband wireless networks," in *Proceedings of Packet Video*, Nov. 2007, pp. 26–35.

[9] WiMAX Forum. (2008) WiMAX system evaluation methodology V2.1. [Online]. Available: http://www.wimaxforum.org/resources/documents/

[10] E. Haghani, S. Parekh, D. Calin, E. Kim, and N. Ansari, "A quality-driven cross-layer solution for MPEG video streaming over WiMAX networks," *IEEE Transactions on Multimedia*, vol. 11, no. 6, pp. 1140–1147, Oct. 2009.