Exploiting MAC Flexibility in WiMAX for Media Streaming

Shamik Sengupta and Mainak Chatterjee
Electrical and Computer Engineering
University of Central Florida
Orlando, FL 32816
{shamik, mainak}@cpe.ucf.edu

Samrat Ganguly and Rauf Izmailov
Broadband Mobile Communications
NEC Laboratories America
Princeton, NJ 08540
{samrat,rauf}@nec-labs.com

Abstract

The IEEE 802.16 standard (commonly known as WiMAX) which has emerged as a broadband wireless access technology, is capable of delivering very high data rates. However, providing performance guarantees to delay sensitive applications like streaming media is still a challenge. In this paper, we study the media access control (MAC) layer of WiMAX and exploit its flexible features to dynamically construct the MAC packet data units (MPDU). The sizes of the MPDUs are constantly modified based on the channel state information. The desired payload is obtained either by aggregation or fragmentation of the upper layer data units. The robustness of MPDUs is also made tunable by means of cyclic redundancy code bits. We consider both the scenarios- with and without feedback. We adhere to the 802.16 specifications and propose adapting the MPDU length for streaming media for better performance. Three metrics are defined- restore probability, goodput and dropping probability. Simulation experiments are conducted which show the performance enhancements of the proposed ARQ-enabled adaptive algorithm in terms of these three metrics.

1 Introduction

The success of streaming applications (such as audio and video) in the Internet domain is creating a demand for streaming services over wireless networks. Such demands are being addressed by a variety of wireless access technologies such as, cellular, wireless LANs, and wireless MANs. However, the widespread use and bandwidth demands of multimedia applications are far exceeding the capacity of current technologies. Moreover, most access technologies do not have the option to differentiate specific application demands or user needs. Though the IEEE 802.11 standard has become the most prevailing technology for indoor access for mobile devices, the outdoor broadband wireless access is still being researched. To that extent, IEEE has defined a new standard for broadband wireless access- the 802.16 family of protocols (commonly known as WiMAX), which can provide data rates of about 100 Mbps and more [2]. Because of such high data rates, WiMAX promises to bring delay sensitive streaming media applications to our mobile devices.

Media files such as video are usually enormous in size and reside in servers in the core network. Since the network poses a bottleneck in the transmission of such amount of data, encoders like MPEG-2 and MPEG-4 are used to compress the data by taking advantage of the temporal and spatial redundancies, making video transmission easier. Inspite of such encodings, the end-to-end performance still suffers due to the unreliable wireless links. Retransmissions and forward error correction schemes have been proposed to shield the effects of the lossy medium [4]. Even with such techniques, it is still difficult to achieve high data rates ($>$ 100 Mbps).

WiMAX is a potential access technology that uses multiple channels for a single transmission and provides bandwidths of upto 350 Mbps [5]. The use of orthogonal frequency division multiplexing (OFDM) increases the bandwidth and data capacity by spacing channels very close to each other and still avoids interference because of orthogonal channels. Using this scheme, WiMAX can transmit up to 30 miles requiring line of sight (LOS). However, cell sizes have a small radius of around 5 miles or less for providing high data rate and uniform coverage throughout the cell. The 802.16 std. can be used in a point-to-point or mesh topology, using pairs of directional antennas, which can be used to increase the effective range of the system relative to what can be achieved in point-to-multipoint mode. But the disadvantage of 802.16 is its frequency of operation. It operates at high frequency (10-66 GHz) requiring LOS and thus multipath is negligible. As a result, in urban areas, where multipath is inevitable, 802.16 might not be that effective. To address this issue, an amendment 802.16a [3] has been standardized that overcomes the difficulties of the original 802.16 standard [2]. 802.16a operates in the spectrum range of 2-11 GHz. Due to longer wavelength, LOS is not necessary and multipath is significant. Data rates of the order of 100 Mbps or more can still be achieved.

Proceedings of the Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM'05)
0-7695-2342-0/05 $20.00 © 2005 **IEEE**

In this paper, we explore the efficiency of streaming media over WiMAX. Though, the medium access control (MAC) layer of WiMAX has been standardized, there are certain features which can be tuned and made application and/or channel specific. We exploit the flexibility present in the MAC layer of 802.16a for construction and transmission of *MAC packet data units* (MPDU) for streaming media over unreliable wireless channel and maintain a high throughput. We design a feedback mechanism at the receiver MAC layer and depending on the feedbacks, the MAC layer at the transmitting side modifies its MPDU payload size and/or cyclic redundancy code (CRC). The dynamic manner in which the MPDUs are changed to match the channel conditions helps in increasing the packet restore probability, thereby reducing MPDU dropping probability and increasing goodput. The reduction in the number of retransmission of dropped or corrupted packets in turn lowers the delay, which is very crucial for streaming applications. Simulation experiments are conducted to verify our proposed scheme. We assume a four-state Markovian channel model and consider both with and without feedback. We show that our proposed feedback technique and variable length MPDUs are effective and significantly increases the goodput and lowers MPDU dropping probability.

The rest of the paper is organized as follows. In section 2, we present the system model. We exploit the MAC layer features and consider aggregation and fragmentation of MAC service data units in section 3. Feedback design schemes to increase packet restore probability are presented in section 4. In section 5, we present the simulation model and the results. Conclusions are drawn in the last section.

2 System Model

The main components of a WiMAX system consist of a base station and multiple subscriber stations which form a cell with a point-to-multipoint structure as shown in figure 1. The base station controls the activity within the cell including admission to the network, access to the medium by subscriber stations, and allocations to achieve the desired quality-of-service (QoS).

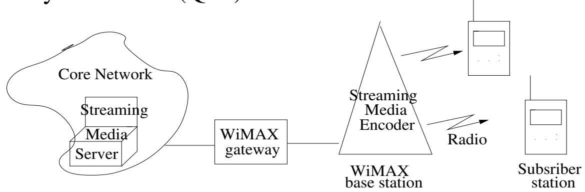

Figure 1. Network architecture

Upon request from the subscriber station, the media server located in the core network starts transmitting the raw data. By media server, we mean the host which has the original application data such as a video. This raw data is routed to the destination base station via the WiMAX gateway. and is fed to the encoder residing at base station. The encoder encodes the raw streaming data and pushes it to the MAC layer in the form of *MAC service data units* (MSDU). The scheduler at the MAC layer uses the channel state information of the subscriber stations obtained through the on-air timing and feedback mechanism, and manipulates the MSDUs to construct the MPDUs.

3 Tweaking the MAC of WiMAX

The MAC layer of 802.16a layer comprises three sublayers which interact with each other through the service access points (SAPs) as shown in figure 2. The service specific convergence sublayer provides the transformation or mapping of external network data, with the help of the service access point. The MAC common part sublayer receives this information in the form of MSDUs which are packed into the payload fields of MPDUs. Privacy sublayer provides authentication, secure key exchange and encryption on the MPDUs formed from the MSDUs and passes them over to the physical layer. Of the three sublayers, the common part sublayer is the core functional layer which provides bandwidth, establishes and maintains connection. Moreover, as the 802.16a MAC provides a connection-oriented service to the subscriber stations, the common part sublayer also provides a connection identifier to identify which connection the MPDU is servicing.

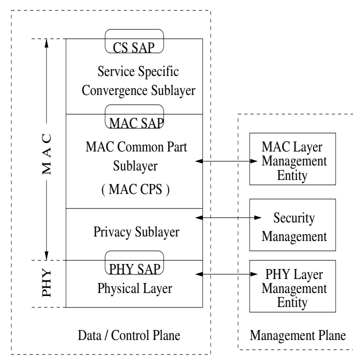

Figure 2. 802.16a MAC layer with SAPs

3.1 Aggregation and fragmentation of MSDU

We focus on the common part sublayer to explore its rich set of features. This sublayer controls the on-air timing based on consecutive frames that are divided into time slots. The size of these frames and the size of the individual slots within these frames can be varied on a frame-by-frame basis. This allows effective allocation of on-air resources and we apply this mechanism on the MPDUs that are to be transmitted. Depending on the feedback received from the receiver and on-air physical layer slots, We exploit the feature of the common part sublayer that modifies the size of the MPDUs by changing the size of the payload.

3.1.1 Aggregation

The common part sublayer is capable of packing more than one complete or partial MSDUs into one MPDU. In figure 3, we see that the payload of the MPDU can accommodate more than two complete MSDUs, but not three. There-

0-7695-2342-0/05 $20.00 © 2005 IEEE

fore, a part of the third MSDU is aggregated with the previous two MSDUs to fill up the remaining payload field preventing wastage of resources. The payload size is determined by on-air timing slots and feedback received from subscriber station.

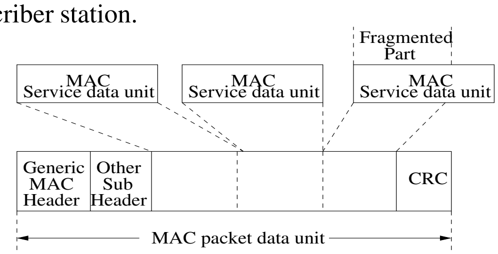

Figure 3. Multiple MSDUs form a MPDU

3.1.2 Fragmentation

The common part sublayer can also fragment a MSDU into multiple MPDUs. In figure 4, we see how a portion of a single MSDU is used to fill the entire payload field of a MPDU. Here, the payload field of the MAC packet data unit to be transmitted is too small to accommodate a complete MSDU. In that case, we fragment a single MSDU and pack the fragmented part into the payload field of the MPDU.

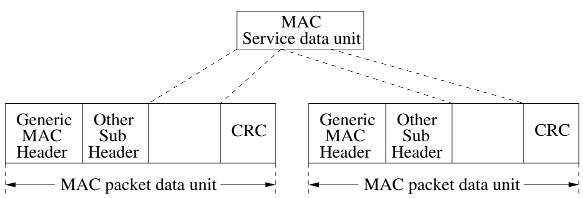

Figure 4. Single MSDU forms multiple MPDUs

We make extensive use of these two features in our algorithm for MPDU construction and transmission discussed later. Since, the MPDUs can be of variable size, the natural question that arises is what the optimal size is.

3.2 Optimal MPDU size

The optimal size of the MPDU must be matched to the channel conditions so as to obtain a desired level of performance. Since packets often get lost being corrupted during transmission in error prone wireless channels, ARQ mechanism is usually used to identify and possibly recover the missing frames. In our case, ARQ will play a crucial role in estimating the channel condition and the fate of the MPDUs that have been transmitted. As a result, the round trip time (RTT) becomes crucial in determining the size of the MPDUs. We define RTT as the time difference between the time the last bit of a MPDU is transmitted and the time the acknowledgement for that MPDU is received. Moreover, we assume zero time interval between two consecutive MPDUs, i.e., the last bit of a MPDU and the first bit of the next MPDU are transmitted back to back.

Let us now show, how the RTT affects the size of the MPDUs. If we assume that t is the time taken to transmit the MPDU and T is the RTT, then the number of MPDUs already transmitted before the acknowledgement of the first MPDU is received is given by $\lceil T/t \rceil$. It can be noted that t depends on the size of the MPDU and thus there is a trade-off between the goodput (information bits/total bits transmitted) and the delay. If a MPDU is large, the transmission time is large but the overhead due to headers is less which helps in maintaining a high goodput. If the MPDU is dropped or corrupted due to bad channel condition, the ARQ mechanism will invoke the retransmission of the large MPDU, which will introduce delay in the transmission. Moreover, by the time the MAC common part sublayer receives the feedback i.e., learns about the bad channel condition, the transmission of the next MPDU would have already started. If the bad channel condition persists, the probability of the subsequent frame being dropped or corrupted is very high. Thus, there will be more retransmissions of large MPDUs under bad channel condition, resulting in severe degradation of goodput compromising the QoS. To clarify, let us consider a simple example.

An illustrative example:

Let the size of a MPDU be 1055 bytes (1000 bytes payload, 50 bytes CRC, and 5 bytes of header). If we consider that the base station transmits at 100 Mbps or more, the transmission time of the MPDU will be 0.084 msec. The RTT will be of the order of 0.05 msec considering a cell of approximately 5 miles radius. Thus the transmission of the second MPDU will be in progress while the acknowledgment of the first MPDU arrives. As a result, the MAC common part sublayer can only respond and take appropriate measures before the third MPDU is transmitted. Now suppose that the channel condition is bad and remains so for 2 to 3 msec. Both the first and second MPDUs will be dropped and 2110 bytes of retransmission will be incurred.

On the other hand, if the MPDU size is small, say 355 bytes (300 bytes payload, 50 bytes CRC, and 5 bytes of header), the transmission time of each MPDU is 0.0284 msec. With the RTT still being in the order of 0.05 msec, the acknowledgement of the first MPDU arrives while the third MPDU is being transmitted. Now appropriate measures can be taken on the fourth MPDU, and even if loss of all previous three MPDU occurs, still 1065 bytes are needed to retransmit, which is almost half of the previous example. But the main disadvantage of having small MPDUs is the low goodput due to low payload/overhead ratio. Thus, we observe that both large and small MPDUs have their advantages and disadvantages. We propose to combine the advantages of both by dynamically changing the MPDU size in response to the channel conditions.

4 Feedback-based Adaptive MAC

Let us first construct the various types of feedback that will be used by the MAC for construction of the MPDUs. Since different video encoders classify video frames in different ways (for example, I, B, P frames in MPEG-2), we use a generic two level classification and do not restrict ourselves to any particular coding methodology. We classify video frames as *important* and *not so important* and propose to treat them differently. Frames obtained from any encoder can always be classified into two levels depending on their importance. The data for the video frames eventually make the MPDUs which when transmitted might or might not be received by the receiver. Even if received, it might not be correct. The exact state in which a MPDU is received depends of the channel conditions it experienced. Hence, it is important to identify the possible states that a received MPDU might be in. We propose six types of feedback- each of which depends on the state of the received MPDU and its importance level. We show the different possibilities of feedback in table 1.

Feedback type	Feedback classification
1	MPDU received correctly
2	MPDU received with errors, but correctable
3	MPDU received with errors, and uncorrectable
4	MPDU dropped, timeout in receiver MAC occurred
5	Receiver MAC buffer full last stored frame is important
6	Receiver MAC buffer full last stored frame is not so important

Table 1. Different feedback possibilities

One question might arise on how the MAC common part sublayer distinguishes the important and not so important frames being streamed. This can be done by reading the application headers where the type of the frame is embedded. We suggest that a single bit in the "other sub-header" field of MPDUs can be used for the same purpose of differentiating the MPDUs. To represent the 6 types of feedback, only three bits would be sufficient. As our mechanism depends on the retrieval of the feedbacks, we assume that they suffer no loss due to robust coding.

4.1 Packet Restore Probability

If a receiver gets a corrupted video frame, it is in no position to correct the errors. However, if some redundant bits in the form of CRC codes are appended to the payload of these video frames before transmission, then there is a probability that the receiver would be able to detect and possibly correct the errors. The correction capability of these codes will depend on the kind of codes and the length of the code used. Since this paper does not deal with CRC codes, we will just discuss in terms of the simplest of codes- *block codes*. In block codes, M redundancy bits are added to the N information bearing bits (payload). (Note that these extra bits are generated using a generator matrix operating on the bits.) If we consider such a MPDU the resulting bit loss probability is given by [1],

$$b = \sum_{i=M+1}^{M+N} C_i^{M+N} {b_p}^i (1-b_p)^{M+N-i} \frac{i}{M+N} \qquad (1)$$

where, b_p is the bit loss probability before decoding and b is the decoded bit error probability. The restore probability of such a MPDU with payload size N bits and CRC M bits is given by, $p = (1-b)^{(M+N)}$. We apply three schemes to manipulate this packet restore probability:

4.1.1 Decreasing payload keeping CRC fixed:

Let b be the resulting bit loss probability after decoding of a MPDU with payload size N and CRC M bits. Now, if we decrease the payload size to N' ($N' < N$) keeping the CRC field fixed, the the resulting bit loss probability after decoding is given by,

$$b' = \sum_{i=M+1}^{M+N'} C_i^{M+N'} {b_p}^i (1-b_p)^{M+N'-i} \frac{i}{M+N'} \qquad (2)$$

Then, it can be argued that with the decrease in payload with the CRC fixed, $b' < b$. Now, let p' be the new packet restore probability. Then p' is given by,

$$p' = (1-b')^{(M+N')} \qquad (3)$$

As b' and b are close to 0, $(1-b)$ and $(1-b')$ are close to 1, though smaller than 1. Then without any loss of generality, it can be said that, for $N' < N$, $p' > p$, i.e., with a decrease in payload, packet restore probability increases.

4.1.2 Increasing CRC keeping payload fixed:

It can be argued the same way as above that if the CRC is increased keeping the payload fixed, the resulting bit loss probability decreases and packet restore probability of MPDUs increases.

4.1.3 Increasing both payload and CRC:

If the MPDU to be transmitted is an important frame and its payload can not be decreased because of play out frame rate restriction at the receiver, the best scheme would be to increase both the payload and CRC. As we know, increasing payload only will increase the resulting bit error probability, so we must also increase the CRC to compensate for the increased payload.

4.2 Connection Set-up and Transmission

Phase 1: *Subscriber station requests connection request:* Subscriber station that wants to receive video from a base station transmits a ranging request (RNG-REQ) packet that enables the base station to identify the initial ranging, timing and power adjustments. Service flow parameters requests (bandwidth, frequency, peak or average rate) are sent next and variable length MSDU indicators are turned on.

Proceedings of the Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM'05)
0-7695-2342-0/05 $20.00 © 2005 **IEEE**

Phase 2: *Base station confirms connection:*
After receiving connection request from a subscriber station, the base station transmits a ranging response which provides the initial ranging, timing and power adjustment information to the subscriber station. Service flow parameters are agreed upon and a basic connection ID is also provided to the subscriber station.
Phase 3: *Base station starts transmission of MPDUs:*
MSDUs are obtained from the MAC convergence sublayer are packed into MPDUs to be transmitted. As needed, MSDUs can either be aggregated or fragmented to form the desired sized MPDUs. Since no feedback is received at the start of transmission, the payload and CRC sizes agreed at the time of connection establishment are maintained. When a feedback is received, the next awaiting MPDU is formed depending on the type of feedback received. On the reception of each of the 6 feedbacks, the payload and CRC sizes are changed. It can be noted that the increase or decrease in payload and CRC field will depend on the ratio of the payload and CRC. The exact values are found through experimentation to be discussed later in section 5. In qualitative terms, we mention the action taken by the base station.
Feedback type 1:
- increase MPDU payload
- decrease CRC for not so important MPDU
Feedback type 2:
- increase CRC for important MPDU
- keep payload and CRC fixed for not-so-imp MPDU
Feedback type 3:
- decrease payload of MPDU
- increase CRC of MPDU
Feedback type 4:
- same as Feedback 3, but the increment/decrement is more
Feedback type 5:
- stall transmission until further request received
Feedback type 6:
- skip transmission of next few *not so important* frames
- important frame(s) is/are transmitted

5 Simulation Model and Results

We conducted simulation experiments to evaluate the improvements achieved by the proposed technique. Evaluations for adaptive and non-adaptive schemes were done under the same channel conditions for a fair comparison.

5.1 Channel Model

We assumed a four-state Markov model for the channel. Four states were used to have more granularities in the channel conditions. Each state was characterized by a certain bit error probability (BER): *good state* had a BER of 0.045, the *fair state* had a BER of 0.06, the *medium state* had a BER of 0.07, and the *bad state* had a BER of 0.085. By setting appropriate transition probabilities among these four states, we are able to model different channel conditions for our simulation.

5.2 Simulation Parameters

First, it is necessary to figure out the exact increase/decrease in payload and CRC if the goodput is to be optimized. Of course, the increase/decrease will also depend on the ratio of the payload to CRC. In figure 5, we show one such example, where we use three different ratios for payload/CRC. It can be observed that, with initial $payload : CRC :: 500 : 50$ the best goodput is obtained for 1 byte payload increase. Similarly, for $payload : CRC :: 500 : 100$, increase of 5 bytes gives the best goodput. As a result, it can be concluded that the change in payload and CRC depends on the payload CRC ratio.

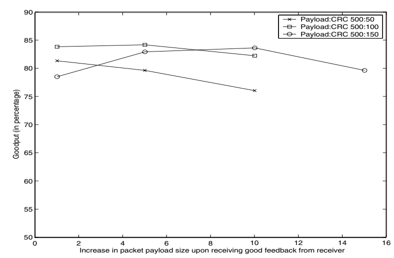

Figure 5. Goodput with payload increment

We assumed an initial packet size of 555 bytes: payload 500 bytes, CRC 50 bytes and header 5 bytes. Upon receiving feedback from the subscriber station, base station fragments or aggregates the MSDUs adaptively as per the proposed algorithm to form the MPDUs. For different types of feedback, the change in size of payload and CRC are shown in table 2. Since for feedback types 5 and 6, the MPDUs are either stalled or skipped, we do not show them in the table. It can be noted that these values apply to the initial payload/CRC ratio. Any other ratio will yield a different set of values.

Feedback type	Payload (imp)	Payload (not-imp)	CRC (imp)	CRC (not-imp)
1	+1 Byte	+1 Byte	+0 Byte	-1 Byte
2	+0 Byte	+0 Byte	+2 Byte	+0 Byte
3	-50 Byte	-50 Byte	+5 Byte	+1 Byte
4	-75 Byte	-75 Byte	+10 Byte	+1 Byte

Table 2. Change in payload and CRC

5.3 Simulation Results

In figure 6, we show how the packet restore probabilities (PRP) varies over time for non-adaptive and adaptive MPDUs construction respectively. The simulated channel was repeated for both cases since we dealt with a finite window, not allowing statistical averaging of the channel. Here, we assumed that the probability of having bad channel state was 15%, i.e., on average the BER was 0.085 for 15% of the time of the total duration of time. It is evident from the figures, that the proposed adaptive scheme outperforms the non-adaptive scheme.

To have a better insight, we magnified a 10 msec window as shown in figure 7. It can be observed that the packet

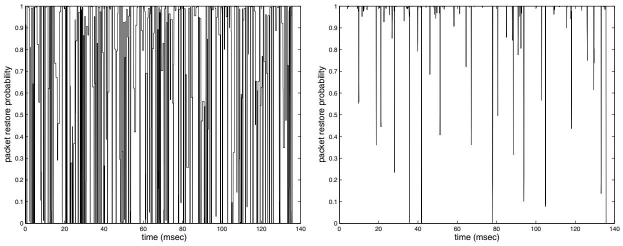

Figure 6. Packet restore probability for non-adaptive and adaptive schemes

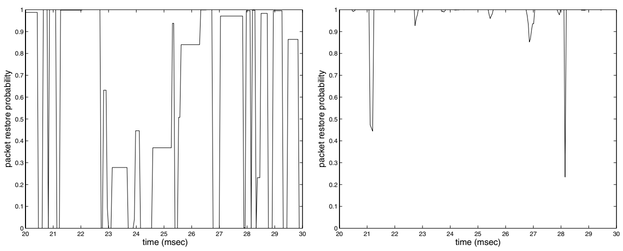

Figure 7. Magnification of 20-30 msec window for non-adaptive and adaptive schemes

restore probability is very low most of the times in non-adaptive scheme compared to the adaptive scheme where it is maximum almost all the time. Though there are drops (in 4 to 5 instances), they hardly reaches zero signifying that the packets *may* still be rectified.

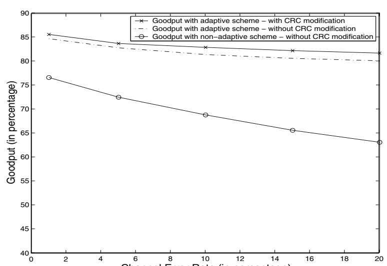

Figure 8. Goodput comparison

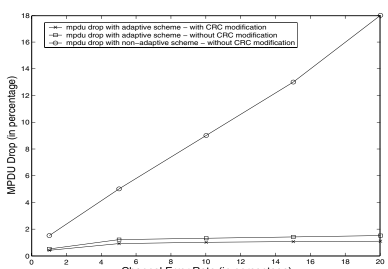

Figure 9. MPDU drop comparison

We compare the goodput and MPDU drop percentage for both schemes which are shown in figures 8 and 9. For the non-adaptive scheme, payload and CRC were kept fixed. In case of adaptive scheme, we first kept the CRC field fixed, modifying only the payload field. Then in the adaptive scheme, we changed both the payload and CRC field. From the figures, it is evident that, with the adaptation in both the payload and CRC field, the improvement is more.

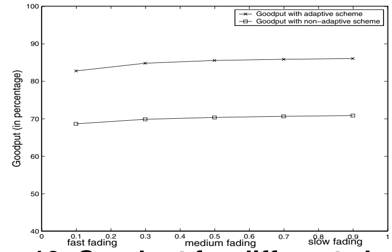

Figure 10. Goodput for different channels

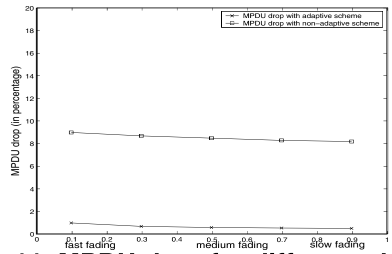

Figure 11. MPDU drop for different channels

In figures 10 and 11, we show the effect of fast, medium and slow fading on the goodput and MPDU drop probability for both schemes. Irrespective of the channel conditions, the adaptive scheme performs better.

6 Conclusions

In this paper, we studied the problem of streaming media over WiMAX and exploited the flexible features present in the MAC layer of 802.16a. We proposed that the size of MAC packet data units be made adaptive to the instantaneous wireless channel state condition. Based on the type of feedback received, variable size MPDUs were constructed either by aggregation or fragmentation of MAC service data units. We conducted simulation experiments to verify the validity of our proposed scheme. Packet restore probability, goodput, and dropping probability of MPDUs were defined as the performance metrics. Simulation results demonstrate the effectiveness and performance improvement of the proposed scheme.

References

[1] B. Sklar, Digital Comn., 2nd ed. Prentice Hall.

[2] IEEE Std. 802.16-2001 IEEE Standard for Local and MAN Part 16:Air Interface for Fixed Broadband Wireless Access Systems, IEEE Std 802.16-2001, 2002.

[3] IEEE Std. for Local and MAN Part 16:Air Interface for Fixed Broadband Wireless Access Systems-Amendment 2: MAC Modifications and Additional Physical Layer Spec. for 2-11 GHz, Std 802.16a-2003 (Amendment to IEEE Std 802.16-2001), 2003.

[4] I. Rhee, “Retransmission-based error control for interactive video applications over the Internet”, IEEE Proceedings of International Conference on Multimedia Computing and Systems, 1998, pp. 118-127.

[5] S.J. Vaughan-Nichols, “Achieving wireless broadband with WiMax”, IEEE Computer, Volume: 37, Issue: 6, June 2004, pp. 10-13.