

# VoIP Performance with Aggressive AMC and SINR Feedback for WiMAX Downlink

Xiangning Fan<sup>1</sup> and Zhu Dengkui<sup>2</sup>

<sup>1</sup> Institute of RF-&-OE-ICs, School of Information Science and Engineering,  
Southeast University, Nanjing, 210096, China

<sup>2</sup> National Mobile Communications Research Laboratory,  
Southeast University, Nanjing, 210096, China  
xnfan@seu.edu.cn

**Abstract.** WiMAX system is based on OFDMA, and it is very suitable for VoIP traffic. At the same time, aggressive AMC and frame bundling are two important techniques to improve the performance of VoIP, and so they are applied to WiMAX system in this paper. The simulation results show that when the system is in short of bandwidth, these two techniques indeed reduce the PER and delay jitter of the VoIP users, and also enhance the spectrum efficiency.

**Keywords:** WiMAX, VoIP, Aggressive AMC, Bundling.

## 1 Introduction

IEEE802.16e is put forward based on IEEE802.16d, and is modified to support the mobility of users. For the time being, WiBro (Wireless Broadband Service) in Korea and the rapidly growing WiMAX (World Interoperability for Microwave Access) are main commercial systems based on OFDM (Orthogonal Frequency Multiplexing Modulation). While VoIP (Voice over Internet Protocol) is important for future wireless mobile networks and there are many papers which have evaluated VoIP performance for 3GPP 1xEV-DO revision A mobile cellular system [1,2]. However, little evaluation has been made for WiMAX system. In this paper, both aggressive AMC (Adaptive Modulation and Coding) technique and voice frame bundling technique are applied to VoIP traffic in WiMAX system, and the analysis and simulations are given.

The organization of this paper is as follows: In Section 2, the system model of WiMAX downlink and the VoIP traffic model are given. In Section 3, voice frame bundling and aggressive AMC are discussed. Simulation conditions are summarized in Section 4. And in Section 5, the performance simulation results are presented and analyzed. Section 6 concludes the paper.

## 2 System Model

### 2.1 WiMAX Downlink Frame Structure

For the PMP (Point to Multi-Point) structure based WiMAX system, the resource information allocated to SS (Subscriber Station) in the DL (DownLink) and UL (UpLink) are broadcasted through MAP message by BS (Base Station). In each TDD (Time Division Duplex) DL frame structure of WiMAX system (see Fig. 1), the first OFDM symbol is the preamble, which is used for timing, frequency synchronization and channel estimation. Following the preamble is FCH (Frame Control Header) and two MAP messages: UL-MAP and DL-MAP. DL-MAP is composed of DL-MAP-IE, which informs SS the property of data such as the position in frame, size, and modulation and coding scheme etc. After receiving the DL-MAP, SS decodes the message and learns from this decoded message the corresponding properties, and thus SS could correctly receive the data transmitted by BS. In WiMAX system, data belong to different SSs but using the same modulation and coding scheme are filled together into a same burst.

In the DL frame, sub-carriers in OFDM symbols are arranged as permutation of PUSC (Partial Usage Sub-Carriers), thus the slots in DL frame is  $N_{slot} = (N_{sym} - 1) / 2 * N_{sch}$ , where  $N_{sym}$  and  $N_{sch}$  are the number of OFDM symbols and sub-channels in DL frame respectively. As compared to normal MAP message, compressed MAP message is also used for smaller size message in WiMAX.

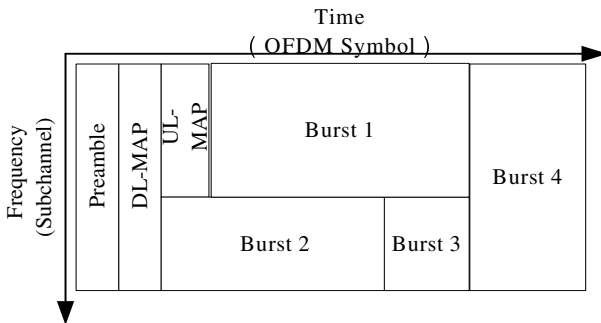


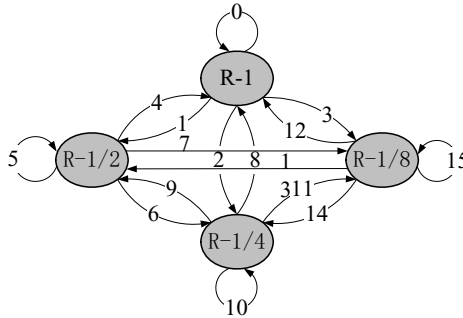
Fig. 1. DL Frame structure in WiMAX

### 2.2 VoIP Traffic Model in WiMAX

Typical phone talk is composed of active talking or talk spurt (ON) period and silence spurt (OFF) period. In the EVRC (Enhanced Variable Rate Codec) model [4], VoIP traffic arrivals at one of the four rates in each 20ms interval. Table 1 presents the value of the four data rates and their corresponding proportion. The state transfer of data transmission obeys Markov model as shown in Fig. 2.

**Table 1.** Parameters of VoIP traffic rate in WiMAX

Item	Parameter
Voice Active Factor	0.4
Full Rate (R-1)	29%
Half rate (R-1/2)	4%
Quarter Rate (R-1/4)	7%
Eighth Rate (R-1/8)	60%



**Fig. 2.** Markov model of state transfer of VoIP data rate in WiMAX

Various protocol headers (such as RTP, UDP, and IP headers) are added to the data output from voice codec to form a voice frame, and then the frame is sent to the MAC (Medium Access Control) layer of BS. In the VoIP frame (packet), the packet size of the protocol is 40Bytes, and after header compressing, it reduces to 2Bytes. Therefore, the average rate of VoIP traffic in MAC is 5.71kb/s, and the average frame (packet) size is 98bits for WiMAX downlink frame.

### 3 VoIP with Aggressive AMC and SINR Feedback in WiMAX

#### 3.1 Bundling of Voice Frame (Packet)

As we know, VoIP traffic arrivals in periods, and the data size is relatively small. If protocol header compression is used, the largest size of VoIP packet is 203 bits and the average is 98 bits for the EVRC VoIP model. In WiMAX system, VoIP has the highest priority, so almost each packet is scheduled timely and forms a single PDU (Protocol Data Unit). PDU in WiMAX system contains a MAC header and CRC (cyclic redundancy check), and their size is 48 bits and 32 bits respectively (see Fig. 3). If we consider the average size of VoIP packets, the proportion of overhead in a VoIP PDU is:  $(48+32) / (48+32+98)=45\%$ . So, quite a part of the bandwidth is used to transmit overhead which is useless for SSs, and therefore, the frequency efficiency is lowered. In order to increase the proportion of voice data in a VoIP PDU and make use of the bandwidth more efficiently, multiple VoIP frames

(packets) should be bundled to form a super VoIP frame (packet) in the CS (convergence sub-layer) layer of WiMAX, and this technique is usually called bundling of VoIP frame (packet) (see Fig. 4) [5,6].

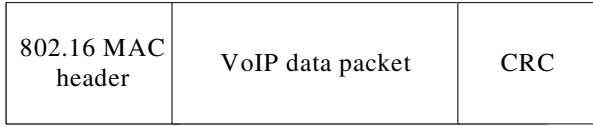


Fig. 3. PDU data packet in WiMAX system

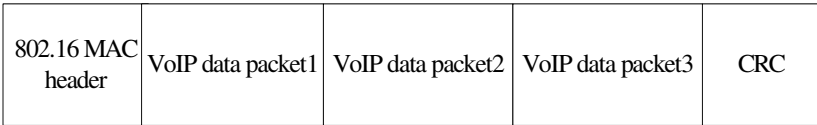


Fig. 4. Bundling VoIP packets in WiMAX system

Suppose that  $K$  VoIP packets form a super packet. Then, its average overhead ratio  $OHR_K$  can be expressed as follows:

$$OHR_K = \frac{L_{MacHead} + L_{CRC}}{L_{MacHead} + L_{CRC} + K * L_{Packet}} = \frac{80}{80 + K * L_{Packet}} \tag{1}$$

where,  $L_{MacHead}$  is the MAC header size,  $L_{CRC}$  is the size of CRC in PDU,  $L_{Packet}$  is the average VoIP packet size (i.e., 98 bits in this paper). Suppose the average bandwidth (Bw) usage rate is  $R_1$  when no bundling is used, then  $R_k$  of  $K$  VoIP packets bundling should satisfy:

$$R_1 \geq R_K \geq R_1(1 - OHR_1)/(1 - OHR_K) \tag{2}$$

According to (1) and (2), as the number  $K$  increases, i.e., as more packets are bundled together to form a super packet, the proportion of the overhead decreases gradually, and the average bandwidth efficiency is boosted consequently. However, one problem comes forth at the same time: the delay of the super VoIP packet also becomes larger. As the voice packets is generated each 20ms in the voice codec, if two packets are bundled, the first packet in the super VoIP packet has already been delayed 20ms when it enters the MAC of BS. So for  $K$  VoIP packets bundling, the first packet in the super VoIP packet will be delayed  $20(K - 1)$  ms, thus little time will be left for MAC scheduler because VoIP traffic is strictly time constrained. As a result, in order to guarantee the QoS (Quality of Service) of end to end delay for each VoIP packet in a super VoIP packet, delay sensitive MAC scheduling algorithm must be used.

### 3.2 Feedback of SINR

WiMAX is a system using multiple carriers. Data of each SS are modulated on multiple sub-carriers. At the receiver, the  $i^{\text{th}}$  SS first calculates SINR (signal to interference and noise ratio) of each sub-carrier, then, the SS maps all of the SINR to an effective value  $SINR_i(t)$  using EESM (Exponential Effective SNR Mapping) [7, 8]. The effective value  $SINR_i(t)$  is filtered through a smooth filter and finally the average effective SINR  $\overline{SINR_i(t)}$  can be gotten as following:

$$\overline{SINR_i(t)} = (1 - \alpha)\overline{SINR_i(t-1)} + \alpha SINR_i(t) \tag{3}$$

where  $1/32 \leq \alpha \leq 1$ .  $\overline{SINR_i(t)}$  is feed back to BS through CQICH (Channel Quality Identity Channel) in period, and BS uses this information as an input parameter to perform AMC and scheduling.

### 3.3 Aggressive AMC

For VoIP user in WiMAX system, the aggressive AMC [9] is implemented as follows: BS first gets the average effective SINR ( $\overline{SINR_i(t)}$ ) from the CQICH, and then it adds an aggressive factor  $\beta$  ( $\beta > 0$ ) to the  $\overline{SINR_i(t)}$ , i.e., :

$$SINR_i(t)_{AMC} = \overline{SINR_i(t)} + \beta \tag{4}$$

where  $SINR_i(t)_{AMC}$  is used to perform AMC in the following way. BS first looks up the BLER-SINR (Block Error Ratio-SINR) curve, and then accordingly determines the highest data rate MCS (modulation and code scheme) that this  $SINR_i(t)_{AMC}$  could support with the constraint of BLER satisfying less than 1%. And the highest MCS is selected as the MCS of the user.

Because of the aggressive factor  $\beta$ , the MCS of VoIP user may be a little higher than what should be for the user’s channel condition. However, if we select the  $\beta$  elaborately, system spectral efficiency could be enhanced, while HARQ (Hybrid Automatic Repeat-Request) retransmission will not largely occur for the wrongly received data at the first transmission.

### 3.4 Scheduling Algorithm

Ref. [10] presents an M-LWDF (modified largest weighted delay first) scheduling algorithm to calculate the priority of SS, for any  $\gamma_i > 0$ , the formula of priority is:

$$s = \arg \max_i \{ \gamma_i D_i(t) W_i(t) \} \tag{5}$$

where  $\gamma_i$  is a parameter represents traffic priority. Ref. [10] has proved that the scheduling algorithm is throughput optimal, and it gives an optimal choice of  $\gamma_i$ . For only VoIP traffic is considered in this paper, so  $\gamma_i$  is neglected.  $D_i(t)$  represents the

maximum data transmitting rate of  $i^{\text{th}}$  SS;  $W_i(t)$  is the delay of data belonged to  $i^{\text{th}}$  SS in BS MAC buffer. Generally,  $D_i(t)$  and  $W_i(t)$  are normalized in practical system, and because only VoIP traffic is considered here, for each service flow, (5) could be rewritten as follows:

$$\text{Priority}(i) = \frac{D(i)}{D_i(t)} \frac{W_i(t)}{T_i} \quad (6)$$

where,  $\overline{D_i(t)}$  is the average data rate of  $i^{\text{th}}$  SS,  $T_i$  is the maximum delay SS could endure which is thought to be the maximum delay of VoIP MAC. Because the maximum data rate of SS is closely correlative to its channel condition, we use  $\frac{SINR(i)}{SINR_i(t)}$  to replace  $\frac{D(i)}{D_i(t)}$  and finally get:

$$\text{Priority}(i) = \frac{SINR_i(t)}{SINR_i(t)} \frac{W_i(t)}{T_i} \quad (7)$$

This algorithm tries to get a balance between the delay performance of data and the bandwidth spectral efficiency. It can maximize the bandwidth spectral efficiency when it satisfies the delay requirements of user's VoIP data. Therefore, it is a proper choice for real time VoIP traffic.

## 4 Simulation Configurations

### 4.1 Simulation Parameters

In this paper, three layered 19 cells 57sectors cellular model is used to simulate practical WiMAX wireless network. Each SS is uniformly located in the whole simulation area. Each SS in the center 3 sectors is static only, and such assumption doesn't affect the final results [11]. Because 1:1 frequency reusing scheme is used, so each SS is interfered by all other cells. Other basic simulation parameters are listed in Table 2.

### 4.2 Simulation Statistics

Statistic variables used in our simulation include PER (Packet Error Rate), average delay, average delay jitter, and bandwidth (Bw) usage rate.

When counting PER, error packets include: (1) the packet dropped in the BS MAC buffer because its maximum delay expires; and (2) the packet not being received rightly after maximum HARQ retransmission.

VoIP delay is the time between VoIP packets entering the convergence sub-layer, and it is rightly received by SS [8].

Bandwidth usage rate is the ratio of bandwidth used in fact with that allocated.

**Table 2.** Simulation parameters

Item	Parameter	Item	Parameter
Cell radius	866m	smooth factor ( $\alpha$ )	0.25
BS transmission power	43dBm	HARQ mode	Chase combing
Path loss model	Costa231	ACK feedback delay	4frames
System bandwidth	10MHz	Maximum HARQ retransmission	3
Frequency reuse factor	1:1	Channel model	ITU PedA ITU Veh B & Veh A
Sub-carrier permutation mode	PUSC	Modulation scheme	QPSK, 16QAM, 64QAM
Symbols in DL frame	24	Channel code	Convolutional turbo code (CTC)
CQICH feedback period	3frames	Code rate	1/2, 2/3, 3/4, 5/6

## 5 Results and Analysis

### 5.1 Performance of Frame Bundling for VoIP in WiMAX

Fig. 5 to Fig. 8 respectively shows the performance of PER, average delay, average jitter, and average bandwidth usage ratio with different bundling number  $K$ . At a certain  $K$  value, three cases are evaluated:

- (1) Bw is enough (120 users each sector);
- (2) Bw is proper (135 users each sector);
- (3) Bw is deficient (150 users each sector).

Fig. 5 shows the PER performance with different bundling number  $K$  and bandwidth (or user number). When Bw is enough, i.e., 120 users in each sector, the PER of no bundling is lower than that of bundling policy being used. But the case changes as the user number increases and Bw becomes deficient. For 135 users in each sector case (i.e., Bw is proper), when the maximum delay is set as 110ms, the PER of 2 VoIP packets bundling is close to that of no bundling. As for 150 users in each sector case (i.e., Bw is deficient), when the maximum delay is set as more than 90ms, the PER performances of bundling 2 or 3 VoIP packets become better than that of no bundling being used. The reason is that when Bw is not enough and bundling is not used, more Bw is occupied by overhead, and as a result, some packets are dropped for expiring the maximum delay, and PER increases rapidly as compared with that of bundling used.

In contrast, if bundling is used, overhead is reduced and more Bw is used for data transmission which improve the PER performance when Bw is not enough. In this case, even if the user number increases, the Bw usage rate is smaller than 1. The main reasons of PER increasing in this case are: (1) as PDU size is bigger for bundling being used, the probability of errors increase; and (2) as the bundling adds a fixed delay to VoIP packet, the HARQ retransmission time of a super VoIP packet is reduced with the strict delay constrain, which increases the probability of those packets being wrongly received.

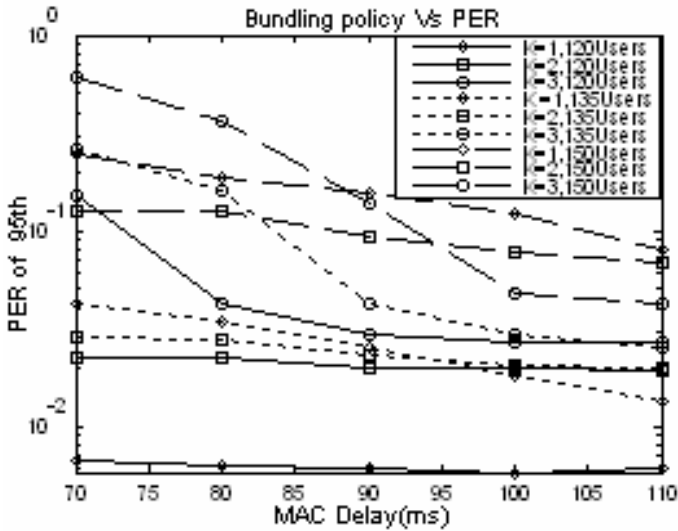


Fig. 5. SS number, bundling policy Vs PER

Fig. 6 shows the average delay performance with different bundling number  $K$  and bandwidth (or user number). When bundling  $K$  VoIP packets is used, Fig. 6 shows that the delay of super VoIP packet is always longer than that of no bundling being used. This is intuition because the first packet in each super VoIP packet has been delayed ( $20(K-1)$  ms) when it enters the MAC layer of BS. But the delay gap between no bundling and with bundling gradually decreases as the user number in

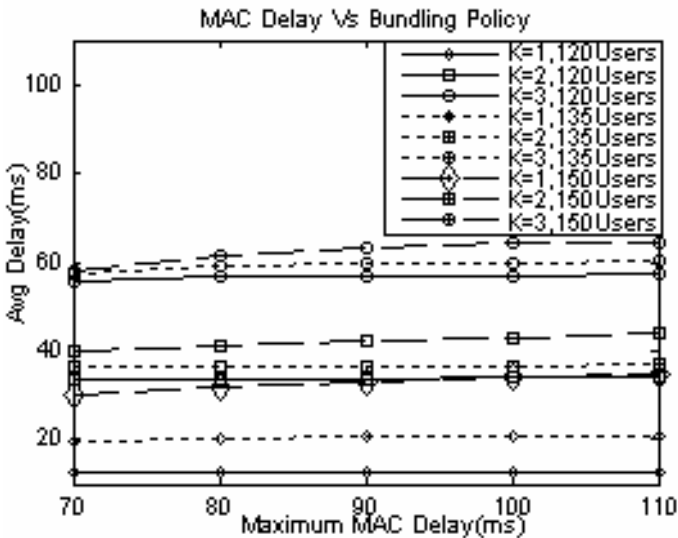


Fig. 6. SS number, bundling policy Vs average delay



each sector increases, and it indicates that super VoIP packets stay with a relatively short time in MAC layer.

Fig. 7 shows the average jitter performance with different bundling number  $K$  and bandwidth (or user number). When the user in each sector is 120 (i.e., Bw is enough), the jitter of no bundling is obviously smaller than that of bundling being used. However, as the user number increases to 135 (i.e., Bw is proper), the jitter of bundling 2 packets becomes smaller than that of no bundling used, and when the user number further increases to 150 (i.e., Bw is deficient), the jitters of bundling 2 and 3 packets are both smaller than that of no bundling used.

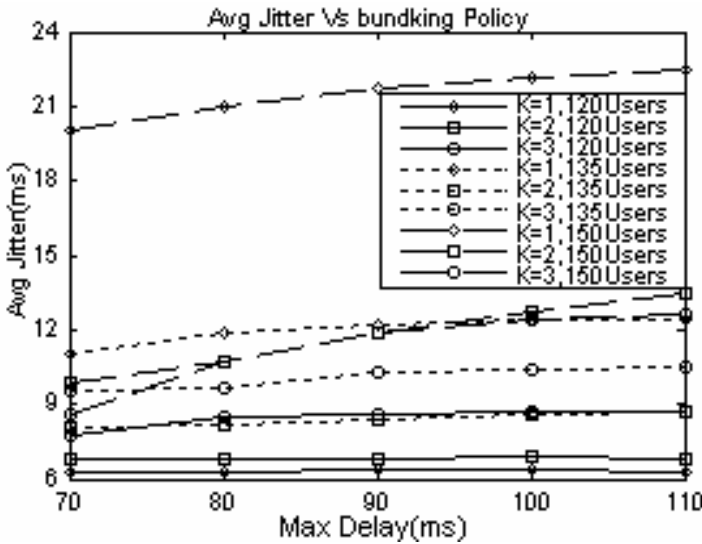


Fig. 7. SS number, bundling policy Vs average jitter

Fig. 8 shows the average Bw usage rate performance with different bundling number  $K$  and bandwidth (or user number). The average Bw usage rate of no bundling, 2 packets bundling and 3 packets bundling are  $OHR_1 = 45%$ ,  $OHR_2 = 29%$ , and  $OHR_3 = 21%$  respectively. According to the simulation results (see Fig. 8), when the number of users in each sector is 120 (i.e., Bw is enough), Bw usage rate are  $R_1 = 89%$ ,  $R_2 = 68%$  and  $R_3 = 60%$  for no bundling, 2 packets bundling and 3 packets bundling respectively. When the number of users in each sector increases to be 135 (i.e., Bw is proper), Bw usage rate are improved to be  $R_1 = 99%$ ,  $R_2 = 71%$  and  $R_3 = 79%$  for no bundling, 2 packets bundling and 3 packets bundling respectively. And when the number of users in each sector is 150 (i.e., Bw is deficient), Bw usage rate are further improved to be  $R_1 = 100%$ ,  $R_2 = 92%$  and  $R_3 = 83%$  accordingly for no bundling, 2 packets bundling and 3 packets bundling respectively.

From the above results, bundling 2 packets is reasonable for predicting the Bw usage rate, and the simulation results also justify that VoIP packet bundling could save much Bw.

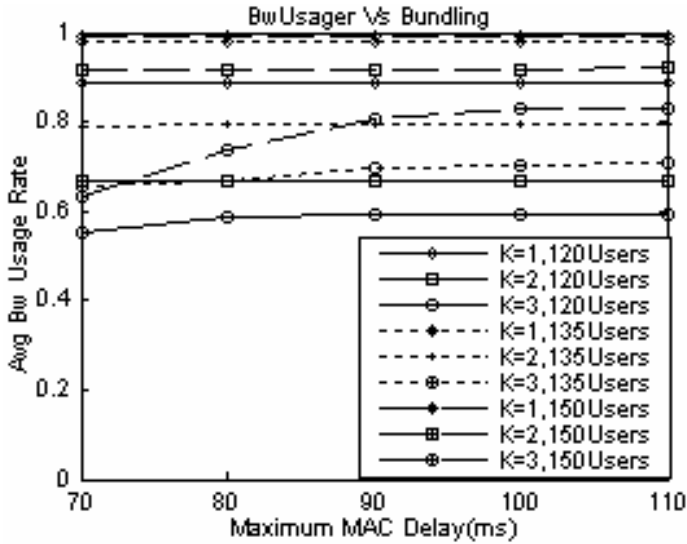


Fig. 8. SS number, bundling policy Vs average Bw usage rate

From the above simulation results and analysis, we can conclude that the bundling of VoIP packets not only lowers the PER and average jitter, but also saves much Bw and enhances the Bw spectral efficiency. The only shortage of frame (packet) bundling is that it may cause relatively longer delay. Combing all the above considerations and analysis, bundling 2 VoIP packets is a proper choice for VoIP traffic in WiMAX.

### 5.2 Performance of Aggressive AMC for VoIP in WiMAX

Fig. 9 to Fig. 12 respectively shows the performance of PER of 95% users, average delay, average jitter, and average spectral efficiency with different aggressive factor  $\beta$  and maximum delay constraint.

Fig. 9 indicates when the aggressive factor  $\beta$  increases from 0dB to 3dB, the PER decreases at first and begins to increase after coming to a minimum. Average delay (see Fig. 10), and average jitter (see Fig. 11) show similar variety trends while the average spectral efficiency shows a reverse trend (see Fig. 12).

As the maximum delay becomes larger, the reduction of PER of aggressive AMC becomes smaller. The reason is that if maximum delay constraint is relatively small, the majority of the error packets are those that are dropped because of maximum delay expiration. The SINR threshold of neighbour MCS is usually about 2-3dB, so if aggressive AMC is used and proper  $\beta$  (0.5dB) is selected, most users will keep their

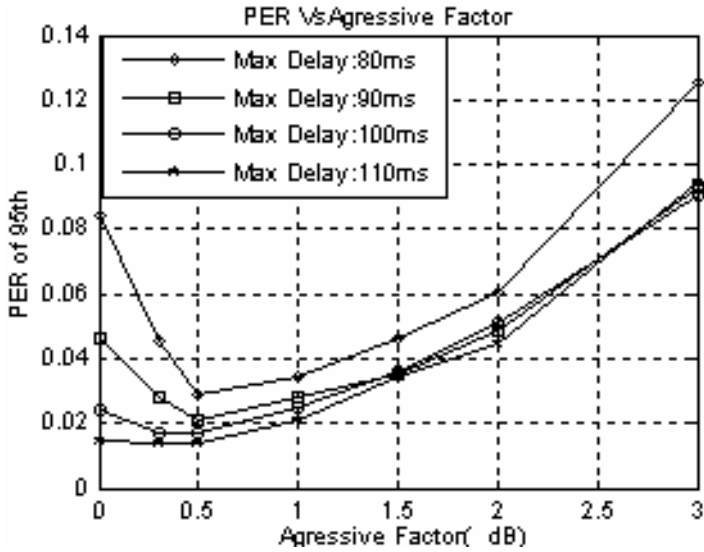


Fig. 9. Aggressive AMC performance: PER of 95% users

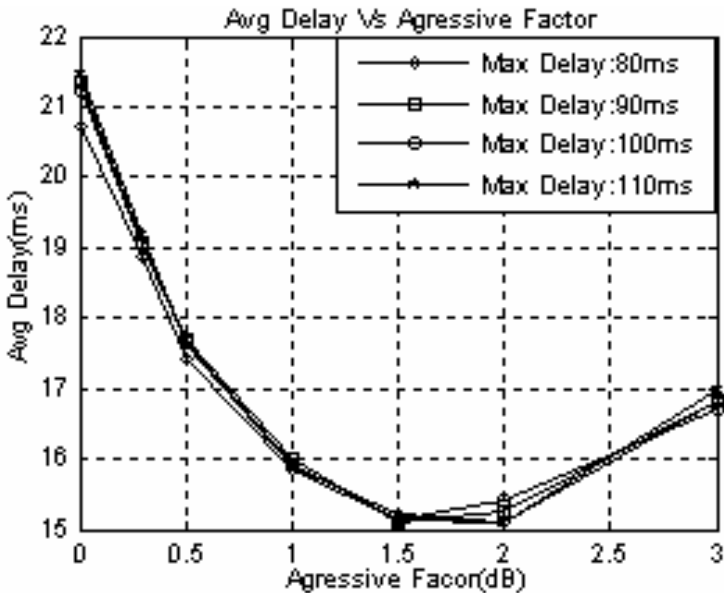


Fig. 10. Aggressive AMC performance: average delay

MCS unchanged while only those whose SINR are already close to the next higher SINR threshold will chosen a higher MCS, and they will use little bandwidth than when aggressive AMC is not used. As a result, the VoIP packets of all users could be scheduled in a relatively shorter time and little of the packets will be dropped, and

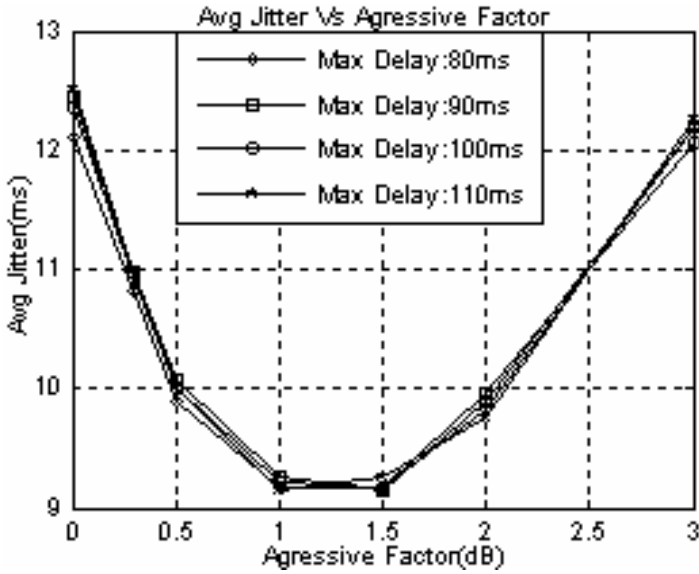


Fig. 11. Aggressive AMC performance: average jitter

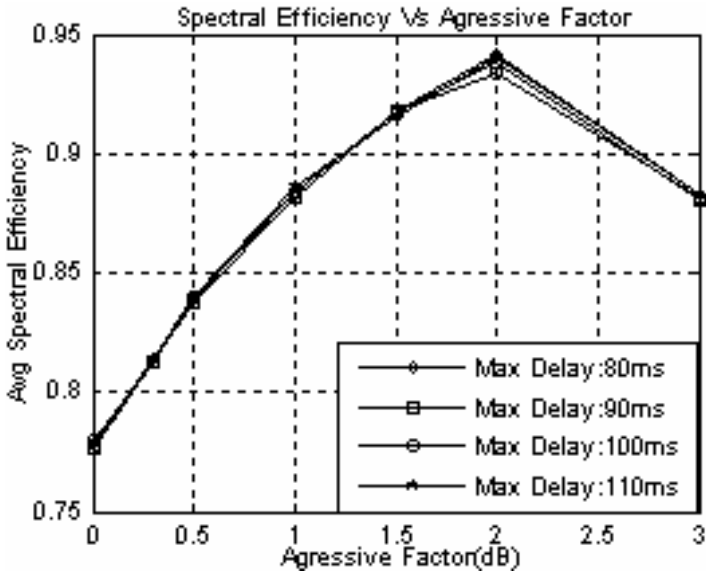


Fig. 12. Aggressive AMC performance: average spectral efficiency

thus the system spectral efficiency is enhanced. At the same time, their SINR are close to next higher threshold, so the transmit error probability is low when a higher MCS is selected and its number is smaller than that of dropped because of maximum delay expiration when aggressive is not used. However, when a larger maximum delay

constraint is selected, aggressive AMC brings a few benefits. So, the main factors which restrict the system performance are users' channel conditions and their corresponding MCS.

If a too large  $\beta$  (3dB for example) is selected, the system performance becomes bad. The PER increases for majority of users selects a higher MCS which couldn't be supported at its channel condition and much HARQ retransmission is need. As a result, packet delay and jitter increase, bandwidth resource is wasted, and spectral efficiency decreases.

Furthermore, when  $\beta$  increases to the point where PER comes to the minimum, the average delay and jitter still decrease, which indicates that we may get the best delay and jitter performance at the cost of a relatively higher PER.

According the above analysis, a proper selected aggressive factor is important to VoIP traffic in WiMAX system.

## 6 Conclusions

This paper mainly studies and analyses how the two techniques of frame (packet) bundling and aggressive AMC effect the performance of VoIP in the downlink of WiMAX system. The simulation results indicates that if these techniques are used properly, the performance of VoIP could be enhanced to a certain degree. The results show that when the system is in short of bandwidth, these two techniques indeed reduce the PER and delay jitter of the VoIP users, and also enhance the spectrum efficiency.

**Acknowledgments.** This paper is supported by the National High Technology Research and Development Program of China (863 Program) No.2007AA01Z2A7.

## References

1. 3GPP (3rd Generation Partnership Project) TS 25 Series: Technical Specification Group Radio Access Network (2006)
2. 3GPP Technical Report (TR) 25.848: Physical Layer Aspects of UTRA High Speed Downlink Packet Access (Release (2000)
3. IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands. IEEE Press, Los Alamitos (December 2005), <http://ieee802.org/16>
4. TIA/EIA/IS-127: Enhanced Variable Rate Codec, Speech Service Option 3 for Wideband Spread Spectrum Dgital Systems (1996)
5. Zheng, H., Rittenhouse, G.: Providing VoIP Service in UMTS-HSDPA with Frame Aggregation. In: IEEE ICASSP, March 2005, vol. 2, pp. 1157–1160 (2005)
6. Zheng, H., Rittenhouse, G., Recchione, M.: The Performance of Voice over IP over 3G Downlink Shared Packet Channels under Different Delay Budgets. In: IEEE VTC2003-Fall, October 2003, vol. 4, pp. 2501–2505 (2003)

7. Motorola: EESM Link Error Prediction for EUTRA System Evaluation. 3GPP TSG RAN WG1 #42, R1-050718 (2005)
8. Liu, H., et al.: EESM Based Link Error Prediction for Adaptive MIMO OFDM System. In: IEEE VTC2007-Spring, pp. 559–563 (April 2007)
9. Talukdar, A., et al.: Aggressive Modulation/Coding Schedule for Maximum System Throughput in a Multi-Carrier System. In: IEEE Int. Conf. Commun., vol. 3, pp. 1823–1828 (2002)
10. Kumaran, M., et al.: Providing Quality of Service Over a Shared Wireless Link. IEEE Comm. Mag. 39, 150–154 (2001)
11. Srinivasan, R., Zhuang, J., Jalloul, L.: Draft IEEE 802.16m Evaluation Methodology. IEEE 802.16 Broadband Wireless Access Working Group (August 2007)