

A New Resource Allocation Scheme for IEEE 802.16-based Networks

Chung-Wei Lin[†], Yu-Cheng Chen[†], Ai-Chun Pang^{†‡}

[†]Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan

[‡] Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei 106, Taiwan

Email: {b90100, b89100, acpang}@csie.ntu.edu.tw

Abstract—IEEE Standard 802.16 sets the stage for broadband wireless access system. This standard includes the medium access control layer (MAC) and physical layer with higher data rates and longer transmission range compared to those supported by the existing wireless technologies. It also provides QoS (Quality of Service) for the applications such as Voice over IP, video streaming and high bandwidth file transfer. With the ability of broadband wireless access of an IEEE 802.16 system, a proper resource allocation scheme for packet transmissions is imperatively needed. In this paper, we present a new resource allocation scheme, called *extra bandwidth granting* (EBG), to improve transmission efficiency of an IEEE 802.16-based network. Our proposed scheme can be adopted along with the existing scheduling algorithms and the multi-priority scheme without any modification. The experimental results show that by using our EBG, the packet queuing delay could be significantly improved, especially for the service flows of higher-priority classes.

I. INTRODUCTION

IEEE 802.16, also called worldwide interoperability for microwave access (WiMAX), is a standard for metropolitan area networks (MANs). IEEE 802.16 [1], [2] originally specifies the air interface, including the medium access control (MAC) layer and physical layer, of a fixed point-to-multipoint broadband wireless technology. It is well suited for point-to-multipoint data transmission with data rates up to 120 Mbps. In addition to fixed wireless accesses, it also specifies mobility functions to support mobile users. Figure 1 shows a typical architecture for a WiMAX-based network [3] where fixed subscriber stations (SSs) and mobile subscriber stations (MSSs) communicate with base stations (BSs) via air interfaces. In this paper, we focus on fixed point-to-multipoint wireless access, and mobility issues for MSSs are not considered.

In IEEE 802.16, BSs are responsible for allocating radio resources for SSs under its covering area. With the ability of broadband wireless access of IEEE 802.16 networks, a proper resource allocation scheme for packet transmissions is imperatively needed, which has a great influence on system performance and quality of service (QoS) provided by IEEE 802.16 networks. In several papers [4], [5], several scheduling algorithms were proposed, such as adaptive uplink and downlink bandwidth adjustment, to achieve higher transmission performance. Even with these methods and algorithms, the packet queuing delay and radio link utilization of an IEEE 802.16-based network cannot be greatly improved due to its frame structure and bandwidth requesting/granting procedure (the details will be described in the following sections). To deal

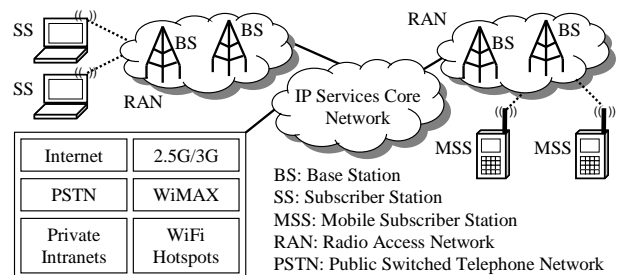


Figure 1. The architecture of a WiMAX network.

with this problem and make the best use of precious resources in IEEE 802.16-based wireless environments, we propose a new resource allocation scheme, called *extra bandwidth granting* (EBG), to boost the performance of a IEEE 802.16 system. The EBG scheme can work well with any type of the physical layer design and the existing scheduling algorithms without any modification of IEEE 802.16 protocols.

The remainder of this paper is organized as follows. The MAC protocol of IEEE 802.16 is introduced in Section II. In Section III, we define the problem under investigation, and present the *extra bandwidth granting* (EBG) to improve transmission efficiency for WiMAX. Section IV presents our developed analytical and simulation models, and summarizes the experimental results to demonstrate the capability of the proposed scheme. Section V is the conclusion.

II. IEEE 802.16 MEDIUM ACCESS CONTROL

The design of IEEE 802.16 MAC is similar to that of IEEE 802.14, the Data Over Cable Service Interface Specifications (DOCSIS) [6]. The standard defines “downlink” (DL) as the transmission direction from a BS to SSs and “uplink” (UL) as the transmission direction from SSs to a BS. It also provides two modes of duplexing, including the frequency division duplexing (FDD) and the time division duplexing (TDD).

The downlink subframe structure is shown in Figure 2 (a). It starts with a frame control section [6] that includes a downlink map message (DL-MAP) and an uplink map message (UL-MAP). The DL-MAP and UL-MAP contain information that describes the usage of a specific interval in the downlink and uplink frame, respectively. Following the frame control section, the time division multiplexing (TDM) portion is used to broadcast data from the BS to SSs. Figure 2 (b)

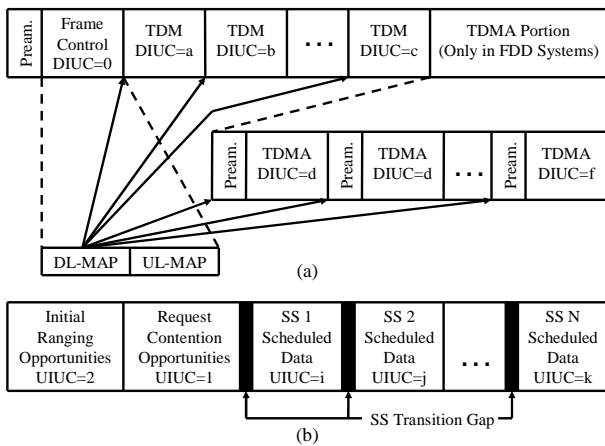


Figure 2. (a) The downlink subframe structure and (b) the uplink subframe structure. The Downlink/Uplink Interval Usage Code (DIUC/UIUC) is a code identifying a particular burst profile that can be used by a downlink/uplink transmission interval.

shows the uplink subframe structure, which indicates that the uplink subframe starts with contention-based initial ranging opportunities and request contention opportunities. Following these contention-based opportunities, there are the intervals for SSs to transmit data. Unlike the downlink subframe, each interval is assigned to a specific SS. The SSs transmit data in their respective allocation based on the previously received UL-MAP in the downlink subframe.

For the uplink scheduling services, Four QoS uplink service classes are defined, and provided by IEEE 802.16 BSs.

- Unsolicited Grant Service (UGS) is used to support real-time service flows that generate fixed size data packets on a periodic basis, such as Voice over IP applications without silence suppression.
- Real-Time Polling Service (rtPS) is used to support real-time service flows that generate variable size data packets, such as MPEG video.
- Non-Real-Time Polling Service (nrtPS) is used to support non-real-time service flows that require variable-size transmissions, such as high bandwidth FTP.
- Best Effort (BE) is used to provide best-effort services.

III. EXTRA BANDWIDTH GRANTING

A. Problem Formulation

The formulated frame structure is shown in Figure 3. In a TDD system, a frame with a fixed length is partitioned into two subframes, i.e. the downlink subframe and the uplink subframe. The downlink subframe starts with the DL-MAP and UL-MAP describing the upcoming downlink and uplink subframes. The remaining downlink and uplink subframes are used for the data transmission according to the DL-MAP and UL-MAP, respectively. It should be mentioned that the bandwidth request of each service flow is transmitted to the BS in the uplink subframe. For simplification, we focus on the uplink channel allocation and assume that any request or data of each service flow is transmitted successfully without error

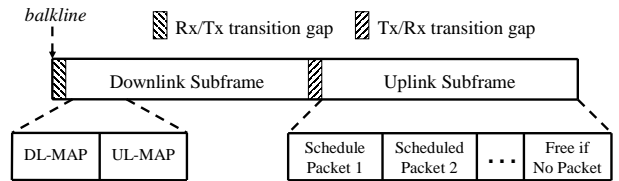


Figure 3. The formulated frame structure.

occurrence. Besides, we define “*balkline*” as the beginning of a frame, which will be used in the following paragraphs.

In a typical WiMAX system, there are several uplink service flows which generate packets to be transmitted. An uplink service flow belongs to one of uplink service classes, i.e. UGS, rtPS, nrtPS and BE. The packet size and packet inter-arrival time of a service flow are random variables with specific distributions. The BS grants the bandwidth for SSs via the UL-MAP carrying the scheduling information before the *balkline*. We try to minimize the *packet queuing delay* defined as the total time that a packet waits in a queue until it could be transmitted over the radio link. Note that if the subframe is not enough for all packets transmission, the remaining packets will be scheduled to the next subframe.

B. Extra Bandwidth Granting

SSs requests the bandwidth for packets that arrive before the *balkline*, and the BS grants the bandwidth via the UL-MAP. In this way, SSs transmit data to the BS according to the UL-MAP. Then the queuing delay for each packet consists of two periods:

- The first period, which starts from the arrival of the packet to the next coming *balkline*, may contain the downlink subframe and the uplink subframe.
- The second period starts from the next coming *balkline* and ends at the beginning of data transmission if the bandwidth is enough for the data transmission. In contrast, if the bandwidth is not enough, the packet will wait for the next frame until there is sufficient resource for data transmission.

In the first period, a free interval may exist because the bandwidth is enough for the current frame. In the current version of IEEE 802.16 standard, even with a free interval, an arriving packet could not be transmitted during the interval because the BS has not granted the bandwidth for it. Thus we propose *extra bandwidth granting* (EBG) to make the best use of these free intervals to minimize the queuing delay of packets (especially for high-priority pack transmission) and enhance the performance of an IEEE 802.16 system.

Figure 4 shows an example of the uplink subframe structure for the scenario of our *extra bandwidth granting* scheme. When the BS has scheduled the packets of the current frame according to the requests of service flows, there may be some bandwidth remaining at the end of the downlink subframe and uplink subframe. This scheme is triggered by the BS to schedule transmissions of these free intervals and put the scheduling information into the UL-MAP. If a service flow

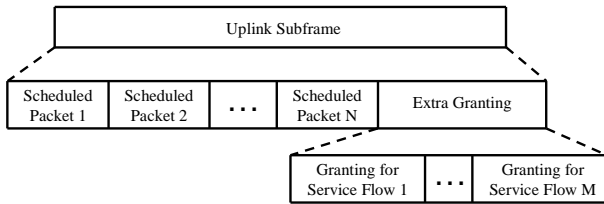


Figure 4. The uplink subframe structure for EBG.

has some extra bandwidth and some packets which have arrived and not yet be considered into requesting, the packets will be transmitted by the extra granting time slots. If the bandwidth is not enough for the packets which have requested the bandwidth, this scheme will not be activated.

Because packets which obtain bandwidth via normal request procedure are surely in the queue, they are scheduled in the front of those extra granting time slots. These packets may be scheduled by the first-come-first-serve (FCFS) scheduling or the shortest-job-first (SJF) scheduling. Because neither of the two scheduling algorithms the uplink QoS class of service flows, a multi-priority scheme is usually incorporated to differentiate the transmission levels for different classes of service flows. Note that all these algorithms can work well with our *extra bandwidth granting* scheme.

C. Bandwidth Arrangement of Extra Bandwidth Granting

Regarding detailed arrangements of our *extra bandwidth granting*, how to adequately divide the bandwidth for the service flows is important, which will be discussed as follows.

- *Equality* method is to divide the bandwidth equally as shown in Figure 5 (a) and (b). If there are M uplink service flows and the current frame has the capacity of B bytes to transmit more packets, each uplink service flow will get more B/M bytes than that it requested.
- *Proportion* method is to divide the bandwidth according to the request of each service flow during the current frame as shown in Figure 5 (c) and (d). If the M uplink service flows respectively request R_1, R_2, \dots, R_M bytes and $R = R_1 + R_2 + \dots + R_M$, then they will get more $BR_1/R, BR_2/R, \dots, BR_M/R$ bytes than that they requested.

In Section IV, the experimental results for the *Equality* and *Proportion* methods will be shown and compared.

D. Scheduling of Extra Bandwidth Granting

Another problem of the *extra bandwidth granting* is how to schedule these extra time slots for the service flows with different QoS classes. Two alternatives will be presented as follows.

- *Non-Inversion* Method:
The packets of the high-priority service flows, e.g. the UGS class, are allocated from the beginning of the extra time slots as shown in Figure 5 (a) and (c). For a service flow of the UGS class, if a packet is transmitted in those extra time slots, the earlier transmission can result in a

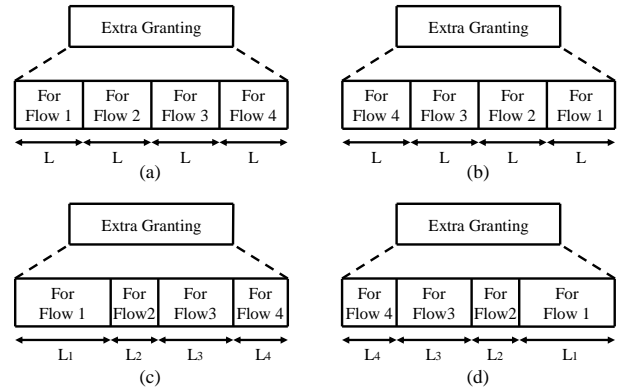


Figure 5. Suppose that there are four service flows, numbered from 1 to 4, with the class UGS, rtPS, nrtPS and BE respectively. The detailed arrangements of EBG when using (a) the *Equality* and non-*Inversion* methods, (b) the *Equality* and *Inversion* methods, (c) the *Proportion* and non-*Inversion* methods and (d) the *Proportion* and *Inversion* methods.

smaller delay for this packet. For this reason, the non-*Inversion* method is proposed.

- *Inversion* Method:

The packets of the high-priority service flows are allocated from the end of the extra time slots as shown in Figure 5 (b) and (d). For a service flow of the UGS class, the later extra time slots in the *extra bandwidth granting* could result in more packets being transmitted in those extra time slots. For this reason, the *Inversion* method is proposed. The details will be discussed in the following paragraphs.

In a normal situation, for a service flow of the UGS class, the non-*Inversion* method schedules the extra time slots of the service flow earlier than the *Inversion* method. An instinct is that the non-*Inversion* method can result in a smaller delay for the service flow. However, when using the *Inversion* method, the queuing delay of the service flow is possibly smaller than that the non-*Inversion* method results in. In Section IV, the experimental results will be shown and discussed.

IV. NUMERICAL RESULTS AND DISCUSSION

A. Experimental Parameters

The model of an IEEE 802.16 network is implemented in an event-driven simulation, and its assumptions are described as follows.

- The transmission rate of an uplink channel is set to 100 Mb/s. Frame duration is set at 5 msec, including 0.2 msec for transition gaps, 2.4 msec for downlink subframe and 2.4 msec for uplink subframe.
- There are four packet generators as four service flows existing in the system. Except for the UGS class, the packet inter-arrival times of other three classes are modeled as exponential distributions with the rate—3.4965 packets per msec. In average, half of the system workload results from packet arrivals of these three classes. The packet inter-arrival times of the UGS class are constant, and they can be 0, 5, 10, 15 and 20 packets per msec

Table 1

The comparison of queuing delays (in millisecond) when using the FCFS/SJF scheduling with/without QoS consideration

System Workload	Service Class	FCFS w/o QoS	SJF w/o QoS	FCFS with QoS	SJF with QoS
0.5	UGS	—	—	—	—
	rtPS	5.723	5.440	5.325	5.229
	System	5.724	5.433	5.724	5.628
0.6	UGS	5.784	5.317	5.173	5.173
	rtPS	5.840	5.593	5.564	5.471
	System	5.822	5.500	5.708	5.645
0.7	UGS	5.954	5.501	5.356	5.356
	rtPS	5.965	5.753	5.804	5.708
	System	5.960	5.626	5.791	5.743
0.8	UGS	6.062	5.600	5.447	5.448
	rtPS	6.093	5.917	6.042	5.949
	System	6.075	5.730	5.863	5.826
0.9	UGS	6.285	5.829	5.600	5.601
	rtPS	6.313	6.162	6.278	6.187
	System	6.295	5.947	6.029	5.998

to reach system workloads to 0.5, 0.6, 0.7, 0.8 and 0.9, respectively.

- The packet sizes of rtPS, nrtPS and BE service flows are modeled with Pareto distribution [7], and the UGS packet size is constant, i.e., 120 bytes according to the G.711 standard. In the Pareto distribution, we set the parameters for α , the shape and the maximum packet size as 81 bytes, 1.1 and 1500 bytes, respectively.

The simulation time is 10 seconds. The results are the average of 30 simulation rounds to achieve that 99% confidence intervals for all measured results do not exceed 1% of the measured values. Another experiment with an unstable channel condition is also finished. The results are similar to those in the following paragraphs.

B. FCFS and SJF Scheduling without QoS Considerations

Our proposed EBG scheme can work with any of the existing scheduling algorithms such as FCFS and SJF without any modification. It is well-known that the SJF scheduling gives the optimal solution to minimize the system queuing delay (average of all packets). The third and fourth columns of Table 1 show the comparison of queuing delays without QoS considerations between the FCFS scheduling and the SJF scheduling. Due to the lack of space, only the queuing delays of the UGS class, the rtPS class, and the system are listed. These values indicate the system queuing delay of SJF is smaller than that of FCFS, validating the analytic results. Among them, the UGS class produces a greater improvement because the average packet sizes are smaller than the sizes of other classes.

C. FCFS and SJF Scheduling with QoS Considerations

For QoS considerations, a multi-priority scheme is implemented. If two packets are in different classes, the one with the higher priority is scheduled before the other. If two packets are in the same class, the FCFS scheduling and the SJF scheduling are both applicable, but the multi-priority scheme

Table 2

The comparison of system queuing delays (in millisecond) when using the extra bandwidth granting

System Workload	SJF with QoS	Equality and non-Inversion	Equality and Inversion	Proportion and non-Inversion	Proportion and Inversion
0.5	5.628	1.862	2.005	3.060	3.058
0.6	5.645	2.012	1.988	2.977	3.180
0.7	5.743	2.070	2.036	3.075	3.409
0.8	5.826	2.260	2.811	3.290	3.546
0.9	5.998	3.492	3.808	3.697	3.849

produces contrary effects on them as shown in the fifth and sixth columns of Table 1.

- For the FCFS scheduling, the multi-priority scheme decreases the system queuing delay because the average packet sizes of the UGS class are smaller than those of other classes.
- For the SJF scheduling, the multi-priority scheme increases the system queuing delay because there are some larger packets being scheduled before smaller packets due to QoS priority.

Furthermore, a service flow of the higher priority has a smaller queuing delay. It is indeed the main purpose of the multi-priority scheme.

D. Extra Bandwidth Granting

The *extra bandwidth granting* can be adopted along with the existing scheduling algorithms. In the experiment, it is implemented with the SJF scheduling and QoS considerations. From the experimental results from the Table 2, we can see that the queuing delay can be enormously reduced when our EBG is adopted. The comparison for each class of service flows is shown in Figure 6. In the situation where the system workload is below 0.8, the queuing delay of the *Equality* method is smaller than that of the *Proportion* method for each class. As the system workload is up to 0.8 and 0.9, the queuing delay of the UGS class increases rapidly, implying the *Proportion* method has more flexibility in dealing with different traffic loads.

If the *Proportion* method is used, the performances of the UGS class and the rtPS class with the non-*Inversion* method are better than that with the *Inversion* method, except one case—the rtPS class with the system workload 0.9. On the contrary, the performance of the nrtPS class and the BE class with the *Inversion* method are all superior to that with the non-*Inversion* method. The superiority is attributed to the fact that the classes are scheduled earlier. In fact, considering the case of the rtPS class with the system workload 0.9, the UGS class averagely gets more extra bandwidth than the total extra bandwidth of the nrtPS class and the BE class, so the *Inversion* method actually schedules the extra time slots of the rtPS class earlier than the non-*Inversion* method. Conclusively, among the non-*Inversion* method and the *Inversion* method, the one schedules the extra time slots of a service flow earlier can result in a smaller queuing delay for the service flow.

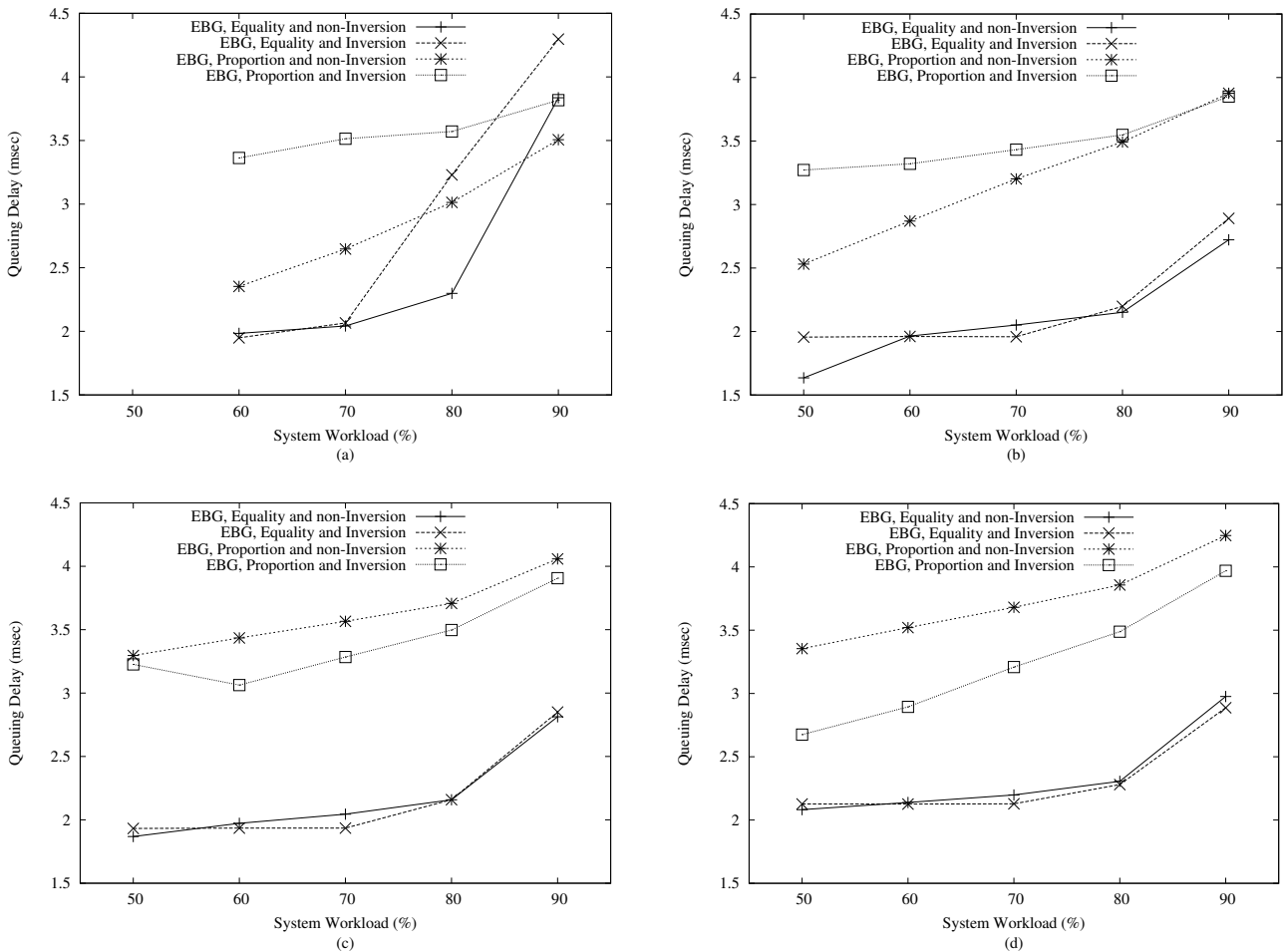


Figure 6. The comparison of queuing delays for the (a) UGS, (b) rtPS, (c) nrtPS and (d) BE class when using the *extra bandwidth granting* (EBG).

However, if the *Equality* method is applied, the argument is not always held. By the results as shown in Figure 6 (a), when the system workload is 0.6, the queuing delay of the *Inversion* method is smaller than that of the non-*Inversion* method, implying that transmitting more packets (the *Inversion* method) in the extra time slots is sometimes (especially when the system workload is low) more effective to decrease the average delay, compared with transmitting packets earlier (the non-*Inversion* method) in the extra time slots.

V. CONCLUSIONS

In this paper, we proposed a new resource allocation scheme, called *extra bandwidth granting* (EBG), to improve transmission efficiency of an IEEE 802.16-based network. Our proposed scheme can be adopted along with the existing scheduling algorithms and the multi-priority scheme without any modification. The experimental results demonstrated the capability of the proposed scheme and showed that by using the EBG, the queuing delay can be significantly improved, especially for the service flows of higher-priority classes.

ACKNOWLEDGMENT

This work was supported in part by National Science Council under the contract 94-2219-E-002-019 and Intel.

REFERENCES

- [1] *IEEE Standard for Local and metropolitan area networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Std. 802.16, Apr. 2002.
- [2] *Draft IEEE Standard for Local and metropolitan area networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Std. 802.16, May 2004.
- [3] E. Agis, H. Mitchel, S. Ovidia, S. Aissi, S. Bakshi, P. Iyer, M. Kibria, C. Rogers, and J. Tsai, "Global, interoperable broadband wireless network: Extending WiMAX technology to mobility," *Intel Technology Journal*, vol. 8(3), pp. 173–187, 2004.
- [4] K. Wongthavarawat and A. Ganz, "Packet scheduling for QoS support in IEEE 802.16 broadband wireless access systems," *International Journal of Communication Systems*, vol. 16, pp. 81–96, 2003.
- [5] O. Gusak, N. Oliver, and K. Sohraby, "Performance evaluation of the 802.16 medium access control layer," in *Proc. Computer and Information Sciences (ISCIS'04)*, 2004, pp. 228–237.
- [6] C. Eklund, R. B. Marks, K. L. Stanwood, and S. Wang, "IEEE standard 802.16: A technical overview of the wirelessman air interface for broadband wireless access," *IEEE Commun. Mag.*, vol. 40(6), pp. 98–107, 2002.
- [7] A. Jamalipour, *The Wireless Mobile Internet-Architectures, Protocols and Services*. John Wiley and Sons, 2003, ch. 6, pp. 149–176.