

# Optimal Radio Resource Partition for Joint Contention- and Connection-Oriented Multichannel Access in OFDMA Systems

Li-Chun Wang, *Senior Member, IEEE*, and Anderson Chen, *Student Member, IEEE*

**Abstract**—The IEEE 802.16e World Interoperability for Microwave Access (WiMax) system plays an important role in the future wireless metropolitan area network (WMAN). Orthogonal frequency division multiple access (OFDMA), adopted in the IEEE 802.16e WiMax system, has many advantages in the physical layer and also poses many challenges for radio resource allocation. One of interesting radio resource allocation issue in the OFDMA system is to partition the overall radio resource (bandwidth and time duration) into two portions: one for random access and the other for connection-oriented access. In the IEEE 802.16e WiMax system, a truncated binary back-off algorithm is adopted to resolve the contention in random access, while the time-division OFDMA is used for the connection-oriented access. The main contribution of this paper is to design an analytical approach to determine the optimal amount of reserved radio resource in both time and frequency domains for random access, with the objective of maximizing the overall efficiency of radio resource while satisfying the delay requirements for supporting real-time services. Furthermore, an analytical model for calculating the access latency and the efficiency of the reserved radio resources is developed.

**Index Terms**—OFDMA, WiMax, radio resource partition, channel access.

## 1 INTRODUCTION

THE IEEE 802.16e World Interoperability for Microwave Access (WiMax) system supports both fixed and mobile point-to-multipoint (PMP) networks, and it will become an important wireless metropolitan area network [1], [2]. This system adopts the orthogonal frequency division multiple access (OFDMA) technique, which has many advantages in the physical layer such as overcoming the multipath fading.

### 1.1 Problem Statement

Further enhancing the overall performance of the OFDMA system, radio resource allocation is an important aspect that should be considered. In the OFDMA system, radio resources can be allocated in either time domain or frequency domain. Both the connection-oriented and contention-based medium access control (MAC) protocols are adopted for allocating radio resource in the IEEE 802.16e WiMax system. The purpose for the connection-oriented MAC protocol is to allow users to transmit data in a reserved chunk of time slots and subchannels. However, before transmitting data, mobile stations (MSs) must send requests for the access permission from the base station (BS). The access permissions are granted based on a certain contention procedure in a reserved chunk of time slots and subchannels. Intuitively, increasing the reserved resources for contention can shorten the access latency since there are fewer collisions, but it also

decreases the overall spectrum efficiency. Therefore, determining how long and how many subchannels an OFDMA system should reserve for the contention-based random access scheme is an important issue.

This paper presents an analytical approach to determine the optimal number of reserved subchannels and time slots for the random access scheme in the OFDMA-based IEEE 802.16e system. We formulate an optimization problem to maximize the efficiency of the reserved radio resources while meeting the delay requirement for supporting real-time applications. It is noteworthy that we consider a multichannel random access scheme based on the truncated binary backoff algorithm. Because multiple channels are reserved for contention, the multichannel random access scheme can allow multiple users to access the subchannels concurrently, unlike the carrier-sense multiple access (CSMA) MAC protocol. Moreover, to reduce collisions during the contention period, the truncated binary backoff algorithm is implemented in the IEEE 802.16e WiMax system to adjust the user's transmission time according to the number of collisions in his previous attempt. Based on the proposed approach, performance enhancement can be achieved without changing the specifications of the IEEE 802.16e WiMax. To our knowledge, there have been few reports on the performance analysis of this kind of multichannel random access scheme with the truncated binary backoff and its impact on the design of the optimal number of the reserved subchannels and time slots in the OFDMA systems.

### 1.2 Related Work

Several previous studies have reported reducing access latency in the IEEE 802.16e WiMax system. Choi et al. [3] suggested a method to simplify the conventional network entry/reentry process. In [4], a new contention resolution

• The authors are with the Department of Communication Engineering, National Chiao-Tung University, 1001 Ta Hsueh Road, Room 900, Engineering Building IV, Hsinchu 300, Taiwan, R.O.C.  
E-mail: lichun@cc.nctu.edu.tw.

Manuscript received 14 Mar. 2007; revised 12 Dec. 2007; accepted 2 Apr. 2008; published online 20 June 2008.

For information on obtaining reprints of this article, please send e-mail to: [tmc@computer.org](mailto:tmc@computer.org), and reference IEEECS Log Number TMC-2007-03-0080. Digital Object Identifier no. 10.1109/TMC.2008.96.

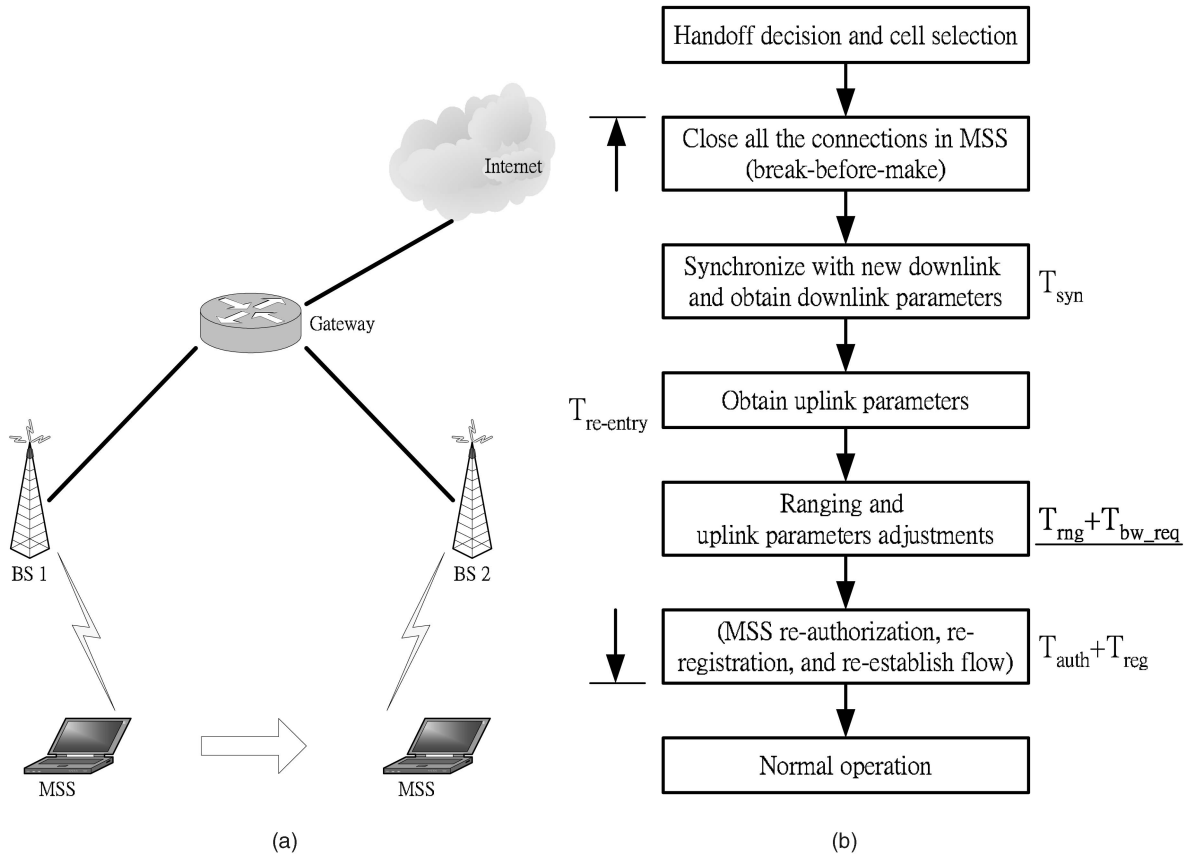


Fig. 1. The IEEE 802.16e reentry process. (a) Example of a network model for hand-over process. (b) Summary of the network reentry process in IEEE 802.16e.

algorithm was proposed for the multichannel random access in OFDMA systems, in which data are transmitted at abundant subchannels so that collisions and access delay are decreased. In [5] and [6], a new set of orthogonal signals were designed to reduce the number of attempts for ranging, i.e., adjusting the uplink transmission power and timing. In [7], the authors suggested a subcarrier permutation algorithm to improve the link reliability and to decrease the number of retransmissions. It seems that most of the aforementioned approach may require modifying the specifications of the current IEEE 802.16e WiMax system.

Some analysis techniques have been reported to investigate the MAC performance of the IEEE 802.16e WiMax system. The throughput of the connection-oriented MAC protocol of the IEEE 802.16e WiMax system was investigated by considering the overhead of the control channels [8], [9]. In addition, the frame access delay in the contention and polling process of the IEEE 802.16e networks have been studied in [10] and [11]. Nevertheless, in [10] and [11], it was assumed that the slotted ALOHA is used to resolve the contentions in the channel access. As for the performance analysis of exponential backoff, the authors in [12] and [13] provide a simple but accurate model to detail the behaviors and performances of the single channel random access scheme. To our knowledge, the access latency analysis of a joint contention- and connection-oriented multichannel access scheme combined with the binary backoff algorithm used in the IEEE 802.16e WiMax system is an open issue.

### 1.3 Organization of the Paper

The rest of paper is organized as follows: In Section 2, we first introduce the connection setup procedure in IEEE 802.16e and present the role of random access in connection setup. Next, we describe the IEEE 802.16 MAC protocol and its contention resolution mechanism in Section 3. The parameter design problem and the performance analysis are given in Section 4. Finally, we give the numerical results and our concluding remarks in Sections 5 and 6, respectively.

## 2 BACKGROUND

### 2.1 Frame Structure and Link Establishment Procedures in the OFDMA-TDD Mode

### 2.2 IEEE 802.16e Reentry Process

As an MS moves across BSs, the network reentry process is initiated to establish and reserve a channel. During the reentry process, the access latency to the new target BS should be reduced to meet the requirements for supporting real-time services. Fig. 1a is an illustrative example, where an MS moves from the serving BS, BS1, to the target BS, BS2, while Fig. 1b shows the key step specified in the reentry process of the IEEE 802.16e WiMax system.

We first briefly introduce the handoff process in the IEEE 802.16e WiMax system. The serving BS periodically broadcasts the information on its neighboring BSs to all the MSs within the cell. Meanwhile, the MS measures the strengths of the pilot signals from the neighboring BSs. If the signal

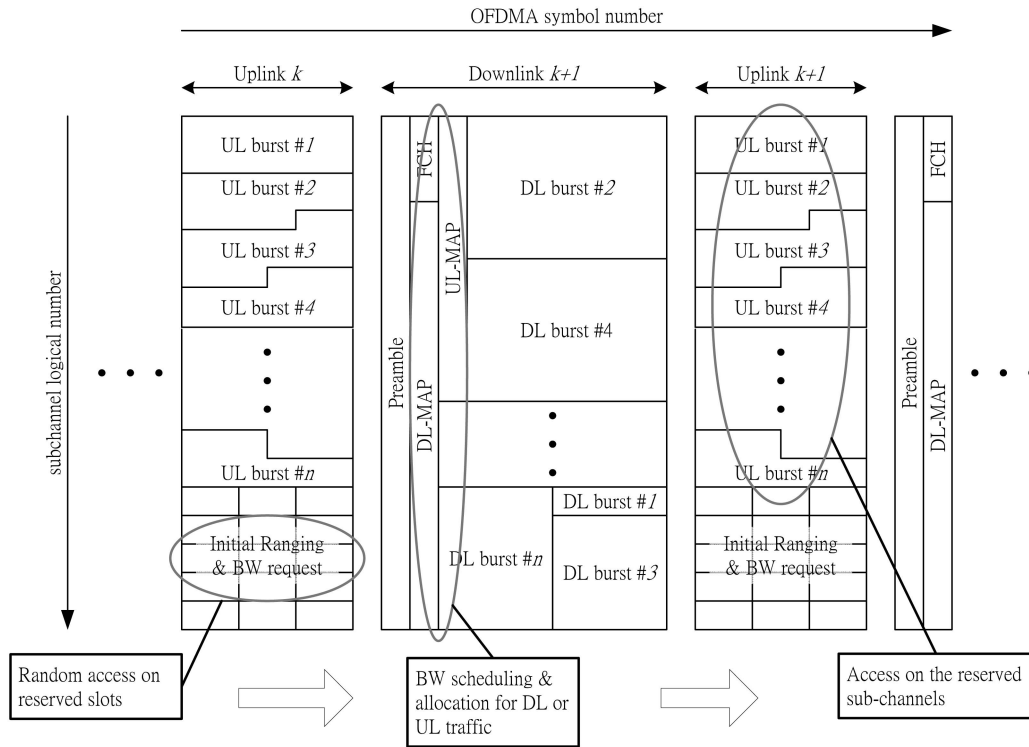


Fig. 2. Frame structure and Link Establishment of TDD-OFDMA mode in IEEE 802.16.

strength from the serving BS is too weak, the MS will send a handoff request message to inform the surrounding BSs. The serving BS negotiates with its neighboring BSs and selects a target BS. Next, the serving BS replies the handoff response message to the MS.

After receiving the handoff response message from the serving BS, the MS releases all the current connections and starts the network reentry process to the target BS. The key steps in the network reentry process are downlink/uplink synchronization, ranging, reauthorization, reregistration, and bandwidth request, as shown in Fig. 1b. During the reentry process, the MS first synchronizes with the target BS to obtain the downlink and uplink transmission parameters. Next, the MS performs the ranging process to acquire the accurate timing offset and to adjust power such that its transmissions can be aligned with the received frames from other MSs. Then, the target BS executes the reauthorization and reregistration processes to identify whether the handoff MS can establish connections. Finally, the MS sends the bandwidth request message to obtain the information of the allocated subchannels.

The total latency  $D_{re-entry}$  of the network reentry process can be expressed by

$$D_{re-entry} = D_{syn} + D_{rng} + D_{auth} + D_{reg} + D_{bw\_req}, \quad (1)$$

where  $D_{syn}$ ,  $D_{rng}$ ,  $D_{auth}$ ,  $D_{reg}$ , and  $D_{bw\_req}$  represent the delays in the steps of downlink/uplink synchronization, ranging, reauthorization, reregistration, and bandwidth request, respectively. Among these steps, the required time for the synchronization  $D_{syn}$  is only in the scale of the OFDMA symbols, which duration is much shorter than a frame period. The authorization and registration process can be carried out in the backbone network of BSs before breaking

the connection. Thus, the time of the above three steps ( $D_{syn}$ ,  $D_{auth}$ , and  $D_{reg}$ ) can be negligible for the calculation of the handoff latency. However, the initial ranging and bandwidth request are not only executed in real time but also need to resolve the collisions due to the random access. Unfortunately, a frame only reserves the limited number of slots for the random access; and thus, the station may require wait for multiple frames to successfully send the request. Therefore, the access latency of the initial ranging and bandwidth request is the dominating factor for the performance of the network reentry process.

Fig. 2 illustrates the frame structure in the IEEE 802.16e OFDMA-TDD mode. As shown in the figure, a frame consists of two portions: the downlink and uplink subframes. A downlink subframe starts with preamble symbols followed by downlink map (DL-MAP) and uplink map (UL-MAP), which are used to inform the users of their allocated subchannels and time slots for downlink and uplink transmissions, respectively. Next, downlink data are transmitted at allocated subchannels and time slots. After an additional transmit/receive transition gap (TTG), the ranging and bandwidth request as well as users' data are transmitted in the uplink direction according to the UL-MAP. After a receive/transmit transition gap (RTG), the same procedures are repeated for another pair of downlink and uplink subframes.

The IEEE 802.16e WiMax system adopts a connection-oriented MAC protocol. In this kind of MAC protocol, MSs must make a reservation with the BS before transmitting data. Consider the uplink transmission first. In Fig. 2, with packets are ready to send in the *uplink subframe k*, MSs first transmit bandwidth requests to the BS at the reserved ranging/bandwidth request subchannels. Receiving these

requests, the BS allocates the subchannels in the UL-MAP of the next *downlink subframe*  $k + 1$ . In the downlink case, the BS negotiates the QoS requirements by using the dialog of dynamic service addition/change (DSA/DSC) messages on the MS's basic connection identifier (CID). After receiving the grant from the MS, the users' data are transmitted at the reserved subchannels and time slots according to the DL-MAP.

To summarize, the BS coordinates the frame transmissions of all MSs in both the uplink and downlink based on the bandwidth request procedures and DSA/DSC dialog. However, a certain amount of resources must be reserved for the bandwidth request procedure and ranging in the network reentry process, in which the MSs contend with each other. Thus, we must determine how large a chunk of subchannels and time slots should be reserved for MSs to send the messages of bandwidth request. In Section 3, we will formulate an optimization problem to design these parameters.

### 3 PROBLEM FORMULATION

#### 3.1 Multichannel Random Access Protocol with Collision Avoidance

To transmit the initial ranging and bandwidth request messages, the IEEE 802.16e WiMax system adopts the multichannel random access with the truncated binary exponential backoff algorithm [1], [2], which is similar to the backoff procedures in the IEEE 802.11 wireless local area network (WLAN). This backoff algorithm is used to resolve the contentions between the MSs in the process of random channel access. For a given initial backoff window size, an MS first selects a random number of time slots smaller than the backoff window size. It is implied that an MS defers its transmission until the end of the selected time period. After waiting for this period, the MS randomly selects subchannels and sends the bandwidth request and the ranging request (RNG\_REQ) frames. Once the BS successfully receives the request, it allocates the subchannels and time slots for the new connection in the next DL-MAP or UL-MAP, as shown in Fig. 3a. However, as shown in Fig. 3b, collisions may occur when the bandwidth/ranging request frames are transmitted. If the BS fails to respond to the bandwidth/ranging request frames after a time-out duration,  $T_{rng}$  or  $T_{bw\_req}$ , the MS doubles its backoff window size and repeats the aforementioned procedures until reaching the maximum number of transmission attempts. Here,  $T_{rng}$  or  $T_{bw\_req}$  are defined as the waiting periods for the ranging response (RNG\_RSP) frame and the bandwidth grant frame, respectively.

For clarity, we explain the difference of the multichannel random access MAC protocol with collision avoidance used in the IEEE 802.16e WiMax system and the CSMA/CA MAC protocol used in the IEEE 802.11 WLAN. Unlike the CSMA/CA MAC protocol, the BS in the IEEE 802.16e WiMax system is capable of receiving the requests from multiple MSs at the same time since the MSs can transmit at different subchannels. In this case, a frame collision occurs when more than one station accesses at the same chunk of subchannels and time slots.

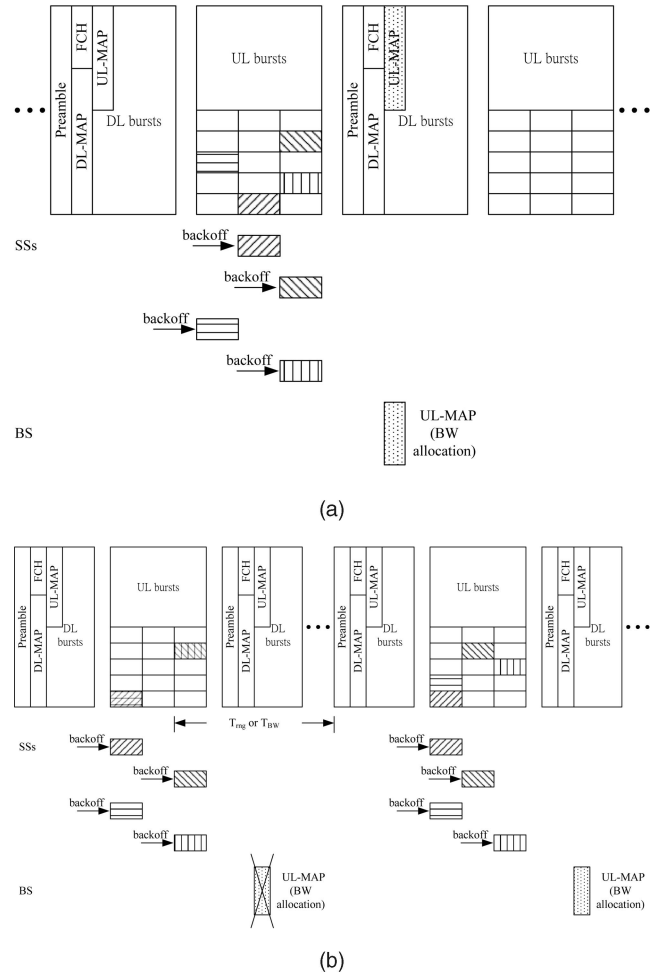


Fig. 3. Illustrations of the multichannel random access for the initial ranging and bandwidth request in IEEE 802.16e. (a) The transmission succeeds at the first attempt. (b) Retransmission occurs when the previous attempt has a collision.

#### 3.2 Resources Partition Optimization

Now, we formulate an optimization problem with the objective of maximizing the efficiency of the reserved chunk of subchannels and time slots for random access subject to the delay constraint of supporting real-time services. Intuitively, reserving more radio resources for random access can yield shorter access latency due to fewer collisions. However, it also leads to more idle radio resources. With the optimal numbers of reserved subchannels and time slots for random access, we can optimize radio resource partitioning for the time-division OFDMA system.

The above optimization problem can be solved by a mixed-integer nonlinear programming (MINLP) technique. Define the efficiency  $\eta$  as the ratio of the reserved chunk of subchannels and time slots on which bandwidth request and ranging requests frames are successfully transmitted over the total reserved radio resources. Denote  $r$  and  $m$  as the numbers of the reserved subchannels and time slots for random access, respectively. It will be shown later that the considered efficiency  $\eta$  is a nonlinear function of  $r$  and  $m$ , as well as the number of mobile users (denoted by  $k$ ). With the above definitions, an MINLP optimization problem is

formulated to maximize the efficiency  $\eta(r, m, k)$  by determining the optimal values of  $r$  and  $m$  for a given  $k$ , i.e.,

$$\text{MAX}_{r,m} \eta(r, m, k), \quad (2)$$

subject to the following constraints:

1.  $r \leq N_f$ ;
2.  $m \leq N_t$ ;
3.  $D(r, m, k) \leq D_{max}$ .

Later, in Section 5, we will find that the maximization of the efficiency  $\eta$  is equivalent to the optimization of the reserved resources left for the data transmission.

Clearly, both integers  $r$  and  $m$  are constrained by the total number of subchannels in the frequency domain ( $N_f$ ) and the total number of time slots in the time domain ( $N_t$ ). Furthermore, to support the delay-sensitive services, the access delay  $D$  of sending the bandwidth request and ranging request frames is required to be shorter than the predefined maximal delay ( $D_{max}$ ). Note that the access delay  $D$  is a continuous value and is a nonlinear function of  $r$ ,  $m$ , and  $k$ . Since both  $\eta(\cdot)$  and  $D(\cdot)$  are nonlinear functions of  $r$ ,  $m$ , and  $k$ , and conditions 1, 2, and 3 are the constants for both the discrete and continuous values, (2) becomes an MINLP optimization issue. The next important step to solve the above optimization problem is to derive  $\eta(r, m, k)$  and  $D(r, m, k)$ , which will be discussed in Section 4.

## 4 ANALYSIS

Now, we develop the analytical models for the average access latency and the spectrum efficiency for random access. It is assumed that each MS always has frames to send, and collisions occur when more than one station access the same subchannel and time slot.

### 4.1 Efficiency of Reserved Radio Resources

Denote  $N_s$  as the number of successfully transmitted bandwidth and ranging requests during one uplink subframe, and  $N'_s$  as that during one time slot. It is assumed that each request asks for a chunk of one time slot at one subchannel. Then, based on the previous definition, the average spectrum efficiency  $\eta$  of the reserved radio resources for random access can be expressed as

$$\eta(r, m, k) = \frac{E[N_s]}{mr}, \quad (3)$$

$$= \frac{mE[N'_s]}{mr}, \quad (4)$$

where  $m$  and  $r$  are the numbers of reserved time slots and subchannels for random access. According to the random access procedure in the standard, it is worth to note that the contention at each time slot and subchannel is independent to the others.

Given  $\overline{W}$  as the average backoff time slots that an MS defers before each transmission attempt, the transmit

probability  $\tau$  that a station transmits a request in a reserved slot can be expressed as

$$\tau = \frac{1}{\overline{W}}. \quad (5)$$

Note that based on the random access procedure in the standard, the backoff count is decreased by one for each reserved slot in a frame, while it is frozen in the other slots of a frame and continues to count down in the next frame. Therefore, the transmit probability  $\tau$  is equal to the inverse of the average backoff counts because both of these factors only consider the reserved slots instead of the whole of a frame.

Then, let  $W_{min}$  and  $f$  be the initial backoff window size and the maximum number of transmission attempts, respectively. Because the users with colliding frames double their backoff window sizes in the next transmission in order to avoid collisions, the average number of backoff slots for the  $i$ th retransmission  $W_i$  can be written as

$$W_i = \frac{2^i W_{min} + 1}{2}, \quad (6)$$

for  $i = 0, \dots, f$ . Given  $T_o$  as the time-out duration,<sup>1</sup> a station has to defer an extra duration of  $m \lceil T_o / T_f \rceil$  slots before the retransmission according to the random access procedure. With the random backoff time slots  $W_i$  for the  $i$ th retransmission, the average backoff slots  $\overline{W}$  then becomes

$$\begin{aligned} \overline{W}(m) &= (1 - p_c)W_0 + p_c \left( (1 - p_c)(W_1 + m \lceil \frac{T_0}{T_f} \rceil) \right) \\ &\quad + \dots + p_c^f (1 - p_c) \left( W_f + m \lceil \frac{T_0}{T_f} \rceil \right) \\ &\quad + p_c^{f+1} \left( W_f + m \lceil \frac{T_0}{T_f} \rceil \right) \\ &= \left( \frac{1}{2} + m \lceil \frac{T_0}{T_f} \rceil \right) [(1 - p_c)(1 + p_c + \dots + p_c^f) + p_c^{f+1}] \quad (7) \\ &\quad - m \lceil \frac{T_0}{T_f} \rceil (1 - p_c) + \frac{W_{min}}{2} \\ &\quad \times \left[ (1 - p_c)(1 + 2p_c + \dots + (2p_c)^f) + p_c(2p_c)^f \right] \\ &= \frac{1}{2} + m \lceil \frac{T_0}{T_f} \rceil p + \frac{W_{min}}{2} \frac{1 - p_c - p_c(2p_c)^f}{1 - 2p_c}, \end{aligned}$$

where  $p_c$  is the collision probability, which is defined as the probability that more than one station transmit frames at the same subchannel and time slot given that a station transmits. Assume that each station randomly selects one of  $r$  reserved subchannels to access with the probability  $1/r$ . Because  $1/r$  and  $\tau$  are independent, the collision probability  $p_c$  can be written by

$$p_c = 1 - \left( 1 - \frac{\tau}{r} \right)^{k-1}. \quad (8)$$

Note that in (7),  $p_c^i(1 - p_c)$  represents the success probability for the  $i$ th retransmission. Thus, the transmission probability  $\tau$  can be obtained by jointly solving (5)~(8).

1. The time-out duration  $T_o$  for the initial ranging request is different from that of the bandwidth request. As specified in the standard,  $T_o = 10$  and 200 msec for the bandwidth request and initial ranging, respectively.

In addition, because each station has the probability  $\frac{\tau}{r}$  to access one chunk of  $r$  reserved subchannels and  $m$  time slots, the average number of successful transmissions for the initial ranging or bandwidth request within one slot becomes

$$E[N'_s] = r \binom{k}{1} \frac{\tau}{r} \left(1 - \frac{\tau}{r}\right)^{k-1}. \quad (9)$$

By substituting (9) into (3), the efficiency  $\eta$  of the reserved radio resources is given by

$$\eta(r, m, k) = k \frac{\tau(m)}{r} \left(1 - \frac{\tau(m)}{r}\right)^{k-1}. \quad (10)$$

## 4.2 Average Access Latency

As defined previously in Section 3.2, the average access latency  $D$  is equal to the overall duration that an MS waits until successfully transmitting a bandwidth/ranging request frame. Given the average number of overall slots  $E[W_D]$  before a successful request transmission and the frame duration  $T_f$ , the average access latency  $D$  can be written as

$$D = \left\lceil \frac{E[W_D]}{m} \right\rceil \cdot T_f. \quad (11)$$

Here, we only count the latency in terms of frame duration because the slot time is much shorter than a frame duration.

Suppose that each bandwidth/ranging request frame transmission only occupies one reserved slot in a frame. Then, considering the collision probability  $p_c$ , the average backoff slots at the  $i$ th retransmission  $W_i$  and the time-out duration  $T_o$ , the average overall backoff slots  $E[W_D]$  given the request is successfully transmitted is equal to

$$\begin{aligned} E[W_D] &\triangleq E[\text{the total number of the required slots for a} \\ &\quad \text{successful transmission | the request is} \\ &\quad \text{successfully transmitted}] \\ &= \underbrace{W_0}_{\text{successful at the 1st attempt}} + \\ &\quad \underbrace{\left(W_1 + m \left\lceil \frac{T_o}{T_f} \right\rceil\right)}_{\text{successful at the 2nd attempt}} p_c + \dots \\ &\quad + \underbrace{\left(W_{f-1} + m \left\lceil \frac{T_o}{T_f} \right\rceil\right)}_{\text{successful at the } f\text{-th attempt}} p_c^{f-1} \\ &\quad + \underbrace{\left(W_f + m \left\lceil \frac{T_o}{T_f} \right\rceil\right)}_{\text{successful after the } f\text{-th attempt}} \frac{p_c^f}{1 - p_c} \\ &= \frac{(1 - 2p_c)(1 + W_{\min}) + W_{\min} p_c (1 - (2p_c)^f)}{2(1 - 2p_c)(1 - p_c)} \\ &\quad + m \left\lceil \frac{T_o}{T_f} \right\rceil \frac{p_c}{1 - p_c}. \end{aligned} \quad (12)$$

By substituting (8) and (12) into (11), the average access latency  $D$  can be obtained. Note that the collision probability  $p_c$ , as shown in (8), is affected by the

TABLE 1  
System Parameters

|  |          |
|--|----------|
| Frame duration, $T_f$                      | 5 msec   |
| Timeout duration for the bandwidth request | 10 msec  |
| Timeout duration for the initial ranging   | 200 msec |
| Initial backoff window size, $W$           | 16       |
| Maximum transmission attempts, $f$         | 7        |
| Number of stations, $k$                    | 100      |
| Default number of reserved time slots      | 3        |
| Default number of reserved sub-channels    | 6        |
| Maximum number of reserved time slots      | 48       |
| Maximum number of reserved sub-channels    | 70       |
| VoIP delay constraint, $D_{max}$ [14]      | 25 msec  |

number of reserved subchannels  $r$ . Thus, the average overall slots  $E[W_D]$  and the access latency  $D$  are also the functions of  $r$ .

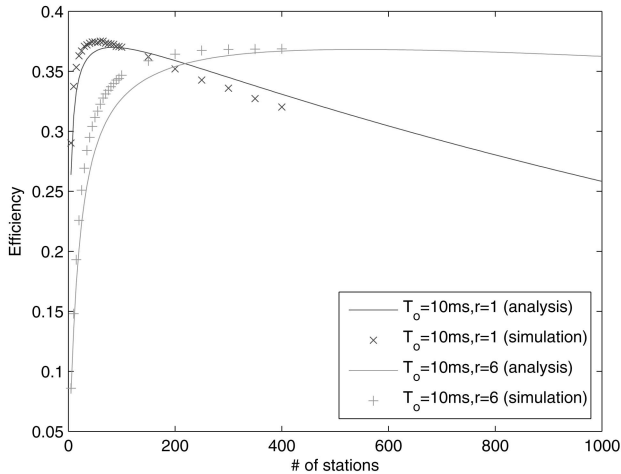
## 5 NUMERICAL RESULTS

In this section, we investigate the effectiveness of the proposed radio resource partition technique to determine the optimal number of reserved subchannels and time slots for random access in the time division OFDMA system. Referring to the IEEE 802.16e WiMax system, we consider the following parameters, including an uplink subframe with 70 subchannels and 48 time slots, frame duration  $T_f = 5$  msec, the bandwidth request and initial ranging time-out duration 10 and 200 msec, respectively. Other related system parameters are listed in Table 1.

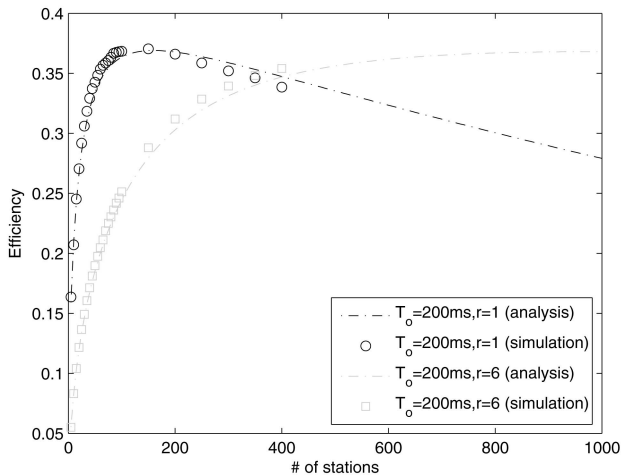
### 5.1 Performances of Default Radio Resource Partition

First, we evaluate the performance of an OFDMA system with the parameters listed in Table 1, in which one and six subchannels, as well as three time slots, are reserved for the random access. According to (10), the efficiency of the reserved radio resources for the random access versus various numbers of stations is plotted in Figs. 4a and 4b. From the figures, the highest efficiency of the reserved random access radio resources occurs for the number of users  $k = 85$  and 160 when  $r = 1$  and  $T_o = 10$  and 200 msecs, respectively. This phenomenon can be explained by two factors. On one hand, when the number of stations is smaller, the reserved radio resources are more likely to become idle and thus the efficiency decreases. On the other hand, as the number of station increases, the contentions among users result in more collisions and reduce the efficiency. Thus, there exist an optimal number of stations to achieve the maximum efficiency for various time-out values. Similar results can also be observed in the multi-channel case.

Furthermore, Figs. 4a and 4b also show the difference between the single and multiple channel random access. Interestingly, one can observe that the multichannel random access has slowly increasing efficiency, as the number of stations grows compared to the single channel case. This phenomenon results from the abundant reserved resources for resolving the contentions so that the system is capable of accepting more contentions among stations.



(a)



(b)

Fig. 4. Efficiency of the reserved radio resources based on the multichannel random access MAC protocol with three time slots per frame and one/six subchannels to send bandwidth/ranging requests in an OFDMA system by simulation and analysis. (a)  $T_o = 10$  msec. (b)  $T_o = 200$  msec.

Nevertheless, its payoff is the low efficiency of the reserved resources for the case of a few stations due to the wasted subchannels.

Next, based on (11), Fig. 5 shows the average latency of the multichannel random access MAC protocol in the OFDMA system for various numbers of stations and time-out values. Consider the delay constraint  $D_{max} = 25$  msec for supporting real-time services. Subject to the above delay requirement, the system can support only up to 25 stations sending the bandwidth request based on the default partition scheme, i.e., three subchannels and six time slots are designated for random access out of the total 70 subchannels and 48 time slots. As for the initial ranging transmission, the system with the default partition scheme cannot satisfy the delay constraint due to the long time-out duration before any retransmission. Clearly, the more radio resources reserved for the random access the shorter the access latency because of the less retransmission. However, from the figure, only one partition scheme can satisfy the delay requirement for

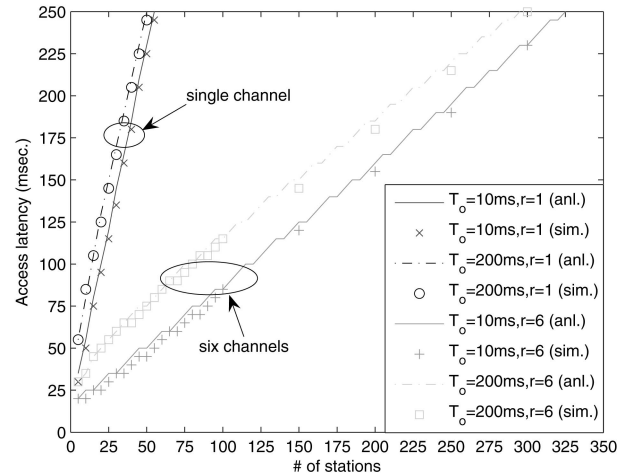


Fig. 5. Average latency of the reserved radio resources based on the multichannel random access MAC protocol with three time slots per frame and one/six subchannels to send bandwidth/ranging requests in an OFDMA system by simulation and analysis.

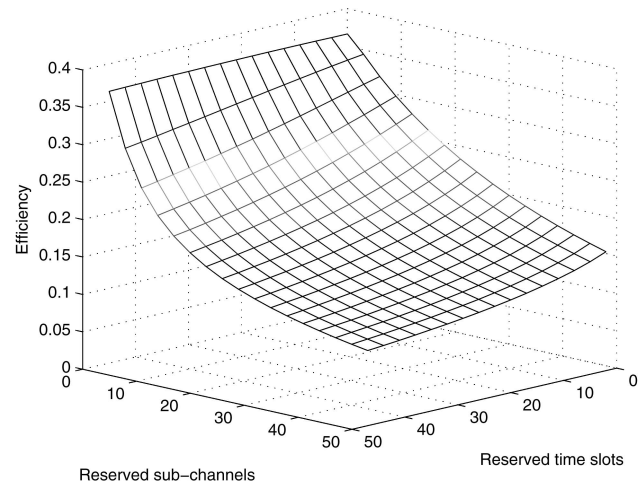


Fig. 6. Effects of various subchannels and time slots on the efficiency of the reserved radio resources based on the multichannel random access MAC protocol in an OFDMA system with 100 stations.

real-time services. Therefore, it is necessary to investigate some alternative resource partition schemes through the presented analytical models for the efficiency and access delay of the multichannel random access MAC protocol in the OFDMA system, which will be discussed next.

## 5.2 Performances of Alternative Radio Resource Partitions

In Fig. 6, we investigate the effects of various numbers of subchannels  $r$  and time slots  $m$  on the efficiency of the reserved radio resources based on the multichannel random access MAC protocol in the considered OFDMA system with 100 stations. We have the following two observations:

- As shown in the figure, the efficiency of the reserved radio resources is monotonously decreasing as the number of the reserved time slots or subchannels in a frame grows. This is because too many reserved radio resources may cause them staying idle. For example, when the number of reserved slots is

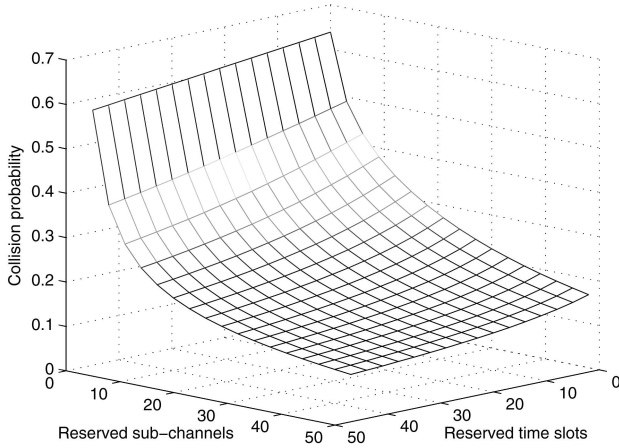


Fig. 7. Effects of various reserved subchannels and time slots on the collision probability of the multichannel random access MAC protocol in the considered OFDMA system with 100 stations.

$m = 3$ , the efficiency  $\eta$  for the number of reserved subchannels  $r = 1 \sim 50$  decreases from 0.37 to 0.15. Similarly, as  $r = 6$ ,  $\eta = 0.35 \sim 0.3$  for  $m = 1 \sim 48$ . The phenomenon of monotonously decreasing efficiency with the increasing number of reserved subchannels represents that the maximization of efficiency performance is also to optimize the reserved resources left for the data transmissions for the given access latency requirement. Thus, although we aim to maximize the efficiency by adjusting the numbers of reserved subchannels and time slots, this problem is also an optimization problem for the reserved resource left for the data transmission.

- Furthermore, as shown in the figure, the efficiency of the reserved radio resources seems to be more sensitive to the changes of the number of reserved subchannels in the frequency domain than that of the number of reserved time slots in the time domain. We will try to explain this situation by investigating the effect of the reserved radio resources in the frequency and time domains on the collision probability of the multichannel random access MAC protocol next.

By means of (8), Fig. 7 shows the effects of various reserved subchannels and time slots on the collision probability of the multichannel random access MAC protocol in the considered OFDMA system with 100 stations. Clearly, reserving more radio resources for random access can yield a lower collision probability. However, it is interesting to investigate in which domain of the OFDMA system, frequency domain or time domain, is more effective in improving the efficiency and latency performances for the system adopting the multichannel random access MAC protocol. As shown in the figure, increasing the number of the reserved subchannels  $r$  can reduce the collision probability more effectively than changing the number of reserved time slots  $m$ . For  $m = 6$ , changing  $r = 6$  to 48 can lower the collision probability  $p_c$  from 0.64 to 0.17. By contrast, the collision probability  $p_c$  decreases from 0.49 to

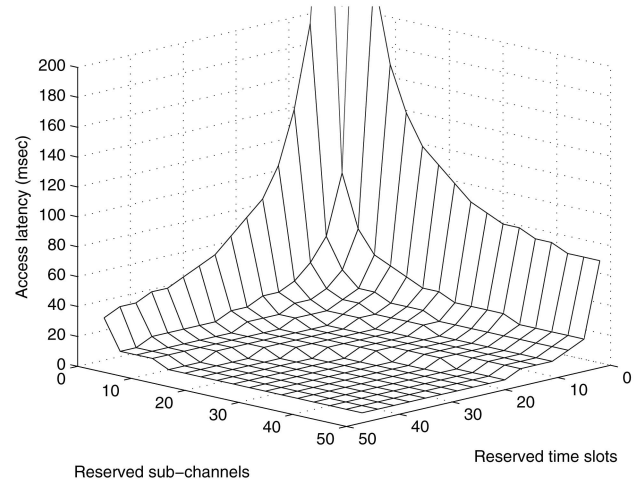


Fig. 8. Effects of various reserved subchannels and time slots on the average latency of the multichannel random access MAC protocol in the considered OFDMA system with 100 stations.

only 0.4 when  $m$  changes from 2 to 48 for  $r = 6$ . By comparing (8) and (10), one can easily see that a lower collision probability leads to a higher efficiency for the reserved random access radio resources. Hence, in Fig. 7, one can explain why in Fig. 6 changing the amount of reserved radio resource for random access in the frequency domain of the OFDMA system can more significantly affect the efficiency than in the time domain of the OFDMA system.

Fig. 8 illustrates the effects of various reserved subchannels and time slots on the average latency of the multichannel random access MAC protocol in the OFDMA system with 100 stations. In contrast to the phenomenon observed in Fig. 6, the number of the reserved time slots dominates the latency performance compared to that of the reserved subchannels. For example, the access latency for  $r = 6$  decreases from 95 msec to 10 msec as  $m$  changes from 2 to 48. On the contrary, when  $r$  increases from 2 to 50, the latency for  $m = 6$  is reduced only from 95 msec to 20 msec. Intuitively, the more reserved time slots help stations to resolve the contentions earlier and to send the request in the prior frame. Although, from (11), increasing the number of reserved subchannels can lower the collision probability, it is not as effective as increasing the time slots from the aspect of reducing the access latency because the station waits more frames to count down the backoff slots in the case of few reserved slots. Notably, it is important to combine Figs. 6 and 8 to determine a set of  $(r, m)$  with the maximum efficiency and satisfactory delay performance. For the case  $k = 100$  and  $T_o = 10$  msec,  $(r, m) = (2, 29)$  can achieve the highest efficiency under the constraint of  $D_{max} = 25$  msec.

### 5.3 Performance of Optimum Radio Resources Partitions

Now we apply the systematic MINLP optimization approach introduced in Section 3.2 to determine the optimal values of  $(r^*, m^*)$ , which represent the subchannels and the time slots reserved for random access in the frequency



TABLE 2  
Optimum Parameters Values for Bandwidth Request  
Transmissions ( $T_o = 10$  msec)

| No. of stations | No. of reserved time slots | No. of reserved sub-channels |
|-----------------|----------------------------|------------------------------|
| 10              | 7                          | 1                            |
| 30              | 18                         | 1                            |
| 50              | 29                         | 1                            |
| 100             | 29                         | 2                            |
| 150             | 42                         | 2                            |
| 200             | 38                         | 3                            |

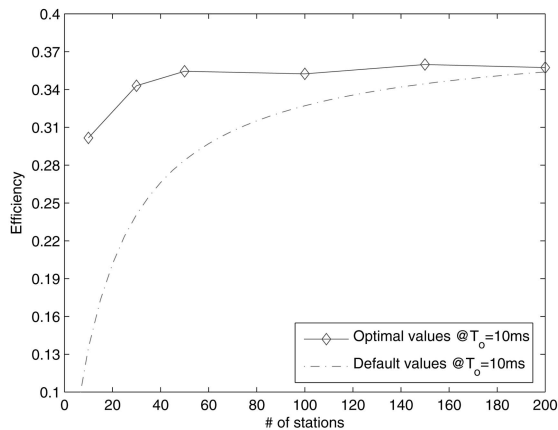
TABLE 3  
Optimum Parameters Values for Ranging Request  
Transmissions ( $T_o = 200$  msec)

| No. of stations | No. of reserved time slots | No. of reserved sub-channels |
|-----------------|----------------------------|------------------------------|
| 10              | 20                         | 1                            |
| 30              | 30                         | 2                            |
| 50              | 34                         | 3                            |
| 100             | 40                         | 5                            |
| 150             | 43                         | 7                            |
| 200             | 40                         | 10                           |

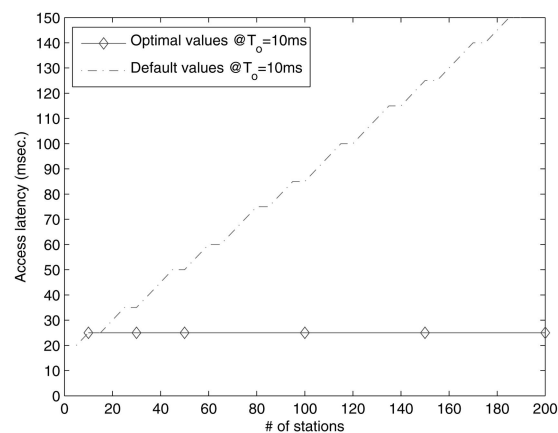
domain and the time domain of the OFDMA system, respectively. In the previous section, an exhaustive search method is used to find the optimal  $(r^*, m^*)$  for the number of users  $k = 100$ . According to (2), Tables 2 and 3 list the optimal  $(r^*, m^*)$  for  $T_o = 10$  msec and 200 msec with an objective of maximizing the efficiency of the reserved subchannels and time slots for random access in the considered OFDMA system subject to the constraint of  $D_{max} = 25$  msec, respectively. For  $k = 100$  and  $T_o = 10$  msec, the optimal  $(r^*, m^*) = (2, 29)$  obtained by the suggested MINLP optimization approach is the same as that obtained by the exhaustive search method. Once the optimum  $(r^*, m^*)$  for various numbers of MSs is generated offline and stored into a table, the BS can dynamically adjust the reserved radio resources for random access to enhance the performance of the OFDMA system.

Furthermore, from the optimal values shown in Tables 2 and 3, we can infer that the system rather reserves the more time slots than the subchannels to meet the delay requirement for real-time applications. Although both subchannels and time slots are used to resolve the contentions among users, the more reserved time slots let the station wait fewer frames to send the request. Thus, the access latency can be easily satisfied the delay requirement. Despite the excess, reserved subchannels cannot be used to lower the waiting time due to the backoff before the transmission, the number of subchannels is still the dominating factor to reduce the collisions among users and to decide the system capacity.

Fig. 9 compares the efficiency of the reserved radio resources for random access and the access latency of the considered MAC protocol according to the above optimal radio resource partition scheme in Table 2 to that according to the default partition scheme  $(r, m) = (6, 3)$  in the case of  $T_o = 10$  msec. Fig. 10 also illustrates the same comparison according to Table 3 for  $T_o = 200$  msec. As shown in Figs. 9b and 10b, the proposed optimum resource partition scheme



(a)



(b)

Fig. 9. Performance comparisons of the optimal radio resource partition scheme according to Table 2 and the default partition scheme  $(r, m) = (6, 3)$  in terms of (a) efficiency and (b) access latency, where  $T_o = 10$  msec.

can satisfy the delay requirement with  $k$  up to 200 for both the bandwidth request and the initial ranging transmissions. Furthermore, for  $T_o = 10$  msec, Fig. 9a shows that the efficiency of the reserved radio resource based on the optimal partition scheme is significantly higher than the default partition scheme especially for the case of few contending stations.

## 6 CONCLUSIONS

In this paper, based on a multichannel random access MAC protocol with a collision avoidance binary backoff algorithm, we formulated an MINLP optimization problem to determine the most appropriate numbers of subchannels and time slots that should be reserved in the frequency and the time domain of the OFDMA system for sending the bandwidth/ranging requests. The objective of the optimization issue is to maximize the efficiency of the reserved resource under the delay constraint of supporting the real-time services. Our numerical results also show that this maximization problem is equivalent to the optimization of the reserved resources left for the data transmission of the established link.

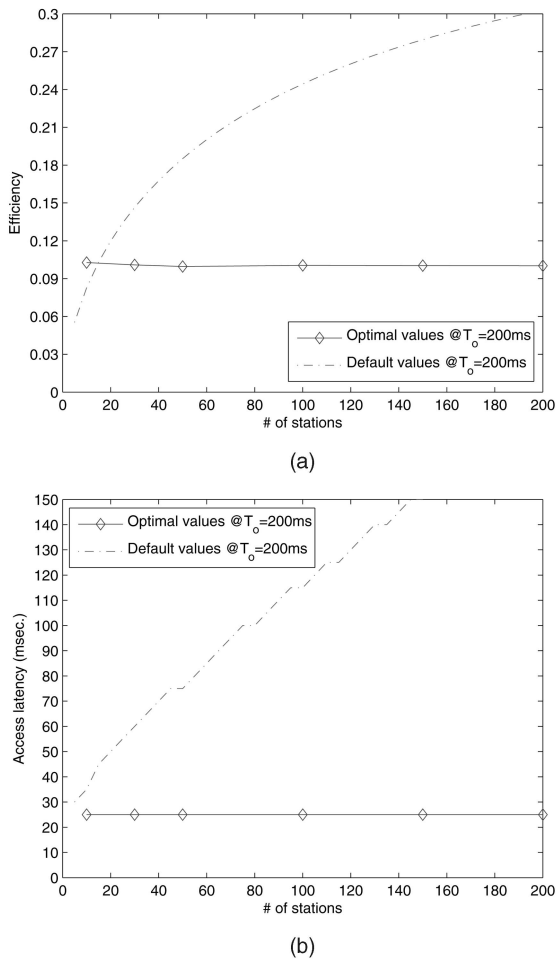


Fig. 10. Performance comparisons of the optimal radio resource partition scheme according to Table 3 and the default partition scheme  $(r, m) = (6, 3)$  in terms of (a) efficiency and (b) access latency, where  $T_o = 200$  msec.

We derived the analytical expressions for the efficiency of the reserved resources and the latency of the contention-based multichannel random access MAC protocol. Based on the developed analytical models, we evaluated the impact of the number of reserved subchannels and time slots on the efficiency and latency performances of the OFDMA systems using the contention-based MAC for link establishment. From the efficiency standpoint, we find that the number of reserved subchannels is a more dominant factor than the number of reserved time slots. From the latency viewpoint, however, the number of the reserved time slots becomes a dominating factor instead.

By using both exhaustive and the proposed MINLP optimization approaches, we have discussed how to find the optimal partition scheme, in which  $r^*$  subchannels and  $m^*$  time slots are reserved for random access during link setup process. It is shown that with the optimal  $(r^*, m^*)$ , the efficiency of the reserved radio resources can be maximized regardless of the number of contending users while still satisfying the delay requirement of real-time services.

## ACKNOWLEDGMENTS

This research was jointly supported by the MoE ATU plan, Industrial Technology Research Institute of (ITRI),

Universal Scientific Industrial (USI), and the National Science Council, Taiwan, under the Contracts NSC-96-2221-E-009-061 and 96-2628-E-009-004-MY3.

## REFERENCES

- [1] IEEE Std. 802.16-2004, *IEEE Standard for Local and Metropolitan Area Networks-Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, Oct. 2004.
- [2] IEEE Std. 802.16e, *IEEE Standard for Local and Metropolitan Area Networks-Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, Dec. 2005.
- [3] S. Choi, G.H. Hwang, T. Kwon, A.R. Lim, and D.H. Cho, "Fast Handover Scheme for Real-Time Downlink Services in IEEE 802.16e BWA System," *Proc. IEEE Vehicular Technology Conf. (VTC '05)*, vol. 3, pp. 2028-2032, May 2005.
- [4] Y.J. Choi, S. Park, and S. Bahk, "Multichannel Random Access in OFDMA Wireless Networks," *IEEE J. Selected Areas in Comm.*, vol. 24, pp. 603-613, Mar. 2006.
- [5] X. Fu and H. Minn, "Initial Uplink Synchronization and Power Control (Ranging Process) for OFDMA Systems," *Proc. IEEE Global Telecommunications Conf. (GLOBECOM '04)*, vol. 6, pp. 3999-4003, Dec. 2004.
- [6] H. Minn and X. Fu, "A New Ranging Method for OFDMA Systems," *Proc. IEEE Global Telecomm. Conf. (GLOBECOM '05)*, vol. 3, pp. 1435-1440, Dec. 2005.
- [7] J. Lee, C. Kim, and J. Kim, "On Subcarrier Allocation for Soft Hand-Over in OFDMA-Based Cellular Systems," *Proc. IEEE Vehicular Technology Conf. (VTC '05)*, vol. 5, pp. 3087-3091, May 2005.
- [8] S. Redana, M. Lott, and A. Capone, "Performance Evaluation of Point-to-Multi-Point (PMP) and Mesh Air-Interface in IEEE Standard 802.16a," *Proc. IEEE Vehicular Technology Conf. (VTC '04)*, vol. 5, pp. 3186-3190, Sept. 2004.
- [9] A.E. Khafa, S. Kangude, and X. Lu, "MAC Performance of IEEE 802.16e," *Proc. IEEE Vehicular Technology Conf. (VTC '05)*, vol. 1, pp. 685-689, Sept. 2005.
- [10] S.M. Oh and J.H. Kim, "The Analysis of the Optimal Contention Period for Broadband Wireless Access Network," *Proc. IEEE Int'l Conf. Pervasive Computing and Comm. Workshops (PerCom '05)*, pp. 215-219, Mar. 2005.
- [11] R. Iyengar, P. Iyer, and B. Sikdar, "Delay Analysis of 802.16 Based Last Mile Wireless Networks," *Proc. IEEE Global Telecomm. Conf. (GLOBECOM '05)*, vol. 5, pp. 3123-3127, Dec. 2005.
- [12] G. Bianchi, "Performance Analysis of IEEE 802.11 Distributed Coordination Function," *IEEE J. Selected Areas in Comm.*, vol. 18, no. 3, pp. 535-547, Mar. 2000.
- [13] B.-J. Kwak, N.-O. Song, and L.E. Miller, "Performance Analysis of Exponential Backoff," *IEEE/ACM Trans. Networking*, vol. 13, no. 2, pp. 343-355, Apr. 2005.
- [14] Y. Amir, C. Danilov, S. Goose, D. Hedqvist, and A. Terzis, "An Overlay Architecture for High Quality VoIP Streams," *IEEE Trans. Multimedia*, vol. 8, pp. 1250-1262, Dec. 2006.



**Li-Chun Wang** received the BS degree in electrical engineering from the National Chiao-Tung University, Hsinchu, Taiwan, in 1986, the MS degree in electrical engineering from the National Taiwan University, Taipei, in 1988, and the MSc and PhD degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1995 and 1996, respectively. From 1990 to 1992, he was with Chunghwa Telecom. In 1995, he was affiliated with Northern Telecom, Richardson, Texas. From 1996 to 2000, he was with AT&T Laboratories, where he was a senior technical staff member in the Wireless Communications Research Department. Since August 2000, he has been with the Department of Communication Engineering, National Chiao-Tung University, Taiwan, as an associate professor and has been promoted to a full professor since August 2005. He was a corecipient of the Jack Neubauer Best Paper Award from the IEEE Vehicular Technology Society in 1997. His current research interests include cellular architectures, radio network resource management, crosslayer optimization for cooperative, and cognitive wireless networks. He is the holder of three US patents with three more pending. He is a senior member of the IEEE.



**Anderson Chen** received the BS, MS, and PhD degrees in communication engineering from the National Chiao Tung University, Hsin-Chu, Taiwan, in 1998, 1999, and 2007, respectively. He is currently working on his national service. His research interests include wireless networks, crosslayer design, and cognitive wireless systems. He is the holder of a US patent with three more pending. He is a student member of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at [www.computer.org/publications/dlib](http://www.computer.org/publications/dlib).**