ARTICLE IN PRESS

Journal of Network and Computer Applications ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at SciVerse ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

A dynamic strategy for packet scheduling and bandwidth allocation based on channel quality in IEEE 802.16e OFDMA system

Feng-Ming Yang [a], Wei-Mei Chen [a,c,*], Jean-Lien C. Wu [b]

[a] Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan
[b] Department of Computer and Communication Engineering, St. John's University, Taipei, Taiwan
[c] Department of Electronic Engineering, National Taiwan University of Science and Technology, #43, Sec.4, Keelung Rd., Taipei, Taiwan

ARTICLE INFO

Article history:
Received 26 April 2012
Received in revised form
21 February 2013
Accepted 10 April 2013

Keywords:
Packet scheduling
Adaptive bandwidth allocation
WiMAX
PMP
IEEE 802.16e
BW-REQ

ABSTRACT

Because the bandwidth usage of broadband wireless access networks is limited, a crucial goal in this field is to improve the system throughput and provisioning of QoS by supporting multiple users based on limited bandwidth resources. This study proposes a two-stage mechanism, called PSBA, to process the operations of packet scheduling and bandwidth allocation based on channel quality information. The proposed approach is adaptive for dynamic channel conditions to improve system throughput and to reduce packet delay in mobile WiMAX. The service flow simulations in the IEEE 802.16e system indicate that PSBA achieved a lower delay time and a higher throughput among multiple users. For the purpose of performance evaluation, a Markov chain with bulk service was built to analyze the WiMAX Point-to-Multipoint (PMP) network based on channel quality by using the bandwidth request message (BW-REQ).

1. Introduction

The Worldwide Interoperability for Microwave Access (WiMAX) is gradually becoming available for the creation of the next-generation wireless internet because of the rapid development of this technology. This emerging technology is used in Broadband Wireless Access (BWA) to provide wireless broadband communications for various types of applications. This dynamic network connects mobile handset devices with a roaming capability, such as laptops, personal digital assistants (PDAs), and cordless phones. Initially, IEEE 802.16 defined the protocols for fixed subscriber stations (SSs) and for mobile subscribers (MSs). The 802.16 committee extended 802.16 to support both mobile and fixed location users, creating the IEEE 802.16e specification for automobile application requirements (Ghosh et al., 2005; IEEE 802.16e-2005, 2006). This extension presented new challenges and opportunities in the resource allocation problem. The efficiency of the dynamic bandwidth allocation mechanism plays a vital role in system performance because the channel condition changes when mobile users pass through the modulation level change and the connection process is triggered. In a WiMAX Point-to-Multipoint (PMP) network, a base station (BS) regulates all communication among the BS and MSs in the radio coverage cell of the BS. To meet QoS requirements, MS sends a bandwidth request message (BW-REQ) to the BS, and the BS attempts to inform MSs of the same cell by multicast and broadcast for polling to contend the slots. For all MSs to send the bandwidth request separately, the service flows of the traffic class must contend to send their requests. For all MSs to send bandwidth requests separately, the service flows that belong to the traffic class must perform contention to send their request messages. This situation results in users being unable to receive proper bandwidth, which causes the request delay for traffic flows, and therefore, its quality of service (QoS) is degraded (Fallah et al., 2008; He et al., 2007).

The architecture of this paper provides superior QoS performance to that of the distributed star model. After receiving the BW-REQ from an MS, the centralized BS scheduler determines a packet transmission opportunity in time slots to meet the requirements from all authorized MSs that are using the available channel resources. In the time division duplex (TDD) mode, a deterministic signal is transmitted from one MS to the BS. Figure 1 illustrates the frame-based data transmission in TDD mode, which consists of the preamble, Frame Control Header (FCH), uplink channel descriptor (UCD), downlink channel descriptor (DCD), DL map (DL-MAP), DL map (UL-MAP), bandwidth request message (BW-REQ), and data burst at the MAC layer. These management messages can be periodically transmitted in the subframe. The sounding signal is

* Corresponding author at: Department of Electronic Engineering, National Taiwan University of Science and Technology, #43, Sec.4, Keelung Rd., Taipei, Taiwan. Tel.: +886 2 27376379.
E-mail addresses: D9602206@mail.ntust.edu.tw (F.-M. Yang), wmchen@mail.ntust.edu.tw, wmchen@mail.ntust.edu.tw (W.-M. Chen), jcw@mail.sju.edu.tw (J.-L. Wu).

1084-8045/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jnca.2013.04.008

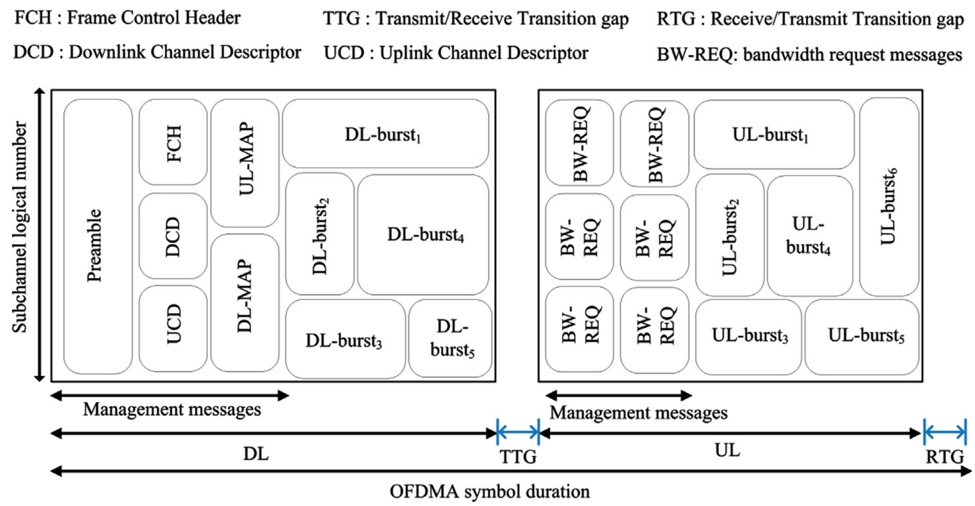

Fig. 1. The frame-based data transmission in TDD mode structure.

transmitted for uplink in the channel sounding zone. The sounding channel is only defined for the uplink quality of signal, not for the received signal quality downlink (IEEE C802.16e-04/103r2, 2004; Zwick et al., 2005).

1.1. Contributions of this paper

In this paper, we propose a simple efficient packet scheduling and bandwidth allocation (PSBA) to guarantee a specified delay constraint and to meet the high throughput requirement. We focused on a centralized PMP architecture, with which the BS shares a common uplink or downlink channel to multiple MSs. Specifically, the contributions of this paper are listed as follows:

- The proposed approach simultaneously considers the user QoS and system capacity expansions of the bandwidth allocator.
- This study maximizes the bits rate for users, and achieves a spectral efficiency that is close to the optimal solution.
- This proposed approach is to guarantee QoS requirements and to provide high traffic throughput in IEEE 802.16e.
- Considering non-starvation of BE traffic, real time traffic can experience the lower packet delay.

1.2. Related work

We considered an Orthogonal Frequency Division Multiplexing (OFDM) physical layer specification with a TDD mode, in which a frame is divided into downlink and uplink subframes. In the downlink subframes, the uplink traffic from MSs to BS is scheduled according to uplink map messages. The BS controls the assignments on the uplink channel and determines a number of slots for the MSs to transmit BW-REQs and data bursts. OFDM technology dynamically allocates slots to various users in a multi-user environment to fully utilize the resources from various symbol times, various frequencies, and various users with distinct channel responses.

The resource allocation mechanism is crucial for QoS requirements in WiMAX networks. Two types of transmission opportunities are available for MSs to send their BW-REQs, which are *centralized polling* and *contention based random access* (Qiang et al., 2007). In centralized polling, each MS transmits a BW-REQ when it is polled by the BS. If the BW-REQ is successfully received, the MS receives bandwidth grant to transmit its data packet; and in contention based random access, all MSs contend to obtain transmission opportunities for sending requests by using contention resolution mechanisms. The previous study in Ma et al. (2010) investigated the methodology of bandwidth allocation for various traffic classes. In the Q-Aware model (Niyato and Hossain, 2006), the BS first examines the queue status and subsequently dynamically assigns bandwidth to the various types of traffic. This approach guarantees QoS for higher-priority traffic, but it can over assign bandwidth to real-time traffic. Previous research presents an adaptive bandwidth allocation (ABA) for meeting the minimal throughput requirements and end-to-end delay constraints to discuss non-starvation of BE traffic (Sheu and Huang, 2011). However, real time traffic can experience the higher packet delay because optimal channel quality can obtain the lower bandwidth. The Adaptive Modulation and Coding (AMC) (Kwon et al., 2005) scheme not only improves the cell throughput but also allocates bandwidth efficiently through user scheduling and bandwidth allocation. The signal-aware dynamic channel allocation (SDCA) scheme, as proposed in Chen and Tan (2007), allocates subchannels resources based on the distance between the BS and MSs. With the calculations of distance constraint, SDCA cannot allocate resources to MSs dynamically based on AMC.

The ABA scheme created a new adaptive bandwidth allocation model for multiple classes. The contribution of ABA scheme not only can meet the minimum throughput requirement of non-real-time packet, but it also avoids any possible starvation of BE traffic. The real time packet delay time of our PSBA mechanism exhibits a higher performance in the WiMAX system on account of the poor packet scheduling policy. Compared to the ABA scheme, the PSBA dynamic manner in which the frames are changed to match the channel conditions, and packet scheduling helps in increasing the packet allocation probability, thereby increasing the number of real time traffic and their quality.

The IEEE 802.16e standard proposes that the Downlink Carrier to Interference and Noise Ratio (DL-CINR) report its operation, which enables MS to determine the channel states. The BS obtains the DL-CINR channel report according to the report response (REP-RSP) message on Channel Quality Information CHannel (CQICH) (Zhang and Letaief, 2004; Bchini et al., 2009). The REP-RSP message is sent by MS in response to the REP-REQ message from BS to report the DL channel status. WiMAX achieves TDD and Frequency Division Duplex (FDD) modes. In the downlink subframe, the synchronization is based on UL-MAP and DL-MAP messages, which comprise the bandwidth for data burst in both downlink and uplink directions. The lengths of uplink and downlink subframes are allocated dynamically by UL-MAP and DL-MAP

Please cite this article as: Yang F-M, et al. A dynamic strategy for packet scheduling and bandwidth allocation based on channel quality in IEEE 802.16e OFDMA system. Journal of Network and Computer Applications (2013), http://dx.doi.org/10.1016/j.jnca.2013.04.008

messages (Jeon and Cho, 2006; Huang et al., 2011). The uplink subframe contains transmission opportunities that are scheduled for the purpose of sending BW-REQs. The bandwidth allocation and request mechanism is responsible for providing a function of the QoS of a service and of the link quality.

Bandwidth allocation mechanism operates by a BS in a WiMAX network. Effectively bandwidth allocation for various traffic classes is an important issue. The challenge of this utility maximization problem is to determine the bandwidth of each subscribed user to be served and the amount of resources to be allocated to each packet of all subscribed user. To resolve the above problem, the proposed approach schedules an appropriate order of packet types and examines the channel status, and subsequently dynamically assigns bandwidth to the various service requirements. The proposed scheme comprises the following two design components: a packet scheduler to guarantee QoS requirements, and a bandwidth allocator to provide high traffic throughput.

2. The proposed algorithm for packet scheduling and bandwidth allocation framework

In this paper, we propose the PSBA mechanism, including packet scheduling and bandwidth allocation, to guarantee a specified delay constraint and to meet the high throughput requirement. We consider a Markovian channel model and study the performance. The communication between MSs and the BS occurs through two directions, that is, uplink and downlink. As packets travel through a BS, a number of packet losses occur for various traffic types and suffering the inaccessible bandwidth after bandwidth allocation.

2.1. Queueing model for traffic class

The following queuing model principles were considered to conform to QoS requirements and to design an efficient service architecture based on delayed latency of real time services and channel state. From the CQICH allocation method, the channel quality information is relayed back to the CQICH reports through the Adaptive Modulation and Coding (AMC) channel controller within BS. The AMC channel selector also determines the Modulation and Coding Scheme (MCS) to maximize user throughput. The IEEE 802.16 medium access control (MAC) protocol defines several bandwidth request allocation mechanisms and five types of traffic characteristics based on priority, including unsolicited grant service (UGS), extended real-time polling service (ertPS), real-time polling service (rtPS), non-real-time polling service (nrtPS), and best effort (BE) (Cicconetti et al., 2006). Figure 2 illustrates the schematic diagram of the system model. The BS scheduler determines the uplink or downlink packets scheduling. Traffic of the same type from a wired and wireless network is aggregated into a traffic class queue following a single Poisson arrival process in the BS buffer. All type of packets are classified and aggregated according to service flows. The UGS traffic is assigned a fixed departure rate to meet its stringent delay requirements. The ertPS traffic changes the allocated bandwidth dynamically, depending on the traffic characteristics. The rtPS traffic type guarantees a specified delay constraint. The nrtPS type is the delay-tolerant service, and the BE data stream does not have any specific QoS requirements. The system assigns initial bandwidth for various types of traffic as λ^{UGS}, λ^{ertPS}, λ^{rtPS}, λ^{nrtPS} and λ^{BE}, based on the requested bandwidth of UGS and ertPS, the required minimum bandwidth of rtPS and nrtPS, and the queue length of BE traffic, respectively. The allocated bandwidth of design parameter for the various traffic classes are μ^{UGS}, μ^{ertPS}, μ^{rtPS}, μ^{nrtPS} and μ^{BE}. The allocated bandwidth $\mu^{UGS}+\mu^{ertPS}+\mu^{rtPS}+\mu^{nrtPS}+\mu^{BE}\leqq\mu$, where μ is the upper bound available bandwidth within the BS (Sheu and Huang, 2011). The average packet allocated bandwidth by five traffic class is less than the available bandwidth within the BS, which implies that the theoretical queuing model can serve as an upper bound to all of traffic.

Model assumptions: The assumptions of the availability channel model are provided as follows:

1. The arrival of the packets to the BS buffer follows the Poisson process with the arrival rate λ, and the service time for each user to complete a packet is exponentially distributed with parameter μ.
2. The state n represents the number of packets within the frame duration.
3. Assume that N homogeneous users are running in parallel channels during a single WiMAX frame time, and each channel has a fixed coding rate r_k ($1\leqq k\leqq N$).
4. Let r_{max} and r_{min} be the MCS of the maximum coding rate and the minimum coding rate, respectively. (where $r_{max}=r_1>r_2>r_3>\ldots>r_N=r_{min.}$)
5. Denote the probability mass function for K to obtain the value of k by $g(k)=\Pr(K=k)$, where $k=1, 2,\ldots, N$.
6. The service rate is μ_k, which is exponential to the coding rate r_k and $g(k)=\mu_k/\mu$.

For the first, second and third assumptions, the packets of the BS buffer are assumed to follow a Poisson process, which may be explained as either in the operational phase or in a steady state. Their assumptions of independent channels in the BS are an efficient approximation to reality because the packets that are served by the various channels are uncorrelated. The service time, in accordance with the exponential distribution, has also been widely accepted. A frame is generally allowed to contain packets for various users rather than for a single user (Ma et al., 2010; Sheu and Huang, 2011). Assumption 4 for channel quality information and queue state information is suggested to achieve high throughput while balancing the fairness among multiple users. The packet departure time with a higher coding rate is ahead of those with a lower coding rate channel. Hence, a higher coding rate channel can serve a lower number of waiting packets in a process time. The fifth and sixth assumptions of the actual number of packets in any departure module is also a random variable K, which can obtain any positive integral value less than N with probability $g(k)$. If μ_k is the departure rate of batches of size K, then $g(k)=\mu_k/\mu$, where μ is the composite departure rate of all batches and is $\sum_{i=1}^{N}\mu_i$.

The conventional Markov model was implemented to analyze the traffic class queue model. The Markov model was constructed for the packets queue. The process of the packets queue assumes that N channel servers are working. The Markov process is illustrated in Fig. 3, based on the assumptions in our study. The state n represents the number of packets in the BS buffer queue. According to the assumption, the packet completing rate is μ and the probability distribution is $g(k)$ for the number of departure packets in the channel of coding rate r_k. Subsequently, the transition completing rate from state n to state n-k is $\mu g(k)$. Denote P_n as the steady probability for the system remaining at state n. It is easy to derive P_n by solving the following Chapman–Kolmogorov:

$$\begin{cases}(\lambda+\mu)P_n=\lambda P_{n-1}+\mu\sum\limits_{i=1}^{N}P_{n+i}g(i) & \text{if} n\geq 1\\ \lambda P_0=\mu P_1 & \text{if} n=0.\end{cases} \tag{1}$$

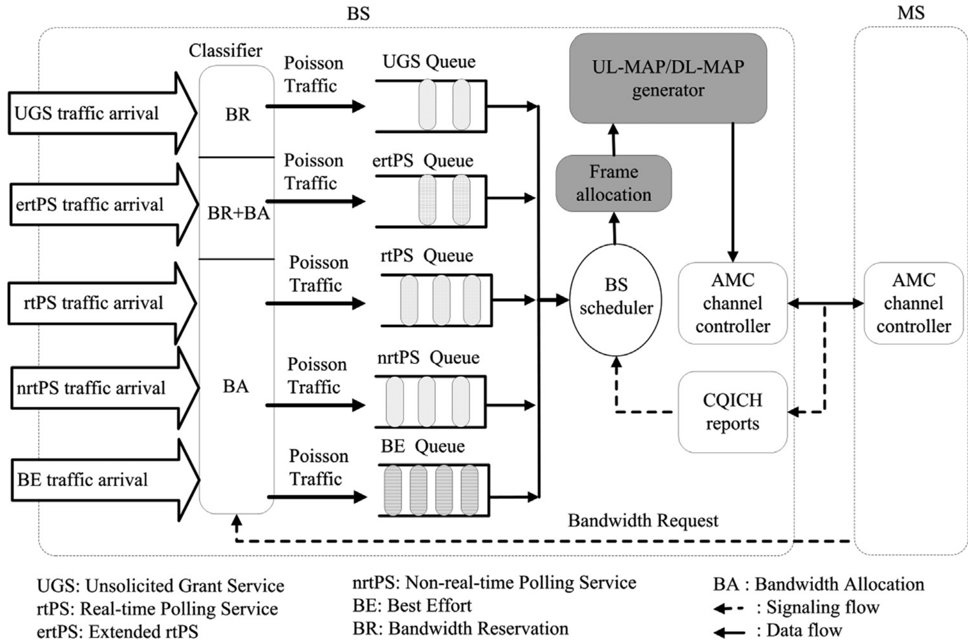

Fig. 2. Schematic design diagram of system model.

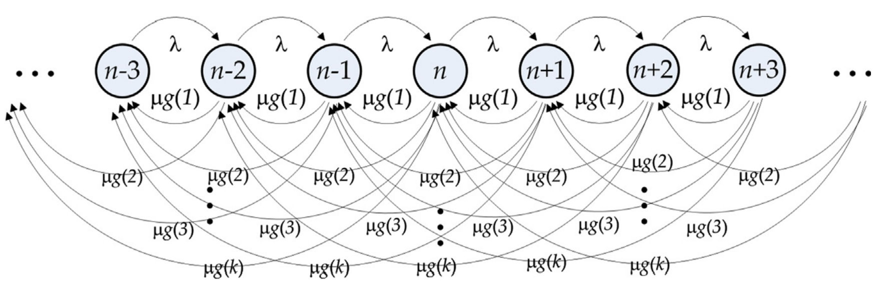

Fig. 3. Markov model for the packets queue.

For each state n, there is also Eq. (1), which can be rewritten in operator notation as

$$\left[\mu\left(\sum_{i=1}^{N} g(i)D^{i+1}\right) - (\lambda+\mu)D + \lambda\right]P_n = 0 \ \text{where} n \ge 0, \ \ D > 0.$$

We determined the characteristic equation $f(r)$ by using the boundary condition$\Sigma P_n = 1$, then

$$f(r) = \mu\left(\sum_{i=1}^{N_q} g(i)r^{i+1}\right) - (\lambda+\mu)r + \lambda = 0$$

Thus, the optimization solving for the designing problem can be obtained by

$$r_s = \min\{r|0 < r < 1 \ \ \text{s.t.} f(r) = 0\}.$$

The mean queuing length in the system (that is, L_q), which should be smaller than the BS buffer length, can be calculated from $L_q = r_s^2/(1-r_s)$. Let W_q be the mean of queuing time of the system, which can be obtained by Little's formula, Gross and Harris (1998) expressed as

$$W_q = \frac{r_s}{\lambda(1-r_s)} - \frac{1}{\mu} \tag{2}$$

Specifically, Eq. (2) computes the delay time of the queue by considering the packet arrivals and the packet departures. Increasing the number of high channel coding rate operation decreases the system waiting time. The BS has notable flexibility in controlling the downlink and uplink by the bandwidth allocator. Let T_r and T_d be the size of the request slots and data slots (both in frames), respectively. We defined bandwidth efficiency as the ratio of the average time utilized by the MSs for data transmissions in a frame to the total time allocated for N_r request slots and N_d data slots in a frame. The probability of receiving request messages in a frame is a binomial distribution. Let N_s denote the average number of MSs that receive bandwidth grant from the BS, which can be computed from:

$$N_s = \sum_{j=0}^{N_r} \min(j, N_d)\binom{N_r}{j}\tau^j(1-\tau)^{N_r-j}, \ \ j \in [0, N_r],$$

Where τ is the probability that a MS transmits a data burst in a frame. The ratio of the average number of slots that are used to transmit data to the total number of slots allocated to transmission bursts for the requesting mechanism in a frame is the bandwidth efficiency (η), which is obtained as

$$\eta = \frac{N_s \times T_d}{N_r \times T_r + N_d \times T_d}$$

Similarly, we defined throughput of a BS (denoted by θ) as the average number of bits transmitted from MSs to the BS in one cell. The throughput of a BS is markedly dependent on the physical channel coding rate profile. Subsequently, the throughput θ of a channel can be computed by

$$\theta = \frac{(\lambda^{UGS} + \lambda^{ertPS} + \lambda^{rtPS} + \lambda^{nrtPS} + \lambda^{BE})}{\sum_{i=1}^{N} g(i)\mu_i} \times N \times \eta$$

2.2. Packet scheduling and bandwidth allocation mechanism

A more robust modulation is used as the channel state deteriorates. The delay latency and channel state of real time services conform to QoS requirements (Fattah and Leung, 2002; Liu and Layland, 1973). The real-time queue at the transmitter must provide a queue delay state for the Earliest Deadline First (EDF) calculator to schedule real-time packets. For non-real time services, the delay constraint is insensitive and is in the acceptable range. The experiments guarantee the quality of real time applications and the high throughput requirement. However, a number of traffic classes experience the arrival of the packets after the arrival time of deadline constraint because packets with superior channel quality obtain the most resources to access slots. The QoS guarantee is the main factor in the PSBA design based on the channel quality information and deadline constraint. The proposed approach divides the PSBA mechanism into the packet scheduler and bandwidth allocator of the BS. The PSBA mechanism stores the channel quality information of the duration of the frame time.

This mechanism also uses the following parameters:

- N_q: the total number of packets for a buffer within a BS in a queuing system;
- $L_{i,j}$: the packet length of the ith packet in bytes at jth MS;
- C_j: the CQI value at the jth MS, which can be obtained from the channel quality information measured from CQICH and MCS;
- d_j: the maximum packet access delay for jth MS;
- $d_{i,j}$: the deadline of packet access delay of the ith packet at jth MS.

Given $g(k)$, the expected system waiting time of the frame allocation process is minimized. Increasing the number of high channel coding rate operation decreases the system waiting time. When possible, various channel coding rates allocation stream types are combined if the packets are allocated for the frame. This is the main reason that the various locations of users have various channel coding rates. However, the high channel coding rate accesses a large number of slots, causing starvation for the lower channel coding rate applications in the IEEE 802.16e environment.

2.2.1. Packet scheduling

The packet scheduling contains two parts, which represent the stream type transmission traffic application that is contributed by real time traffic and non-real time traffic. Let T be the transmission stream type of packet. The channel state part of one packet is C_j/N_q. Considering the deadline of traffic classes, the following equation schedules the stream type as

$$T(i,j) = \underset{j}{\operatorname{argmax}} \left\{ \frac{C_j \times T_d}{N_q \times d_{i,j}} \right\} \tag{3}$$

By solving the equality in Eq. (3), we obtain the stream type transmission traffic application for the user. The delay time can be minimized with the optimal values, which are obtained by the C_j and $d_{i,j}$ parameters. However, the solution of the optimization Eq. (3) is not explicit because C_j and $d_{i,j}$ are independent and non-synchronous. Thus, it is necessary to determine the relationship between channel quality information and the deadline of traffic classes, and provide a queuing aware suboptimal frame allocation solution.

Our mechanism allocates packets to build a frame allocation model. A channel contains only one coding rate. Figure 4 depicts the flowchart of the frame allocation process. After the priority decision on the stream type in the WiMAX standard was conducted, the highest priority was the stream type that had the earliest deadline and optimal channel state. The allocation process assigns the stream types that occupy subframe traffic to the BS

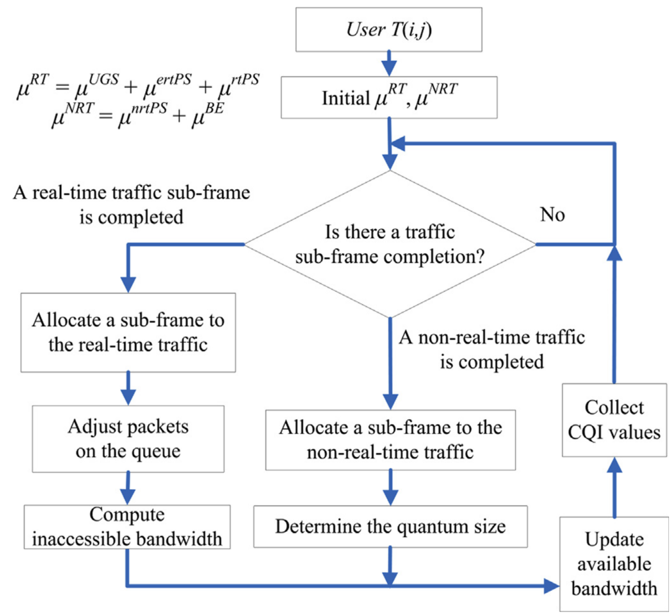

Fig. 4. Flowchart of the dynamic packet scheduling.

based on their CQI values. For a gain of update BS buffer, the real-time stream type must adjust their packets from the frames in order. The allocation frame priority is also based on channel quality order; the lower priority for non-urgent stream type serves as the inferior channel state of the frame. The user allocation process ensures that the non-real-time stream types assume that the FCFS scheduling is based on fairness. The frame allocation process guarantees a low packet delay time between several stream types. Consequently, users with a superior channel condition experience higher perceived quality among several stream types. To combine the ability of providing fair queuing and the simplicity in implementation, non-real-time stream frame allocation must deficit the quantum size for all users. Each user can have various packets, and the stream types of these packets can have various transmission coding rates. These streams arrive at the frame with a Poisson distribution at a certain rate. The processing depicts a queuing model based on a preemptive system. Thus, the type of unused packets among MSs, which is denoted as $T(i,j)$, is represented as

$$T(i,j) = \begin{cases} \operatorname{argmin} j(L_{i,j}/c_j) & \text{if } d_{i,j} < d_j \\ j & \text{if } d_{i,j} \geq d_j \end{cases} \tag{4}$$

The proposed frame allocation process assumes that the 802.16e system class agrees with the $M/M/c$ model in allocating the frame process at the MAC layer. The WiMAX environment includes our stream types with various frames. The objective of the algorithm is to reduce the packet delay and improve the bandwidth utilization while achieving the determined target delay.

The overall throughput can be higher if we reduce the packet delay for the mechanism because more bandwidth may be assigned for a shorter size of the packet. The detailed frame allocation of real time packet scheduling is presented in Fig. 5. Assume that T_d=0.7 (millisecond) is the given frame size with a fixed number of slots for both FCFS and channel condition mechanisms. However, the packet scheduling can divide one frame into several packets. This example illustrates the FCFS case when three transmission coding rates (64QAM, 16QAM and QPSK) have its own user, the first packet at user3 (that is, $T(1,3)$) arrives at the BS buffer before the arrival of the third packet at user 2 (that is, $T(3,2)$), however, $T(3,2)$ will be blocked since the FCFS will ensure

Please cite this article as: Yang F-M, et al. A dynamic strategy for packet scheduling and bandwidth allocation based on channel quality in IEEE 802.16e OFDMA system. Journal of Network and Computer Applications (2013), http://dx.doi.org/10.1016/j.jnca.2013.04.008

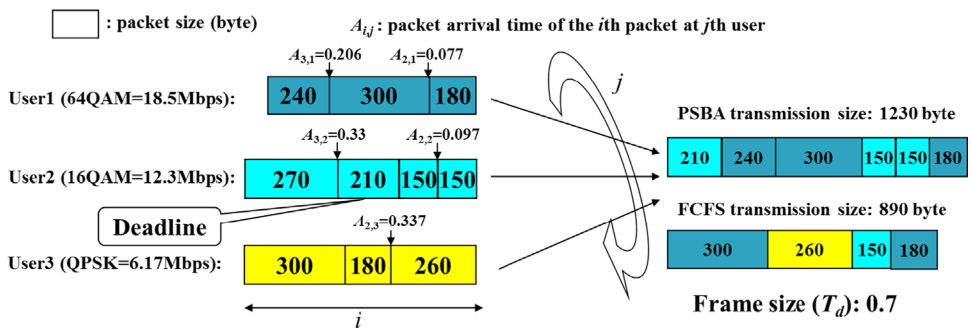

Fig. 5. Example of the dynamic real time packet scheduling.

the first packet arrival time arrives before the second or third packet arrival time, hence $T(3,2)$ will not be served even it reaches the deadline of traffic classes. In PSBA mechanism of this system, $T(3,2)$ will have preemptive priority over $T(1,3)$. The results of bandwidth utilization are similar with various numbers of users. This indicates that the bandwidth utilization is markedly related to the channel condition in the system by Eq. (4). Therefore, if multiple connection users result in improved bandwidth utilization, the user with the type of unused packets (that is, $T(i,j)$) is selected. Our results consistently revealed superior bandwidth utilization that is produced by either unicast polling only or by contention resolution only.

2.2.2. Bandwidth allocation

The design goal of the request protocol is for MSs to periodically aggregate bandwidth requests. The polling-based bandwidth request is focused on how to effectively allocate bandwidth to various traffic classes. The remaining bandwidth, if any, is allocated to the lower traffic classes that still have traffic in the BS buffer. This paper also allows a polling-based bandwidth request on data frame transmissions for traffic class. The bandwidth request opportunity size model is a type of subcarrier permutation defined in IEEE 802.16e OFDMA PHY. The subcarriers are distributed pseudo-randomly per sub-channel, and multiple users access multiple subcarriers. The set of used subcarriers is partitioned into logical clusters, and the group gathers their clusters (Stiakogiannakis and Kaklamani, 2009).

The overall throughput can be higher if we reduce the packet delay for the mechanism because more bandwidth can be assigned for a shorter size of the packet. However, the packet scheduling can divide one frame into several packets. The results of bandwidth utilization are similar with various numbers of users. This indicates that the bandwidth utilization is markedly related to the channel condition in the system. For non-real time connections, the MS requests additional bandwidth without piggybacking its amount on the packet header. To guarantee the bandwidth request, we used deficit round-robin servicing with a quantum size assigned to each packets (Shreedhar and Varghese, 1996). If the packet size was too large, the remainder from the previous quantum was added to the quantum for the next round. The quantum size of non-real time traffic class is defined as

$$S_u = \frac{\sum_{j=1}^{N} C_j}{N} \times T_d \tag{5}$$

This indicates that the low channel quality or high channel quality is dependent on the value of the S_u variables. When the first user is serviced, the quantum size is assigned to the bandwidth of the first user. The remainder, after servicing the first user, is the remaining quantum size. If the requirement bandwidth of next user is carried over to the amount of remaining quantum size, it is assumed that the additional bandwidth is equal to the quantum size.

The coding rate of ertPS traffic is adaptive according to BW-REQ (Sabry et al., 2009). Figure 6 illustrates that the BW request mechanism of ertPS when the rate changes as well as the BW granting scheme. If the transmission rate must be changed, explicit BW requests are sent by piggybacked BW-REQ headers. Otherwise, the coding rate equals the rate that was used in the last BW request. Although the bandwidth allocation of ertPS is similar to that of UGS, some allocated bandwidth is inaccessible because the coding rate changes frame by frame. The inaccessible bandwidth (H^{ertPS}) is expressed as follows:

$$H^{ertPS} = \int_0^{T_f} h^{ertPS}(t)dt = \int_0^{T_f} \sum_{i=1}^{N_c} t \times (R_i/E_i - R_{i+1}/E_{i+1})dt,$$

where N_c denotes the total number of coding rate variations. Let R_i be the ith coding rate, and E_i is the number of packets in R_i coding rate. Let T_f be the duration of the frame. Piggybacking, as defined in the IEEE 802.16 standard, is a method that is used by MSs to transmit BW-REQs, and is optional and not able to transmit all types of bandwidth requests. The inaccessible bandwidth is because ertPS also uses UGS for resource allocation and the video rate changes on frame by frame basis.

In the proposed bandwidth allocator, we first assigned a fixed departure rate to UGS to fulfill its stringent delay requirements. For the least important traffic class (BE), we also prevented it from starvation. The goal of this study was to improve throughput requirement with QoS guarantee. The packets were selected from each stream type upon reaching the frame because priorities of stream types are decided by the packets scheduling. The stream type with the highest data rate acquires the highest priority for accessing a frame. The granted bandwidth of the stream type is a critical parameter in the bandwidth allocator. First, the bandwidth allocator assigns initial bandwidth, r^{UGS}, r^{ertPS} and r^{rtPS}, based on the requested bandwidth of UGS, and the required minimum bandwidth of ertPS and rtPS, respectively. If a non-real time traffic class exists, the bandwidth allocator assigns the quantum size of non-real time traffic class (S_u) to bandwidth for μ^{nrtPS} and μ^{BE}, if any of the traffic classes have packets that are waiting in the buffer. Let A be the inaccessible bandwidth for all stream types. The assignment of bandwidth by the bandwidth allocator in a BS is described as Fig. 7.

The BS allocates bandwidth among various users in the service class. As the nrtPS stream changes its polling size, to obtain extra bandwidth and avoid a sudden request, the bandwidth allocator allocates bandwidth based on the inaccessible bandwidth (H^{ertPS}) and the quantum size (S_u).

3. Experimental results

The experiments in this paper were performed by using OPNET Modeler 14.5 for an educational version with WiMAX Module capability to simulate an IEEE 802.16e environment

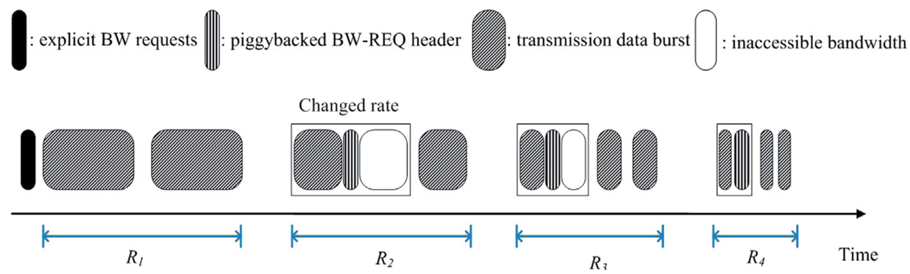

Fig. 6. Example of inaccessible bandwidth in ertPS.

Table 1
Fundamental simulation AMC level parameters.

Modulation	Coding	Bits/symbol/Hz	Required SNR (dB)
QPSK (r_1,r_2)	1/2	1	9.4
	3/4	1.5	11.2
16QAM (r_3,r_4)	1/2	2	16.4
	3/4	3	18.2
64QAM (r_5,r_6)	2/3	4	22.7
	3/4	4.5	24.4

Input : All variables specified in packet scheduling

Output : μ^{UGS}, μ^{ertPS}, μ^{rtPS}, μ^{nrtPS} and μ^{BE}

for u=1 **to** N **do**

$\mu^{UGS} \leftarrow r^{UGS}$;

$\mu^{ertPS} \leftarrow \min(r^{ertPS}\text{-}H^{ertPS},\ \mu - \mu^{UGS})$;

$\mu^{rtPS} \leftarrow \min(r^{rtPS},\ \mu - \mu^{UGS} - \mu^{ertPS})$;

  if ($L_{i,j} \leqq S_u$)

  $\mu^{nrtPS} \leftarrow \min(L_{i,j},\ \mu - \mu^{UGS} - \mu^{ertPS} - \mu^{rtPS})$;

    else if ($L_{i,j} > S_u$)

    $\mu^{nrtPS} \leftarrow \min(L_{i,j}+S_u,\ \mu - \mu^{UGS} - \mu^{ertPS} - \mu^{rtPS})$;

    end if

  end if

end for

$A \leftarrow \mu - \mu^{UGS} - \mu^{ertPS} - \mu^{rtPS} - \mu^{nrtPS}$;

for u=1 **to** N **do**

  if ($A \geqq 0$)

  $\mu^{rtPS} \leftarrow A\text{-}\min(\mathrm{BE})$;

  $\mu^{BE} \leftarrow \min(\mathrm{BE})$;

  end if

  Modify S_u using equation (5);

end for

Fig. 7. Dynamic bandwidth allocator.

(http://www.opnet.com/WiMAX, 2012). To investigate the stable state result for users, the simulation was conducted for 3600 s. We conduct simulation experiments to verify our proposed scheme. The experimental results revealed that the throughput of PSBA is 21% higher than that of the IEEE 802.16e standard model.

Table 2
Traffic class service parameters.

Service class	Type	Maximum sustained traffic rate	Minimum reserved traffic rate	Deadline constraint
Gold	ertPS	64 kbps	64 kbps	60 ms
Silver	rtPS	1 Mbps	0.5 Mbps	25 ms
Bronze	nrtPS	384 Mbps	384 kbps	N/A
Platinum	UGS	2.5 Mbps	2.5 Mbps	N/A

3.1. Simulation environment

We measured the performance in a typical network, which was composed of seven cells and 35 MS nodes that were randomly distributed. The experimental environment is based on a simulation OPNET WiMAX network topology which contains seven cells. Each cell has a BS, and the BS locates five MSs on a cell. The Transmission Powers of MS and BS Node were set to 0.5 W and 5 W, respectively. All the transceiver antennas were used in the omni-direction antenna. These simulations assumed that the moving mode of the MSs follows the random waypoint mode. The movement speed of each MS was assumed to be uniformly distributed between 1 and 5 m/s. The OPNET node models of BS and MS are WiMAX_bs_ethernet4_slip4_router and WiMAX_ss_workstation, respectively. The fixed BS node model featured router functionality. The BS node had four Ethernet interfaces, four Serial Line Internet Protocol (SLIP) interfaces, and one WiMAX interface. The MS node model featured workstation functionality. The global configuration object was used to configure the parameters, such as service classes and PHY profile, in the WiMAX_Config node.

This study used the path loss and log-normal shadowing of the channel model strategy (IEEE 802.16j-06/013r3, 2007), and the AMC level was six. Table 1 summarizes the AMC level parameters in the OPNET simulator. The supported modulation and coding rates are as follows: QPSK (1/2, 3/4), 16QAM (1/2, 3/4), 64QAM (2/3, 3/4). The modulation and coding rate changed dynamically depending on the signal-to-interference-noise ratio. The simulator also implemented HARQ and ARQ messages, because real-time service is the delay sensitive type. If the departure time of real-time packets occurred after the deadline time, the packet retransmission mechanism was implemented. Table 2 presents the experimental data for traffic class services. The WiMAX traffic was generated between the BS and the users. The application server was the source of download for all users, going through the BS and MSs. The deadline of real time traffic is critical to the system operation, which is the maximum delay latency guideline. The deadline values of ertPS and rtPS were assigned as 60 ms and 25 ms, respectively (Kim and Hwang, 2009; Chakchai et al., 2009).

3.2. Simulation results

Figure 8 illustrates the system throughput for the IEEE 802.16e standard, ABA (Sheu and Huang, 2011) and PSBA mechanism.

The simulation results in Fig. 8 indicate that saturation occurs when the simulation time reaches 1200 s. Saturation indicates that all time slots of one frame have been scheduled. The ABA scheme reserves the extra bandwidth and reallocates the remaining bandwidth. By using the ABA scheduling scheme, the maximal average throughput reaches 5.55 Mbps before the system becomes saturated. The PSBA mechanism with a packet scheduling and bandwidth allocation process achieves higher performance because it has a higher average throughput. This is the main reason for obtaining the channel quality information and the inaccessible bandwidth. After the channel condition reaches an improved state due to the transmission coding rate, the system can transmit more data with fewer time slots. Because PSBA considers the method of allocation of the packets, the frames can select the appropriate packet with a superior modulation technique to the wireless channel transmission, thereby achieving a maximal throughput of 6.35 Mbps. This phenomenon demonstrates that the proposed PSBA mechanism markedly improves the throughput compared to the IEEE 802.16e standard and ABA mechanism because of the extra bandwidth consumed and the inferior channel conditions in the IEEE 802.16e standard and the ABA mechanism. Compared to the IEEE 802.16e, the PSBA mechanism achieves 21% higher throughput in this scenario.

Figure 9 demonstrates the relationship among the packet delay times of the ertPS application type for three scheduling schemes. In an IEEE 802.16e and ABA mechanism, the ertPS applications have the highest priority for accessing the frames. High priority applications cause the packet arrival interval to decrease in conjunction with the delay time. Moreover, the ABA mechanism allocates more bandwidth based on the actual bandwidth requests after reserving a portion of the bandwidth for ertPS traffic by piggybacking. According to Eq. (2), MS uses the high channel coding rate to transmit packet, it decreases the packet delay times. Compared to the ABA mechanism, the ertPS packet delay time of the PSBA mechanism exhibits a higher performance in the WiMAX system because each MS monitors its channel state continuously. It gathers the channel quality information that the BS receives through the BW-REQs sent by MSs. However, the persisting inferior channel conditions over a number of radio links may delay relevant packets in the BS buffer for a long period of time. To prevent the transmission of delayed packets from wasting bandwidth, it is necessary to periodically verify the inaccessible bandwidth from the BS.

As illustrated in Fig. 10, the standard delay time of rtPS on average is less than 0.15 ms in the ABA scheme. This demonstrates that the delay time decreases slightly as adaptive bandwidth allocation. The PSBA mechanism proposes that packet scheduling and bandwidth allocation is based on the channel quality information and resumes inaccessible bandwidth. When the channel state is optimal or the traffic load of high priority applications (for example, UGS and ertPS traffic) is low, the rtPS application should also be serviced with a higher coding rate. Thus, the PSBA mechanism is suitable for use in the BS of current WiMAX systems because it has a higher average throughput and lower packet delay time than IEEE 802.16e and ABA mechanism.

Jain's fairness index is a conventional method of assessing the quality of the traffic type (Jain et al., 1984). Figure 11 illustrates the Jain's fairness index value at various traffic applications for 35 nodes, indicating that all of the traffic applications in this paper had relatively high Jain's fairness index values ($0.6 < \mathrm{FI}$). From the chart, we determined that the IEEE 802.16e scheme has a higher fairness than the ABA and PSBA mechanisms. For IEEE 802.16e schemes, the fairness of traffic applications depends on the amount of transmitted traffic by the selected packets and the traditional round robin approach, which processes packets immediately when they arrive or depart to the BS. For the ABA mechanism, the fairness of traffic applications depends on the amount of transmitted bandwidth by adaptive bandwidth allocation. For the PSBA transmission model, a number of ertPS traffic applications are reallocated a portion of bandwidth for non-real time traffic accessing. In the real time traffic type, the fairness index of ABA is higher than that of PSBA because PSBA considers

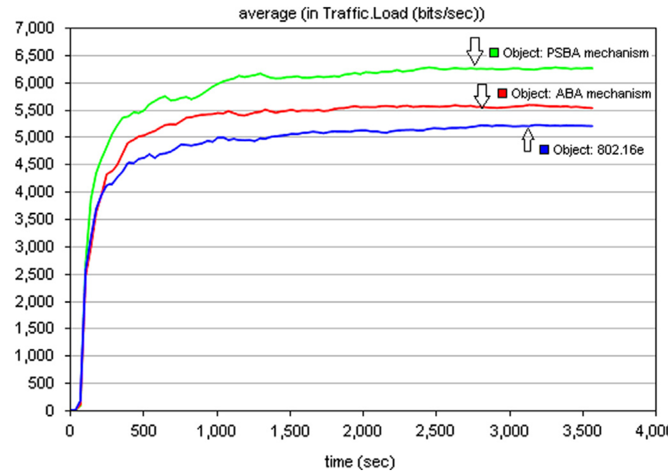

Fig. 8. System throughput.

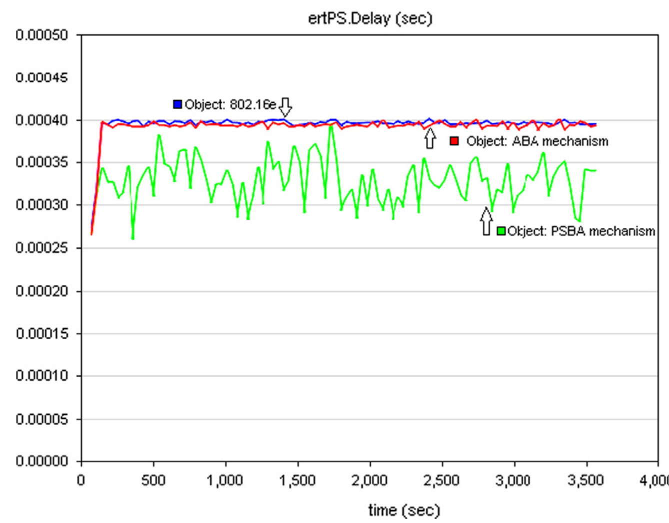

Fig. 9. Delay time of ertPS.

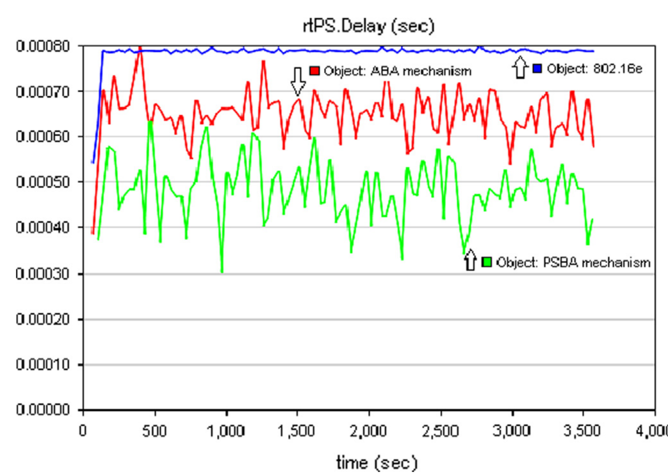

Fig. 10. Delay time of rtPS.

Please cite this article as: Yang F-M, et al. A dynamic strategy for packet scheduling and bandwidth allocation based on channel quality in IEEE 802.16e OFDMA system. Journal of Network and Computer Applications (2013), http://dx.doi.org/10.1016/j.jnca.2013.04.008

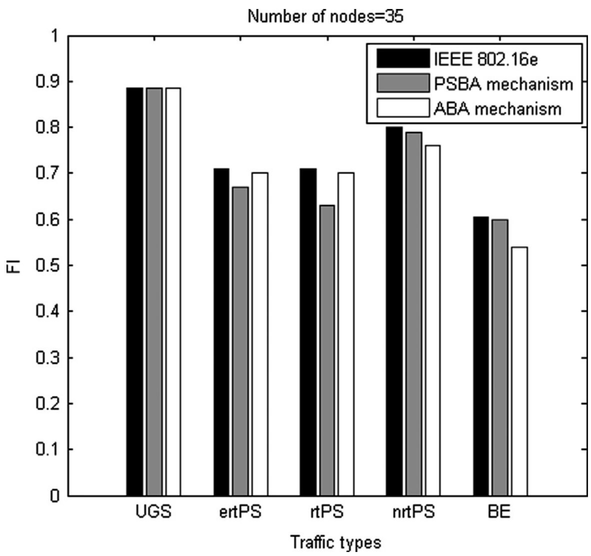

Fig. 11. Jain's fairness index value.

the method of allocating frames from the real time stream based on channel quality and packets of deadline. In the non-real time traffic type, the fairness index of ABA is lower than that of PSBA. This is the main reason that PSBA has more useful and inaccessible bandwidth. This result addresses the fundamental problem of the trade-off between resource efficiency and user fairness in wireless networks that use opportunistic radio resource allocation (Rodrigues and Casadevall, 2011).

4. Conclusions

This paper proposes a two-stage mechanism, called PSBA, to process the operations of packet scheduling and bandwidth allocation based on channel states, QoS, and the bandwidth requirements of multimedia MSs and BSs. This study compares the proposed PSBA mechanism with the system performance of ABA and the standard mechanism that is formed by the IEEE 802.16e in various simulations to evaluate the performance of the packet scheduling as an effective bandwidth allocation approach. The service flow simulations in the OPNET simulator indicated that the PSBA achieved a lower delay time and higher system bandwidth for multiple users. The experimental results indicated that the throughput of PSBA is 21% higher than that of the IEEE 802.16e standard by using the OFDMA technique.

References

Bchini T, Tabbane N, Tabbane S, Chaput E, Beylot, AL QoS in IEEE 802.16e for VoIP and video with mechanism of soft handover – FBSS in highway. In: Proceedings of the international conference on the eighth international conference on networks, Cancun, Mexico; March 2009. p. 88–93.

Chakchai SI, Jain R, Tamimi AK. Scheduling in IEEE 802.16e mobile WiMAX networks: key issues and a survey. IEEE Journal on Selected Areas in Communications 2009;27(2):156–71.

Chen J, Tan W. Predictive dynamic channel allocation scheme for improving power saving and mobility in BWA networks. ACM/Springer Mobile Networks and Applications 2007;12(1):15–30.

Cicconetti C, Lenzini L, Mingozzi E, Eklund C. Quality of service support in IEEE 802.16 networks. IEEE/ACM Transactions on Networking 2006;20(2):50–5.

Fallah YP, Agharebparast F, Minhas M, Alnuweiri H, Leung VCM. Analytical modeling of contention-based bandwidth request mechanism in IEEE 802.16 wireless networks. IEEE Transactions on Vehicular Technology 2008;57 (5):3094–107.

Fattah H, Leung C. An overview of scheduling algorithms in wireless multimedia networks. IEEE Wireless Communications 2002;9(5):76–83.

Ghosh A, Wolter DR, Andrews JG, Chen R. Broadband wireless access with WiMAX/ 802.16: current performance benchmarks and future potential. IEEE Communications Magazine 2005;43(2):129–36.

Gross D, Harris CM. Fundamentals of queueing theory. 3nd edn. New York: Wiley; 99–100.

He J, Guild K, Yang K, Chen HH. Modeling contention based bandwidth request scheme for IEEE 802.16 networks. IEEE Communications Letters 2007;11 (8):689–700.

Huang CY, Chung WC, Chang CJ, Ren FC. An Intelligent HARQ Scheme for HSDPA. IEEE Transactions on Vehicular Technology 2011;60(4):1602–11.

IEEE C802.16e-04/103r2: Signaling methodologies to support closed-loop transmit processing in TDD-OFDMA; 2004.

IEEE 802.16e–2005. IEEE standard for local and metropolitan area networks — Part 16: air interface for fixed broadband wireless access systems; 2006.

Jain, R., Chiu, D.M., Hawe, W. A quantitative measure of fairness and discrimination for resource allocation in shared system. DEC Research Report TR-301; 1984. p. 1–38.

Jeon SY, Cho DH. Channel adaptive CQI reporting schemes for HSDPA systems. IEEE Communications Letters 2006;10(6):459–61.

Kim BJ, Hwang GU. Performance analysis of the ertPS algorithm and enhanced ertPS algorithm for VoIP Services in IEEE 802.16e systems. IEICE Transactions on Communication 2009;E92.B(6):2000–7.

Kwon T, Lee H, Choi S, Kim J, Cho D, Cho S, et al. Design and implementation of a simulator based on a cross-layer protocol between MAC and PHY layers in a WiBro compatible IEEE 802.16e OFDMA system. IEEE Communications Magazine 2005;43(12):136–46.

Liu CL, Layland J. Scheduling algorithms for multiprogramming in a hard real-time environment. Journal of the ACM 1973;20:46–61.

Ma M, Lu J, Fu CP. Hierarchical scheduling framework for QoS service in WiMAX point-to-multi-point networks. IET Communications 2010;4(9):1073–82.

Niyato D, Hossain E. Queue-aware uplink bandwidth allocation and rate control for polling service in IEEE 802.16 broadband wireless networks. IEEE Transactions on Mobile Computing 2006;5(6):668–79.

Qiang N, Vinel A, Yang X, Turlikov A, Tao J. Investigation of bandwidth request mechanisms under point-to-multipoint mode of WiMAX networks. IEEE Communications Magazine 2007;45(5):132–8.

Rodrigues EB, Casadevall F. Control of the trade-off between resource efficiency and user fairness in wireless networks using utility-based adaptive resource allocation. IEEE Communications Magazine 2011;49(9):90–8.

Sabry, A, El-Badawy, H, Shehata, K, Ali, A. A novel resource allocation technique for vbr video traffic in the uplink over WiMAX networks. In: Proceedings of the international conference on information and multimedia technology, Jeju Island, Korea; December 2009. p. 442–8.

Sheu TL, Huang KC. Adaptive bandwidth allocation model for multiple traffic classes in IEEE 802.16 worldwide interoperability for microwave access networks. IET Communications 2011;5(1):90–8.

⟨http://www.opnet.com/WiMAX⟩ [accessed April 2012].

IEEE 802.16j-06/013r3. Multi-hop relay system evaluation methodology (channel model and performance metric); 2007.

Shreedhar M, Varghese G. Efficient fair queuing using deficit round robin. IEEE/ACM Transactions on Networking 1996;4(3):375–85.

Stiakogiannakis, IN, Kaklamani, DI. IEEE802.16e — WiMAX: performance analysis of partial and full usage of sub-channels under fractional frequency reuse. In: Proceedings of the international conference on wireless technology conference (EuWIT 2009), Rome, Italy; September 2009. p. 41–4.

Zhang YJ, Letaief KB. Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems. IEEE Transactions on Communications 2004;3 (5):1566–75.

Zwick T, Beukema TJ, Nam H. Wideband channel sounder with measurements and model for the 60 GHz indoor radio channel. IEEE Transactions on Vehicular Technology 2005;54(4):1266–77.