RESEARCH   **Open Access**

Dynamic admission control and bandwidth reservation for IEEE 802.16e mobile WiMAX networks

Chiapin Wang[*], Wan-Jhen Yan and Hao-Kai Lo

Abstract

The article presents a dynamic connection admission control (CAC) and bandwidth reservation (BR) scheme for IEEE 802.16e Broadband Wireless Access networks to simultaneously improve the utilization efficiency of network resources and guarantee QoS for admitted connections. The proposed CAC algorithm dynamically determines the admission criteria according to network loads and adopts an adaptive QoS strategy to improve the utilization efficiency of network resources. After new or handoff connections enter the networks based on current admission criteria, the proposed adaptive BR scheme adjusts the amount of reserved bandwidth for handoffs according to the arrival distributions of new and handoff connections in order to increase the admission opportunities of new connections and provide handoff QoS as well. We conduct simulations to compare the performance of our proposed CAC algorithm and BR scheme with that of other approaches. The results illustrate that our approach can effectively improve the network efficiency in terms of granting more connections by as large as about 22% in comparison with other schemes, and can also guarantee adaptive QoS for admitted new and handoff connections.

Keywords: IEEE 802.16e WMAN, connection admission control, bandwidth reservation, resource allocation.

1. Introduction

Broadband wireless access networks have rapidly been growing in these years to support the increasing demands of wireless multimedia services, like streaming audio/video, Internet Protocol TV, and video conferencing. Mobile Worldwide Interoperability for Microwave Access (WiMAX), which has been standardized by IEEE 802.16e [1], is one of the most promising solutions to provide ubiquitous wireless access with high data rates, high mobility, and wide coverage. The IEEE 802.16e Media Access Control (MAC) layer provides differential Quality of service (QoS) for various classes of scheduling services, which are 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). Each scheduling class is associated with a set of QoS parameters for quantifying its bandwidth requirement, e.g., maximum/minimum data rates and maximum delays. The radio resources (i.

e., time slots and frequency spectrums) for different scheduling services are centrally controlled by the base station (BS). To provide QoS for data transmissions in WiMAX networks, BS generally applies a Connection Admission Control (CAC) scheme which determines whether a new connection should be established according to the available network resources. Essentially, the effectiveness of CAC schemes can be critical to both the performances of QoS for admitted connections and the utilization efficiency of network resources. However, the IEEE 802.16e standards do not specify how to implement CAC mechanisms and remain that as open issues.

On the other hand, a bandwidth reservation (BR) mechanism is also important to the provisioning of QoS for some prioritized users like users in a handoff process. Handoff occurs when mobile station (MS) transfers its connection from the original serving BS with worse and worse link qualities to a neighboring BS with better qualities. In general, a handoff user will be prioritized over a new incoming user in order to provide better user-perceived satisfaction especially when it is with real-time applications which have specific QoS

* Correspondence: chiapin@ntnu.edu.tw
Department of Applied Electronic Technology, National Taiwan Normal University, Taipei, Taiwan

requirements, e.g., throughput demands and delay/jitter constraints. Since the reserved bandwidth cannot be taken by a new coming user, the design of BR mechanisms can significantly affect the performance of handoff QoS and also the utilization efficiency of network resources.

The CAC and BR problems have largely been investigated in previous study [2-20]. The authors of [2,3] propose to adopt minimum bandwidth requirements as the admission criteria for all classes of scheduling services. The approach can provide more connections admitted into networks but may cause a relatively low QoS performance. The authors of [10] propose to divide the scheduling services into two groups: one group consists of UGS, ertPS, and rtPS which adopt maximum bandwidth requirements for the admission criteria, while another group consists of nrtPS and BE which adopt minimum bandwidth requirements. The approach may over favor the higher-class services and cause a starvation of lower-class services. Instead of using fixed criteria for an admission control as described above, the studies in [11,12] propose to dynamically determine the admission criteria by using a game-theoretic approach. However, it does not take the network load into consideration and may introduce great computational complexities.

With regard to the BR schemes, a fixed guard channel scheme [13] is proposed to reserve a certain amount of bandwidth for upcoming handoff connections to assure seamless handoff processes. When the total bandwidth utilization of existing users reaches the threshold, no more new connections can be admitted into the network. Nevertheless, when a fixed amount of bandwidth can never be used for new connections, a certain portion of network resources may be wasted. The study [10] proposes to dynamically adjust the quantity of reserved bandwidth based on the arrival and departure behavior of handoff connections to make the resource utilization more efficient. However, if handoff connections occur infrequently, the quantity of reserved bandwidth for handoffs is almost fixed and this approach would be similar to the fixed guard channel scheme and cause a waste of network resources as well.

Both CAC schemes and BR mechanisms are important research issues in wireless networks due to scarce radio resources, dynamic channel qualities, and diverse user demands. However, to the best of authors' knowledge, most efforts tackle one of the two problems individually while little work considers the joint design of the two mechanisms. We are thus motivated to present a joint design of CAC and BR mechanisms which aim at simultaneously improving the utilization efficiency of network resources and guaranteeing QoS for admitted new connections and handoff connections. The proposed CAC scheme dynamically determines the admission criteria according to network loads and adopts an adaptive QoS strategy to increase the amount of admitted connections for the network efficiency. The key idea of our CAC scheme is based on the fact that most scheduling services are with adaptive QoS requirements, e.g., maximum and minimum rates. Therefore, the admission criteria can be determined according to the amount of available wireless resources for increasing the number of admitted connections with adaptive QoS. For example, if the network capacity is adequate or sufficient, bandwidth requirements for higher QoS might be adopted as the admission criteria. Alternatively, if the network load is quite heavy, the admission criteria may be degraded to meet lower QoS requirements. After the admission criteria are determined, the proposed BR scheme dynamically adjusts the amount of reserved bandwidth for handoffs according to the arrival distributions of new/handoff connections to increase the connection admission opportunities and also guarantee the bandwidth requirements for handoff QoS. The basic idea of our adaptive BR scheme is a rational inference that generally the occurrences of new incoming connections may be much more frequent than that of handoff connections [21-24]. This observation originates from common BS deployment that the overlap areas of a given BS between its neighboring stations are parts of its coverage area. Since handoffs arise only when users cross through the overlap areas, it is a general situation to observe more new connections occurring than handoff connections. Thus, the optimal BR should take into account the arrival behavior of not only handoff connections, but also new connections in order to avoid a waste of network resource as possible.

We conduct simulations of 802.16e transmission scenarios to evaluate and compare the performances of the proposed CAC algorithm and BR scheme with that of other approaches. Simulations results illustrate that our approach can effectively improve the network efficiency in terms of increasing the number of granted connections by as large as about 22% in comparison with other schemes, and also can guarantee adaptive QoS for admitted new and handoff connections. The remainder of this article is organized as follow. In Section 2, we briefly illustrate the QoS architecture and resource allocation mechanism of IEEE 802.16e networks. Section 3 presents the proposed CAC algorithm and BR scheme. In Section 4, we construct simulation scenarios to demonstrate the effectiveness of our approach. Section 5 draws our conclusions.

2. IEEE 802.16e QoS architecture and resource allocation mechanism
2.1 IEEE 802.16e QoS architecture
The IEEE 802.16e MAC layer provides QoS differentiation for various categories of scheduling services. The

IEEE802.16e uplink scheduling framework is shown in Figure 1. The scheduling of uplink packet transmissions is centrally controlled in the BS. The IEEE 802.16e standards adopt a connection-oriented MAC protocol, i.e., each connection is associated with a connection ID. When a service flow generated at the application layer arrives at the MAC layer, the MS first sends a connection establishment request to the BS. The admission control mechanism at BS then estimates whether the remaining bandwidth can support the QoS requirements of new connections without violating existing users' QoS. If the connection request is accepted, the BS replies with a connection response which indicates the connection IDs for each direction of this connection. After the process of connection establishment is finished, the MS can issue a bandwidth request. The connection classifier then classifies the service data units into different scheduling classes according to their service flow identifier and connection identifier. The uplink bandwidth requests by users are performed on a per connection basis, whereas the BS grants bandwidth on a per subscriber station basis (GPSS). After the BS allocates a certain amount of bandwidth to each of the MSs, the packet scheduler at each MS will redistribute the bandwidth to the corresponding connection. By means of the connection-admission-control mechanism and request-grant bandwidth-allocation scheme, QoS for different scheduling classes can be guaranteed.

The IEEE 802.16e standard divides all service flows into five scheduling classes, each of which is associated with a set of QoS parameters for quantifying its bandwidth requirement. The five scheduling classes are described as follows.

(1) UGS: UGS is designed to support real-time service flows with fixed-size packets generated at periodic intervals (i.e., constant bit rate–CBR), such as T1 services and voice-over-Internet-Protocol (VoIP) applications without silence suppression. This service can grant a fixed amount of bandwidth for CBR real-time applications without any requests.

(2) rtPS: rtPS is designed to support real-time service flows with variable-size packets generated at periodic intervals (i.e., variable bit rate–VBR), such as Motion Pictures Experts Group (MPEG) video. Based on a polling mechanism to request bandwidth periodically, this service can guarantee QoS such as the minimum data rate and maximum latency for VBR real-time applications.

(3) ertPS: The characteristic of this service class is between UGS and rtPS. On detecting that the allocated bandwidth is either insufficient or excessive, ertPS can send a request to change the amount of allocated bandwidth like rtPS does. Otherwise, if the bandwidth demand remains unchanged, ertPS behaves as UGS. ertPS is designed to support VBR real-time data services such as VoIP applications with silence suppression.
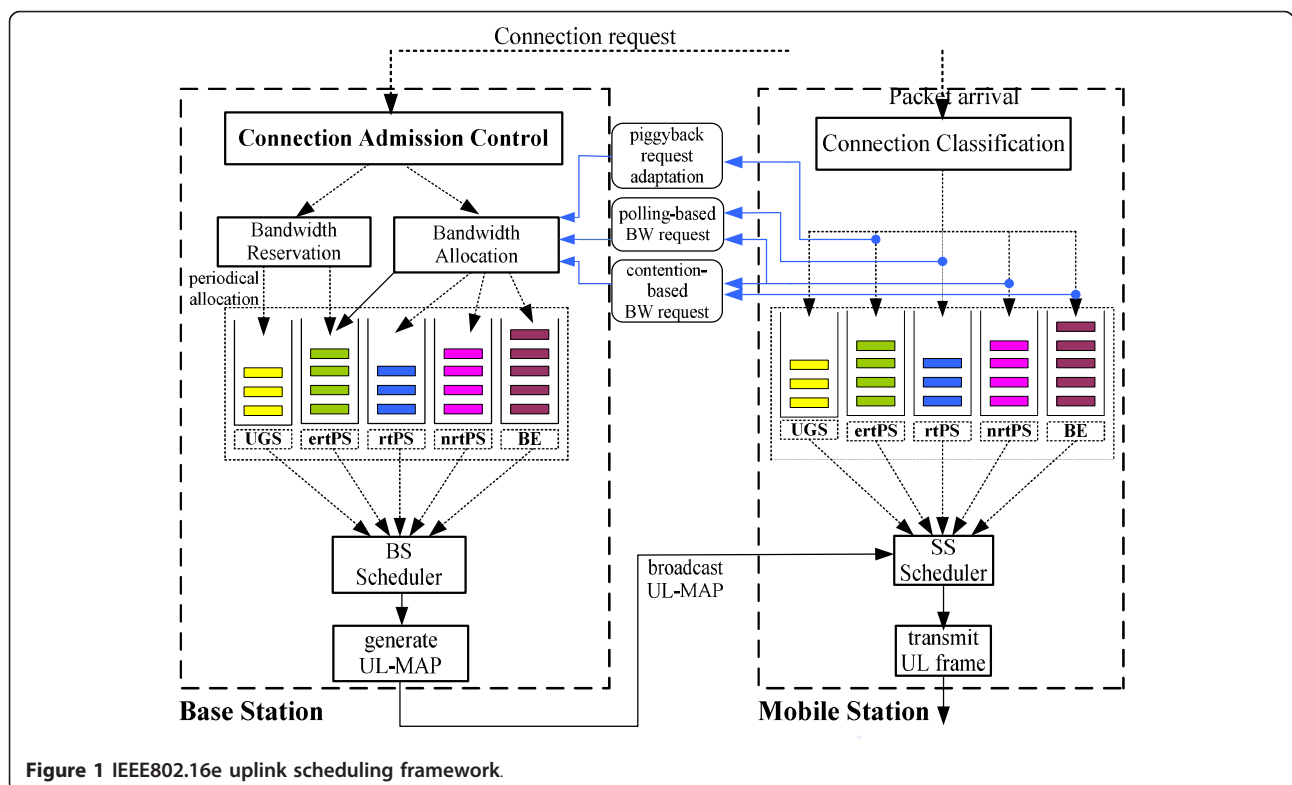

Figure 1 IEEE802.16e uplink scheduling framework.

(4) nrtPS: This service class is to support non-real-time VBR services which require minimum-data-rate guarantees but can be tolerant to delay, such as File-Transfer-Protocol (FTP) applications.

(5) BE: The BE service is designed for best-effort applications which have no explicit QoS requirements, e.g., web services or e-mail.

The QoS parameters and the supporting application types associated with each of the IEEE 802.16e scheduling classes are shown in Table 1.

2.2. Bandwidth allocation mechanism

The IEEE 802.16e physical layer (PHY) adopts an Orthogonal Frequency Division Multiple Access (OFDMA) slot as the minimum possible resource. The IEEE 802.16e PHY supports Frequency Division Duplex (FDD) and Time Division Duplex (TDD) for bandwidth allocation mechanisms. In FDD mode, the uplink (UL) and downlink (DL) channels are located on split frequencies, with which a fixed duration frame is used for both UL and DL transmissions. In TDD mode, the UL and DL transmissions are arranged at different time periods using the same frequency. In this article, we focus on the TDD mode for the IEEE 802.16e bandwidth allocation mechanism.

In TDD mode, Time Division Multiplexing (TDM) is used for DL transmissions and Time Division Multiple Access (TDMA) is used for UL transmissions. As shown in Figure 2, a TDD frame has a fixed duration and contains one DL subframe and one UL subframe whose durations can adapt to the traffic loads of UL and DL transmissions. The DL subframe consists of a preamble, Frame Control Header (FCH), and a number of data bursts. The FCH specifies the profiles of the DL bursts that immediately follow it. The broadcast messages

including downlink map (DL-MAP), uplink map (UL-MAP), DL Channel Descriptor (DCD), UL Channel Descriptor (UCD), etc., are sent at the beginning of these DL bursts. The UL subframe contains a contention interval for initial ranging and bandwidth request and UL PHY protocol data units (PDUs) from different MSs. The DL connections are scheduled by BS in a broadcast manner, while the UL connections apply a request-grant mechanism for bandwidth allocation in a shared manner. The UL bandwidth requests are performed on a per connection basis, whereas the BS grants bandwidth on a per subscriber station basis (GPSS). After the BS allocates a certain amount of bandwidth to each of the MSs, each MS will redistribute the bandwidth to the corresponding connection. The information about bandwidth allocations for DL and UL transmissions is broadcast to the MSs through DL-MAP and UL-MAP messages at the beginning of each frame. Therefore, each MS can receive from and transmit data to BS in the predefined OFDMA slots.

2.3. Packet scheduling mechanism

As shown in Section 2.1, the IEEE 802.16e standard defines five scheduling classes. However, it does not specify the scheduling mechanism for the five classes and the design is left for researchers [25]. The design of a scheduling mechanism must take into account the specific QoS constraints of different applications, e.g. the maximum allowable delay and minimum data rate [3]. A feasible solution is to decide on a service class first according to the characteristics of each class and next choose an appropriate user in the selected class [26]. In the second phase, the packet scheduling of different users among a given class may consider some performance metrics such as throughput and fairness, while the maximum rate scheduling (greedy algorithm) and Proportional Fairness (PF) scheduling can be applied, respectively. The maximum rate scheduling is effective to advance the overall system throughput as it allocates resources to users with relatively good channel qualities among them [27]. On the other hand, the PF scheduling can improve the fairness of channel utilization among users as it distributes resources among them with consideration of their previous records of utilization [28-30].

Throughput and fairness, however, are conflicting performance metrics [31]. To maximize system throughput, more resources should be allocated to the users in good channel conditions. This may cause most radio resources monopolized by a small number of users, leading to unfairness. On the contrary, if resources are allocated in a fair manner, resources may be allocated to the users with weak channel conditions. This can result in the degradation of system throughput. To escape

Table 1 IEEE802.16e QoS classes

QoS classes	Applications	QoS parameters
UGS	T1 services, VOIP without silence suppression	Max Rate
		Min Rate
		Jitter
rtPS	Video Streaming	Max Rate
		Min Rate
		Max Latency
ertPS	VOIP with silence suppression	Max Rate
		Min Rate
		Max Latency
		Jitter
nrtPS	FTP	Max Rate
		Min Rate
BE	Web browsing, e-mail	Max Rate

Figure 2 The TDD frame architecture (*Source*: IEEE [1]).
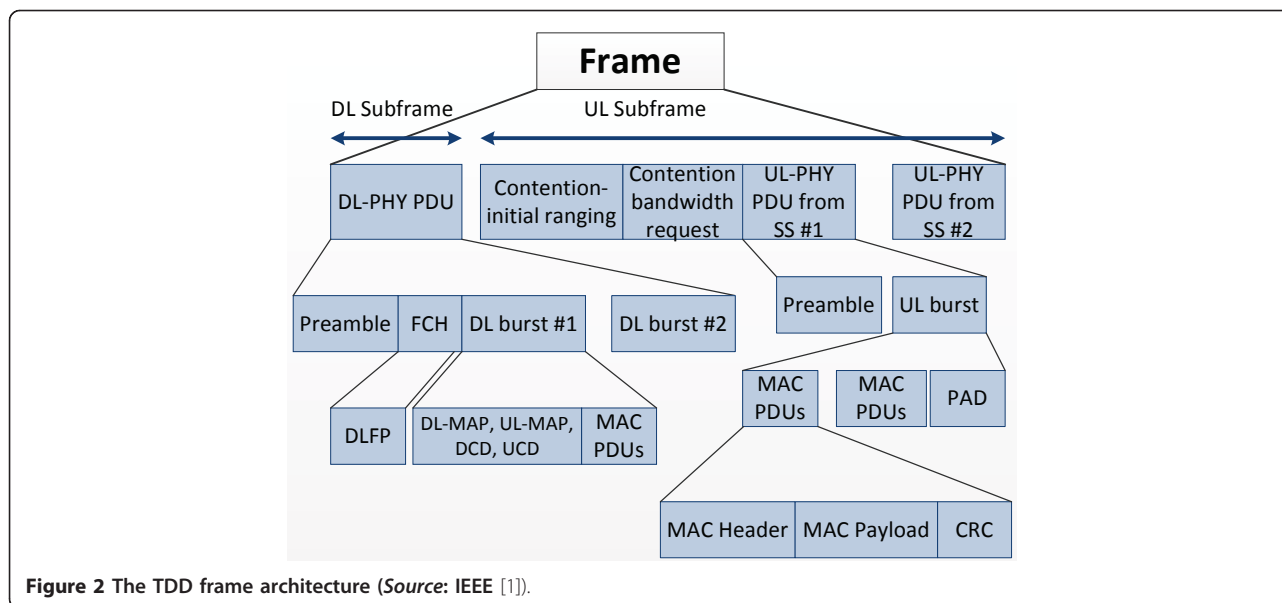
from the "throughput-fairness" dilemma, we can consider "utility" for packet scheduling. Utilities are a performance metric which can fully represent the degree of user satisfaction for a given application [32,33]. Thus, it is a more appropriate metric for packet scheduling since even users with the same service class actually are with various demands for network resources due to their specific characteristics. This fact implies that resources should be allocated to users according to the application performance metric of "satisfaction" rather than network performance metrics such as "throughput" or "fairness" [34].

In this article, we focus on the problems of CAC and BR for mobile WiMAX networks and do not discuss the design of packet scheduling. In the following section, we will propose a joint design of CAC and BR mechanisms as a solution of the addressed problems.

3. Proposed CAC algorithm and BR scheme
3.1. Dynamic CAC algorithm
We design a dynamic CAC algorithm which adjusts the admission criteria depending on the network loads and uses an adaptive QoS provisioning strategy in order to increase the efficiency of bandwidth utilization. The key idea of our dynamic CAC scheme is based on the fact that most scheduling services are with adaptive QoS requirements, e.g., maximum and minimum rates. Therefore, the admission criteria can be adjusted with the amount of available wireless resources to increase the number of admitted connections with adaptive QoS. For example, if the network capacity is adequate or sufficient, higher QoS requirements could be adopted as the admission criteria. Alternatively, if the network load

is rather heavy, the admission criteria may level down for lower QoS.

We can take into account several approaches to design the admission criteria according to network loads. A simple way, as shown in Figure 3a, is to adopt the maximum or minimum bandwidth requirement alternatively as the admission criterion depending on the network load nl with respect to a threshold, γ. That is

$$b_i = \begin{cases} b_{i,\max}, & nl < \gamma \\ b_{i,\min}, & nl \geq \gamma \end{cases}, \qquad (1)$$

where b_i is the admission criteria for connection i; $b_{i,\max}$ and $b_{i,\min}$ are the maximum and minimum bandwidth requirements corresponding to the highest and lowest QoS, respectively, for connection i. Furthermore, we can consider a linear adaptation approach which smoothly regulates the admission criteria b_i in proportion to the change of network load nl as shown in Figure 3b. Accordingly, b_i can be expressed as

$$b_i = \alpha b_{i,\max} + (1 - \alpha) b_{i,\min}, \text{ where } \alpha = \begin{cases} 1, & nl < nl_{th}^{\min} \\ \dfrac{nl_{th}^{\max} - nl}{nl_{th}^{\max} - nl_{th}^{\min}}, & nl_{th}^{\min} \leq nl < nl_{th}^{\max} \\ 0, & nl \geq nl_{th}^{\max} \end{cases} \qquad (2)$$
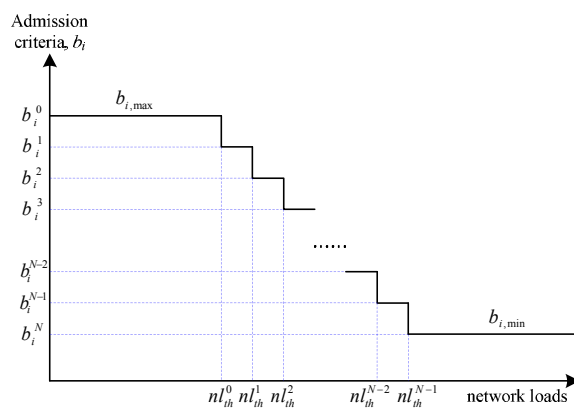
where $nl_{th}^{\min}$ and $nl_{th}^{\max}$ refers to the minimum and maximum threshold of network loads respectively. If nl is less than $nl_{th}^{\min}$, the weighted factor α will be set as 1 and $b_{i,\max}$ would be adopted as the admission criterion to provide the best QoS. On the other hand, if nl is equal to or larger than $nl_{th}^{\max}$, α will be set as 0 and $b_{i,\min}$ would be used to provide the lowest QoS. If nl is
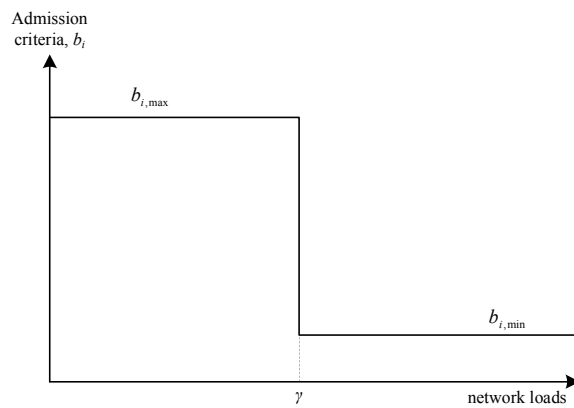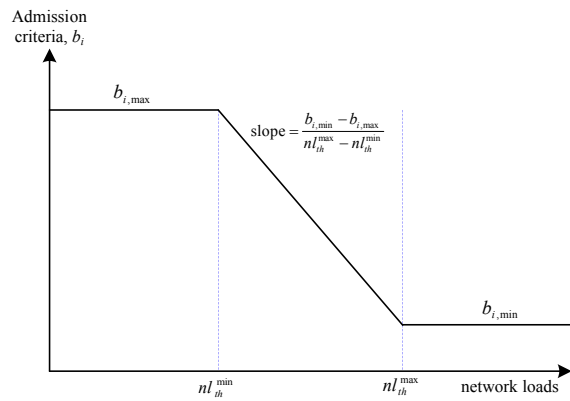
Figure 3 Several functions to adjust the admission criteria according to the network load. (a) Hard-decision function. (b) Linear function. (c) Quantized-step function.

within the range of $[nl_{th}^{\min}, nl_{th}^{\max})$, b_i will be linearly regulated in proportion to the change of nl to provide an adequate QoS. In general, $nl_{th}^{\min}$ and $nl_{th}^{\max}$ can be varied with the network bandwidth B and each type of service. With a given B, $nl_{th}^{\min}$ and $nl_{th}^{\max}$ for a higher service class should be larger in order to provide differential QoS among various service classes. A feasible approach to determine the values of $nl_{th}^{\min}$ and $nl_{th}^{\max}$

with regard to each service class is to set different ratios of $nl_{\text{th}}^{\min}$ and $nl_{\text{th}}^{\max}$ to B appropriately. For example, in the simulations we consider the following setting of ($nl_{\text{th}}^{\min}/B$, $nl_{\text{th}}^{\max}/B$): (3/6, 5/6), (2/5, 4/5), (2/6, 4/6), (1/5, 3/5), and (1/6, 3/6), respectively, for UGS, rtPS, ertPS, nrtPS, and BE in order. Finally, as shown in Figure 3c, this approach is to quantize the bandwidth within [$b_{i,\min}$, $b_{i,\max}$] with $N + 1$ levels, each of which refers to a feasible admission criteria. The adaptive criteria b_i can be expressed as

$$b_i = \begin{cases} b_i^0 (b_{i,\max}), & nl < nl_{\text{th}}^0 (nl_{\text{th}}^{\min}) \\ b_i^n, & nl_{\text{th}}^{n-1} \leq nl < nl_{\text{th}}^n, \ 1 \leq n \leq N-1 \\ b_i^N (b_{i,\min}), & nl \geq nl_{\text{th}}^{N-1} (nl_{\text{th}}^{\max}) \end{cases} \quad (3)$$

where nl_{th}^n refers to the nth threshold of network load ($1 \leq n \leq N - 2$). In particular, nl_{th}^0 and nl_{th}^{N-1} refer to the minimum and maximum thresholds of network loads, $nl_{\text{th}}^{\min}$ and $nl_{\text{th}}^{\max}$, respectively. Here, b_i^k ($1 \leq k \leq N - 1$) refers to the kth feasible admission criteria as the network load nl is within the range of [nl_{th}^{k-1}, nl_{th}^k). In particular, b_i^0 corresponds to $b_{i,\max}$ as $nl < nl_{\text{th}}^0$, and b_i^N corresponds to $b_{i,\min}$ as $nl \geq nl_{\text{th}}^{N-1}$, respectively. The value of $N + 1$ refers to the number of quantization levels. When N is 1 in particular, the quantized step function has two levels and is equal to the hard-decision function as shown in Figure 3a. In contrast, if N is infinite, the quantized step function has countless levels and is equal to the linear adaptation approach as shown in Figure 3b. Normally, the value of N can be determined considering a reasonable number of adaptive criteria b_i between $b_{i,\min}$ and $b_{i,\max}$, (e.g., N = 5-15 in general). When the value of N is determined, the unknown values of nl_{th}^n ($1 \leq n \leq N - 2$) and b_i^k ($1 \leq k \leq N - 1$) can therefore be obtained by using a uniform quantizer which equally partitions the region within (b_i^0, b_i^N) and (nl_{th}^0, nl_{th}^{N-1}). In this article, we consider the linear adaptation approach as shown in Figure 3b for our CAC algorithm and use it for performance evaluations in the following section.

Denote the bandwidth allocated to existing new connections and handoff connections as b_n and b_h, respectively. A handoff connection will be accepted in the network as long as the bandwidth available can satisfy the bandwidth requirement. Thus, the condition to accept a handoff connection is

$$\text{ho\_accepted} = (b_{i,\text{ho}} + b_n + b_h) \leq B, \quad (4)$$

where $b_{i,\text{ho}}$ is the admission criterion of handoff connection i determined in Equation (2). A new connection

will be accepted when the following condition is satisfied:

$$\text{new\_accepted} = ((b_{i,\text{new}} + b_n) \leq th_{\text{ad}}) \cap ((b_{i,\text{new}} + b_n + b_h) \leq th_{\max}), \quad (5)$$

where $b_{i,\text{new}}$ is the admission criterion of new connection i provided in Equation (2). The first term in the rhs of Equation (5) aims for increasing the admitting opportunities for new connections when it only determines whether the resources allocated to new connections exceed the threshold. The last term in the rhs aims for ensuring a minimum BR for handoff (i.e. B - th$_{\max}$). In Section 3.3 we will have a more detailed illustration for the criteria considered in Equation (5).

3.2. Estimation of system capacity

The system capacity B may be dynamic as the IEEE 802.16e standards on PHY support multiple transmission rates by using adaptive modulation and coding (AMC) schemes. The transmitter will determine one from various modulation and coding schemes (MCSs) available according to the channel conditions of packet delivery to provide reliable link qualities, large network coverage, and high data rates as possible. The modulation types supported in the IEEE 802.16e standards include Binary Phase-Shift Keying (BPSK), Quadrature Phase-Shift Keying (QPSK), 16-Quadrature Amplitude Modulation (16-QAM), and 64-QAM. With the Convolutional Turbo Code (CTC) and different code rates, the MCSs provided for WiMAX with 5 and 10 MHz channels are summarized in Table 2[35]. In a DL transmission for example, each MS informs its current perceived channel quality to the BS periodically, and then the BS will choose a specific MCS corresponding to this channel condition. The transmission data rate with a given MCS can be evaluated as

$$R_{\text{MCS}_i} = (n_{\text{Data\_SC}}/T_S) * b_{\text{MCS}_i}, \quad (6)$$

where $n_{\text{Data\_SC}}$ is the number data sub-carriers; T_S is symbol period, and b_{MCS_i} is the amount of information bits per symbol with respect to the ith MCS, MCS$_i$. In addition, adopting a Multiple-Input and Multiple-Output (MIMO) mechanism can further increase the transmission data rates to several-fold the original amount.

Here, our CAC scheme estimates the current system capacity B by consideration of the proportion of used MCSs [36,37]. Take an example for the capacity estimation as follows. Consider a 10-MHz channel spectrum with a 2×2 MIMO mechanism in the downlink transmission. Assume the proportion of used MCSs is QPSK 3/4 = 25%, 16-QAM 1/2 = 25%, and 64-QAM 5/6 = 50%. Thus, the estimated system bandwidth B in this case will be (9.5*2)*25% + (12.67*2)*25% + (31.68*2)*50% = 42.77 Mbps. In general, B in the downlink can

Table 2 The mobile WiMAX PHY data rates (*Source*: WiMAX Forum [35])

Parameter		Downlink	Uplink	Downlink	Uplink
System bandwidth		5 MHz		10 MHz	
FFT size		512		1024	
Null sub-carriers		92	104	184	184
Pilot sub-carriers		60	136	120	280
Data sub-carriers		360	272	720	560
Symbol period		102.9 μs			
Frame duration		5 ms			
OFDM symbols/frame		48			
Data OFDM symbols		44			
Mod	**Coding rate**	**5 MHz channel**		**10 MHz channel**	
		Downlink rate (Mbps)	Uplink rate (Mbps)	Downlink rate (Mbps)	Uplink rate (Mbps)
QPSK	1/2 CTC, 6x	0.53	0.38	1.06	0.78
	1/2 CTC, 4x	0.79	0.57	1.58	1.18
	1/2 CTC, 2x	1.58	1.14	3.17	2.35
	1/2 CTC, 1x	3.17	2.28	6.34	4.70
	3/4 CTC	4.75	3.43	9.50	7.06
16 QAM	1/2 CTC	6.34	4.57	12.67	9.41
	3/4 CTC	9.50	6.85	19.01	14.11
64 QAM	1/2 CTC	9.50	6.85	19.01	14.11
	2/3 CTC	12.67	9.14	25.34	18.82
	3/4 CTC	14.26	10.28	28.51	21.17
	5/6 CTC	15.84	11.42	31.68	23.52

range from 2.12 to 63.36 Mbps widely with AMC in different channel situations. Based on the number of supported users with respect to each MCS and the proportion of adopted MCSs, our CAC scheme can estimate the current system capacity and network load and update the information periodically every frame period in the BS site.

3.3. Adaptive BR scheme

We propose an adaptive BR scheme which dynamically adjusts the amount of reserved bandwidth for handoffs according to the arrival distributions of both handoff and new connections. The objective of our scheme is to simultaneously increase the admission opportunities for new coming users and guarantee QoS for handoff users. The basic idea of our adaptive BR scheme is a rational inference that generally the occurrences of new incoming connections may be much more frequent than that of handoff connections [22-24]. Thus, the optimal BR should take into account the arrival behavior of not only handoff connections, but also new connections in order to avoid a waste of network resource as possible.

The scheme considers two thresholds for BRs, th_{min} and th_{max}, which refer to the minimum and maximum thresholds of reserved bandwidth, respectively. The

adaptive threshold of BR, th_{ad}, is initially set as (th_{min} + th_{max})/2, and will be dynamically adjusted within the range [th_{min}, th_{max}] according to the arrival behavior of new connections and handoff connections to control the reserved bandwidth for handoff connections. When the condition shown in Equation (4) is met and a handoff connection is accepted, the adaptive threshold th_{ad} will be decreased with the amount of allocated resources for the handoff connection without going below th_{min}. That is

$$th_{ad} = \min (th_{ad} - b_{i,ho}, th_{min}) \qquad (7)$$

Alternatively, when the condition shown in Equation (5) is met and a new connection is accepted, th_{ad} will be increased with the amount of allocated resources for the new connection without exceeding th_{max}. That is

$$th_{ad} = \max (th_{ad} + b_{i,new}, th_{max}) \qquad (8)$$

To more clearly illustrate the characteristics of our proposed BR scheme, we conduct a simplified transmission scenario to provide a preliminary performance comparison between our scheme and the fixed threshold (FT) and dynamic threshold (DT) schemes [10]. With the FT scheme, the threshold of reserved bandwidth, th_{fixed}, is fixed; the new connection would be accepted

only if $(b_{i, \text{new}} + b_n + b_h) \leq \text{th}_{\text{fixed}}$. With the DT scheme, the threshold of reserved bandwidth, th_{dyn}, will be adjusted within the range $[\text{th}_{\text{min}}, \text{th}_{\text{max}}]$ depending on the arrival and departure of handoff connections as follows: when a handoff connection is accepted, the threshold th_{dyn} will be increased with the amount of resources allocated to the handoff user; when an existing handoff connection terminates, th_{dyn} will be decreased with the amount of released resources. The new connection will be accepted only when $(b_{i, \text{new}} + b_n + b_h) \leq \text{th}_{\text{dyn}}$. With regard to handoff connections, all the three algorithms have the same admission strategy that a handoff connection will be accepted as long as the amount of bandwidth available can meet its requirement, i.e. $(b_{i, \text{ho}} + b_n + b_h) \leq B$.

We exploit the following simplified scenario to compare the performances of the three BR schemes. Assume that the total amount of network resources B is 100 units. Consider that the network is empty in the beginning. Assume that 80 new connections, 5 handoff connections, and 5 new connections arrive sequentially. Assume that each of the arrival connections requests a unit of resources. For the FT scheme, th_{fixed} is set as 80 units. In the DT scheme, th_{min} and th_{max} are 80 and 90 units, respectively, and th_{dyn} is set as 80 units initially. In our scheme, th_{min} and th_{max} are 0 and 90 units, respectively, and consequently the initial value of th_{ad} is 45 units.

Figure 4 shows the process of connection admissions with the three schemes, respectively. It is shown that with the FT scheme, the first 85 connections (80 new and 5 handoff connections) are accepted. But, the last five new connections are rejected due to a fixed reservation manner that a certain amount of bandwidth can never be used by new connections. The DT scheme has the same performance as the FT scheme that the last five new connections are blocked. The reason is that the DT scheme adjusts the reserved bandwidth according to the arrival and departure behavior of handoff connections only. Thus, when the occurrence of handoff connections is relatively infrequent as the scenario shown above, the DT scheme will be similar to the FT scheme and can cause a waste of network resources as well. With our scheme, the 90 connections totally can be granted into the network. Note that our scheme can still be effective in the situations when the number of new connections is nearly equal to or less than that of handoff connections (this will be further examined in the following section). It is shown that the proposed scheme outperforms the two schemes in terms of increasing the number of admitted connections.

Although the DT scheme and our proposed scheme both apply an adaptive manner for their BR strategies, there are two essential differences which lead to their dissimilar performances. One difference is between their criteria for accepting new connections. For the DT scheme, the new connection will be accepted if the amount of its bandwidth requirement and the bandwidth allocated to existing connections would not exceed the threshold, i.e. $(b_{i, \text{new}} + b_n + b_h) \leq \text{th}_{\text{dyn}}$. In our scheme, one criterion for accepting a new connection as shown in the first term of the rhs in Equation (5) only examines whether the amount of the requirement and the bandwidth allocated to existing "new connections" does not exceed the threshold, i.e. $(b_{i, \text{new}} + b_n) \leq \text{th}_{\text{ad}}$. Thus, with our scheme, the existing handoff connections will not lessen the resource for an incoming new connection. That is, when the DT scheme and our proposed scheme are with the same amount of reserved bandwidth ($\text{th}_{\text{dyn}} = \text{th}_{\text{ad}}$), our scheme can increase a resource space of b_h for accepting more new connections. Meanwhile, our BR scheme can guarantee QoS for handoff connections because of another condition for accepting a new connection as shown in the last term of the rhs in Equation (5). That is, the total amount of required bandwidth for a new connection and allocated bandwidth for existing connections would not exceed the amount of guarded bandwidth for handoff, th_{max}, i.e. $(b_{i, \text{new}} + b_n + b_h) \leq \text{th}_{\text{max}}$. Thus, at last the B - th_{max} bandwidth amount can be reserved for handoff users.

Another difference is between their adjustment strategies for the threshold of BR, th_{dyn} and th_{ad}. With the DT scheme, th_{dyn} is increased or decreased when a handoff connection is established or terminated respectively, and will remain constant whether a new connection concludes or runs its course. When the occurrence of handoff connections is relatively infrequent as the scenario shown above, the DT scheme will be similar to the FT scheme and can cause a waste of network resources too. In our scheme, th_{ad} is increased or decreased when a new or handoff connection is admitted and runs its course, respectively, and will remain constant whether a new or handoff connection concludes. Note that a lower threshold of BR advances the acceptance rate of handoff users whereas a higher threshold increases the admission opportunities for new connections. In the sense, our scheme has potential to grant more new connections into networks especially when the occurrences of new connections are much more than that of handoff connections. As aforementioned, our scheme would not sacrifice handoff QoS for favoring new connections when no less than the bandwidth amount of B - th_{max} will be reserved for handoff users. Thus, the proposed adaptive BR scheme can simultaneously improve the network efficiency by granting more new connections and also guarantee handoff QoS.

To summarize, Figure 5 presents the joint design of the proposed CAC algorithm and BR scheme, and

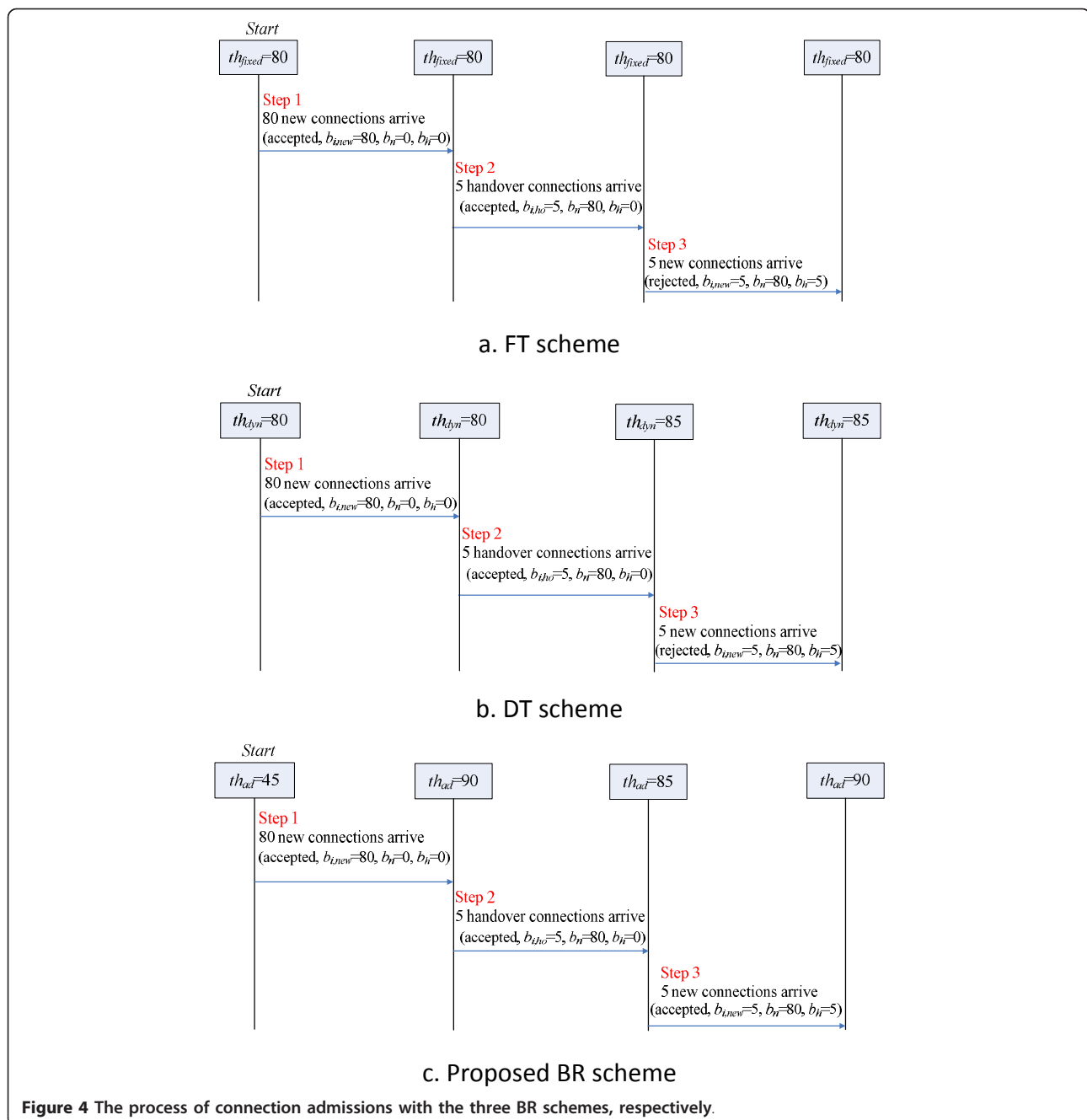**Figure 4 The process of connection admissions with the three BR schemes, respectively**.

Figure 6 shows the corresponding pseudocodes. The block diagram in Figure 5 contains the following three steps.

(1) When a new connection or handoff connection arrives, it will inform the BS of its highest and lowest bandwidth requirements (i.e., $b_{i, \max}$ and $b_{i, \min}$). The proposed dynamic CAC scheme will adjust the admission criterion using Equation (2) according to the currently estimated system capacity and network load.

(2) When the admission criterion is determined, the proposed adaptive BR scheme will accept or reject this handoff or new connection depending on the criterion in Equation (4) or Equation (5), respectively.

(3) If a handoff connection is established, th_{ad} will be decreased with the amount of allocated resources as Equation (7) shows; if a new connection is granted, th_{ad} will be increased with the amount of allocated resources as Equation (8) shows.
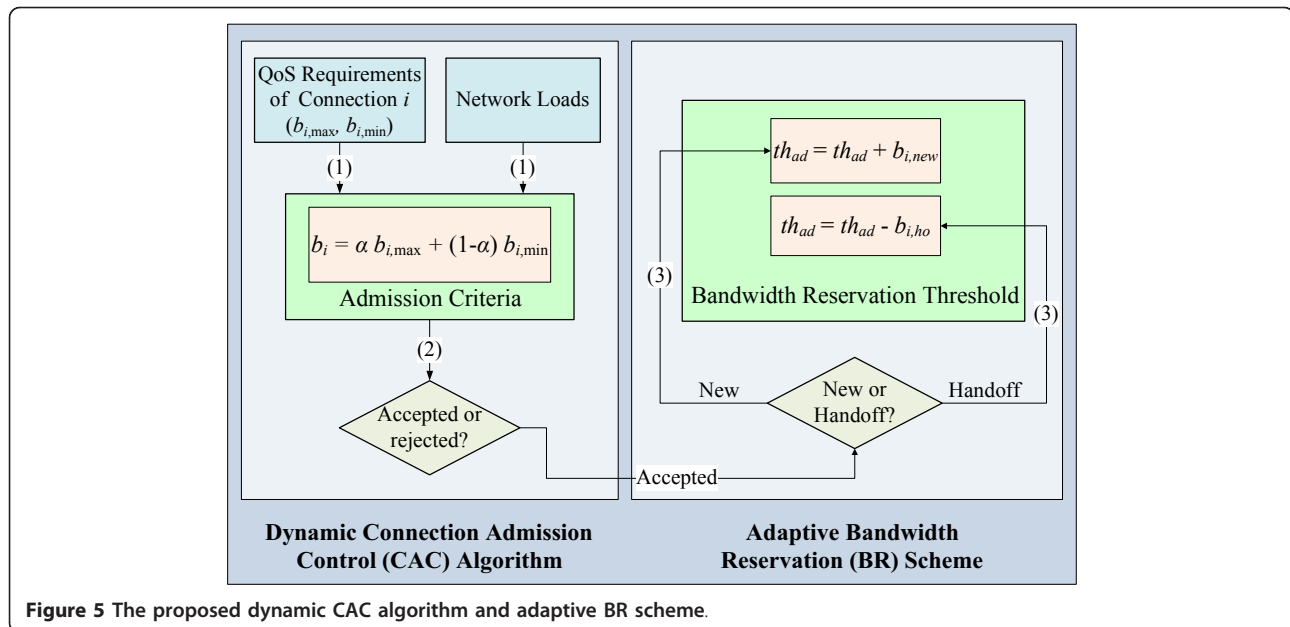
Figure 5 The proposed dynamic CAC algorithm and adaptive BR scheme.

Figure 6 Proposed CAC and BR algorithm.

The implementation of the proposed CAC and BR schemes in practice can involve the overheads as follows.

(1) *The estimate of system capacity*: The system capacity can be evaluated at the BS with the specific PHY characteristics like channel spectrum, the amount of data sub-carriers, supported MCSs, used MIMO mechanisms, etc. The estimation of system capacity can be obtained in the initial phase of a network built-up.

(2) *The estimate of network loads*: In general, the network loads can be evaluated at the BS with the information of currently adopted MCSs and the number of supported users with respect to each MCS. This part may need the exchanges of some context information between BS and SSs periodically, e.g., currently channel condition and used modulation.

(3) *The determination of admission criteria for incoming connections*: When a connection arrives and requests for an admission, it will inform the BS of its specific QoS requirements, e.g., maximum and minimum data rates. Based on the estimated system capacity and network loads and QoS parameters, the BS will compute the admission criteria for the incoming connection with its specific QoS parameters.

(4) *The adaptation of BR for handoff connections*: If a connection is admitted in the network, the BS will therefore adapt the BR threshold depending on the type of connection, i.e., new or handoff.

4. Performance evaluations and results

In this section, we conduct simulations of 802.16e transmission scenarios to demonstrate the effectiveness of the proposed CAC algorithm and BR scheme. The simulator is constructed in C and followed the IEEE 802.16e standard closely [35,38]. The channel spectrum is 10 MHz. The MAC frame duration is 5 ms, which consists of 1024 OFDM subcarriers (840 data and pilot subcarriers). One MAC frame includes 48 OFDM symbols, while the first symbol is used for a preamble. The ratio of the symbols of the uplink subframe to those of the downlink subframe is 18:29. In the uplink, three symbols are used for control signaling, and there are 44 OFDM symbols used for data transmissions in the uplink and downlink in total. The simulation set-up considers a 2 × 2 MIMO mechanism and the AMC schemes. We used the following distribution for MCS levels: QPSK 1/12 = 3.71%, QPSK 1/8 = 12.01%, QPSK 1/4 = 29.10%, QPSK 1/2 = 29.67%, QPSK 3/4 = 9.23%, 16-QAM 1/2 = 12.51%, 64-QAM 1/2 = 0.75%, and 64-QAM 3/4 = 3.02% [39,40]. The OFDMA PHY parameters and their values are listed in Table 2.

We provide different simulation scenarios to examine the proposed CAC algorithm and BR scheme individually or jointly to clearly show the performance effects with the two schemes. For each of the scenarios, we assume the connection arrivals and departures are with the Poisson process with a mean arrival rate λ and a mean departure rate equal to one-tenth of the arrival rate. We assume that the BS is aware of the amount of connections with regard to each kind of MCS and the bandwidth requirement of each connection. The BS can therefore exploit this knowledge to estimate the current system capacity and also network loads. Generally, the system capacity and network loads can therefore be estimated for our scheme with the information of currently adopted MCSs and the number of supported users with respect to each MCS. The total simulation period is 1000 s while the results are provided with the average values over 20 times of simulations.

4.1. The connection blocking rates and achieved throughput with different CAC schemes

In this section, we examine the performance of our CAC algorithm individually by the comparisons with that of other CAC schemes. Here, we do not take into account the reserved bandwidth for handoff users. Thus, each of the incoming connections will be admitted into the network as long as the bandwidth available can satisfy the admission criteria determined by various CAC schemes. The simulation set-up for this scenario considers three types of traffic classes which are rtPS, ertPS, and nrtPS. To clearly examine the achieved QoS performances regarding the three types of service classes with different schemes, we separately evaluate the performance for each class with its specific QoS requirement, i.e., maximum and minimum data rates as shown in Table 3.

We compare the performance of our CAC algorithm with that of the static maximum (Static-max), static minimum (Static-min), and bandwidth adaptation (Adapt) scheme considered in [11,12]. The Static-max and Static-min schemes adopt the highest and lowest QoS requirements, respectively, as the admission criteria. The Adapt scheme considers the highest QoS criteria for a new coming user. If the resources available are insufficient to meet the bandwidth requirement of a new connection, the bandwidth allocated to existing

Table 3 The maximum and minimum rates associated with different scheduling services
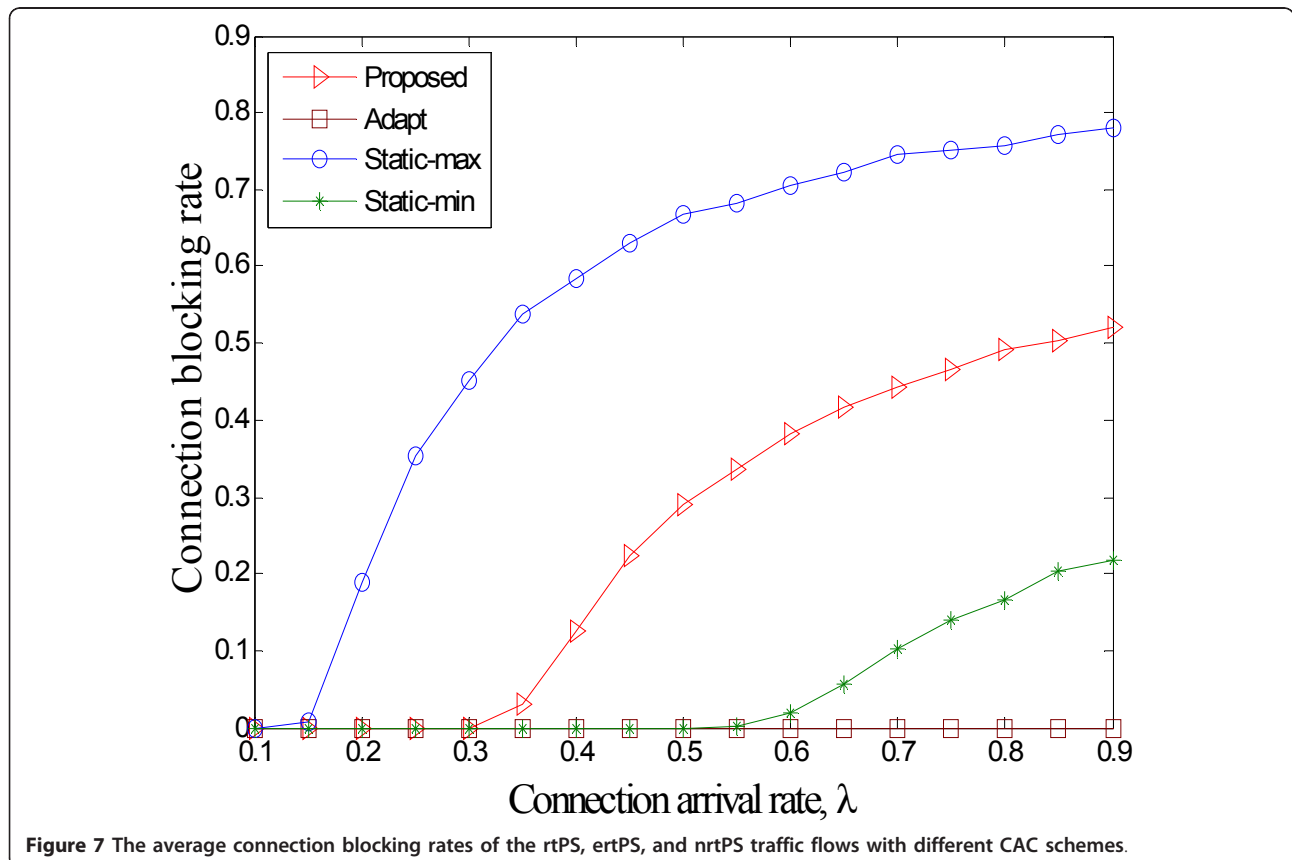
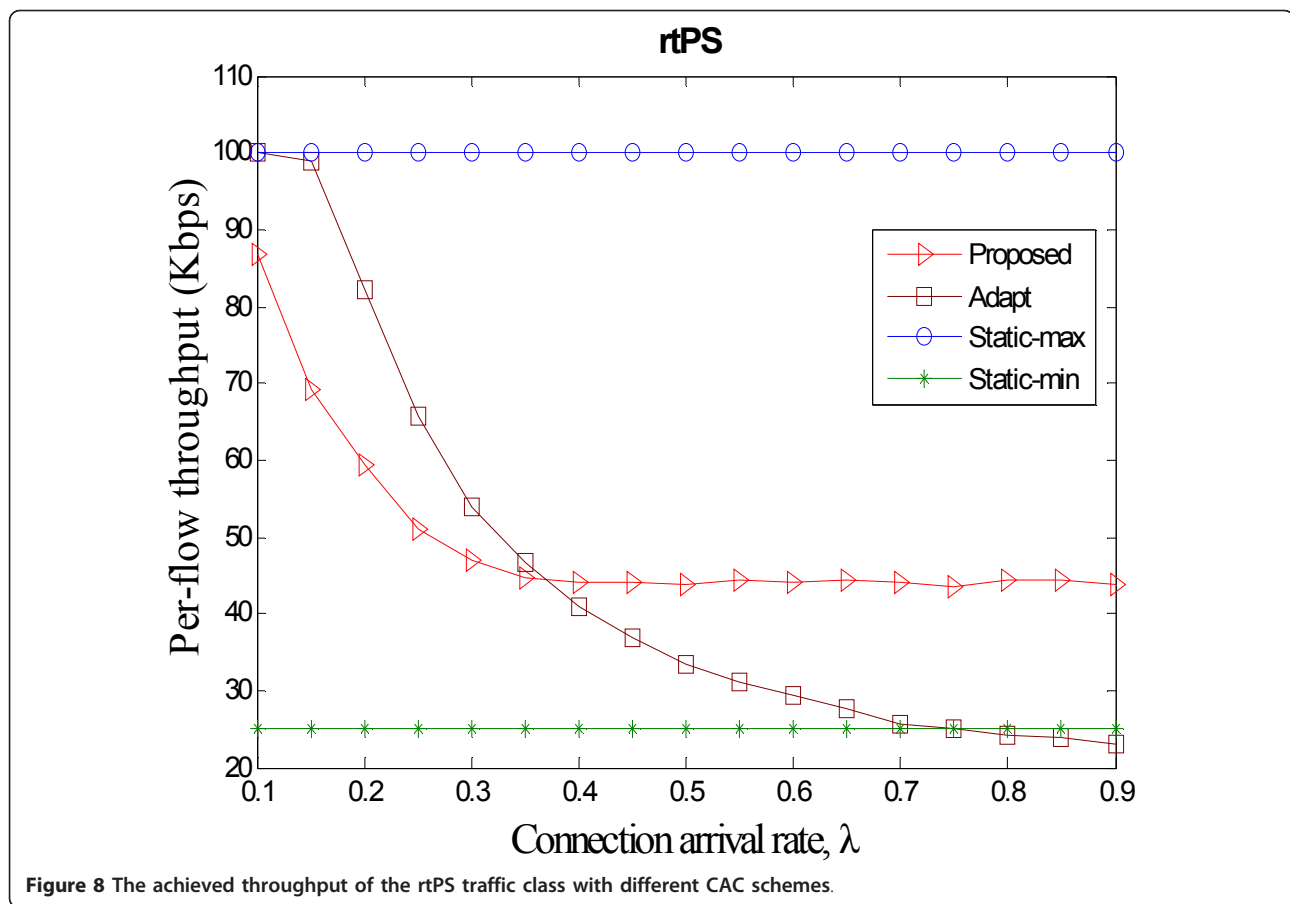| Scheduling services	Maximum rate (kbps)	Minimum rate (kbps)
UGS	100	100
rtPS	100	25
ertPS	75	25
nrtPS	60	20
BE	20	0

users will be decreased to satisfy the QoS requirement of the new connection. In our scheme, we consider a linear adaptation approach as shown in Figure 3b which adjusts the admission criteria according to the variation of network loads. We use the following ratios of (nl_{th}^{min}/B, nl_{th}^{max}/B): (2/5, 4/5), (2/6, 4/6), and (1/5, 3/5), for rtPS, ertPS, and nrtPS, respectively, as we consider in Section 3.1. Then, the minimum and maximum thresholds of network loads (nl_{th}^{min}, nl_{th}^{max}) in Equation (2) for rtPS, ertPS, and nrtPS can be derived, respectively, depending on the currently estimated system capacity B. The performances are indexed as the connection blocking rate and the achieved QoS in terms of per-flow throughput, i.e., the average data rates supported per established connection.

Figure 7 presents the average connection blocking rates of the rtPS, ertPS, and nrtPS flows with different CAC schemes. It is shown that the blocking rates rise with the increase of connection arrival rates λ for all schemes except the Adapt scheme. The blocking rate with the Adapt scheme is almost 0 steadily, which is the lowest among all schemes when the incoming connections will always be established regardless of the amount of bandwidth available. The blocking rate with the

Static-max scheme is the highest among all due to its strictest admission criterion. With the proposed scheme, the blocking rate is between that with the Static-max and Static-min schemes since the admission criteria are dynamically adjusted according to network loads (i.e., connection arrival rates).

Figures 8, 9, and 10 show the achieved per-flow throughput of the rtPS, ertPS, and nrtPS service classes, respectively, with different CAC schemes. It is shown in these figures that the Static-max and Static-min schemes provide steady throughput corresponding to the highest-fidelity and minimum QoS, respectively. It is shown from Figures 7, 8, 9, and 10 that the Static-max and Static-min schemes cannot properly balance the tradeoff between network efficiency and user-perceived QoS due to using fixed admission criteria without the consideration of traffic loads. Consider the case of ertPS traffic as shown in Figure 9 for example. With the Adapt scheme, the per-flow throughput keeps at the highest level when the network load is relatively light (λ is less than 0.2); but, it degrades rapidly as the traffic loads get heavier (λ is larger than 0.2) because the resources allocated to existing users are taken away to satisfy new connections' QoS requirements. As the traffic load get much heavier such that λ is larger than about 0.5, the achieved

Figure 7 The average connection blocking rates of the rtPS, ertPS, and nrtPS traffic flows with different CAC schemes.

Figure 8 The achieved throughput of the rtPS traffic class with different CAC schemes.

throughput can no more keep above the value for the lowest QoS (25 kbps). With the proposed scheme, the achieved throughput is varying from 38 to 66 kbps around with the change of network loads, and remains at about 38 kbps when λ is larger than 0.3. It is shown that the per-flow throughput can keep higher than the minimum QoS level. In Figures 8 and 10, we have similar observations regarding the rtPS and nrtPS traffic classes. The results demonstrate that the proposed dynamic CAC algorithm can efficiently improve utilization of network resources and also assure adaptive QoS above the minimum requirements for admitted connections.

4.2. The new connection blocking rate and handoff dropping rate with different BR schemes

In this section, we conduct simulation scenarios which consider new connections and handoff connections to examine the performance of the proposed BR scheme without CAC. The compared BR schemes for performance evaluations are the FT and DT schemes [10] described in Section 3.2. The simulation set-up in this scenario considers five classes of scheduling services, UGS, rtPS, ertPS, nrtPS, and BE, each of which is

associated with specific maximum rates and minimum rates as shown in Table 3. To simply and fairly compare the performances of different BR schemes, our approach here considers the same fixed admission criteria as that used by the FT and DT schemes: the maximum rate is adopted as the admission criteria for UGS, rtPS, and ertPS, and the minimum rate is adopted for nrtPS and BE [10]. Assume that the occurrences of the five scheduling services are with uniform probabilities, with which the arrival rates of new connections and handoff connections are λ_n and λ_h, respectively. We refer to [10] and adopt the following threshold setting. For the FT scheme, th_{fixed}/B is equal to 80%. In the DT scheme, th_{min}/B and th_{max}/B are set as 80 and 90%, respectively; th_{dyn}/B is set as 80% initially. In our scheme, th_{min}/B and th_{max}/B are equal to 0 and 90%, respectively, and consequently the initial value of th_{ad}/B is 45%. We examine the performance of different BR schemes with different ratios of handoff arrival rates to new-connection arrival rates which can be 1:1 or 1:20 in particular. The performance metrics are indexed as the new connection blocking rates and handoff dropping rates.

Figure 11a, b shows the new-connection blocking rate and handoff dropping rate, respectively, with different

Figure 9 The achieved throughput of the ertPS traffic class with different CAC schemes.
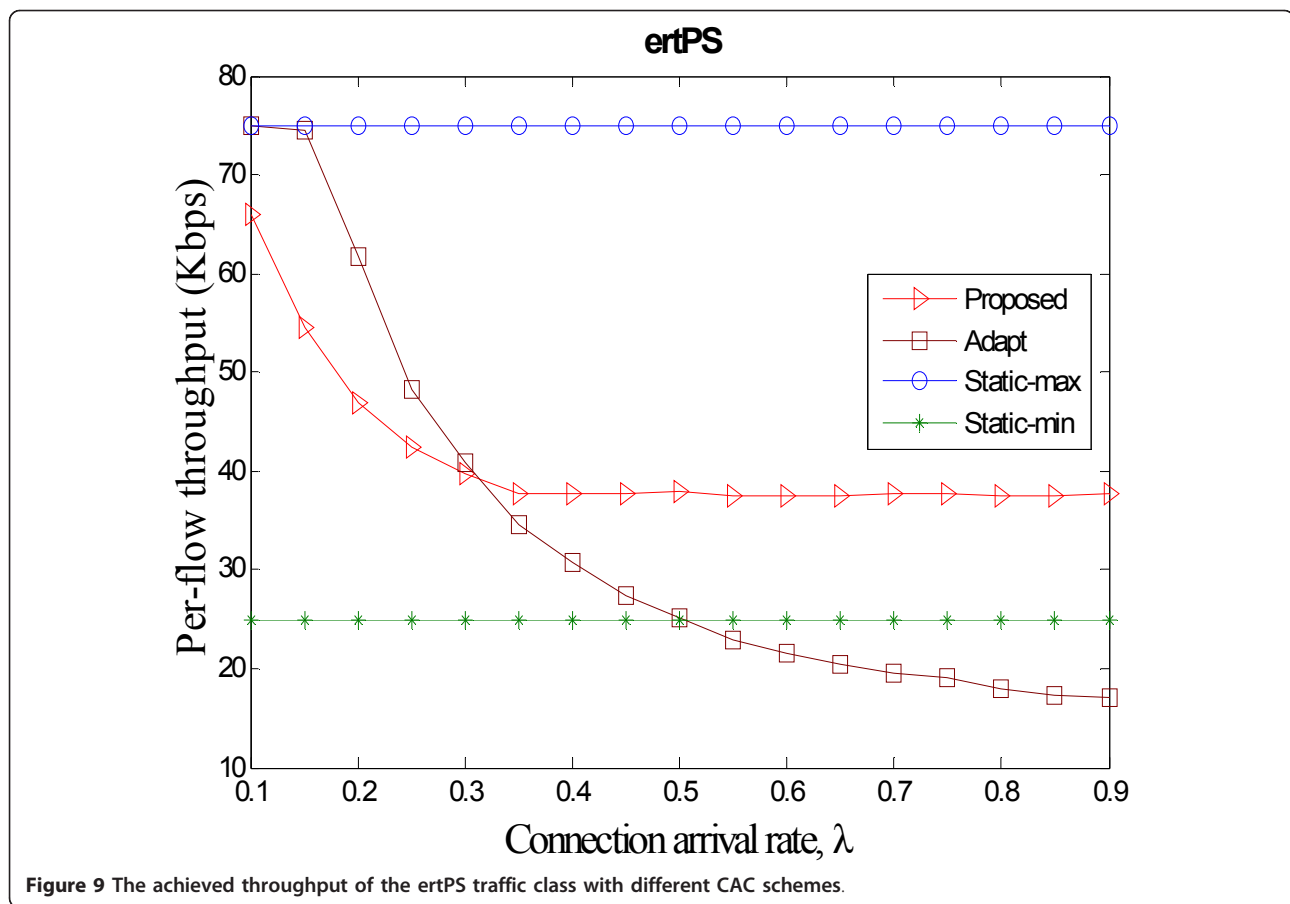
BR schemes as the ratio of handoff arrival rates λ_h to new-connection arrival rates λ_n is 1:1 (the total connection arrival rate $\lambda = \lambda_h + \lambda_n$). It is shown that the FT scheme provides the lowest handoff dropping probability and highest blocking rate when it adopts an FT for the BR. The DT scheme and our proposed scheme have similar performances in terms of both blocking and dropping rates as they determine the BR in an adaptive sense. Generally, the total amount of admitted connections with the three schemes is nearly the same. By comparison, our proposed scheme can admit a few more connections into networks than the FT and DT schemes as its increasing number of new connections is slightly larger than the decreasing amount of handoff connections. For example, when the arrival rate λ is equal to 0.8 ($\lambda_n = \lambda_h = 0.4$) in particular, the total number of admitted new and handoff connections with FT, DT, and the proposed scheme is 219, 218, and 225, respectively. In general with various values of arrival rates λ, our proposed scheme averagely can slightly increase the network efficiency by 0.75 and 1.47% in comparison with the FT and DT schemes, respectively.
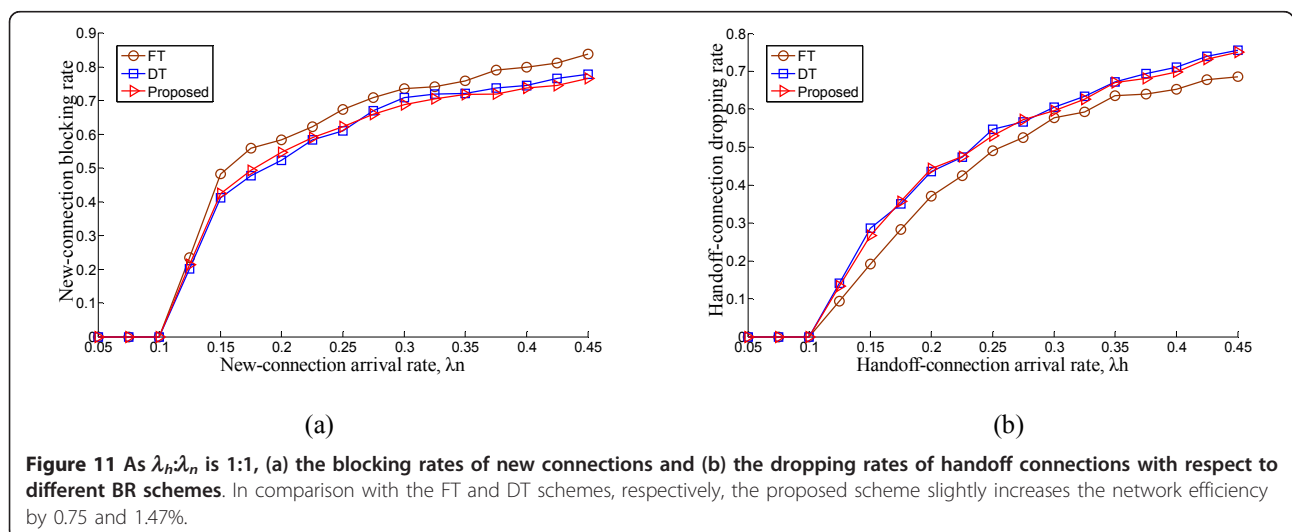
Figure 12a, b shows the new-connection blocking rate and handoff dropping rate, respectively, when the ratio
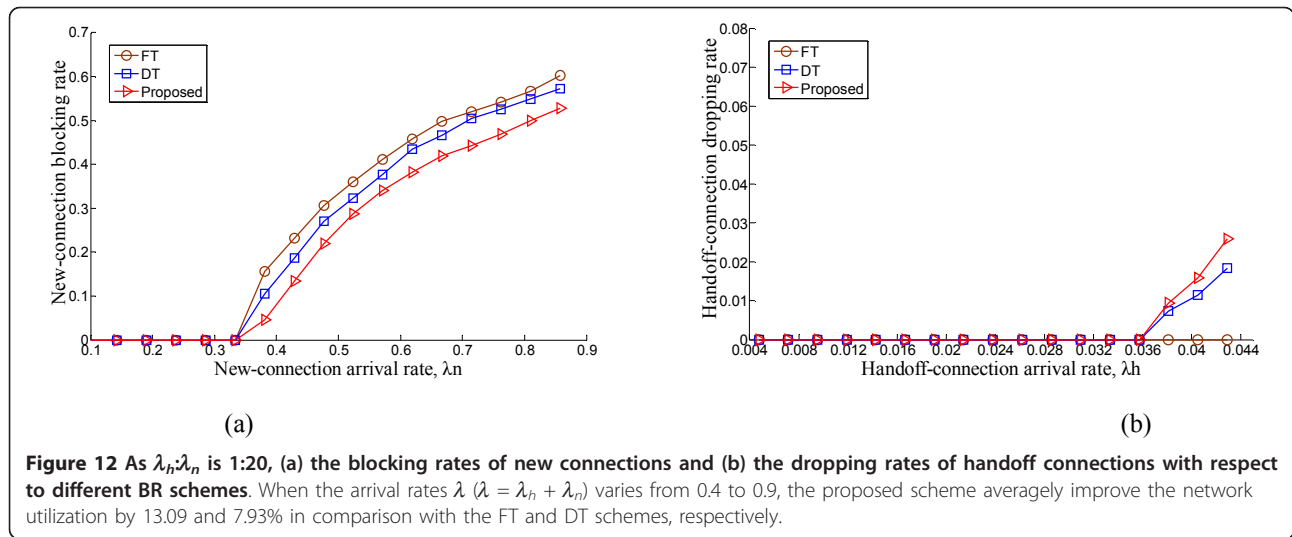
of handoff arrival rates λ_h to new-connection arrival rates λ_n is 1:20. The performance of the FT scheme is similar to that shown in Figure 11a, b that it has the lowest dropping and highest blocking rate among all. With regard to the DT scheme and our proposed scheme, it is shown that their performances are quite dissimilar here, unlike what is shown in Figure 11a, b. We can observe that our proposed scheme provides a much lower blocking probability and a slightly higher handoff dropping rate than the FT and DT schemes (notice that the scales of the vertical axes in Figure 12a, b are different). Regarding this transmission scenario in which the occurrences of new connections are much more frequent than those of handoff connections, the proposed scheme can significantly outperform the two schemes in terms of the amount of admitted connections. For example, when the arrival rate λ is equal to 0.8 ($\lambda_n = \lambda_h = 0.4$) in particular, the blocking rate with FT, DT, and the proposed scheme is about 0.54, 0.53, and 0.47, respectively, while the dropping rate with the three schemes is about 0, 0.007, and 0.009, respectively. Consequently, the total number of admitted new and handoff connections with FT, DT, and the proposed scheme is 387, 399, and 442, respectively. The DT

Figure 10 The achieved throughput of the nrtPS traffic class with different CAC schemes.
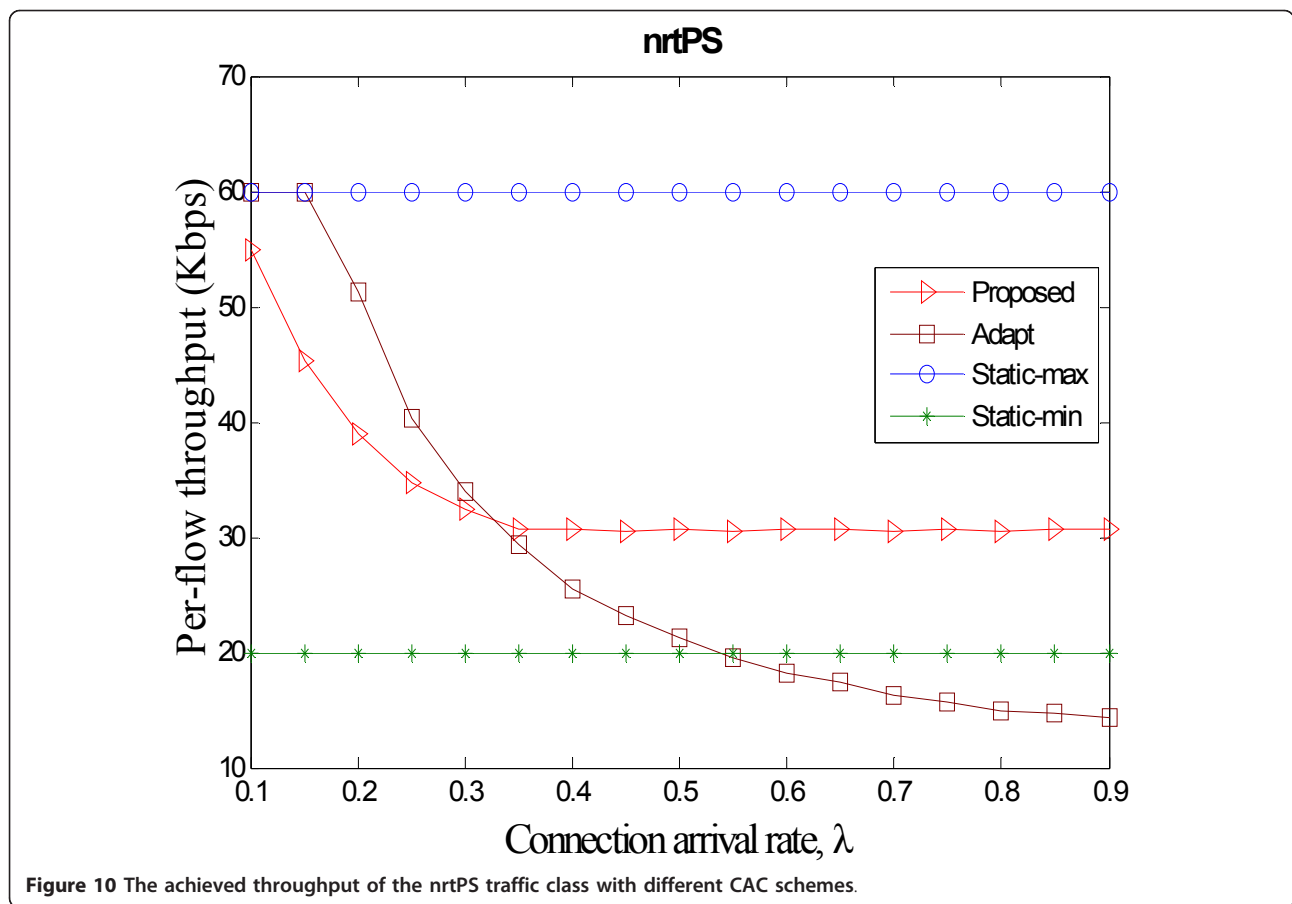
scheme grants 12 connections more than the FT scheme, providing a slight increase of network efficiency by 3.20% over the FT scheme. It is shown that the DT scheme provides a rather marginal improvement over the FT scheme when the occurrences of handoff connections are relatively sparse. In this situation, the DT scheme may cause a waste of network resources as much as the FT scheme does. With the proposed scheme, the number of granted handoff connections is almost equal to that of the FT and DT schemes since a bandwidth guard is used in our scheme to ensure a minimum BR for handoff users. On the other hand, the

(a)

(b)

Figure 11 As $\lambda_h{:}\lambda_n$ is 1:1, (a) the blocking rates of new connections and (b) the dropping rates of handoff connections with respect to different BR schemes. In comparison with the FT and DT schemes, respectively, the proposed scheme slightly increases the network efficiency by 0.75 and 1.47%.

Figure 12 As λ_h:λ_n is 1:20, (a) the blocking rates of new connections and (b) the dropping rates of handoff connections with respect to different BR schemes. When the arrival rates λ ($\lambda = \lambda_h + \lambda_n$) varies from 0.4 to 0.9, the proposed scheme averagely improve the network utilization by 13.09 and 7.93% in comparison with the FT and DT schemes, respectively.

number of admitted new connections with the proposed scheme is much more than that with the FT and DT schemes. Totally, our scheme can grant 55 and 43 connections more than the FT and DT schemes, providing a significant improvement of network efficiency by as large as 14.15 and 10.61% over the FT and DT schemes, respectively. In general when the arrival rates λ ($\lambda = \lambda_h + \lambda_n$) varies from 0.4 to 0.9, our proposed scheme averagely can increase the network efficiency by 13.09 and 7.93% in comparison with the FT and DT schemes, respectively. The simulations demonstrate that our adaptive BR scheme can simultaneously grant more new connections to improve utilization efficiency of network resources and also guarantee BRs for handoff users.
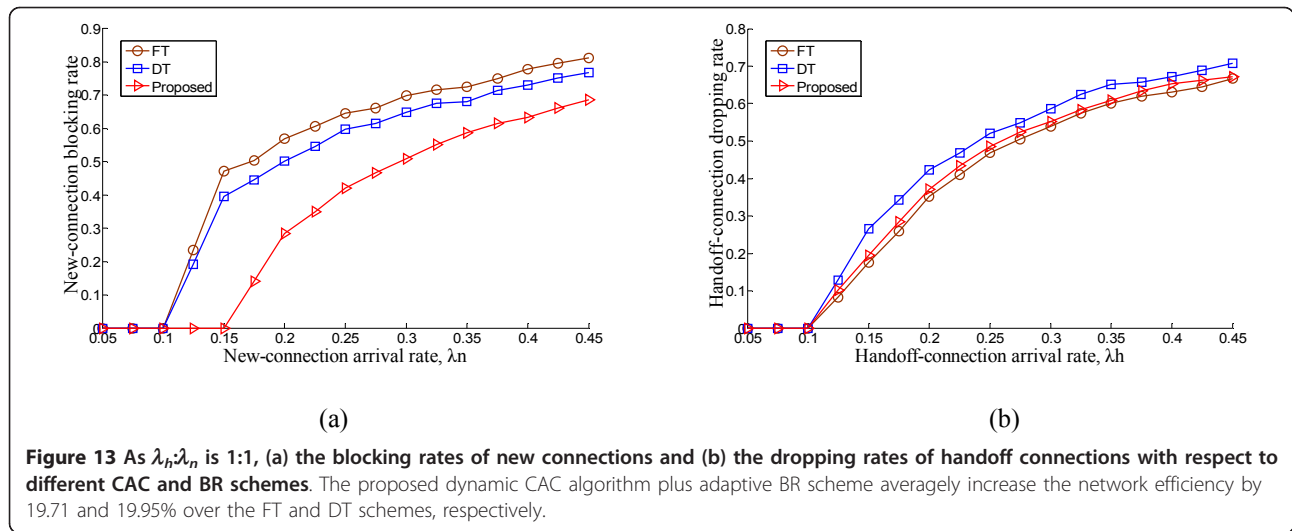
4.3. The new connection blocking rate and hadoff dropping rate with different CAC and BR schemes
In this section, we conduct simulation scenarios to examine the joint performance of our dynamic CAC algorithm and adaptive BR scheme. The simulation set-up for this scenario is similar to that described in Section 4.2, except that our scheme further adopts the proposed CAC algorithm to dynamically adjust the admission criteria for the five scheduling classes. We use the following ratios of (nl_{th}^{min}/B, nl_{th}^{max}/B): (3/6, 5/6), (2/5, 4/5), (2/6, 4/6), (1/5, 3/5) and (1/6, 3/6), for UGS, rtPS, ertPS, nrtPS, and BE, respectively, as we consider in Section 3.1. Then, the minimum and maximum thresholds of network loads (nl_{th}^{min}, nl_{th}^{max}) in Equation (2) for UGS, rtPS, ertPS, nrtPS, and BE can be derived, respectively, depending on the currently estimated system capacity B. The FT and DT schemes use the same admission strategy aforementioned that the scheduling classes of UGS, rtPS, and ertPS adopt the maximum

rate for the admission criteria while the nrtPS and BE classes use the minimum rate for that.

Figure 13a, b presents the new-connection blocking and handoff dropping rate, respectively, when the ratio of handoff arrival rates λ_h to new-connection arrival rates λ_n is 1:1. It is shown that the performance of blocking rates using our approach is much better than that with the FT and DT scheme, while the handoff dropping rate with our scheme is not far from that with the other two schemes. In particular, we compare the performances of our approach shown in Figure 11a with that shown in Figure 13a here (both in which the ratio of handoff arrivals to new-connection arrivals is 1:1), and it is observed that the later provides a great improvement over the former (as the blocking rate is decreased by 46.10% in average). The performance improvement is contributed by the joint use of the proposed CAC algorithm which dynamically determines the admission criteria depending on network loads to increase the connection admission opportunities. As a result, our approach (dynamic CAC algorithm and adaptive BR scheme) averagely can increase the network efficiency by as large as 19.71 and 19.95% over the FT and DT schemes, respectively, with various values of arrival rates.

Finally, Figure 14a, b presents the blocking and dropping rate, respectively, when the ratio of handoff arrival rates λ_h to new-connection arrival rates λ_n is 1:20. Similarly, we can clearly observe the performance improvement with the joint design of our CAC algorithm by comparing the results shown in Figures 12a and 14a particularly (both in which the ratio of handoff arrivals to new-connection arrivals is 1:20). It is shown in Figure 14a, b that our approach can significantly improve the network utilization by totally admitting 22.45 and 17.32% connections more than the FT and DT schemes, respectively. From the results shown in Figures 13 and
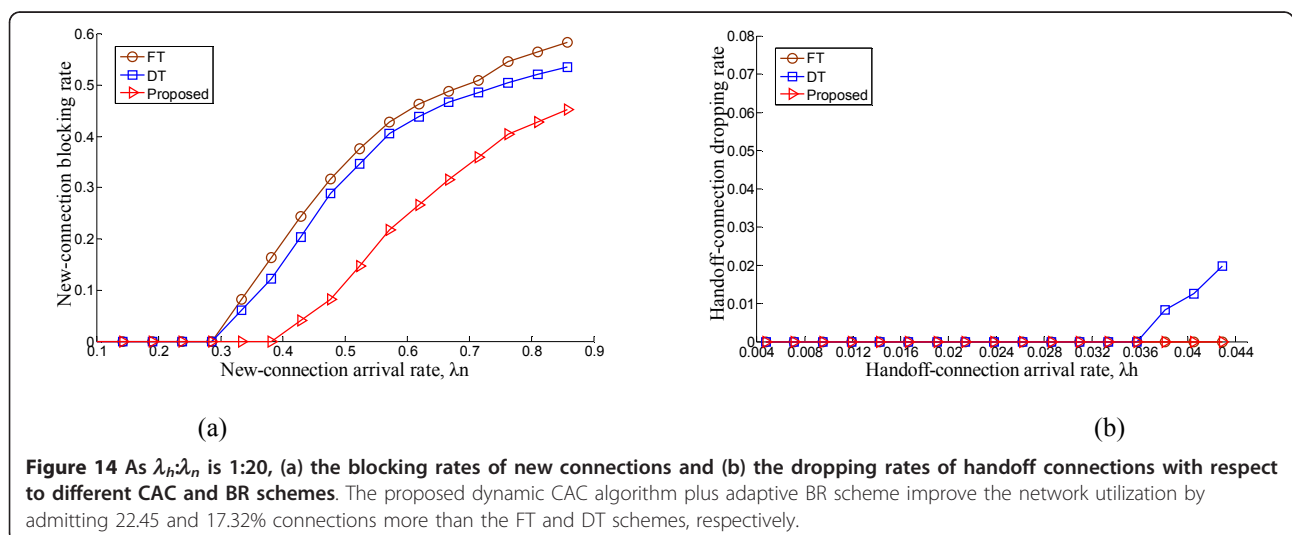
Figure 13 As λ_h:λ_n is 1:1, (a) the blocking rates of new connections and (b) the dropping rates of handoff connections with respect to different CAC and BR schemes. The proposed dynamic CAC algorithm plus adaptive BR scheme averagely increase the network efficiency by 19.71 and 19.95% over the FT and DT schemes, respectively.

14, we can observe that our approach generally performs better in the latter where the occurrences of new connections are much more frequent than that of handoff connections. The performance improvement is attributed to our adaptive BR scheme which increases the network efficiency by dynamically adjusting the amount of reserved bandwidth for handoff users according to the arrival distributions of both handoff and new connections. The simulation results demonstrate the effectiveness of our dynamic CAC algorithm and adaptive BR scheme to increase the network efficiency in terms of granting more connections and also guarantee adaptive QoS for admitted new and handoff connections.

5. Conclusion

The mechanisms of CAC and BR are important research issues in wireless networks due to scarce radio resources, dynamic channel qualities, and diverse user demands. The article presents a dynamic CAC algorithm and an adaptive BR scheme for IEEE 802.16e mobile WiMAX networks to simultaneously improve the utilization efficiency of network resources and guarantee QoS for admitted new connections and handoff connections. Our CAC algorithm dynamically determines the admission criteria according to network loads to provide adaptive QoS for admitted users and improve the utilization efficiency of networks. When new connections and handoff connections enter the network based on the adaptive admission criteria, the proposed BR scheme further dynamically adjusts the amount of reserved bandwidth for handoffs according to the arrival distributions of new/handoff connections to increase the connection admission opportunities and also guarantee the bandwidth requirements for handoff QoS. Simulations

Figure 14 As λ_h:λ_n is 1:20, (a) the blocking rates of new connections and (b) the dropping rates of handoff connections with respect to different CAC and BR schemes. The proposed dynamic CAC algorithm plus adaptive BR scheme improve the network utilization by admitting 22.45 and 17.32% connections more than the FT and DT schemes, respectively.

results demonstrate that our approach can simultaneously improve the efficiency of resource utilization by granting more connections in the network and provide adaptive QoS for admitted new and handoff connections.

Competing interests
The authors declare that they have no competing interests.

References
1. IEEE 802.16 WG, IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems. Amendment 2, IEEE (December 2005)
2. S-C Lo, Y-Y Hong, A novel QoS scheduling approach for IEEE 802.16 BWA systems, in *Proceedings of the 11th IEEE International Conference on Communication Technology (ICCT)*, Hangzhou, China, pp. 46–49 (Nov 2008)
3. A Sayenko, O Alanen, J Karhula, T Hamalainen, Ensuring the QoS requirements in 802.16 scheduling, in *Proceedings of the 9th ACM international symposium on Modeling analysis and simulation of wireless and mobile systems (MSWiM)*, Torremolinos, Malaga, Spain, pp. 108–117 (Oct 2006)
4. B Rong, Y Qian, K Lu, H-H Chen, M Guizani, Call admission control optimization in WiMAX networks. IEEE Trans Veh Technol. **57**(4), 2509–2522 (2008)
5. B Rong, Y Qian, H-H Chen, Adaptive power allocation and call admission control in multiservice WiMAX access network. IEEE Wirel Commun. **14**(1), 14–19 (2007)
6. WS Jeon, DG Jeong, Call admission control for mobile multimedia communications with traffic asymmetry between uplink and downlink. IEEE Trans Veh Technol. **50**(1), 59–66 (2001). doi:10.1109/25.917873
7. D Niyato, E Hossain, Call admission control for QoS provisioning in 4G wireless networks: issues and approaches. IEEE Netw. **19**(5), 5–11 (2005). doi:10.1109/MNET.2005.1509946
8. B Rong, Y Qian, K Lu, Integrated downlink resource management for multiservice WiMAX networks. IEEE Trans Mob Comput. **6**(6), 621–632 (2007)
9. X Chen, B Li, Y Fang, A dynamic multiple-threshold bandwidth reservation (DMTBR) scheme for QoS provisioning in multimedia wireless networks. IEEE Trans Wirel Commun. **4**(2), 583–592 (2005)
10. SB Chaudhry, RK Guha, Adaptive connection admission control and packet scheduling for QoS provisioning in mobile WiMAX, in *Proceedings of IEEE International Conference on Signal Processing and Communication (ICSPC)*, Dubai, United Arab Emirates, pp. 1355–1358 (Nov 2007)
11. D Niyato, E Hossain, A game-theoretic approach to bandwidth allocation and admission control for polling services in IEEE 802.16 broadband wireless networks, in *Proceedings of the 3rd international conference on Quality of service in heterogeneous wired/wireless networks (QShine'06)*, Waterloo, Ontario, Canada, (August 2006)
12. D Niyato, E Hossain, Radio resource management games in wireless networks: an approach to bandwidth allocation and admission control for polling service in IEEE 802.16. IEEE Wirel Commun. **14**(1), 72–83 (2007)
13. L Wang, F Liu, Y Ji, N Ruang-chaijatupon, Admission control for non-preprovisioned service flow in wireless metropolitan area networks, in *Proceedings of the Fourth European Conference on Universal Multiservice Networks (ECUMN'07)*, Toulouse, France, pp. 243–249 (February 2007)
14. K Yu, X Wang, S Sun, L Zhang, X Wu, A statistical connection admission control mechanism for multiservice IEEE 802.16 network, in *Proceedings of IEEE Vehicular Technology Conference (VTC 2009-Spring)*, Barcelona, Spain, pp. 1–5 (April 2009)
15. H Wang, W Lei, DP Agrawal, Dynamic admission control and QoS for 802.16 wireless MAN, in *Proceedings of the Wireless Telecommunications Symposium (WTS2005)*, Cal Poly Pomona, CA, USA, pp. 60–66 (April 2005)
16. TC Tsai, CH Jiang, CY Wang, CAC and packet scheduling using token bucket for IEEE 802.16 networks. J Commun. **1**(2), 30–37 (2006)
17. D Niyato, E Hossain, Queue-aware uplink bandwidth allocation and rate control for polling service in IEEE 802.16 broadband wireless access networks. IEEE Trans Mob Comput. **5**(6), 668–679 (2006)
18. DS Shu'aibu, SK Syed-Yusof, N Fisal, Partition-based bandwidth managements for mobile WiMAX IEEE802.16e. Int Rev Comput Softw (IReCoS). **5**(4), 445–452 (2010)
19. E-C Park, H Kim, J-Y Kim, H-S Kim, Dynamic bandwidth request-allocation algorithm for real-time services in IEEE 802.16 broadband wireless access networks, in *Proceedings of IEEE INFOCOM 08*, Phoenix, Arizona, USA, pp. 1526–1534 (August 2008)
20. E-C Park, Efficient uplink bandwidth request with delay regulation for real-time service in mobile WiMAX networks. IEEE Trans Mob Comput. **8**(9), 1235–1249 (2009)
21. Y Zhang, D Liu, An adaptive algorithm for call admission control in wireless networks, in *Proceedings of the IEEE Global Communication Conference (GLOBECOM)*, San Antonio, TX, USA, pp. 3628–3632 (Nov 2001)
22. J Choi, T Kwon, Y Choi, M Naghshineh, Call admission control for multimedia services in mobile cellular networks: a Markov decision approach, in *Proceedings of IEEE ISCC*, Antibes, France, pp. 594–599 (July 2000)
23. R Ramjee, R Nagarajan, D Towsley, On optimal call admission control in cellular networks, in *Proceedings of IEEE INFOCOM '96*, San Francisco, CA, USA, pp. 43–50 (March 1996)
24. Y Fang, Y Zhang, Call admission control schemes and performance analysis in wireless mobile networks. IEEE Trans Veh Technol. **51**(2), 371–382 (2002). doi:10.1109/25.994812
25. S Nazari, H Beigy, A new distributed uplink packet scheduling algorithm in WiMAX networks, in *Proceedings of IEEE ICFCC*, Wuhan, China, pp. 232–236 (May 2010)
26. Eunyoung Lee, Hyung-Kun Park, Packet scheduling scheme for multiple services in mobile WiMAX system, in *Proceedings of 2nd International Conference on Computer and Network Technology (ICCNT)*, Bangkok, Thailand, pp. 60–63 (April 2010)
27. B Al-Manthari, N Nasser, H Hassanein, Packet scheduling in 3.5 G high speed downlink packet access networks: breadth and depth. IEEE Netw Mag. **21**(1), 41–46 (2007)
28. M Nicolaou, A Doufexi, S Armour, Y Sun, Scheduling techniques for improving call capacity for VoIP traffic in MIMO-OFDMA networks, in *Proceedings of IEEE Vehicular Technology Conference (VTC 2009-Fall)*, Anchorage, Alaska, USA, pp. 1–5 (Sept 2009)
29. T Bonald, A score-based opportunistic scheduler for fading radio channels, in *Proceedings of the European Wireless Conference(EWC '04)*, Barcelona, Spain, pp. 283–292 (Sept 2004)
30. J Holtzman, G Barriac, Introducing Delay Sensitivity into the Proportional Fair Algorithm for CDMA Downlink Scheduling, in *Proceedings of IEEE Seventh International Symposium on Spread Spectrum Techniques and Applications(ISSSTA)*, Parsippany, NJ, USA, pp. 652–656 (Sept 2002)
31. W-H Kuo, W Liao, Utility-based resource allocation in wireless networks. IEEE Trans Wirel Commun. **6**(10), 3600–3606 (2007)
32. FP Kelly, Charging and rate control for elastic traffic. Eur Trans Telecommun. **8**(1), 33–37 (1997). doi:10.1002/ett.4460080106
33. FP Kelly, AK Maulloo, DHK Tan, The rate control for communication networks: shadow prices, proportional fairness and stability. J Oper Res Soc. **409**, 237–252 (1998)
34. G Song, Y Li, Utility-based resource allocation and scheduling in OFDM-based wireless broadband networks. IEEE Commun Mag. **43**(12), 127–134 (2005)
35. Forum™ WiMAX, Mobile WiMAX - Part I: A Technical Overview and Performance Evaluation. http://www.wimaxforum.org/technology/downloads/Mobile_WiMAX_Part1_Overview_and_Performance.pdf (2006)
36. MR Ashayeri, H Taheri, Mobile WiMAX capacity estimation in various conditions, in *Proceedings of IEEE Iranian Conference on Electrical Engineering (ICEE)*, Isfahan, Iran, pp. 483–488 (May 2010)
37. A Belghith, L Nuaymi, WiMAX capacity estimations and simulation results, in *Proceedings of IEEE Vehicular Technology Conference (VTC 2008-Spring)*, Marina Bay, Singapore, pp. 1741–1745 (May 2008)
38. IEEE 802.16e 2005, IEEE Standard for local and metropolitan area networks. Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems. (February 2006)

39. IEEE C802.16e-04/16, Coverage/capacity simulations for OFDMA PHY in ITU-T channel model including MRC, STC, AAS results. (November 2004)
40. H Lee, H-D Kim, D-H Cho, Smart resource allocation algorithm considering voice activity for VoIP services in mobile-WiMAX system. IEEE Trans Wirel Commun. **8**(9), 4688–4697 (2009)