

RESEARCH

Open Access

Discretionary bandwidth granting scheme for homogenous real-time applications

Kuokkwee Wee^{1*}, Mardeni R², S W Tan² and S W Lee²

Abstract

IEEE 802.16e is an advanced wireless access technology that provides high-speed data transmission in long distance and offers quality of service (QoS) to subscribers. The provisioning of QoS is one of the great features by IEEE 802.16 to support both real-time and non-real-time applications. In IEEE 802.16, the common part sublayer in the MAC layer is responsible for maintaining the QoS services. There are many functions in the common part sublayer; the most popular topics discussed by researchers are the uplink and downlink scheduling algorithms. Many discussions had been made and focused on these two classes of schedulers. Another equally important component but overlooked so far is the bandwidth request and grant module. Bandwidth request and grant module arbitrates the amount of bandwidth to be granted, besides handling the bandwidth requests. This bandwidth request and grant process has always been developed in a conventional way, and its importance has been underestimated. In addition, the bandwidth distribution within a same service class or category also attracted little attention thus far. Many algorithms for interclass scheduling have been studied and proposed but not as much for intraclass scheduling. However, in bandwidth request and grant process, constraints on the required knowledge by the schedulers limit the intraclass scheduling algorithms to be applied onto them. We view the bandwidth request and grant process as an important part of the QoS architecture. In this paper, we proposed a new bandwidth granting scheme for the bandwidth request and granting process, which enables bandwidth to be fairly granted based on the necessity to all the requests from the same service class or category. By applying our scheme, significant improvements have been observed and recorded. Experiment results have proven and confirmed the effectiveness of our proposed scheme as compared to the conventional scheme.

Keywords: Broadband wireless access networks, Bandwidth control, IEEE 802.16, Quality of service, WiMAX, Scheduling, Bandwidth request and grant

1. Introduction

IEEE 802.16 is a set of standards for broadband wireless access (BWA). It was established in 1999 with the aim to deploy broadband wireless metropolitan area networks (WMAN) worldwide. It is aimed to deliver high-speed data services with QoS and prominent security features in several large geographical areas. Today, IEEE 802.16e has been widely deployed to extend the network penetration and coverage where other wired technologies are not available or too costly to be installed.

The WMAN specifications are defined in IEEE standard 802.16-2009 [1]. This new release supersedes the previous IEEE standard 802.16-2001 [2] which defines the

specifications for 2 to 11 GHz band applications to support orthogonal frequency division multiplexing (OFDM) and orthogonal frequency division multiple access (OFDMA). The new standard not only provides methodologies to achieve reliable and link adaptive high rate transmission over the wireless link but also QoS with a full set of parameters that permits service differentiation up to the connection level [3]. Thus, IEEE 802.16 is able to serve hundreds of different subscribers with different requirements in the network.

Today, IEEE 802.16e is deployed to support a wide variety of multimedia applications and Internet services. Triple play service, i.e., the ability to run voice service, online video, and Internet surfing simultaneously, is one of them. Voice service and online video are real-time applications while web browsing is a non-real-time application. This combination of the different services causes the

* Correspondence: wee.kuokkwee@mmu.edu.my

¹Faculty of Information Science & Technology, Multimedia University, Jalan Ayer Keroh Lama, Bukit Beruang, Melaka 75450, Malaysia

Full list of author information is available at the end of the article

network requirements become complicated and sophisticated. On top of that, some specific applications are also relied on IEEE 802.16e systems, for instance, closed-circuit television (CCTV) application and military surveillance system. These types of applications usually are sensitive in delay, and the QoS requirements are the same. Therefore, a prominent broadband access technology must be able to undertake different network traffics from different types of applications and fulfill their network requirements.

In IEEE 802.16, convergence sublayer maps upper layer packets to IEEE 802.16 protocol data unit when packets arrived from internet protocol (IP) or asynchronous transfer mode networks; in vice versa process, packets are received from physical layer (PHY). Packets from upper layer are mapped and categorized into five different service classes: 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). UGS is for voice service with constant bit rate, and ertPS is targeted for voice-over IP without silence suppression that has variable data rate. In addition, rtPS is intended for real-time video applications. Delay-insensitive applications with large data transfer and web browsing are grouped under nrtPS and BE service class, respectively. As stated in [1], UGS has the highest precedence, followed by ertPS, rtPS, nrtPS, and lastly BE. QoS implementation however is not specified in the standards and thus for vendors to design and optimize.

This paper focuses on the bandwidth request (BW REQ) and grant process for real-time applications in the IEEE802.16e network. It is observed that the bandwidth request and grant process is overlooked by many researchers. A typical priority-based mechanism is usually used by many researchers who proposed the QoS structure in the IEEE 802.16 networks. Also, this study only focuses on homogenous real-time application, i.e., real-time video streaming. This is because QoS requirements of real-time traffic are more challenging to maintain than non-real-time applications. Likewise, requests of the same service class or category are more difficult to be handled since they have very similar requirements. Hence, we do not consider other applications in our experiment but homogeneous real-time traffic. Besides, this study is to address the issue of internal bandwidth contention in a same service class or category. It is understood that the internal bandwidth contention also happens even in heterogeneous applications where more than one requests are having the same requirements. This study proposes a scheme to address the fundamental issues and problems of the conventional approach used in bandwidth request and grant process. It also defines detailed parameters affecting the bandwidth granting process at the base station (BS). Moreover, the proposed algorithm is able to improve the network performance (throughput, delay, and jitter) in

two different network scenarios. Particularly, the contributions of this work can be summarized as follows:

1. The issues that only consider the bandwidth request and service class as factors in bandwidth granting process are addressed in this work. The amount of bandwidth request and service class should not be the only two consideration factors in granting bandwidth process.
2. This work discovers that network performance could be further improved by distributing the bandwidth fairly among all the requests. At the meantime, strict restrictions imposed on bandwidth granting in a network scenario where the wireless channel conditions are varied may not improve the network performance in overall.
3. This work proposes and verifies an efficient bandwidth granting scheme for the point-to-multipoint (PMP) operation mode of the IEEE 802.16e system. The proposed scheme takes into consideration all of the factors that could affect bandwidth granting process and compiles value of the factors to be used in decision making later. The proposed scheme is also applicable to other OFDM/OFDMA-based broadband wireless access (BWA), i.e., long-term evolution. Furthermore, the proposed scheme had been tested and its efficiency verified in two different modulation network scenarios.

The rest of this paper is organized as follows: Section 2 briefly reviews some related research works. An overview of medium access control (MAC) layer, bandwidth request and grant process, challenges, and motivation are presented in Section 3. Section 4 elaborates our proposed bandwidth granting scheme. Simulation results are presented in Section 5, and Section 6 concludes this study.

2. Related research works

QoS architecture in IEEE 802.16e networks consists of a number of important components that work together to achieve the QoS provisioning goals. Call admission control (CAC) is the first level of QoS components to deal with packets. To accept or reject a service flow for admission into the network, CAC considers the required and available resource in the system. Authors from [4] introduced a two-dimensional (2-D) CAC framework. The 2-D framework consists of a dimension for downlink (DL) and a dimension for uplink (UL) where each has a utility and fairness constrained optimal revenue policy that operates on approximation algorithms. Another proposal is the latency rate scheduler [5] which is based on proposed optimal frame size that includes the time needed to transmit a maximum size packet and separation gaps of downlink DL and UL subframes. It also guarantees the maximum

delay required by the user. The author in [6] looked into not only the CAC but also into bandwidth allocation module and bandwidth control protocol for time division duplex (TDD) mode worldwide interoperability for microwave access (WiMAX). The bandwidth is assigned according to the priority of service class in its first step, followed by weighted round robin (WRR) in the second step. Bandwidth is given to higher priority classes if the maximum sustained rates (MSR) are fulfilled. The paper introduced auto adjustment for maximum sustained rate by two approaches, adaptive bandwidth borrowing algorithm general (ABB-General) and adaptive bandwidth borrowing algorithm reactive (ABB-Reactive). Both of the approaches are based on the relationship between bandwidth and the number of the bandwidth requests being fulfilled to adjust the MSR. ABB-General showed better performance in light load network while ABB-Reactive performs better in heavy load network. In addition, some studies on the downlink and uplink bandwidth allocation are also found in [7,8]. In [7], the study focuses on the adaptive ratio between uplink and downlink streams for BE traffic only. The adjustments for the downlink and uplink are by the current traffic profiles. It also cooperates with the BS scheduler to throttle the transmission control protocol source when acknowledgements are infrequent to prevent one-way bottleneck. The DL and UL allocations are further enhanced by adjusting the bandwidth ratio of DL to UL adaptively according to the current traffic profile, wireless interference, and transport layer parameters in [8].

Uplink bandwidth requests were studied in [9-11]. In [9], the focus is on the parameters that will affect the bandwidth request instead of the scheduling algorithms. The author introduced target delay and dual feedback in bandwidth request process. Target delay is a key to determine the amount of bandwidth request. Target delay is up to its maximum tolerable value. The dual feedback consists of loops that return queue length and arrival rate. The target delay and dual feedback are incorporated in well-known scheduling algorithms. In [10], the delay performance of the two bandwidth request mechanisms proposed in the IEEE 802.16 standard, random access and polling, was studied. The results showed that random access is more efficient than polling when the request rate is low. However, its performance degrades rapidly when channel load increases. Adaptive switching between random access and polling can improve system performance claimed by the authors [10]. However, more comprehensive proposal that demonstrates the influence of channel noise on the BW REQ mechanism is not negligible [11]. In [11], the channel noise degrades the performance of both throughput and delay; thus, a predictive approach that considers the collision probability, throughput, and mean delay of BW REQs under various traffic and channel conditions was introduced.

In order to guarantee the QoS requirements for the service classes in IEEE 802.16, some priority-based scheduling algorithms were proposed in [12-15]. The algorithms follow the QoS precedence prepared in [2]. Other than that, the comparison on the common weight-based scheduling algorithms was presented in [16,17]. WRR that works on the minimum reserved traffic rate and weighted fair queue (WFQ) which is associated with processing finishing time were compared in IEEE 802.16 network in [16,17]. In [18], the authors adopted the earlier deadline first (EDF) scheduling algorithm for uplink rtPS packets to avoid the delay violation. New design of an adaptive bandwidth scheduling scheme is based on proportional method of bandwidth for deadline and for request. Besides, in [19], a fixed departure rate is assigned to UGS; specified delay constraint is composed on rtPS traffic, and a minimum throughput requirement is imposed on nrtPS traffic. First, the adaptive bandwidth allocation assigns initial bandwidth to all the service classes by estimating their required initial bandwidths. If there is remaining bandwidth, it then assigns extra bandwidth to rtPS, nrtPS, and BE. The combination of various algorithms was found in [20]. Hybrid algorithms are vital in the distribution of bandwidth among the diverse traffic classes and evaluated hybrid (EDF + WFQ + FIFO) and hybrid (EDF + WFQ) schemes [20]. Some intraclass scheduling algorithms are studied in this study. Among the intraclass scheduling algorithms, look-ahead deficit round robin with weight adjustment in [21] utilizes (1) the packets from the flows that have yet to set up quanta or still backlogged after the service round and (2) maximum packet size and frame size. Besides, in [22], four variants of throughput-based intraclass scheduling are examined. The proposed utilitarian solution is based on the highest throughput, the Nash solution (successor of utilitarian solution) that uses the product of the throughput instead of summation in utilitarian solution, Kalai-Smorodinsky solution which uses weighted max-min, and the egalitarian solution follows the max-min throughput.

Unlike the wired network, channel condition is an important factor that may affect the performance of a wireless network. In IEEE 802.16, the channel condition of a subscriber station (SS) is described in the downlink interval usage code (DIUC) and uplink interval usage code (UIUC) messages. As the name implied, DIUC is the downlink channel condition while UIUC represents the uplink channel condition. These messages are transmitted through downlink media access protocol and uplink media access protocol (ULMAP), respectively. A better channel condition, for example, 64 quadrature amplitude modulations (QAM), could lead to higher throughput than a poor channel, for example, quadrature phase shift keying (QPSK) modulation scheme. Hence, there are a lot of scheduling proposals that consider the network

condition as one of the factors in scheduling and integrate it with MAC layer, named cross-layer design (PHY + MAC), as proposed in [17] and [23-27].

3. IEEE 802.16 MAC layer

The MAC layer of IEEE 802.16 is the most prominent protocol that distinguishes itself from other broadband wireless systems. The primary task of MAC layer is to act as a coordinator between the higher layers and physical layer. Packets from the upper layer, known as MAC service data units (MSDUs), are passed to MAC layer to be organized into MAC protocol data units and then forwarded to PHY layer for transmission over the air link. In receiving transmission from PHY layer, MAC performs the reverse steps. MAC also handles the management messages exchanged between SSs and BS.

One of the three sublayers in MAC, the convergence sublayer, is designed to classify and map the MSDUs to the designated service class and connection identifier (CID). Other processing, for instance, optional payload header suppression and packet header restoring, are also the functionalities of the convergence sublayer. Security sublayer, another sublayer in MAC, is responsible for providing authentication, secure key exchange, encryption, and integrity in IEEE 802.16 networks. It is formed to prevent all well-known security attacks that may cause denial of service and other threats to the networks. Lastly, the common part sublayer represents the core of the MAC protocol, and its responsibilities are bandwidth allocation, connection establishment, and maintenance. Besides, frame construction, multiple access, bandwidth request and grant, scheduling, radio resource management, and QoS management are also part of the common part sublayer. The common part sublayer is thus the core of MAC layer in maintaining the network QoS structure.

3.1. IEEE 802.16 bandwidth request and grant process

As in [1,2], there is no bandwidth request and grant process at the downlink; bandwidth request and grant process only happens at uplink, before the uplink transmissions commence. This process involves BS and the SS that desire to perform uplink transmission. In bandwidth request and grant process, the SS initializes the process by building the bandwidth request message. The bandwidth request message is composed by counting the size of the queue or by predicting the new packet arrival [8]. The bandwidth request message that contains the information of bandwidth needed for next cycle and the

corresponding CID are shown in Figure 1. CID is an essential component in IEEE 802.16 as it is used to identify a connection and transmission between BS and SS. In bandwidth request and grant, it is used to identify which SS is requesting bandwidth.

Upon the construction of bandwidth request message, SS will process ULMAP message and identify the dedicated time to send the bandwidth request message. Bandwidth request sent may be incremental or aggregate. An incremental bandwidth request is based on the additional amount of bandwidth needed while an aggregate bandwidth request is based on the absolute bandwidth needed. Besides, SS can send its bandwidth demand on per connection or per station basis. For per connection basis, SS passes its bandwidth request for each connection (one connection may have more than one service flows) to BS by its data CID. This bandwidth request only consists of the bandwidth needed for that particular connection, not for other connections. On the other hand, SS accumulates all the bandwidth requests of all its service flows and sends to BS through its basic CID in per station approach. By this method, the management messages needed are much fewer than per connection basis.

Bandwidth request can be sent in several ways by the SS, either implicitly or explicitly. The implicit method is bandwidth stealing where bandwidth request message is sent instead of data message during the uplink transmission. Besides, a method which is the piggybacking bandwidth request on other MAC packets being sent to the BS is also widely used as implicit bandwidth request mechanism. Contention-based mechanism which uses similar approach in IEEE 802.11 is also allowed in IEEE 802.16 network. For explicit mechanism, a polling-based mechanism is used. The bandwidth request could be demanded by using unicast or multicast polling explicitly. Unicast polling is only used when the bandwidth is sufficient for polling all SSs individually.

Otherwise, SSs may be polled by multicast polling to minimize the collision. Somehow, not all the service classes are eligible to apply all the bandwidth requests mechanisms. See Table 1 for all the service classes and their eligibilities for various bandwidth requests mechanisms.

On the BS side, all the bandwidth request messages from SSs are located and kept in a queue. After the reception of bandwidth request, bandwidth request and grant manager at the BS will start to allocate the available bandwidth accordingly. The assignment of the available bandwidth depends on the algorithms adopted by

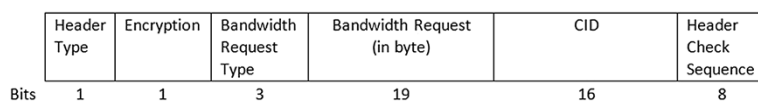


Figure 1 Bandwidth request message.

Table 1 Bandwidth request mechanisms implicit and explicit

Scheduling type	Piggyback request	Bandwidth stealing	Contention region	Polling
UGS	No	No	No	PM bit is used to request a unicast poll.
rtPS	Yes	Yes	No	Unicast polling
nrtPS	Yes	Yes	Yes	No restriction unless prohibited by scheduler.
BE	Yes	Yes	Yes	No restriction.

the bandwidth request and grant manager. Although the bandwidth requests could be per connection basis or per station basis, the BS allocates bandwidth to SSs on per station basis. The allocated bandwidth is represented in physical slot (PS) unit for each SS (if any) and then incorporated in ULMAP message which will be broadcasted to every SS in the network.

Each SS analyzes all the management messages from BS. Upon the receiving of its ULMAP message, SS will extract the information from uplink time slots. The number of these time slots determines the number of packets a SS can send in that cycle. A SS will remain in idle mode until the assigned time slot arrives and commences uplink transmissions soon after. The entire process flow is presented in Figure 2.

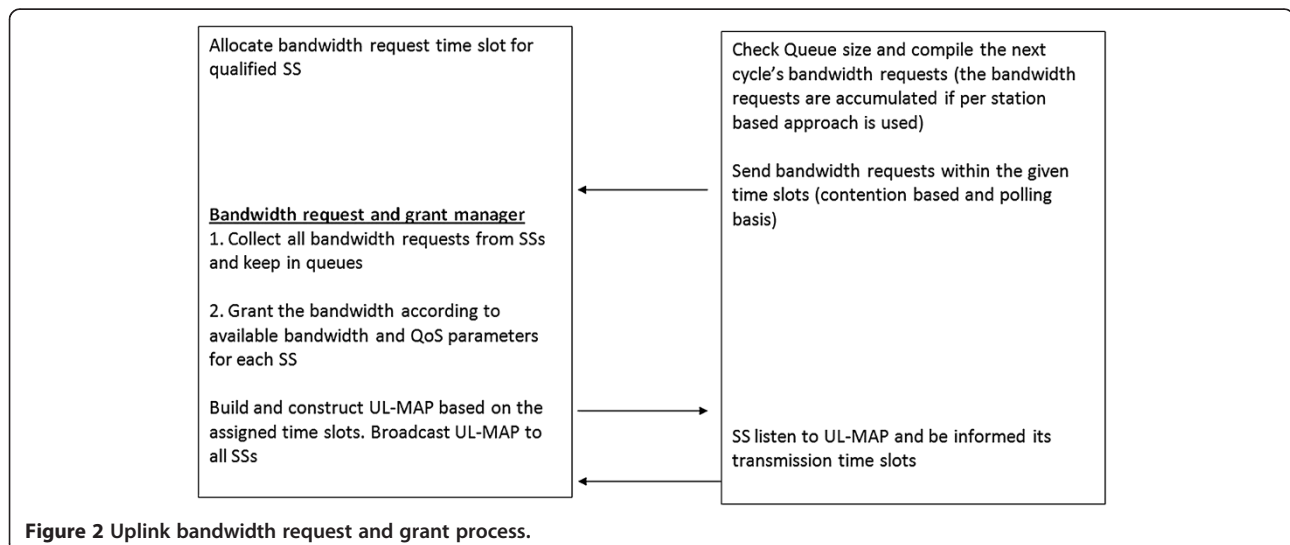
3.2. Challenges and motivations in bandwidth request and grant process for the same service class

In bandwidth request and grant module, there are only three sets of parameters a BS may have from SS. They are the amount of bandwidth request for the next frame, the wireless network channel condition, and the QoS precedence. The impact on the QoS service class precedence becomes negligible if the bandwidth competition is between the requests from the same service class. This scenario is common in data communications that have several requests in a service class by different SSs at the

same time even though there may be only one real-time application running in the entire network, for example, CCTV application that uses IEEE 802.16 to stream video. Moreover, the BS has no other local information from SS such as the packet size and time stamp except for the three parameters mentioned above. Bandwidth request and grant module is bounded by these constraints, and thus, not all the scheduling algorithms can be applied in it.

BS retrieves a SS bandwidth demand by extracting bandwidth request message header that is received from the SS. This bandwidth request message only informs the BS the needed amount of bandwidth for the next cycle even though there are proposals to incorporate some extra information in the bandwidth request message, but the problem to assign bandwidth fairly among requests of the same service class still persists. In [3], the bandwidth request message is partitioned into two categories, guaranteed and non-guaranteed; but the effort will be in vain if all the bandwidth requests are from only one category, either guaranteed or non-guaranteed type. The bandwidth request and grant module will still need to face the need to manage bandwidth requests within the guaranteed category itself or within non-guaranteed category internally. Our proposed scheme is designed to look into this issue and offer a solution.

Intraclass or intracategory request serving in bandwidth request and grant process has not been investigated



seriously by researchers thus far. One of the problems is that the available bandwidth is always being assigned to one or two greedy requests within the same service class which might cause bandwidth starvation and high packet latency to other service flows. Secondly, there could be some service flows which are not in critical condition but they are getting bandwidth. In fact, the mechanism should reduce some of the bandwidth and redistribute to others. Our proposed scheme takes into consideration of this issue and views it as a motivation to reallocate the bandwidth. The third issue is that the bandwidths are being assigned to some service flows while other service flows do not get any in a cycle. This is a shortcoming in bandwidth request and granting process. All the above issues will result in high delay and jitter to those service flows not allocated with any bandwidth. Hence, our proposed scheme applies bandwidth redistribution on some qualified requests. The qualified requests are decided by a factor point system which will be discussed later. With such approach, it will be able to relieve bandwidth starvation in the network and subsequently lower the latency of packets.

Our objective is to propose an efficient algorithm in bandwidth request and grant process which is often overlooked by many researchers. The primary intention in the proposed scheme is to reduce the latency and jitter for real-time service class by redistributing the bandwidth fairly. Secondly, the proposed scheme looks into and resolves the problems within the one same service class during bandwidth granting process. It is proven that the redistribution of bandwidth successfully reduces the latency and jitter for rtPS service flow as shown in simulation results to be presented later.

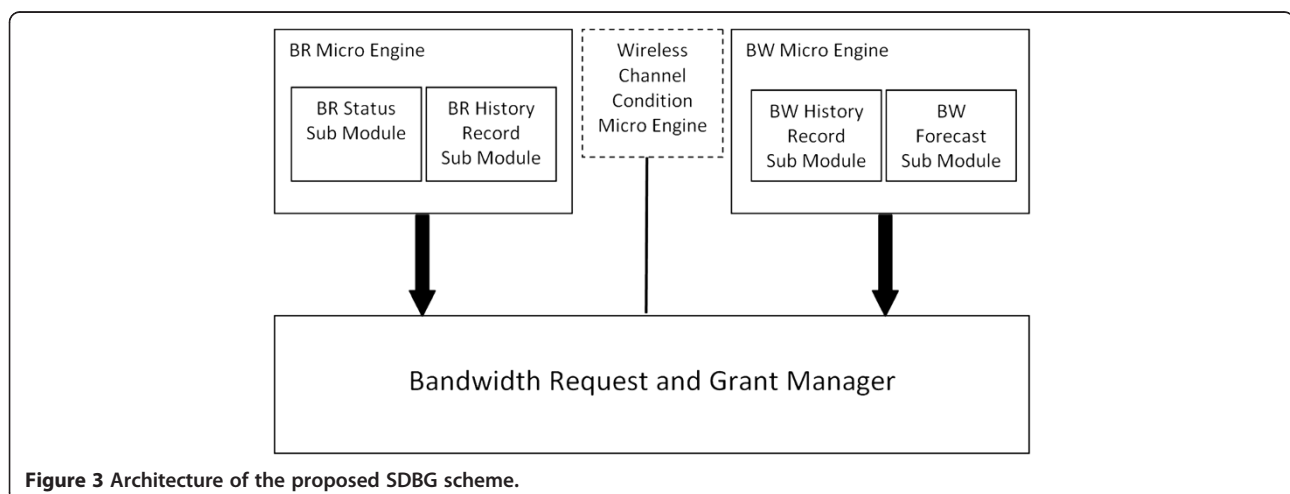
4. Self-discretionary bandwidth granting scheme

Our proposed self-discretionary bandwidth granting scheme (SDBG) is designed to grant the bandwidth to each SS in a fair manner. More importantly, the granted

bandwidth must be based on actual need and must be precise and adequate. In order to commit and assign accurate bandwidth assigned to a SS, SDBG utilizes the three parameters, i.e., QoS precedence, bandwidth request, and wireless channel condition obtained from SS, and an element is generated by the BS itself, which is the bandwidth allocation history. SDBG does not require any new extra information exchanged between BS and SS except those stated in [1]. This prevents additional bandwidth to be consumed for the extra information exchange. With additional bandwidth consumption, it reduces the network performance in terms of throughput and delay. Besides the abovementioned parameters, the BS self-generated granted bandwidth history is also included as an input.

The architecture of the proposed SDBG scheme is illustrated in Figure 3, and it can be adapted for other OFDMA systems easily. The bandwidth request information from SS is fed into the bandwidth request (BR) microengine which consists of two submodules, BR status submodule, and BR history record submodule. Information on wireless channel condition is utilized by wireless channel condition (WCC) microengine while the BS self-generated bandwidth history is utilized by bandwidth (BW) microengine. BR microengine and BW microengine will feed a cumulative factor point to bandwidth request and grant manager to be considered in bandwidth granting process. WCC microengine plays its role when the total granted bandwidth for a group of SSs is found out of the plan where a group of SSs overconsumes the available bandwidth. The overconsumption scenario is discussed in later part. WCC microengine is essential in maintaining the network hierarchy structure and the provisioning of QoS in our proposed scheme.

Typically, there are three outcomes from SDBG, i.e., the bandwidth request of a SS is (1) fully granted, (2) partially granted, and (3) not granted.



4.1. Bandwidth request and grant manager

Bandwidth request and grant manager relies on three components in order to complete the bandwidth granting process. They are BR microengine, BW microengine, and WCC microengine.

Bandwidth request and grant manager first computes the amount of overdemanded bandwidth by all the SSs in the current cycle. Bandwidth request and grant manager will not activate the three microengines if the total requested bandwidth is less than the total available bandwidth. In this case, it means that the demand is less than the supply, so every request can be fulfilled and no special arrangement is needed. However, when the total requested bandwidth is more than the total available bandwidth, a mechanism is needed to distribute the bandwidth fairly among all the SSs. This mechanism also helps to prevent overassigning bandwidth to any SSs. Hence, bandwidth request and grant manager activates the three components to undertake the bandwidth granting process.

BR microengine and BW microengine are equally important in our SDBG scheme, and therefore, each of the microengines contributes inputs of equal weight to the bandwidth granting process. The accumulated inputs are independent from each other, and it is used to regulate the amount of bandwidth to be granted to a SS.

Factor point, a decimal point number ranged from 0.00 to 0.25 (from 0% to 25% of contribution in overall), is the input returned by each of the submodules in BR microengine and BW microengine. WCC microengine does not provide any factor point to bandwidth request and grant manager, but it assists in maintaining the network hierarchy and QoS structure from the channel condition perspective. The factor points are then accumulated and subtracted from 1 (the highest value) to determine the amount of requested bandwidth to be reduced or sacrificed in order to comply with the total available bandwidth. The total of factor points that accumulated from all the submodules is called cumulative factor point (CFP). For example, a scenario where the total overdemanded bandwidth from all SSs is y , and the total available bandwidth is z . The n th SS has 0.50 as CFP, and the total CFP is 1; thus, this SS will have its granted bandwidth determined by cutting 50% off the requested bandwidth request. The calculation is presented in (1):

$$\text{GrantedBW}_n = \text{BwRequest}_n - \left[\left(\frac{\text{CFP}_n}{\sum \text{CFP}} \times (y - z) \right) \right] \quad (1)$$

where CFP_n denotes the CFP of the n th SS, and, BwRequest_n is the bandwidth request while GrantedBW_n is the granted amount of bandwidth for the n th SS. GrantedBW is taken from (1) because the newly calculated bandwidth request of a SS has been determined after

considering and removing the amount of overdemanded bandwidth request. High value in factor point indicates that the requesting SS is sensitive in losing bandwidth. On the other hand, low factor point value signals that the SS is more flexible in its bandwidth demand and adjustments could be carried out.

The total granted bandwidth is then grouped according to the channel condition. This information is essential since WCC microengine is invoked by the percentage of granted bandwidth to the channel condition. There are two situations where WCC microengine will be invoked:

1. The total granted bandwidth for the SSs with the best channel condition exceed 80% of the total available bandwidth or
2. The total granted bandwidth for the SSs from not the best channel condition consumed more than 20 percent of the total available bandwidth.

4.2. Bandwidth request microengine

BR microengine is designed to learn the bandwidth request behavior for every SS in the network. This is because knowing the bandwidth request behavior will able to diminish bandwidth starvation since the bandwidth request amount is one of the factors to determine the amount of granting bandwidth. BR microengine consists of two submodules which are the BR status and BR history record. Both submodules work independently from each other, and they are based on different sets of input parameter. Overall, BR microengine returns up to a maximum 0.50 factor points (50% of inputs) to bandwidth request and grant manager.

4.2.1. BR status submodule

The BR status submodule considers the amount of bandwidth request in the current cycle only. Usually, a SS accumulates both the current arrival packets and previous remaining packets as a new bandwidth request message as shown in Figure 4. Hence, the bandwidth request messages in current cycle are a mixture of the old and new packets. BR status submodule utilizes the newly arrived packets to know the incremental or decremental of the actual bandwidth request by a SS. By doing this, the actual amount of bandwidth request for the new packets in the current cycle could be identified. This information is useful in order to do flexibility adjustment on the bandwidth for a SS. New packets are more delay tolerable compared to old packets. With this information, it gives flexibility in bandwidth granting process to bandwidth request and grant manager. Although the description above assumed the bandwidth request is based on the queue size, other mechanisms used in building the bandwidth request may also applicable as long as there is a mixture of new and old portions in the bandwidth request.

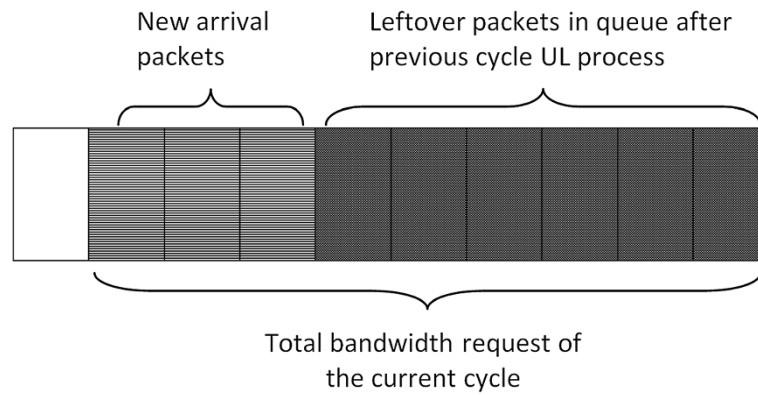


Figure 4 The mixture of new and old bandwidth request.

In BR status submodule, a factor point will be assigned to reflect the changes of the current bandwidth request. This submodule contributes 25% in overall in determining the number of bandwidth which should be granted for a SS. Generally, higher point is given if an increase of bandwidth request is detected while lower factor point is allocated if there is a decrease of bandwidth request. A maximum 0.25 factor point is awarded if the increase is 100% more than the previous bandwidth request. On the other hand, 0 factor point is awarded if there is no increment compared to previous bandwidth request. The calculation of factor point for this submodule is presented in (2) and (3):

$$f(x) = \begin{cases} \frac{(\text{BWRequest}_x - (\text{BWRequest}_{x-1} - \text{BWAllocated}_{x-1}))}{\text{BWRequest}_x}, & x > 1 \\ \text{BWRequest}_x, & x = 0 \end{cases} \quad (2)$$

$$\text{FP}_{\text{BRstatus}} = \begin{cases} 0.25, & f(x) \geq 1 \\ f(x) \times \text{maxFP}, & f(x) > 0 \text{ and } f(x) < 1 \\ 0, & f(x) \leq 0 \end{cases} \quad (3)$$

where x is number of cycle, and $f(x)$ denotes the changes of the bandwidth request for x cycle. BWRequest_{x-1} and BWAllocated_{x-1} are the bandwidth request and bandwidth allocation for $x-1$ cycle, respectively. However, the $f(x)$ is the amount of bandwidth request for the first cycle of bandwidth granting process. $\text{FP}_{\text{BRstatus}}$ indicates the computed factor point of BR status submodule while maxFP is the maximum factor point with 0.25 as its value.

4.2.2. BR history record submodule

The BR history record submodule takes consideration on the bandwidth request history record of a SS. The history record is used as a reference to understand the tendency of bandwidth demand. BR history record submodule keeps track on the average bandwidth request and its standard

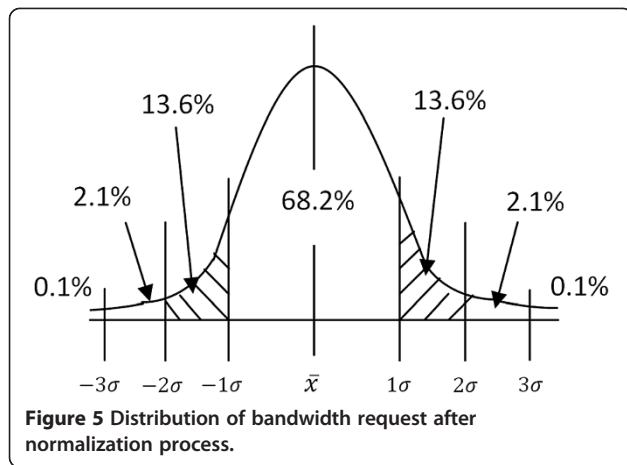
deviation. The archiving process starts from first until the last bandwidth request received. Mean and standard deviation for every SS are stored separately and independently. The mean is the average bandwidth request while standard deviation is a measure of how the bandwidth requests are spread out, during a given time frame. Due to time saving and processing consumption concerns, BR history record submodule only keeps the updated average and standard deviation instead of all the bandwidth request records. Indeed, the new bandwidth request record is added to the previous mean and standard deviation by (4) and (5):

$$\bar{x}_n = \frac{(n-1)\bar{x}_{n-1} + \bar{x}_n}{n} \quad (4)$$

$$\sigma_n = \begin{cases} \sqrt{\frac{(n-2)\sigma_{n-1}^2 + (x_n - \bar{x}_n)(x_n - \bar{x}_{n-1})}{n-1}}, & n > 1 \\ x, & n = 1 \end{cases} \quad (5)$$

where \bar{x}_n and σ_n denote the mean and standard deviation of bandwidth request for the n th cycle, respectively. Anyhow, the mean and standard deviation of each SS are calculated in every cycle to maintain the accuracy.

With the combination of the mean and standard deviation, the current bandwidth request can be classified into several categories. Since the value between 1 standard deviation from mean covers majority of sample (68.2%), also known as inflection points, it is considered as norm case, see Figure 5. Whichever value of bandwidth request falls in this boundary is assumed no much change in bandwidth request or the changes are not significant, and it is awarded 0.10 factor points. However, if a bandwidth request value is more than 1 standard deviation and less than 2 standard deviations from the mean, this bandwidth request is classified as a slight exceed in the norm; 0.15 points are given as its factor point. For the bandwidth



request in between 2 standard deviations and 3 standard deviations more from the mean, this is considered as a far exceed in the norm, so 0.20 is given. Also, for those bandwidth requests values more than 3 standard deviations, these are extreme cases, and 0.25 factor points are assigned.

On lower site, if the bandwidth request falls between -2 standard deviations and -1 standard deviation, it is assumed as low bandwidth request, and 0.05 factor points are assigned. Besides, for any bandwidth below -2 standard deviations, it meant that the bandwidth request is very low, thus 0 as its factor point.

4.3. Bandwidth microengine

Similar to BR microengine, BW microengine consists of two submodules, BW history record and BW forecast. BW microengine is designed to study the co-relationship between bandwidth granting and bandwidth request processes. Also, it assists in estimating the bandwidth needed by a SS in the next cycle. BW microengine does not reserve bandwidth for a SS based on the information gathered from its submodules, but it calculates the factor points which is needed by bandwidth request and grant manager in bandwidth granting process. BW microengine will feed up to a maximum of 0.50 factor points (50% overall) to bandwidth request and grant manager.

4.3.1. BW history record submodule

BW history record submodule is designed to know the relationship between granted bandwidth and bandwidth request of a SS. BW history record submodule obtains the average bandwidth request of a SS by using the equation in (4) as well as the average granted bandwidth. Similar to BR history record submodule, the average of granted bandwidth and bandwidth request for every SS are stored and archived individually and independently. The BW history record submodule employs the average granted bandwidth and the average bandwidth request

as the reference to estimate the sufficiency of allocated bandwidth for the next cycle frame. If the average granted bandwidth is higher than bandwidth request, it meant that sufficient bandwidth is granted most of the time; hence, the SS does not experience bandwidth starvation frequently. Meanwhile, this also indicates that this SS has a higher tolerance in terms of delay and throughput for the current cycle. BW history record submodule will recommend low factor point to this SS in this case. The granted bandwidth to bandwidth request ratio and factor point are calculated by using (6) and (7):

$$BW2BR_n = \frac{\bar{x}_n - \mu_n}{\bar{x}_n} \quad (6)$$

$$FP_{BW\text{history}} = \text{MIN}(\text{maxFP} \times BW2BR_n, 0.25) \quad (7)$$

where $BW2BR_n$ is the granted bandwidth to bandwidth request ratio value for the n th SS, μ_n and \bar{x}_n are the mean for granted bandwidth and bandwidth request, respectively. $FP_{BW\text{history}}$ is the factor point for BW history record submodule, and maxFP is the maximum factor point with 0.25 as its value. Higher factor point is committed by high $BW2BR$ value because high $BW2BR$ value reflects the absence of sufficient bandwidth been allocated.

4.3.2. BW forecast submodule

BW forecast submodule does not apply any probability calculation. BW forecast submodule exercises the unassigned bandwidth request as the consideration aspect. Unassigned bandwidth request is the remaining balance bandwidth request after the bandwidth assignment process is finished. In our approach, BS may not assign full bandwidth to all the requests; some SSs may only acquire certain percentage of their demands. These adjustments are required to minimize the bandwidth starvation and to prevent overallocation of bandwidth to certain greedy requests. Anyhow, the adjustments are properly calculated and figured out by our proposed approach.

A quotient, $BW\text{minReq}$, is computed based on unassigned bandwidth request as dividend and the average of bandwidth request as divisor. The value of $BW\text{minReq}$ implies the minimum bandwidth request will be sent by a SS to compromise the inadequate bandwidth assigned in the previous cycle. The assumption is that there is no packet dropped at SS or bandwidth request for the next cycle which is not less than the unassigned bandwidth request in the current cycle.

With high value of $BW\text{minReq}$, it is expected that the current bandwidth request is large and this request should be entertained urgently. High value of $BW\text{minReq}$ also denotes that this bandwidth request is now less tolerant to the delay; thus, high factor point is dedicated. Highest factor point, 0.25, is assigned if $BW\text{minReq}$ is greater than 1, and 0 point is given wherever there is no unassigned

bandwidth request. The BWminReq and factor point are calculated by using (8) and (9):

$$BWminReq_x = \begin{cases} \frac{BWRequest_{x-1} - BWAllocated_{x-1}}{\bar{\chi}_{x-1}}, & x > 1 \\ 0, & x = 0 \end{cases} \quad (8)$$

$$FP_{BWforecast} = \text{MIN}(\max FP \times BWminReq_x, 0.25) \quad (9)$$

where x is number of cycle, and $BWminReq_x$ denotes minimum bandwidth request will be sent by SS for x cycle. $BWRequest_{x-1}$ and $BWAllocated_{x-1}$ are the bandwidth request and bandwidth allocation for $x-1$ cycle, respectively. However, the $BWminReq_x$ value is set to 0 for the first cycle of bandwidth granting process. $FP_{BWforecast}$ indicates the computed factor point of BW forecast submodule while $\max FP$ is the maximum factor point with 0.25 as its value.

4.4. WCC microengine

WCC microengine extracts the SS UIUC index from the ULMAP message that transmitted between BS and SS on regular basis. The UIUC index gives information on the recent channel condition of an active SS. The channel condition is then passed to bandwidth request and grant manager as one of the three components to arbitrate bandwidth request in the current cycle.

In order to maintain the network hierarchy structure and provision of QoS, our bandwidth request and grant manager always tend to allocate maximum bandwidth for the better channel conditioned SSs. By doing this,

the maximum throughput of a network could be retained. However, there are always some SSs that could not have a better channel condition due to radio frequency fading or multipath effects. Hence, we proposed a simple module to undertake this issue. We also take into the consideration the industrial practice in [6] whereby a maximum of 80% of the bandwidth is always assigned to the best channel conditioned group of SSs unless they do not need all. The remaining 20% of bandwidth and the unutilized bandwidth from the upper groups (if any) are then been shared by the lower groups. In other words, SSs from the best channel condition will only consume up to a maximum of 80% of total bandwidth even though requested more. The other 20% of the total bandwidth is assigned to SSs with lower channel condition. The unutilized bandwidth will be shared in the case that the requested bandwidth is less than the maximum pre-assigned bandwidth. The algorithm for WCC microengine is given in Algorithm 1.

5. Simulation results

Our simulation is conducted by using Qualnet 5 network simulator (Los Angeles, CA, USA) [28] with at least ten experiments conducted for each scenario and each algorithm. The results are then averaged based on the number of experiments. Our simulation model is first presented followed by the two different network scenarios deployed in the experiments to investigate the performance of the proposed algorithm. The

Variable:

BR: Bandwidth Request for SS

MaxBW: 80% of the total available bandwidth

MinBW: 20% of the total available bandwidth

```
//sort from best to worst channel condition
1. Sort all SSs according to UIUC index
2. for best to worst channel condition do
3.   Accumulate BR for the same channel condition,  $\sum BR$ 
4.   Granted bandwidth =  $\text{Min}(\text{MaxBW}, \sum BR)$ 
5.   if MaxBW >  $\sum BR$ 
6.     MaxBW = MinBW + (MaxBW -  $\sum BR$ )
7.   else
8.     MaxBW = MinBW
9.   end if
10.  MinBW = 0
11.  if MaxBW equal to 0
12.    break for loop
13.  end if
14. end for
/*the pre-assigned 20% bandwidth may not be fully taken by other lower groups */
15. if MaxBW > 0
16.  Assign MaxBW to the best channel condition SSs
17. end if
```

simulation results and discussions are at the end of this section.

5.1. Simulation model

Our simulation model is a typical PMP mode that consists of a BS residing at the center of a geographical area with ten SSs surrounding in the service coverage area of the BS. The parameters of PHY and MAC layers used in the simulation experiments are summarized in Table 2. Our proposed scheme is deployed in the bandwidth request and grant module which is only exists in uplink. Therefore, the entire simulation focuses on the uplink process and omits the downlink transmission. In these experiments, we apply the following hypotheses:

1. Uplink scheduling algorithms deployed at SSs which utilized the granted bandwidth at a maximum rate.
2. The variation of any wireless channel is independent to each SS.
3. All connections have been admitted and ready for transmission.

For input traffics, the incoming traffics from [3] are formulated by removing all the downlink traffics, uplink nrtPS, and uplink BE service flows. Our proposed SDBG scheme is targeting on homogenous real-time traffic with bandwidth request mechanism and the bandwidth resource management for the same service class or category. Thus, non-real-time traffics (nrtPS and BE) are omitted. Overall, all SSs are equipped with two rtPS connections with 1.2 and 0.8 Mbps as the traffic loads, respectively. Each of these connections is pinned with a unique CID to conform to the IEEE 802.16 standard in [1], see Table 3. The incoming traffics are generated starting from 0 s until the end of simulation. Each simulation experiment runs for 10 to 90 s.

5.2. Wireless network scenarios

5.2.1. Scenario 1: 64 QAM modulation scheme only

In this scenario, we wanted to evaluate the performance of our proposed SDBG scheme when all SSs use the same and only one modulation and coding scheme. All SSs are located close to the BS with 64 QAM 3/4

modulation and coding scheme implemented. The wireless links between SSs and BS are line of sight. This network configuration allows us to examine the efficiency and the impact of our proposed SDBG scheme in the QoS structure of IEEE 802.16 networks. Via this, the maximum improvement on the QoS requirements in terms of throughput, latency, and jitter can be known.

To evaluate our proposed SDBG scheme, a bandwidth request and grant algorithm that works without regulate and examine the necessary factors had been developed. It is called unregulated bandwidth control (URBC) algorithm. URBC algorithm sorts all the bandwidth requests according to their service classes' precedence in IEEE 802.16 standard from the highest to the lowest. After the service class precedence checking, the bandwidth request and grant manager looks into the arrival time of the bandwidth request message. Bandwidth granting process commences based on service class precedence and the arrival time for bandwidth request message. The algorithm will grant the bandwidth according to the bandwidth request provided that there is enough available bandwidth. The granted bandwidth is the available bandwidth if the request exceeds the available bandwidth, and the algorithm will not assign more than the amount requested. No other factors will be taken into the consideration except service class precedence, the bandwidth demand, and the available bandwidth. This algorithm is simple and typically used by many researchers, e.g., [3,8,16] and [29,30].

A part of that, a round robin (RR) scheduling scheme is developed to study and compare with URBC and SDBG in bandwidth request and granting process. The RR scheduling algorithm is the common algorithm used in intraclass scheduling and it is compliant to IEEE 802.16 standard [1-3]. However, RR has not been implemented as the algorithm in bandwidth request and grant process by other researchers. In this study, the RR has been developed as one of the comparative scheduling algorithms. Other scheduling like WRR, DRR, DWRR, EDF,

Table 2 IEEE802.16e system parameters

Channel bandwidth	20 MHz
Fame duration	10 ms
DL/UL subframe duration	5 ms
Modulation scheme	16 QAM and 64 QAM
UCD/DCD broadcast interval	5 s
TTG/RTG	10 μ s
Transmission scheme	TDD

Table 3 CIDs and traffic load

Subscriber station	CID	
SS 1	201	202
SS 2	203	204
SS 3	205	206
SS 4	207	208
SS 5	209	210
SS 6	211	212
SS 7	213	214
SS 8	215	216
SS 9	217	218
SS 10	219	220
Incoming traffic load (Mbps)	0.8	1.2

Table 4 Summary of scheduling algorithms

Scheduling algorithm	Required information	For bandwidth request/grant protocol?
URBC [3,8,16,29,30]	None	Yes, classical approach. As a benchmark in this study.
RR	None	Yes, but no proposal in bandwidth request/grant protocol yet. It has been developed in this study for benchmarking purpose.
Priority based [12-15,19,30]	QoS precedence	Yes, but no difference in homogeneous traffic. As benchmark in this study.
WRR [16,17]	Peak packet rate and minimum packet rate	No packet information
Sun and Gani [24]	EDF + channel information	Same issues as in EDF
WDRR	HoQ/HoL of the packet	No packet information
WFQ	Packet size and channel capacity, finishing time	No packet information No finishing time info
SCFQ	Packet size and channel capacity, finishing time	No packet information No finishing time info
WF2Q	Time tags of every packet	No packet time info
EDF [18]	Earliest deadline of the maximum latency	No packet time information
Najah et al. [20]	EDF + WFQ + FIFO and EDF + WFQ	Same issues as in EDF and WFQ
Yuan and Duan [21]	Quantum value, maximum packet size and frame size	No information on packet size
Garroppo et al. [22]	Throughput	For non-real-time applications

WFQ, and other time-based intraclass scheduling algorithms are not suitable to be used in bandwidth request and granting process because of two reasons: first, there is no QoS requirement difference (same QoS parameter for the same service class in bandwidth request and granting process) and second, no time information for each packet (bandwidth request message only contains CID and bandwidth request amount). See Table 4 for the summary of the scheduling algorithms and the

reasons why they are not suitable in bandwidth request and granting process.

5.2.2. Scenario 2: mixture of 16 QAM and 64 QAM modulated SS

We evaluate the performance of SDBG scheme, RR scheduling scheme, and URBC approach in a scenario where the SSs are stationed in a combination of different modulation schemes. Twenty percent of the SSs

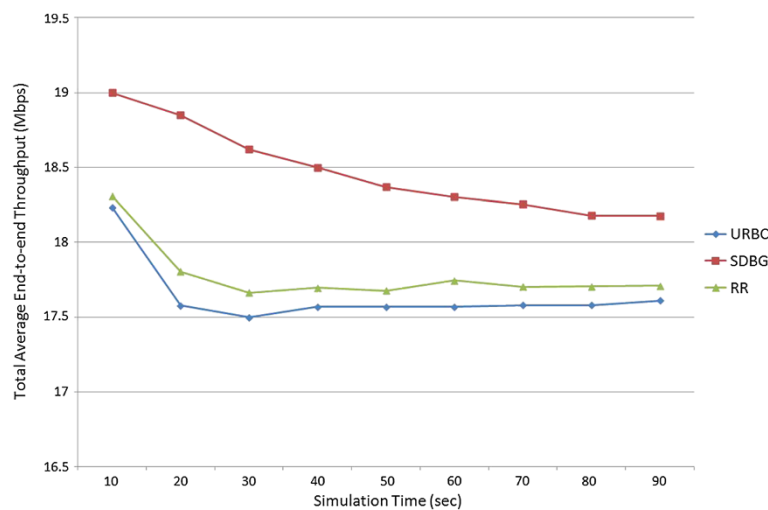


Figure 6 Total average end-to-end throughput for scenario 1.

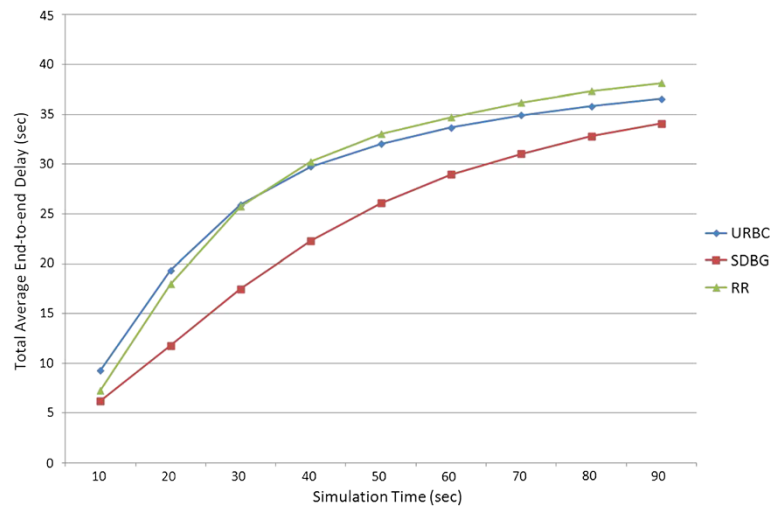


Figure 7 Total average end-to-end delay for scenario 1.

are placed far away from the BS where the 16 QAM modulation schemes could be achieved. Other 80% of the SSs remain unchanged on their last positions as in scenario 1. This is a practical network planning scenario as network planning always recommend maintaining at least 80% of the population in the highest modulation scheme.

We also developed another algorithm with some controls on the bandwidth in order to examine the impact by WCC microengine in SDBG approach. Unlike scenario 1, in practical network deployment where SSs resides in several coding and modulation schemes, the channel condition for a SS in scenario 2 may vary from one to another. Hence, an algorithm named as unregulated but conditioned bandwidth control (URCBC) is proposed. The URCBC algorithm inherited all the

functionalities from URBC algorithm, but it owned a mechanism to ascertain the following conditions:

1. Not more than 80% of total bandwidth is given to SS from the highest modulation scheme when other lower groups need more than 20% bandwidth.
2. Not more than 20% bandwidth occupancy for lower groups if more than 80% of the total bandwidth demand requested by the highest modulation scheme group.

This intention is to prevent the poor channel condition SSs from occupying too much bandwidth. In fact, a PS may carry 27 bytes in the highest modulated network (64 QAM 3/4) but it can just transmit 6 bytes in the poorer modulated network (QPSK 1/2). In order to

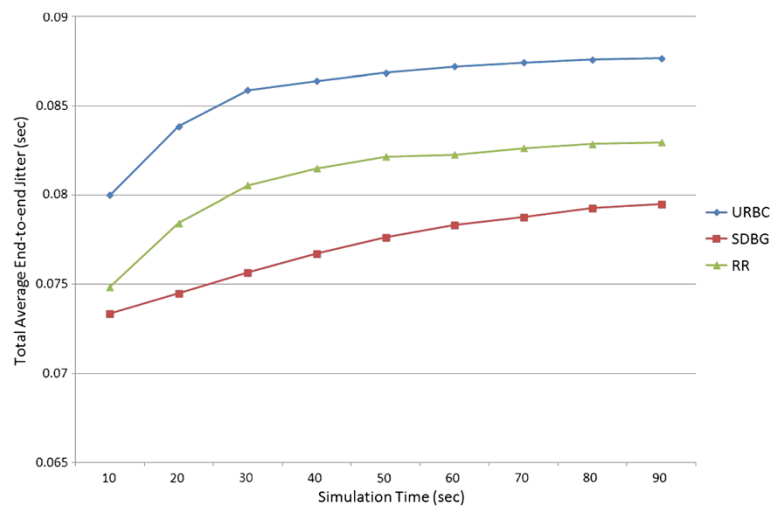


Figure 8 Total average end-to-end jitter for scenario 1.

maintain the network throughput, majority of the bandwidth should be given to the highest modulated SSs, which is also in line with the telecommunication planning and recommendation [31].

5.3. Simulation results and discussions under two different scenarios

Figure 6 presents the total average end-to-end throughput for SDBG, URBC, and RR in scenario 1 where there are only 64 QAM modulated SSs. All schemes and approaches show the decline of total average end-to-end throughput as the simulation time passes. The proposed SDBG scheme always has the better performance than both the URBC approach and RR scheme. SDBG recorded 4.21% higher than URBC and 3.78% higher than RR at 10 s of the simulation time. The SDBG's advantage increased to around 7% if compared to URBC and 5.5% compared to RR in the next 20 s of the simulation time. Volumes of incoming traffics are adequate but did not overload the network queues in the first 30 s; this creates rooms for the proposed SDBG scheme to shift and redistribute the bandwidth. In the first 30 simulation seconds, as high as 7.24% for URBC and 5.88% for RR in the total average end-to-end throughput are achieved. The differences between SDBG scheme and other scheme and approach become smaller from 40 s of simulation time and onwards. It is believed that the packets queue become more, and this resulted to fewer opportunities for SDBG scheme to execute its algorithms effectively. However, we still observed improvement ranged from 5.31% to 3.23% for URBC and from 4.55% to 2.65% for RR during 40 to 90 s of simulation time.

The total average end-to-end delay for scenario 1 is depicted in Figure 7. In total average end-to-end delay, SDBG always has the lower readings, between 3 and 8 s

and between 1 and 8 s, throughout the simulation as compared to URBC approach and RR scheme, respectively. Similar to the result presented in total average end-to-end throughput, SDBG has the best performance at the 20 s simulation time which is 39.09% (7.56 s) lower than URBC approach and 34.49% lower than RR. The sufficiency of history records needed by the algorithm and low queue occupancy are the two main contributions in this best performance case. However, the gaps between SDBG and other approaches become smaller as the simulation time passes. At 90 s, where the queue volumes are high and network is congested, SDBG approach still manages to secure 2.48 s (6.77%) lower than URBC and 4.07 s (10.685) lower than RR in the latency. At this point, the redistribution of bandwidth impact is not as great as in the early simulation, but its result is still very encouraging.

The jitter for scenario 1 which is depicted in Figure 8 to know the instantaneous difference between one and another packet, although it is not a compulsory QoS parameter in [1]. Our SDBG scheme possesses 8% to 11% lower jitter than URBC and 4% to 8% lower jitter than RR during the simulation. It is observed that the jitter for all approaches increase linearly, but the URBC approach shows that its increment tends to be more flat earlier than the proposed SDBG scheme. URBC approach shows that the increment is only 0.5 ms at 30 s, but SDBG scheme achieves such little increment only at the 60 s. The RR approach also increases its jitter as many as 0.5 ms at the first 30 s. This means that both the URBC approach and RR scheme reach its maximum jitter much earlier than our proposed SDBG scheme or, in other words, SDBG postponed the occurrence of high jitter in the network.

Figure 9 visualizes the comparison on the total average end-to-end throughput for URBC, RR, and SDBG schemes in the network scenario 2. As known, channel

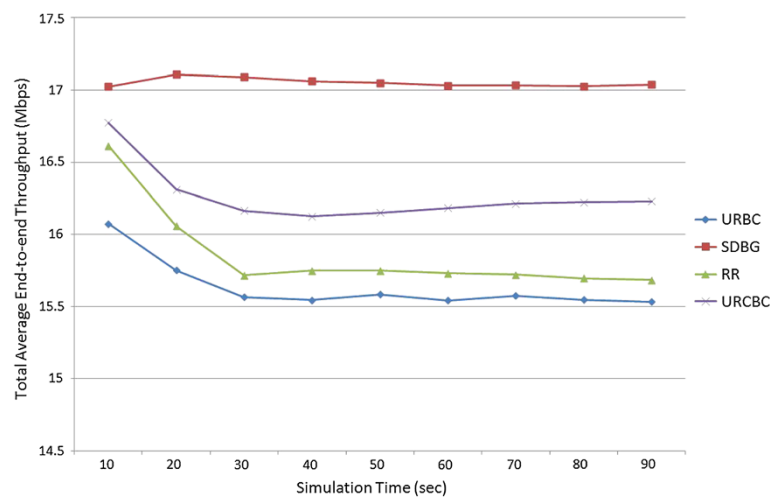


Figure 9 Total average end-to-end throughput for scenario 2.

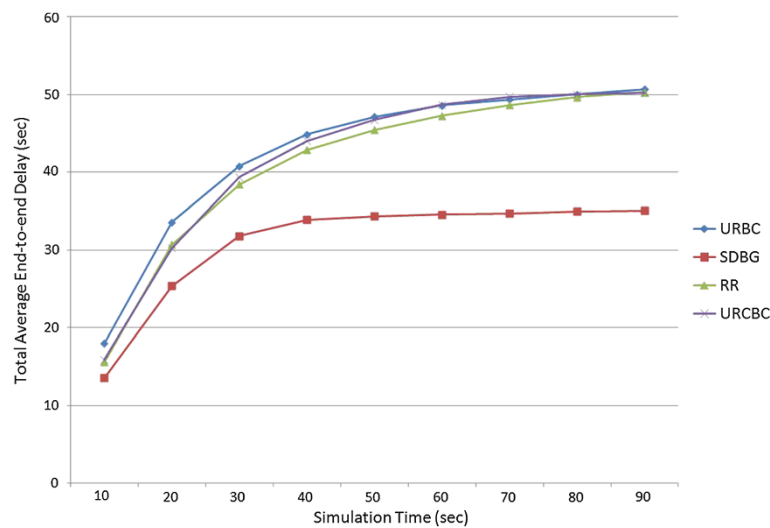


Figure 10 Total average end-to-end delay for scenario 2.

condition is a considerable factor when SSs are modulated in different schemes. From Figure 9 observation, SDBG has the best performance in throughput, followed by URCBC and RR, and the worst is URBC in the entire simulation time. Overall, SDBG throughput outperforms 9.07% and 4.68% on average compared to URBC and URCBC, respectively. Meanwhile, URCBC throughput has a higher average 3.86% than URBC which does not possess any conditions on the bandwidth control. The results reveal that unrestricted bandwidth granting to lower modulated SSs harms the throughput performance. This is caused by the granted time slot to lower modulated SSs that can only carry smaller amount of data as compared to high modulated SSs. It can be concluded that the wireless channel link is also an

important factor in bandwidth granting process. Moreover, the throughput performance could be better if some factors are considered as in our proposed SDBG scheme. With the factors esteemed by SDBG scheme, SSs are granted with the 'necessary' bandwidth which is determined by the microengines found in SDBG. From the results observed from URCBC, any algorithm which has fixed a certain amount of bandwidth to the best modulated SSs but neglecting other factors is not a promising scheme.

Total average end-to-end delay for scenario 2 is shown in Figure 10. Likewise in Figure 9, SDBG scheme is the champion in the total average end-to-end delay among the four schemes. As shown in Figure 10, SDBG recorded a minimum of 22% and a maximum of 31% better latency

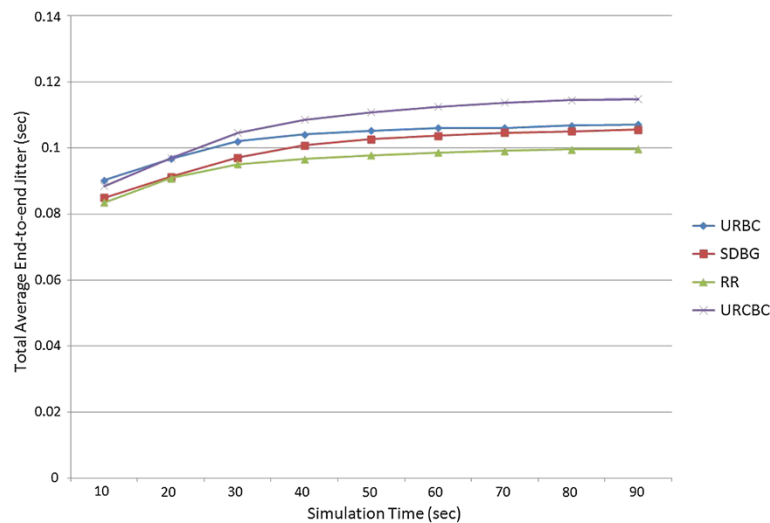


Figure 11 Total average end-to-end jitter for scenario 2.

results than the worst scheme, URBC throughout the simulation. The achievements of URCBC and RR are moderate among the four schemes, but URCBC superseded URBC to become the poorest performer in total average end-to-end delay at the point of 60 to 80 s with tiny variance, around 0.3% on average. URCBC turns to be better than URBC at the end of the simulation. It is observed that variance between URBC and URCBC is small especially after 50 s. This outcome indicates that the rule restriction by URCBC does not help much in latency, as opposed to the phenomenon found in total average end-to-end throughput. In contrast, SDBG is proving good results in both throughput and delay in all the simulation time. It is believed that the main reason of good achievements in latency is the cogitation steps done by the microengines in SDBG.

Figure 11 is the total average end-to-end jitter for scenario 2. RR scheme has the lowest readings for the whole simulation. This is because RR serves every active bandwidth request with equal amount of bandwidth regardless of any other factors. SDBG has the big differences with the URBC and URCBC schemes, but the gap is narrowing between SDBG and URBC as the simulation time passes. The variance starts from 5.91% at 10 s and keep on dropping until 1.45% at 90 s. Opposite to URBC, the variance among SDBG and URCBC becomes bigger till the end of the simulation time. SDBG attained a difference of 4.07% at 10 s, and it increased to 8.06% when the simulation was complete. Moreover, URCBC surpasses the URBC approach to have the highest jitter values since 20 s of the simulation time. These occurrences discover that the restriction rules enforced by URCBC caused long waiting time between one and other packets especially to the lower modulated SSs. This problem becomes worse when more and more packets have been enqueued as seen in Figure 11. Meanwhile, URBC, by its nature design, does not perform as bad as the URCBC. URBC design gives as many as bandwidth to the requester regardless its channel condition. Hence, lower modulated SSs are given the same service as the highest modulated SS. Due to this reason, the time difference between one packet and another packet is smaller. For SDBG, due to the combination on the decisions made by its three microengines, SDBG manages to grant the necessary bandwidth fairly among all the SSs even with different modulation schemes. This is once again proven that SDBG approach had taken the significant and precise decisions in bandwidth request and grant process.

6. Conclusions

In bandwidth request and granting process, bandwidth granting for the same service class or category is not a negligible issue. With a proper mechanism imposed on the

bandwidth granting process, it could improve the network performance. The proposed SDBG scheme not only manages to improve the network performance significantly but also does not create extra management message that will cause extra transmission. The SDBG utilizes the existing information and self-generated data in making bandwidth decision which are not giving burden to the entire network. The factors considered by SDBG are relevant and had assisted bandwidth request and grant manager to grant the bandwidth more efficiency and accurately. Extensive simulation results confirm that our scheme is able to reduce the delay and jitter for real-time applications besides the throughput.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Faculty of Information Science & Technology, Multimedia University, Jalan Ayer Keroh Lama, Bukit Beruang, Melaka 75450, Malaysia. ²Faculty of Engineering, Multimedia University, Persiaran Multimedia, Cyberjaya, Selangor 63100, Malaysia.

Received: 12 January 2013 Accepted: 13 May 2013

Published: 25 May 2013

References

1. IEEE Std 802.16-2009, *IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems* (IEEE, 2009)
2. IEEE Std 802.16-2001, *IEEE Std. 802.16-2001 IEEE, Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems* (IEEE, 2002)
3. XF Bai, A Shami, Y Ye, Robust QoS control for single carrier PMP mode IEEE 802.16 systems, *IEEE Trans. Mobile. Computing* **7**(4), 416–429 (2008)
4. R Bo, Q Yi, L Kejie, C Hsiao-Hwa, M Guizani, Call admission control optimization in WiMAX networks, *IEEE T Veh Technol* **57**(4), 2509–2522 (2008)
5. D Eden Ricardo, Walter Godoy Junior and Augusto Foronda, Scheduling mechanisms with call admission control (CAC) and an approach with guaranteed maximum delay for fixed WiMAX networks, in *Quality of Service and Resource Allocation in WiMAX* (InTech, Croatia, 2012), pp. 59–84
6. H Chia-Hsing, W Tsung-Ju, C Hann-Tzong et al., Bandwidth control protocol in WiMAX network, in *International Symposium on Parallel and Distributed Processing with Applications* (ISPA, Taipei, 6–9 Sept 2010), pp. 342–349
7. C Chih He, L Wanjiun, L Tehuang et al., Adaptive downlink/uplink bandwidth allocation in IEEE 802.16 (WiMAX) wireless networks: a cross-layer approach, in *IEEE Global Telecommunications Conference* (Washington, DC, 26–30 Nov 2007), pp. 4775–4779
8. C-H Chiang, W Liao, T Liu, I Chan, H-L Chao, Adaptive downlink and uplink channel split ratio determination for TCP-based best effort traffic in TDD-based WiMAX networks, *IEEE J Sel Area Comm* **27**(2), 182–190 (2009)
9. E-C Park, Efficient uplink bandwidth request with delay regulation for real-time service in mobile WiMAX networks, *IEEE T Mobile Comput* **8**(9), 1235–1249 (2009)
10. Q Ni, A Vinel, X Yang, A Turlikov, J Tao, Investigation of bandwidth request mechanisms under point-to-multipoint mode of WiMAX networks, *IEEE Commun Mag* **45**(5), 132–138 (2007)
11. Q Ni, H Ling, A Vinel, X Yang, M Hadjinicolaou, Performance analysis of contention based bandwidth request mechanisms in WiMAX networks, *IBM Syst J* **4**(4), 477–486 (2010)
12. KK Wee, SW Lee, Priority based bandwidth allocation scheme for WiMAX systems, in *2nd IEEE International Conference on Broadband Network & Multimedia Technology* (Beijing, 18–20 Oct 2009), pp. 15–18
13. Y Wang, S Chan, M Zukerman, RJ Harris, Priority-based fair scheduling for multimedia WiMAX uplink traffic, in *IEEE International Conference on Communications* (Beijing, 19–23 May 2008), pp. 301–305
14. LFM De Moraes, PD Maciel et al., Analysis and evaluation of a new MAC protocol for broadband wireless access, in *International Conference on*

Wireless Networks, Communications and Mobile Computing, vol. 1 (Hawaii, 13–16 June 2005), pp. 107–112

15. W Lilei, X Huimin et al., A new management strategy of service flow in IEEE 802.16 systems, in *3rd IEEE Conference on Industrial Electronics and Applications* (Singapore, 3–5 June 2008), pp. 1716–1719
16. I Khirwar, Y Anjulata, T Preeti, Comparative assessment of WiMAX scheduler in fixed and mobile WiMAX networks for VoIP using QualNet, in *International Conference on Computer and Communication Technology (ICCT)* (Allahabad, 17–19 Sept 2010), pp. 15–21
17. A Arash, W Tan Su, An enhanced cross layer downlink scheduling algorithm for IEEE 802. 16 networks, in *International Conference on Information Networking (ICOIN)* (Barcelona, 26–28 Jan 2011), pp. 212–217
18. W Nie, H Wang, N Xiong, Low-overhead uplink scheduling through load prediction for WiMAX real-time services. *IBM Commun* **5**(8), 1060–1067 (2011)
19. T-L Sheu, K-C Huang, Adaptive bandwidth allocation model for multiple traffic classes in IEEE 802.16 worldwide interoperability for microwave access networks. *IBM Commun* **5**(1), 90–98 (2011)
20. A Najah Abu, D Pratik, H Hossam et al., A performance study of uplink scheduling algorithms in point-to-multipoint WiMAX networks. *Computer Communications* **32**(3), 511–521 (2009)
21. X Yuan, Z Duan, Fair round-robin: a low-complexity packet scheduler with proportional and worst-case fairness. *IEEE T Comput* **58**(3), 365–379 (2009)
22. G Garroppo Rosario, G Stefano, I Davide et al., Radio-aware scheduler for WiMAX systems based on time-utility function and game theory, in *IEEE Global Telecommunications Conference* (Honolulu, 30 Nov–4 Dec 2009), pp. 1–6
23. H Chenn Jung, C Yi Ta, G Chih Tai, Y Dian Xiu, C You Jia et al., A self-adaptive bandwidth reservation scheme for 4g mobile WiMAX, in *International Workshop on Cross-Layer Design* (Jinan, 20–21 Sept 2007), pp. 55–59
24. T Sun Zhen, A Gani, Intelligent uplink bandwidth allocation based on PMP mode for WiMAX, in *International Conference on Computer Technology and Development*, vol. 1 (Kota Kinabalu, 13–15 Nov 2009), pp. 86–90
25. L Qingwen, W Xin, GB Giannakis, A cross-layer scheduling algorithm with QoS support in wireless networks. *IEEE T Veh Tech* **55**(3), 839–847 (2006)
26. L Jia Ming, C Jen Jee, W You Chiun, T Yu Chee, A cross-layer framework for overhead reduction, traffic scheduling, and burst allocation in IEEE 802.16 OFDMA networks. *IEEE T Veh Tech* **60**(4), 1740–1755 (2011)
27. J She, Y Xiang, H Pin Han, Y En Hui, A cross-layer design framework for robust IPTV services over IEEE 802. 16 networks. *IEEE J Sel Area Comm* **27**(2), 235–245 (2009)
28. The Qualnet Network Simulation Tool. <http://www.qualnet.com>. Accessed 08 Jul 2011
29. A Mohammed Sabri, Comparative study of scheduling algorithms in WiMAX. *Int J Sci Eng Res* **2**(2), 1–7 (2011)
30. R Mitul, KH Khandhedia, DN Wandra, N Khandhar, R Khandhedia, Performance evaluation of WiMAX network using Qualnet simulator. *AKGEC Journal of Technology* **1**(2), 10–13 (2010)
31. Symmetry MX CBS5000 Tier 1 Knowledge Course, Issue 3, Module 7: Capacity, Lesson 2, SR Telecom & Co. S.E.C. , 381–386 (2008)

doi:10.1186/1687-1499-2013-135

Cite this article as: Wee et al.: Discretionary bandwidth granting scheme for homogenous real-time applications. *EURASIP Journal on Wireless Communications and Networking* 2013 **2013**:135.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com