Nguyen et al. EURASIP Journal on Wireless Communications and Networking 2014, **2014**:104
http://jwcn.eurasipjournals.com/content/2014/1/104

EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

QoS-aware dynamic resource allocation for wireless broadband access networks

Tri M Nguyen, Taihyung Yim, Youchan Jeon, Yeunwoong Kyung and Jinwoo Park*

Abstract

The development of the advanced wireless access technologies is focusing on the enhancement of mobile user satisfaction in terms of quality of service (QoS). As the number of mobile users increases, the amount of traffic passing through base station (BS) significantly increases so that the preplanned capacity of downlink or uplink can be exceeded from time to time resulting in the degradation of users' QoS satisfaction. A way to utilize the downlink (DL) and the uplink (UL) of an orthogonal frequency division multiple access frame efficiently is therefore needed subject to users' bandwidth demand which changes continuously over the time. In this paper, we propose a QoS-aware dynamic resource allocation (QDRA) scheme which dynamically adjusts the DL/UL ratio to allocate bandwidth for QoS support to users based on the usage statistics. The performance of the proposed QDRA scheme was examined by applying to the WiMAX standard specifications, to show its superiority of maintaining a higher QoS support for various classes of services compared to the pre-reported adaptive method.

Keywords: Packet scheduling; Resource allocation; QoS; OFDMA; WiMAX

1 Introduction

The advanced wireless access technologies such as long-term evolution (LTE) and Worldwide Interoperability for Microwave Access (WiMAX), so called the 4th generation wireless technologies, have begun to be deployed over the world to efficiently support various new services and features such as multimedia services with high data rates and wide coverage area, as well as all-IP with security and QoS support [1-3].

With the wide spread of the broadband wireless networks, the amount of traffic transmitted in data networks grows extremely fast. According to Cisco [4], it has grown eight times during the last 5 years and will grow four times more until 2015, mainly because of an increase in mobile traffic which has grown 26 times between 2010 and 2015. Falaki et al. [5] started to investigate smartphone traffic from the aspect of the ratio of downlink (DL) and uplink (UL) traffic. A wide variation among users is manifested in the work, which is likely caused by diversity in application usage, revealing that the DL traffic becomes ten times greater than the UL traffic. In addition, the studies from [6,7] also pointed out that DL traffic and UL traffic change dramatically over the time even in a day. It leads to a necessity of dynamic regulation between DL traffic and UL traffic to improve the overall system throughput.

To support a variety of multimedia applications, for example, WiMAX standard defines five types of service flow, each of which should satisfy different QoS requirements such as minimum throughput requirement and delay/jitter constraints. Therefore, a scheduling algorithm is a central part in guaranteeing the QoS for multimedia services (both real-time and non-real-time services) while efficiently utilizing the available bandwidth. Several scheduling algorithms have been proposed for QoS-supported bandwidth allocation in WiMAX [8-12]. Although the previous algorithms were designed for QoS satisfaction, they did not take into account the operational characteristics of the mobile WiMAX such as varying channel conditions in time as well as the dramatic change between DL traffic and UL traffic during a service. Such the variation of operational conditions have been found important in maintaining the users' QoS satisfaction thru many practical WiMAX deployment experiences, because a small portion of bandwidth degradation on non-real-time data connections may result in unnoticeable perceived QoS change on end-users, but the same

*Correspondence: jwpark@korea.ac.kr
School of Electrical Engineering, Korea University, Seoul, 136-713, Korea

bandwidth degradation on real-time multimedia connection may cause the connection to be dropped resulting in the serious QoS degradation.

In this paper, we propose a QoS-aware dynamic resource allocation scheme called QDRA which takes into account regulating DL traffic and UL traffic, allocating bandwidth to users based on QoS requirements as well as the wireless channel conditions. QDRA consists of two phases, the first of which dynamically adjusts the DL/UL subframe ratio based on the requested DL and UL bandwidths at the BS, and in the second phase, the DL and UL schedulers independently allocate their own bandwidth to the individual DL and UL users based on QoS requirements and data in queues. In order to evaluate the proposed QDRA scheme, we applied it to the WiMAX standard specifications making the analytic observations on the performance improvement. The rest of this paper is organized as follows. Section 2 discusses the literature on mobile WiMAX and some related works. Section 3 describes our proposed scheme. Section 4 shows simulation results. Finally, we make a conclusion in Section 5.

2 Background

2.1 Mobile WiMAX

Orthogonal frequency division multiple access (OFDMA) is used in almost wireless broadband access technologies including mobile WiMAX. Mobile WiMAX supports both frequency division duplexing (FDD) and time division duplexing (TDD). Figure 1 shows the WiMAX OFDMA TDD frame structure. The WiMAX frame consists of one DL subframe and one UL subframe. The DL subframe and UL subframe are separated by a TTG (transmit/receive transition gap) and RTG (receive/transmit transition gap). The frames are shown in two dimensions with frequency along the vertical axis and time along the horizontal axis. A symbol is the smallest allocation unit in the time domain, and a subchannel is the smallest logical allocation unit in the frequency domain. A slot is the minimal possible bandwidth allocation unit defined in frequency and time domain. It consists of one subchannel and one to three symbols. A preamble is used for time synchronization which accounts for one symbol. The downlink map (DL-MAP) and uplink map (UL-MAP) define the burst-start time and burst-end time, modulation types and forward error control (FEC) for each subscribe station (SS). Frame control header (FCH) defines these MAP's lengths and usable subcarriers. The SS allocation is in term of bursts. In the figure, there is one burst per SS.

Mobile WiMAX uses the request-grant mechanism in MAC layer for the bandwidth allocation. The BS allocates bandwidths and broadcasts the data packets to all SSs on the DL. Conversely, SS needs to send a bandwidth request to BS. There are two methods for bandwidth request including contention-based and contention-free. In the contention-based mode, the SS sends a bandwidth request to BS based on backoff algorithm during contention periods. The BS will poll each SS for bandwidth request in the contention-free mode. Based on bandwidth requests from SSs, the BS will allocate bandwidth to SSs.

There are five standardized QoS classes in mobile WiMAX including unsolicited grant service (UGS), extended real-time polling service (ertPS), real-time polling service (rtPS), non-real time polling service (nrtPS), and best effort (BE) [2]. In UGS, the BS allocates fixed-size grants periodically; UGS connections do not need to send any bandwidth requests. In rtPS, the BS

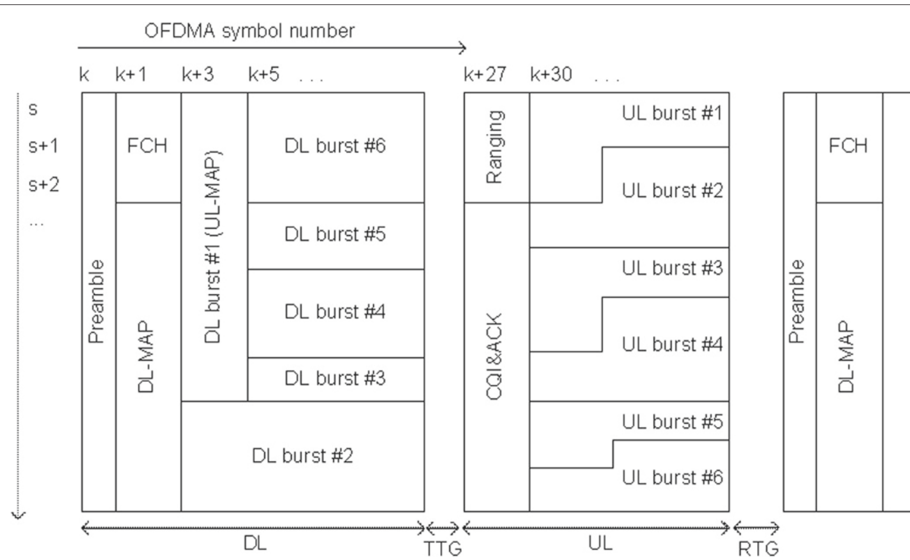

Figure 1 Mobile WiMAX OFDMA frame structure.

periodically polls the SS by granting one slot for sending a bandwidth request, while the goal of ertPS is to combine the advantages of UGS and rtPS. The two remaining QoS classes are aimed to non-real time traffic. NrtPS is similar to rtPS except that connections are polled less frequently and they can use contention request opportunities. BE connections can only receive resources through contention.

2.2 Related works

Chiang et al. presented adaptive DL and UL channel split ratio determination in [13]. In this paper, the authors focus on Transmission Control Protocol (TCP)-based traffic and investigate the impact of improper bandwidth allocation to DL and UL channels on the performance of TCP. They proposed an adaptive split ratio scheme which adjusts the bandwidth ratio of DL to UL adaptively in order to maximize the aggregate throughput of TCP-based traffics.

Rastin Pries et al. [14] focused on the performance of the IEEE 802.16 TDD mode in rural areas with only one cell. They compared different settings for the TDD split and evaluated their strengths and weaknesses for several traffic profiles. They also proposed an algorithm for a dynamic setting of this ratio in a single cell scenario, depending on the current load condition. Adhicandra proposed an adaptive subframe allocation algorithm allowing the UL subframe to borrow resources from the DL subframe [15]. However, in order to adjust the DL/UL ratio, both studies are only based on reports of data usages in DL subframe and UL subframe from previous OFDMA frames without taking into account the overall DL and UL traffic.

Sarigiannidis et al. [16] introduced a mapping scheme for IEEE 802.16 applying horizon mapping and proposed an adaptive prediction-based scheme that is able to adjust the DL subframe capacity. QoS support is not considered in these papers.

In one of the early studies on QoS support in WiMAX networks, Wongthavarawat and Ganz introduced a two-layer scheduling algorithm (priority scheme) for bandwidth allocation [8]. This algorithm allocates the bandwidth to each service class by strict priority in the first layer, and each service class has its own scheduling algorithm in the second layer. The UGS uses fixed bandwidth based on its bandwidth requirement, rtPS uses early deadline first, nrtPS uses weighted fair queue (WFQ), and BE uses round robin (RR). The disadvantage of the strict priority is higher priority classes may starve bandwidth for lower priority classes.

Cicconetti et al. [9] focus on the available QoS support mechanisms in the MAC sublayer and evaluate their effectiveness through simulation. They conduct the performance evaluation based on two common application scenarios conceived by the WiMAX Forum: residential and small to medium-size enterprises (SME). The test case uses 7 MHz channel bandwidth with carrier frequency between 2 and 11 GHz and operating in FDD mode.

Freitag and Da Fonseca [10] proposed an UL scheduling mechanism for IEEE 802.16 networks. They consider QoS requirements in scheduling decisions. The deficit round robin (DRR) and WFQ algorithm are employed for bandwidth allocation. However, the methods provided by these authors do not consider the varying channel quality received by each SS.

Amir and Nasser [11] proposed a bandwidth allocation framework to all classes of service defined by the standard with different QoS requirements. They employ a two-level packet scheduling scheme with a call admission control policy and dynamic bandwidth allocation mechanism. However, they do not consider the dynamic bandwidth allocation between the DL subframe and the UL subframe.

3 QoS-aware dynamic resource allocation

Figure 2 shows a procedure of the proposed QoS-aware dynamic resource allocation scheme. The proposed scheme includes two phases. In the first phase, it dynamically adjusts the DL/UL subframe ratio based on the incoming traffic at BS (DL data traffic and UL request). In the second phase, the DL and UL will dynamically allocate bandwidth to users based on their priority and QoS requirements.

3.1 Phase 1: DL/UL subframe allocation

Since a fixed subframe split ratio of DL/UL may cause DL or UL to be overloaded when diverse multimedia service traffics are severe, the algorithm adjusting dynamically the DL/UL ratio based on the amount of the DL data queues and the UL request queues is needed to be employed, as explained as follows.

Let DL_symbols and UL_symbols be the number of symbols occupied in each DL subframe and UL subframe, respectively. When the total number of symbols for a WiMAX frame is a constant L, we can obtain

$$\text{DL\_symbols} + \text{UL\_symbols} = L \tag{1}$$

For a 10- MHz channel, for example, there are 48.6 symbols in a 5 ms frame when the OFDMA symbol time is 102.8 μs. Of these, 1.6 symbols are used for TTG and RTG leaving 47 symbols for L.

In the partially used subchannelization (PUSC) permutation which is typically used in a mobile wireless environment [17], each DL slot and UL slot consists of 28 subcarriers over two symbol times and 24 subcarriers over three symbols times [3], respectively. So that DL_symbols should have the form of $2k+1$, and UL_symbols should have the form of $3h$. For a 10-MHz channel with the

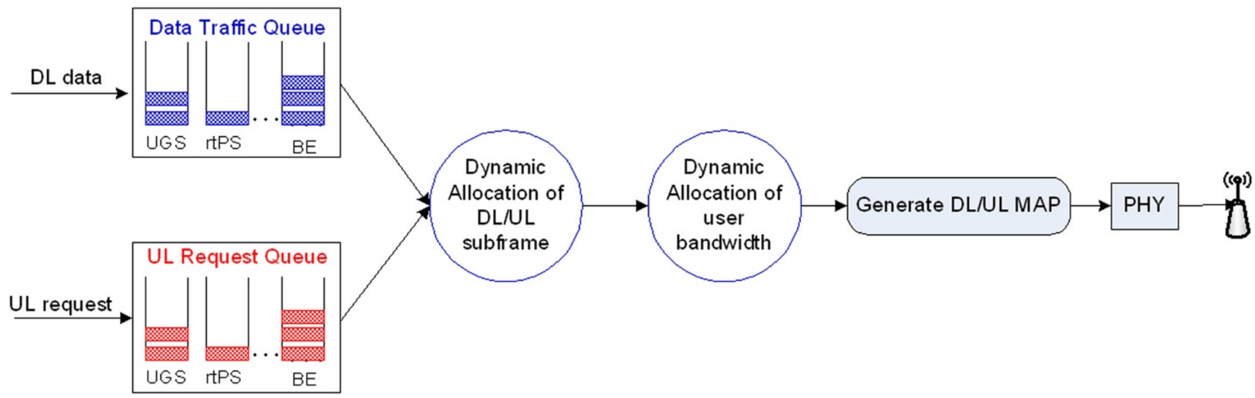

Figure 2 Procedure of QoS-aware dynamic resource allocation scheme.

DL/UL ratio of 2:1, DL_symbols and UL_symbols are 29 and 18, respectively. At 10-MHz channel, the number of used subcarriers for DL and UL is 840. In this case, the number of DL slots (DL_slots) is 14 × 30, or 420 slots. The number of UL slots (UL_slots) 6 × 35, or 210 slots. By considering these conditions, we have

$$\text{DL\_slots} = \alpha \times \text{DL\_symbols} \quad (2)$$

$$\text{UL\_slots} = \beta \times \text{UL\_symbols} \quad (3)$$

At 10 MHz channel, $\alpha = 15$ and $\beta = \frac{35}{3}$.

In addition, DL subframe and the UL subframe are described by the following equations:

$$\text{DL\_slots} = H_{\text{DL}} + \text{DL\_bursts\_slots} \quad (4)$$

$$\text{UL\_slots} = H_{\text{UL}} + \text{UL\_bursts\_slots} \quad (5)$$

Here H_{DL} and H_{UL} are the head parts of the DL subframe and UL subframe, respectively. The head part of the DL subframe consists of Preamble, FCH (frame control header), DL_MAP, and UL_MAP. Preamble and FCH are predefined in the technical standards, and DL_MAP and UL_MAP indicates the frame organization. The head part of UL subframe consists of CQI (channel quality indicator) and ranging, which are predefined in the technical standards. DL_ bursts_ slots is the total number of slots used for all DL bursts except the first DL burst (UL_MAP), and UL_ bursts_ slots is the total number of slots used for all UL bursts. Data expressed in DL_ bursts_ slots consist of DL UGS traffic (UGS_{DL}) and non-UGS traffic, called DL_bursts_slots (non-UGS), and data expressed in UL_ bursts_ slots consist of UL UGS traffic (UGS_{UL}) and non-UGS traffic called UL_ bursts_ slots (non-UGS). They are rewritten as

$$\text{DL\_bursts\_slots} = \text{UGS}_{\text{DL}} + \text{DL\_bursts\_slots (non-UGS)} \quad (6)$$

$$\text{UL\_bursts\_slots} = \text{UGS}_{\text{UL}} + \text{UL\_bursts\_slots (non-UGS)} \quad (7)$$

$$\text{DL\_bursts\_slots (non-UGS)} = \sum_{i=1, ci \neq \text{UGS}}^{n} \left\lceil \frac{d_{ci}}{\text{DS}_{ci}} \right\rceil \quad (8)$$

$$\text{UL\_bursts\_slots (non-UGS)} = \sum_{j=1, cj \neq \text{UGS}}^{m} \left\lceil \frac{u_{cj}}{\text{DS}_{cj}} \right\rceil \quad (9)$$

where d_{ci} and u_{cj} are the total size of the packets in the DL connection *ci* and the UL connection *cj*, respectively, which will be scheduled to be sent in this frame, DS_{ci} and US_{cj} are the slot size (bytes) depending on the modulation and coding scheme (MCS) employed for DL connection *ci* and UL connection *cj*, respectively (Table 1 shows the number of bytes per slot for various MCS values). For each MCS, the number of bytes is calculated as [Number bits per symbols × Coding rate × 48 data subcarriers and symbols per slot)/8 b] [2]), *n* is the number of DL bursts except UL-MAP, and *m* is the number of UL bursts.

In order to achieve efficient utilizations for DL subframe and UL subframe, the queue size of the DL and UL should be reflected on the DL/UL ratio. In this paper, therefore, we propose that the ratio of DL_ data_ busts (non-UGS)

Table 1 Slot capacity for different MCSs

MCS	Bits per symbol	Coding rate	Bytes per slot
QPSK 1/8	2	0.125	1.5
QPSK 1/4	2	0.25	3.0
QPSK 1/2	2	0.50	6.0
QPSK 3/4	2	0.75	9.0
QAM-16 1/2	4	0.50	12.0
QAM-16 2/3	4	0.67	16.0
QAM-16 3/4	4	0.75	18.0
QAM-64 1/2	6	0.50	18.0
QAM-64 2/3	6	0.67	24.0
QAM-64 3/4	6	0.75	27.0
QAM-64 5/6	6	0.83	30.0

and UL_ data_ busts (non-UGS) is set according to the ratio of DL queue size and UL queue size except UGS traffic.

The DL and UL queue size can be calculated as

$$Q_{DL} = \sum_{i=1, ci \neq UGS}^{N} \lceil \frac{d_{ci}}{DS_{ci}} \rceil \tag{10}$$

$$Q_{UL} = \sum_{i=1, cj \neq UGS}^{M} \lceil \frac{u_{cj}}{DS_{cj}} \rceil \tag{11}$$

Here, N and M are the total number of DL connections and UL connections, respectively.

In order to avoid the constant change of DL subframe and UL subframe, the average of the DL queue size and UL queue size during the last X frames should be considered as follows:

$$\overline{Q}_{DL} = \frac{\sum_{i=1}^{X} Q_{DL}(i)}{X} \tag{12}$$

$$\overline{Q}_{UL} = \frac{\sum_{i=1}^{X} Q_{UL}(i)}{X} \tag{13}$$

It is clear that $\overline{Q}_{DL}$ and $\overline{Q}_{UL}$ are both able to be calculated.

We have

$$P = \frac{\overline{Q}_{DL}}{\overline{Q}_{UL}} \tag{14}$$

From (1) to (7) and (14), we obtain:

$$UL\_slots = \frac{L\alpha\beta - (H_{DL} + UGS_{DL})\,\beta + (H_{UL} + UGS_{UL})\,P\beta}{P\beta + \alpha} \tag{15}$$

$$DL\_slots = \frac{L\alpha\beta - \alpha UL\_slots}{\beta} \tag{16}$$

3.2 Phase 2: priority-based user bandwidth allocation

After calculating the DL/UL subframe ratio, the DL subframe size and UL subframe size are defined. They are reflecting the proportion of the DL traffic to UL traffic.

Figure 3 shows the hierarchical structure of bandwidth allocation for UL. Assuming the bandwidth allocation mechanism for DL can be adopted similarly in the UL traffic management, we discuss the bandwidth allocation mechanism for UL in this subsection and then the resulting performance for both UL and DL will be discussed in the Subsection 4.

In the first layer, 2-Round Priority-based algorithm is used for allocating bandwidth among traffic classes, and algorithms allocating bandwidth among users are applied in the second layer.

A strict priority algorithm distributes the bandwidth among service classes following strict class priority [8], from highest to lowest (UGS, ertPS, rtPS, nrtPS, BE). After all users in high-priority class have been served, users in low-priority class are scheduled for transmission. The

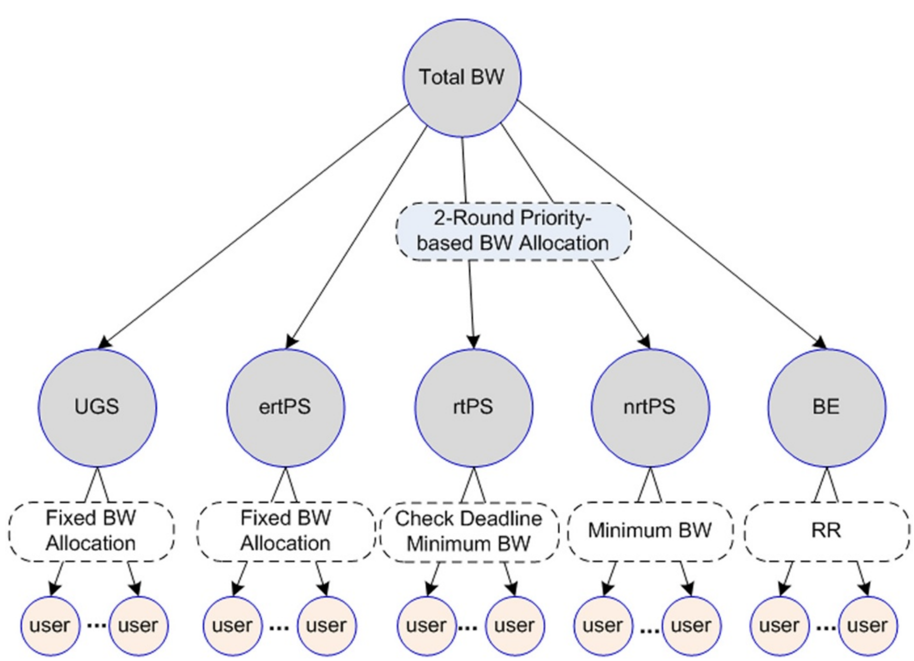

Figure 3 Bandwidth allocation structure.

disadvantage of the strict priority algorithm is that higher-priority classes may starve bandwidth for lower-priority classes.

We propose a modification of strict priority (2-Round Priority-based algorithm) based on the principles: scheduling to send the packets with stringent deadlines in order to meet 'maximum latency' and trying to reserve a minimum bandwidth to users having minimum reserved traffic rate (MRTR).

A flow chart for the algorithm is illustrated in Figure 4:

Step 1: *Allocate bandwidth (BW) to UGS and ertPS.*
UGS and ertPS have fixed bandwidth rates. They require constant bandwidth allocations. The scheduler will allocate fixed bandwidth to UGS traffic classes and then ertPS traffic classes.

Step 2: *Check packets' deadline*
For rtPS traffic, the two important QoS parameters are maximum latency and MRTR. In order to meet the QoS parameter, maximum latency, the earliest deadline first (EDF) algorithm is employed. Packets with the stringent deadlines will be scheduled first. The scheduler determines the packet's deadline based on its arrival time and maximum latency.
The packet's deadline t_{deadline} is its arrival time t_{arrive} plus its maximum delay $t_{\text{delay}}^{\text{max}}$:

$$t_{\text{deadline}} = t_{\text{arrive}} + t_{\text{delay}}^{\text{max}} \tag{17}$$

The packet's timestamp in the queue is denoted as

$$t_{\text{TS}} = t_{\text{arrive}} + t_{\text{delay}}^{\text{queue}} \tag{18}$$

where $t_{\text{delay}}^{\text{queue}}$ is the packet delay time in the queue. The time difference of a packet between the required deadline and timestamp:

$$\Delta t = t_{\text{deadline}} - t_{\text{TS}} \tag{19}$$

If $\Delta t \leq 2f$ (f is the frame length, 5 ms), the packet should be scheduled to send in this frame.

Step 3: *Assign minimum bandwidth to rtPS and nrtPS.*
Suppose B_i is the bandwidth requirement of the

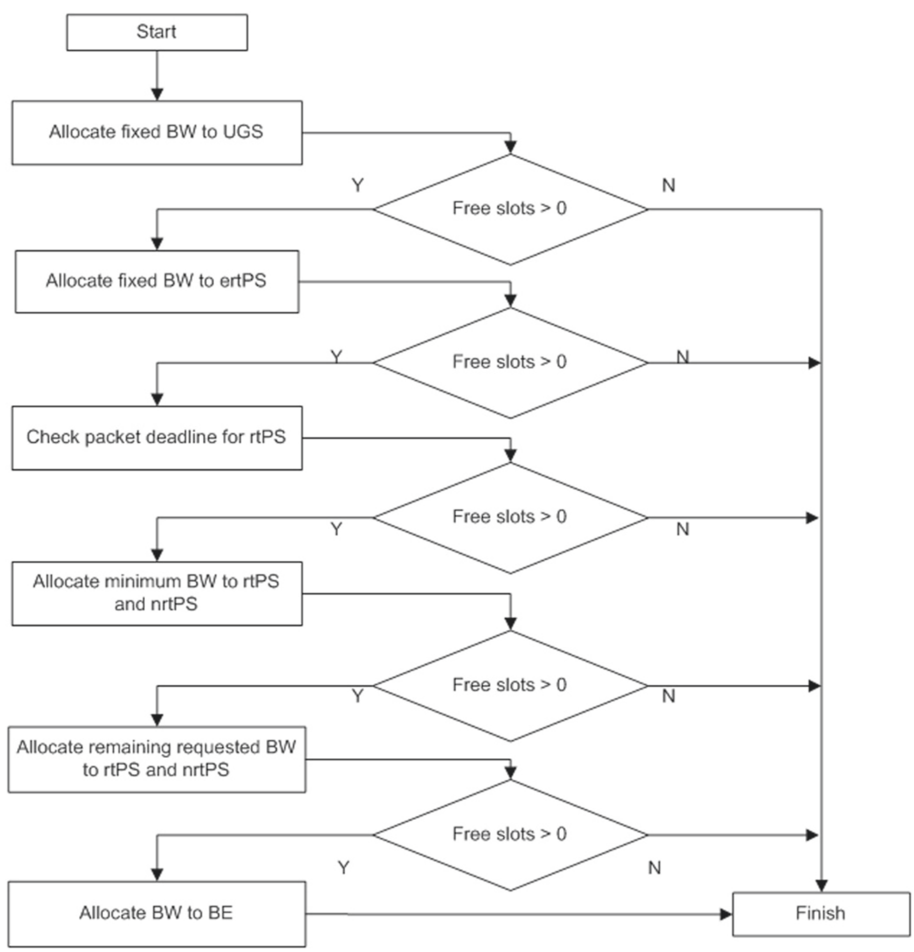

Figure 4 The 2-Round Priority-based algorithm flow chart.

ith user, S_i stands for the slot size (the number of bytes a MS can send in one slot given a specified MCS), FPS is the number of frames the BS sends per second (200 for the WiMAX frame length is 5 ms). The number of slots for the ith user within one WiMAX frame can be calculated as

$$N_i = \frac{B_i}{S_i * \text{FPS}} \quad (20)$$

Based on the minimum bandwidth requirement of a user, we can calculate the minimum slots for users in one frame.

Step 4: *Allocate remaining requested bandwidth to rtPS, nrtPS.*
The free slots is calculated by

$$F^{\text{free}} \leftarrow F - N_{\text{UGS}} - N_{\text{ertPS}} - N^{\min}_{\text{rtPS}} - N^{\min}_{\text{nrtPS}} \quad (21)$$

Here, F is the total UL slots, N_{UGS} and N_{ertPS} are the slots assigned to UGS and ertPS, respectively, $N^{\min}_{\text{rtPS}}$ and $N^{\min}_{\text{nrtPS}}$ are the minimum bandwidth assigned to rtPS and nrtPS, respectively.
The free slots F^{free} are allocated to users in rtPS and then to users in nrtPS, respectively.
The procedure is described as follows:
Suppose that we need to allocate free slots to user $i \in A$, a set with users for bandwidth assigned

$$K_i = S_i \frac{1}{\sum_{\forall j \in A} S_j} \quad (22)$$

$$N^{\text{add}}_i = \min\{N^{\text{req}}_i - N^{\min}_i, K_i * F^{\text{free}}\} \quad (23)$$

$$N_i = N^{\min}_i + N^{\text{add}}_i \quad (24)$$

$$F^{\text{free}} \leftarrow F^{\text{free}} - N^{\text{add}}_i \quad (25)$$

Here, S_i is the slot size (bytes) given MCS of user i (Table 1). $N^{\min}_i$, N^{req}_i, and N^{add}_i are the minimum reserved slots, requested slots, and adding slots for user i, respectively.
This procedure works according to the weighted round robin (WRR) principle in which it considers the user channel condition.

Step 5: *Allocate the free slots to BE.*
WRR is employed to allocating free slots to BE users, and the free slots is defined as

$$F^{\text{free}} \leftarrow F - N_{\text{UGS}} - N_{\text{ertPS}} - N^{\min}_{\text{rtPS}} - N^{\min}_{\text{nrtPS}} - N^{\text{add}}_{\text{rtPS}} - N^{\text{add}}_{\text{nrtPS}} \quad (26)$$

$$K_i = S_i \frac{1}{\sum_{\forall j \in A} S_j} \quad (27)$$

$$N^{\text{allo}}_i = \min\{N^{\text{req}}_i, K_i * F^{\text{free}}\} \quad (28)$$

$$N_i = N^{\min}_i + N^{\text{allo}}_i \quad (29)$$

$$F^{\text{free}} \leftarrow F^{\text{free}} - N^{\text{add}}_i \quad (30)$$

Here, $N^{\text{add}}_{\text{rtPS}}$ and $N^{\text{add}}_{\text{nrtPS}}$ are the adding slots assigned to rtPS and nrtPS, respectively.

Table 2 Simulation settings

Parameter	Values
Duplexing	TDD
System bandwidth (MHz)	10
Initial DL/UL boundary	2:1
Frame duration (ms)	5
Symbol duration (us)	102.86
Type of mapping	PUSC
SS starting interval	Uniform (0,10)
X (Equations 12,13)	5

4 Simulation results

We use OPNET modeler for simulation [18]. The simulation settings are shown in Table 2. It is noted that the vertical bars on the curves in the figures are the 95% confidence intervals.

4.1 Efficiency of the DL/UL subframe allocation

In order to evaluate the efficiency of our proposed DL/UL subframe allocation scheme, we investigate the performance of the proposed scheme and compared the results with the ones of a fixed DL/UL ratio scheme and the adaptive subframe allocation scheme [14]. For the simulation experiments, File Transfer Protocol (FTP) traffic is taken into consideration, and the initial DL/UL ratio setting in the simulation is assumed 2:1. The traffic parameters for FTP we used for the simulations are shown in Table 3.

Figure 5 shows the throughput of QDRA, adaptive scheme, and fixed scheme, respectively. As the number of FTP clients is smaller than 17, both the DL and UL

Table 3 Traffic model parameters

Service class	Packet inter-arrival time	Packet size
UGS (VoIP)	Constant (3.65 ms)	Constant 640 B
ertPS (VoIP with silence suppression)	Constant (3.65 ms)	Constant 640 B
rtPS (video)	Exponential (mean 3 ms)	Exponential (mean 512 B)
nrtPS (FTP)	Exponential (mean 8 ms)	Exponential (mean 1024 B)
BE	Exponential (mean 5 ms)	Exponential (mean 512 B)

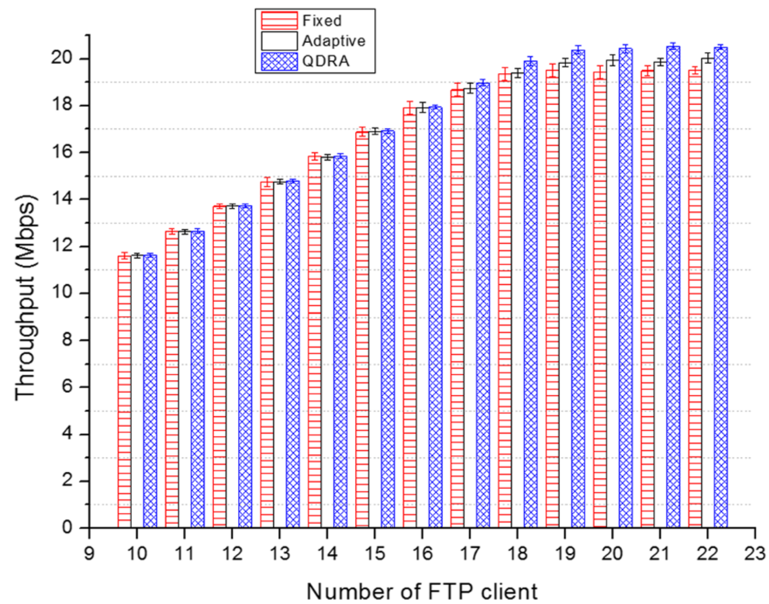

Figure 5 Throughput achieved by fixed, adaptive, and QDRA.

subframe utilizations are below 100%. As the results, all three schemes show identical performances. However, as the number of SSs reaches 17, the DL subframe utilization or UL subframe utilization may approach to its maximum. In case of the fixed scheme, the UL subframe utilization reached its maximum, and as a result, the UL mean delay increased dramatically. From Figure 5, it shows that our proposed scheme outperformed the adaptive and fixed scheme while the fixed scheme showed worse performance. As the number of FTP clients reaches 19, throughputs achieved by QDRA, the adaptive scheme, and fixed scheme are 20.4, 19.8, and 19.5 Mbps, respectively. It is found that QDRA improves 5% of the throughput comparing to the adaptive scheme. As the number of FTP client is from 20, throughputs achieved by all three schemes remain unchanged.

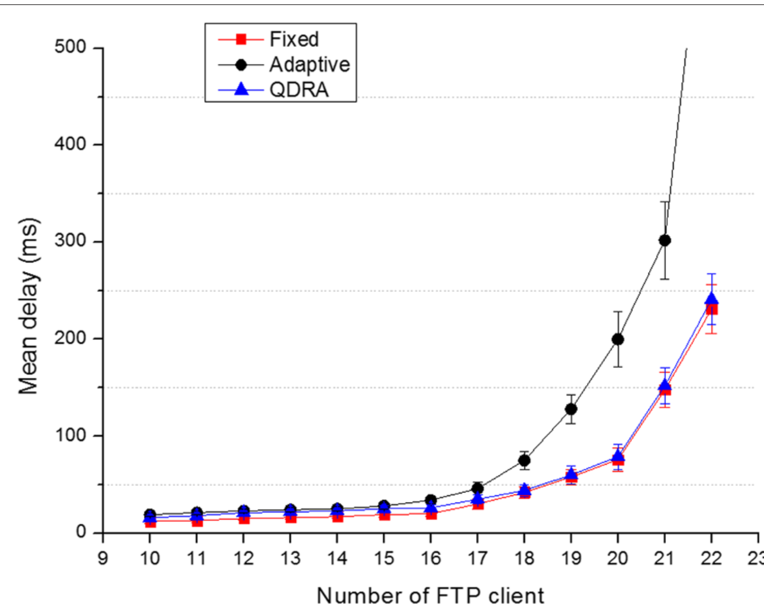

Figure 6 DL mean delay vs. number of FTP clients.

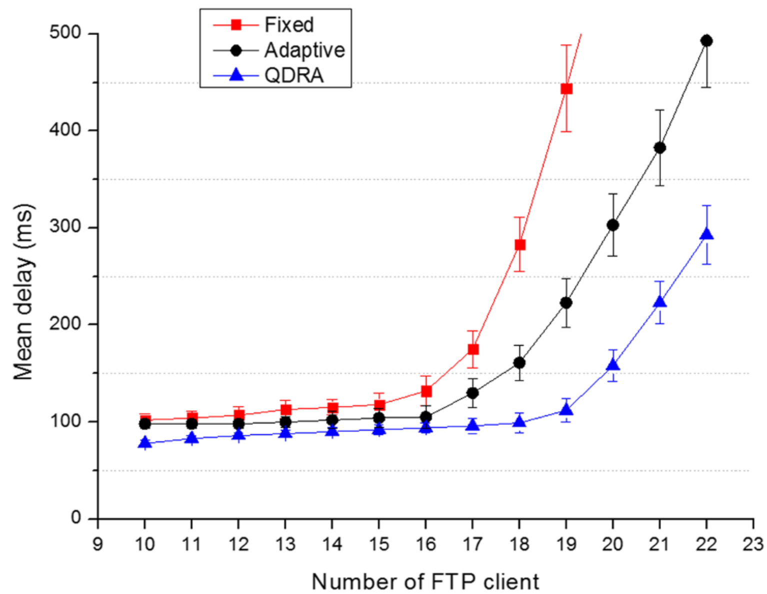

Figure 7 UL mean delay vs. number of FTP clients.

Figure 6 shows DL mean delays achieved by three schemes. The adaptive scheme shows the worst performance, and the fixed scheme performs slightly better than QDRA. As the number of FTP clients reaches 18, DL mean delay in the adaptive scheme increases dramatically. On the other hand, the fixed scheme and QDRA started increasing significantly when the number of FTP clients reaches 21. It can reveal that QDRA regulates DL bandwidth better than the adaptive scheme.

UL mean delays achieved by the three schemes are shown in Figure 7. QDRA shows the best performance following the adaptive and fixed schemes, respectively. As the number of FTP clients is from 17, UL mean delay of the fixed scheme rises dramatically. As the number of FTP clients increases, UL traffic requests are also increased when the UL subframe size is set to a fixed value in the fixed scheme. Consequently, UL mean delay in the scheme is increased dramatically but UL mean delays in the adaptive and QDRA are increased slightly due to the two schemes having regulated the DL and UL traffic. However, the adaptive scheme adjusts the DL/UL subframe ratio based on the observation of the last WiMAX frame while QDRA dynamic allocates DL and UL subframe based on the DL traffic and the UL traffic requests at the BS (Equations 15 and 16). Figures 6 and 7 expose that QDRA performance is better than the adaptive scheme.

4.2 Performance comparison of QDRA with the priority scheme

QoS provisioning capability of the proposed QDRA is discussed in the section. In order to measure the QoS performance, the five types of traffic classes specified in the WiMAX standard are considered. The applications and QoS requirements for the traffic classes are defined in Table 4, and the traffic model parameters for simulations are described in Table 3. The performance of the proposed QDRA scheme is compared with one of the priority scheme [8].

Figure 8a,b illustrates UL throughput achieved by all traffic classes using the priority scheme and QDRA, respectively. When the number of UGS flows is smaller than six, the UL throughputs achieved by all traffic classes using the priority scheme and QDRA are almost identical. All the traffic classes attained the bandwidth requirements described in Table 3. The UL throughputs achieved by ertPS, rtPS, and nrtPS traffic classes using both schemes

Table 4 Applications and QoS requirements

Service class	Class 1	Class 2	Class 3	Class 4	Class 5
Traffic type	UGS	ertPS	rtPS	nrtPS	BE
Application	VoIP	VoIP with silent suppression	Video	FTP	Web browsing
Bandwidth	1.4 Mbps	1.4 Mbps	1.2 Mbps	1.0 Mbps	0.8 Mbps
Delay tolerance	30 ms	50 ms	50 ms	200 ms	500 ms

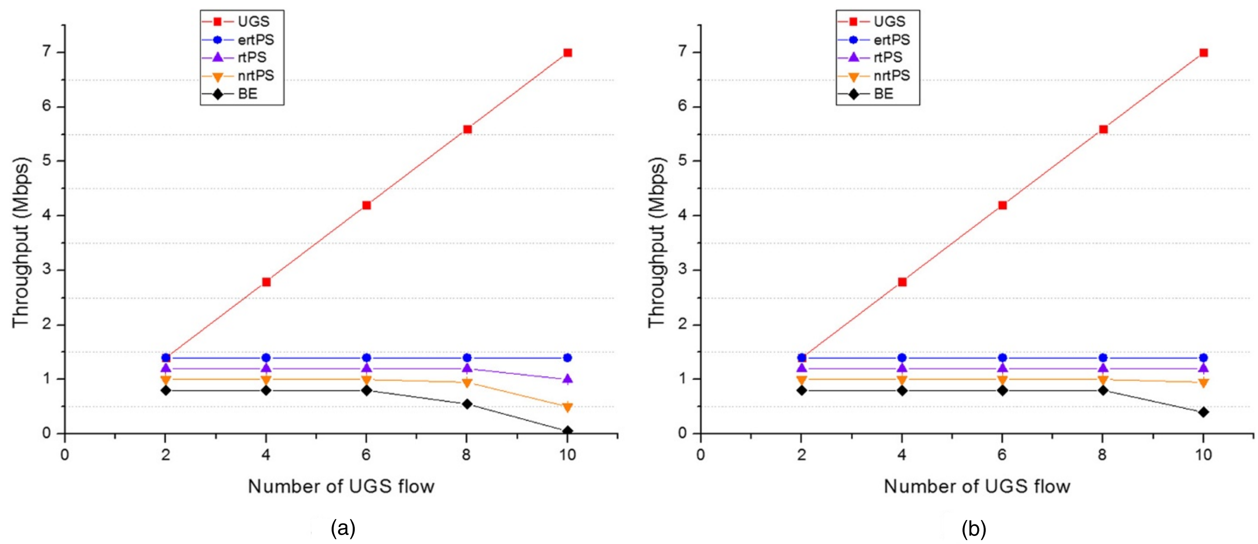

Figure 8 UL throughput achieved by priority scheme (a) and QDRA (b).

are approximately 1.44, 1.2, and 1.0 Mbps, respectively. These results satisfy the QoS requirements for bandwidth allocation. However, as the number of UGS flows is from 8, the nrtPS and BE throughputs of the priority scheme are reduced significantly. These results show that QDRA can satisfy bandwidth requirements for all traffic classes with minimum bandwidth requirements while it also provides a better performance for nrtPS and BE traffic classes.

Figure 9a,b shows UL mean delay achieved by all traffic classes using the priority scheme and QDRA, respectively. For the priority scheme, the UL mean delays of nrtPS and BE traffic classes increase sharply as the number of UGS flows increases. This is due to the downgrade of UL throughput of nrtPS and BE traffics. In addition, as the number of UGS flows reaches to 10, the rtPS traffic class violates its delay requirement. On the contrary, QDRA keeps rtPS delays around 50 ms. This is due to all real-time packets are being checked their deadline in scheduling algorithm in order to meet the delay constraint for the rtPS service class.

Figure 10a,b shows DL throughput achieved for all traffic classes using the priority scheme and QDRA, respectively. In the priority scheme, the DL/UL ratio is set to 2:1. As the number of UGS flows of both DL and

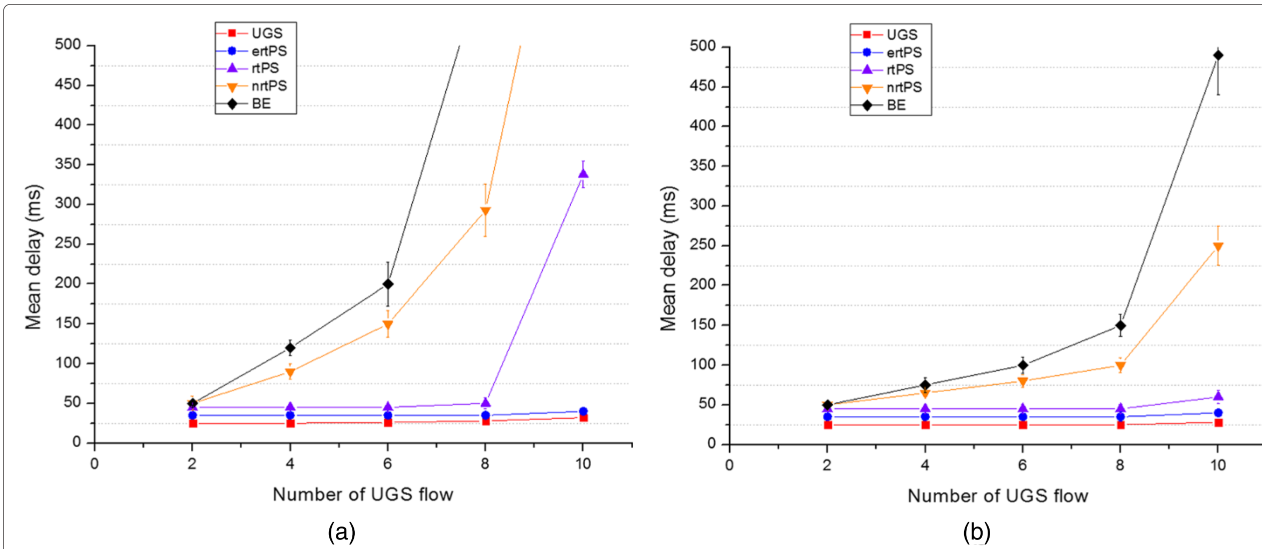

Figure 9 UL mean delay achieved by priority scheme (a) and QDRA (b).

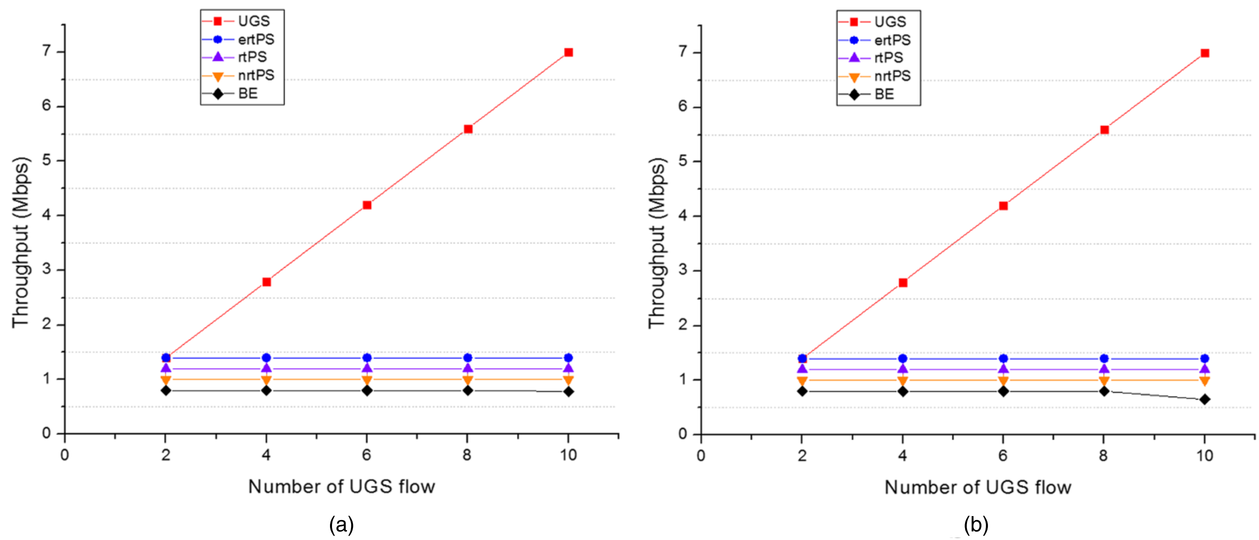

Figure 10 DL throughput achieved by priority scheme (a) and QDRA (b).

UL increases, the proportion of DL and UL traffic is also changed. QDRA dynamically adjust the DL/UL subframe ratio in phase 1 in order to adapt to the current situation, whereas the priority scheme maintains its fixed ratio of DL and UL. As the results, QDRA shows a lower throughput of BE traffic class compared to the priority scheme as the number of UGS flows reaches to 10. QDRA is trying to balance the performance of both DL and UL traffic when its UL throughput is much better than the priority scheme, but its DL throughput is slightly smaller than the priority scheme.

Figures 11a,b shows DL mean delay achieved by all traffic classes using the priority scheme and QDRA, respectively. It is shown that the BE traffic performance of the priority scheme is better than QDRA when the other traffic classes are showing almost the same results for both two schemes. In addition, the DL mean delay of traffic classes such as UGS, ertPS, and rtPS in QDRA all meets the delay requirements.

5 Conclusions

In this paper, we have proposed a scheme called QDRA which provides the QoS improvement for all types of service flows defined by the WiMAX standard. To enhance the bandwidth usage, QDRA dynamically adjusts the DL/UL subframe ratio responding to the requested DL

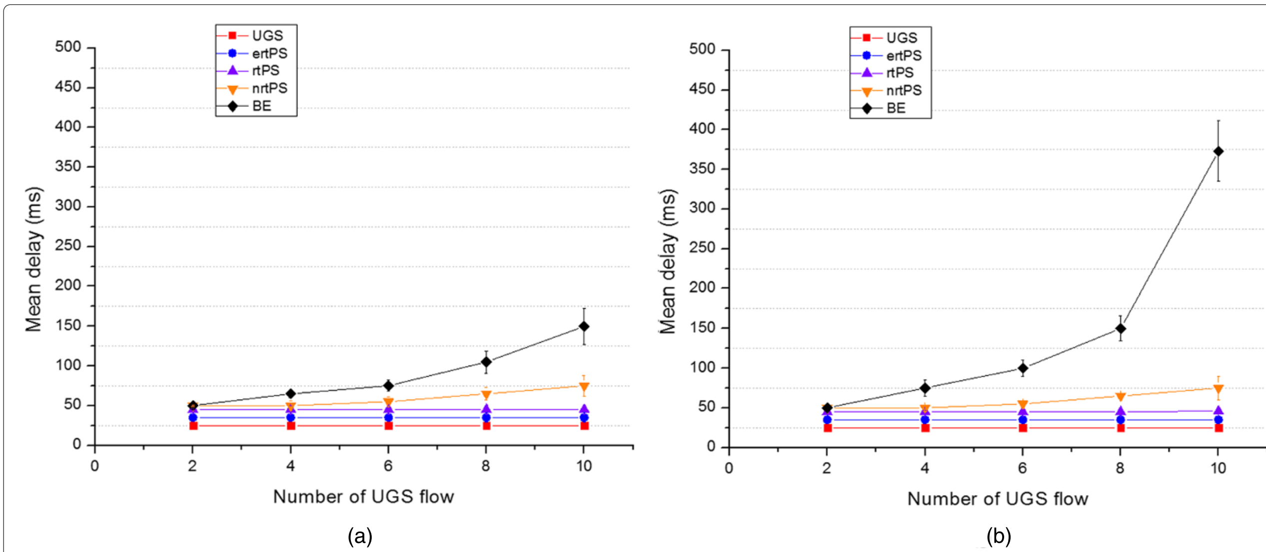

Figure 11 DL mean delay achieved by priority scheme (a) and QDRA (b).

and UL bandwidths while taking account of the priority order of traffic classes. It was also found that QDRA provides a better performance for rtPS, nrtPS, and BE by comparing to the pre-reported priority scheme while satisfying the QoS requirements for all classes of service. In addition, it was found that QDRA improves the throughput and the delay performances compared to the other schemes.

Abbreviations
BE, best effort; BS, base station; CQI, channel quality indicator; DL, downlink; DL-MAP, downlink map; DRR, deficit round robin; EDF, earliest deadline first; ertPS: extended real-time polling service; FCH, frame control header; FDD, frequency division duplexing; FEC, forward error control; IP, Internet protocol; LTE, long-term evolution; MCS, modulation and coding scheme; MRTR, minimum reserved traffic rate; nrtPS, non-real time polling service; OFDMA, orthogonal frequency division multiple access; PUSC, partially used subchannelization; QDRA, QoS-aware dynamic resource allocation; QoS, quality of service; RR, round robin; RTG, receive/transmit transition gap; rtPS, real-time polling service; SS, subscribe station; TCP, Transmission Control Protocol; TDD, time division duplexing; TTG, transmit/receive transition gap; UL, uplink; UL-MAP, uplink map; UGS, unsolicited grant service; WFQ, weighted fair queue; WiMAX: Worldwide Interoperability for Microwave Access; WRR: Weighted round robin.

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

Acknowledgements
This work was supported partly by the IT R&D program of MOTIE/KEIT [10043462, the development of Gigabit Wireless Backhaul Transmission System Connecting EPC Network & Small Cell BS for NLOS/LOS Environment], partly by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2014 and partly by Korea University.

Received: 5 December 2013 Accepted: 12 June 2014
Published: 20 June 2014

References
1. 3GPP, Evolved Universal Terrestrial Radio Access (EUTRA) - Physical Channels and Modulation (Release 10), 3GPP TS 36.211 V10.5.0, Valbonne, France, June 2012
2. IEEE, IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Broadband Wireless Access Systems (IEEE 802.16-2009). New York, May 2009
3. JG Andrews, A Ghosh, R Muhamed, *Fundamentals of WiMAX Understanding Broadband Wireless Networking*, (Prentice-Hall PTR, Upper Saddle River, 2007)
4. Cisco, Cisco visual networking index: global mobile data traffic forecast update, 2010–2015. White Paper, San Jose, CA, February 2011
5. H Falaki, D Lymberopoulos, R Mahajan, S Kandula, D Estrin, A first look at traffic on smartphones, in *Proceedings of the 10th Annual Conference on Internet Measurement* (ACM, New York, 2010), pp. 281–287
6. R Pries, F Wamser, D Staehle, K Heck, P Tran-Gia, On traffic characteristics of a broadband wireless internet access, in *Next Generation Internet Networks* (Aveiro, 1–3 July 2009), pp. 1–7
7. M Kihl, C Lagerstedt, A Aurelius, P Ödling, Traffic analysis and characterization of internet user behavior, in *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)* (IEEE, Moscow, 18–20 October 2010), pp. 224–231
8. K Wongthavarawat, A Ganz, Packet scheduling for QoS support in IEEE 802.16 broadband wireless access systems. Int. J. Commun. Syst. **16**(1), 81–96 (2003)
9. C Cicconetti, L Lenzini, E Mingozzi, C Eklund, Quality of service support in IEEE 802.16 networks. IEEE Netw. **20**(2), 50–55 (2006)
10. JF Borin, NL da Fonseca, Scheduler for IEEE 802.16 networks. IEEE Commun. Lett. **12**(4), 274–276 (2008)
11. E Amir, N Nasser, Dynamic QoS-based bandwidth allocation framework for broadband wireless networks. IEEE Trans. Vehic. Technol. **60**(6), 2690–2700 (2011)
12. C So-In, R Jain, AK Tamimi, Scheduling in IEEE 802.16e mobile WimAX networks: key issues and a survey. IEEE J. Selec. Areas Commun. **27**(2), 156–171 (2009)
13. CH Chiang, W Liao, T Liu, IK Chan, HL Chao, Adaptive downlink and uplink channel split ratio determination for TCP-based best effort traffic in TDD-based WiMAX networks. IEEE J. Select. Areas Commun. **27**(2), 182–190 (2009)
14. R Pries, D Staehle, D Marsico, IEEE 802.16 capacity enhancement using an adaptive TDD split, in *IEEE Vehicular Technology Conference (VTC) 2008* (IEEE, Singapore, May 2008), pp. 1539–1543
15. I Adhicandra, Adaptive subframe allocation in WiMAX networks, in *33rd International Conference on Telecommunications and Signal Processing-TSP,* 17–20 August 2010 (Baden, 2010), pp. 358–362
16. PG Sarigiannidis, GI Papadimitriou, P Nicopolitidis, MS Obaidat, AS Pomportsis, A novel adaptive mapping scheme for IEEE 802.16 mobile downlink framing, in *GLOBECOM 2010* (Miami, FL, 6–10 December 2010), pp. 1–5
17. WiMAX Forum, WiMAX System Evaluation Methodology V2.1. http://www.wimaxforum.org/resources/technical-specifications. Accessed July 2008
18. OPNET Modeler. http://www.riverbed.com

doi:10.1186/1687-1499-2014-104
Cite this article as: Nguyen *et al.*: QoS-aware dynamic resource allocation for wireless broadband access networks. *EURASIP Journal on Wireless Communications and Networking* 2014 **2014**:104.