1. Introduction

The traditional solution to provide high-speed broadband access is to use wired access technologies, such as cable modem, DSL (Digital Subscriber Line), Ethernet, and fiber optic. However, it is too difficult and expensive for carriers to build and maintain wired networks, especially in remote areas. BWA (Broadband Wireless Access) technology is a flexible, efficient, and cost-effective solution to overcome the problems [1]. WiMAX is one of the most popular BWA technologies today, which aims to provide high speed broadband wireless access for WMANs (Wireless Metropolitan Area Network). [2]

WiMAX provides an affordable alternative for wireless broadband access supporting a variety of applications of different types including video conferencing, non-real-time large volume data transfer, traditional voice/data traffic throughput E1/T1 connection, and web browsing.[1]

Each traffic flow requires different treatment from the network in terms of allocated bandwidth, maximum delay, and jitter and packet loss [3], [5]. Traffic differentiation is thus a crucial feature to provide network-level QoS (Quality of Service). The standard leaves QoS support features specified for WiMAX networks (e.g., traffic policing and shaping, connection admission control and packet scheduling). One of the most critical issues is the design of a very efficient scheduling algorithm which coordinates all other QoS-related functional entities.

The key components in WiMAX QoS guarantee are the admission control and the bandwidth allocation in BS. WiMAX standard defines adequate signalling schemes to support admission control and bandwidth allocation, but does not define the algorithms for them. This absence of definition allows more flexibility in the implementation of admission control and bandwidth allocation.

In this study, we focus on evaluating scheduling algorithms for the uplink traffic in WiMAX. We evaluate a number of WiMAX uplink scheduling algorithms in a single-hop network, which is referred to as PMP (Point Multi Point) mode of WiMAX.
2. Overview of WiMAX

In this section, we discuss the WiMAX, the uniqueness of WiMAX uplink scheduling.

2.1 WiMAX standard

BWA technology promises a large coverage and high throughput. Theoretically, the coverage range can reach 30 miles and the throughput can achieve 75 Mbit/s [1]. Yet, in practice the maximum coverage range observed is about 20 km and the data throughput can reach 9 Mbit/s using UDP (User Datagram Protocol) and 5 Mbit/s using FTP (File Transfer Protocol) over TCP (Transmission Control Protocol)[2].

WiMAX standard has two main variations: one is for fixed wireless applications (covered by IEEE 802.16-2004 standard) and another is for mobile wireless services (covered by IEEE 802.16e standard). The 802.16 standards only specify the PHY (Physical) layer and the MAC (Media Access Control) layer of the air interface while the upper layers are not considered.

The IEEE 802.16 standard specifies a system comprising two core components [6]: the SS (Subscribe Station) or CPE (customer premises equipment) and the BS (Base Station). A BS and one or more SS can form a cell with a P2MP structure. Note that the WiMAX standard also can be used in a P2P (Point to Point) or mesh topology. BS acts as a central entity to transfer all the data from MSs (Mobile Station) in a PMP mode. Transmissions take place through two independent channels: downlink channel (from BS to MS) and uplink channel (from MS to BS). Uplink Channel is shared between all MSs while downlink channels is used only by BS.

To support the two-way communication, either FDD (Frequency Division Duplex) or TDD (Time Division Duplex) can be adopted. In the following discussion, we focus on the popular TDD.

The IEEE 802.16 is connection oriented. Each packet has to be associated with a connection at MAC level. This provides a way for bandwidth request, association of QoS and other traffic parameters and data transfer. All data transmissions are connection-oriented and the connections are classified into four types, namely, UGS (Unsolicited Grant Service), also known as CBR (Constant Bit Rate), rtVR (real-time Variable Bit Rate), nrtVR(non real-time Variable Bit Rate), and BE (Best Effort). Each service related to type of QoS class can has different constraints such as the traffic rate, maximum latency and tolerated jitter. In section 4 we will focus more on QoS the WiMAX technologies.

UGS supports real-time service flows that have fixed-size data packets on a periodic basis. RtVR supports real-time service flows that generate variable data packets size on a periodic basis. The BS provides unicast grants in an unsolicited manner like UGS. Whereas the UGS allocations are fixed in size. NrtVR is designed to support non real-time service flows that require variable size bursts on a regular basis. BE is used for best effort traffic where no throughput or delay guarantees are provided. Those service classes are defined in order to satisfy different types of QoS requirements. However, the IEEE 802.16 standard does not specify the scheduling algorithm to be used. Vendors and operators have to choose the scheduling algorithm(s) to be used.
2.2 WiMAX MAC layer

The 802.16 MAC protocol supports transport protocols such as ATM, Ethernet, and IP, and can accommodate future developments using the specific convergence layer. The MAC also accommodates very high data throughput through the physical layer while delivering ATM-compatible QoS, such as UGS, rtVR, nrtVR, and BE. The 802.16 frame structure enables terminals to be dynamically assigned uplink and downlink burst profiles according to the link conditions. This allows for a tradeoff to occur – in real time – between capacity and robustness. It also provides, on average, a 2x increase in capacity when compared to non-adaptive systems.

Fig. 1. Wireless MAC protocol classification [25]

According to the architecture topology, there are two main wireless MAC protocols:

1. Distributed MAC Protocols: these protocols are founded on principles of CS and CA, excluding the distributed ALOHA protocol.
2. Centralized MAC Protocols: these protocols are based on communicating with a central entity – in case of cellular mobile communications: the base station. Thus all communications is organized and supervised according to the BS MAC management protocol. [25]

There are three types of wireless MAC protocol types:

1. Random Access Protocols: according to this access protocol, for a node to be able to access the network it should contend for the medium.
2. Guaranteed Access Protocols: unlike the random access protocols, the communication between nodes is made on some predefined rules. This may be either in the form polling the nodes one by one, or by token exchanging.
3. Hybrid Protocols: these type of protocols are more superior to the previously mentioned other two protocols, since they are made out of the top properties of random access protocols and guaranteed access protocols. [25].

Hybrid protocols can be further subdivided into two categories:

a. Random Reservation Access protocol: these are the protocols by the MAC where a periodic reservation of the bandwidth is granted on the reception of a successful request from nodes supported by the central node.

b. Demand Assignment protocol: The MAC allocated bandwidth according to the need of application of the node.

The hierarchy of the wireless MAC protocols classification could be illustrated as it is shown in figure 1.

Therefore, mobile WiMAX MAC protocol could be classified as the demand assignment protocol; knowing that the mobile WiMAX MAC is designed to support QoS and according to the only MAC protocol that guarantees resources is the DA protocol. [25]

![Fig. 2. The 802.16 protocol stack [6]](image)

WiMAX MAC is subdivided into three sub layers with different functionalities. Figure 2 is a basic illustration of the tasks and services that the MAC sub layers are responsible for.
The upper MAC layer is the CS (Convergence Sub layer). This sub layer is responsible mainly for classification and header suppression of the incoming packets from the network layer. The classification is done according to the QoS parameters of the packet. Then each service flow is assigned a service flow identifier number.

The inner layer is called the MAC CSP (common part sub layer) and it is considered the main sub layer of the MAC layer. Finally, the lower MAC layer is called the MAC security sub layer. Functions like support for privacy, user/device authentication are the responsibility of this sub layer.

Figure 3 illustrates the PHY layer with the three sub layers of the MAC layer. The Figure shows the data/control plane only and it regarded as the scope of the standard.
Quality of Service and Resource Allocation in WiMAX

The 802.16 MAC uses a variable-length PDU (Protocol Data Unit) and other innovative concepts to greatly increase efficiency. Multiple MAC PDUs, for example, may be concatenated into a single burst to save PHY overhead. Multiple SDUs (Service Data Unit) may also be concatenated into a single MAC PDU, saving on MAC header overhead. Fragmentation allows very large SDUs to be sent across frame boundaries to guarantee the QoS. Payload header suppression can be used to reduce the overhead caused by the redundancy within the SDU headers. The 802.16 MAC uses a self-correcting bandwidth request/grant scheme that eliminates any delay in acknowledgements, while allowing better QoS handling than traditional acknowledgement schemes. Depending on the QoS and traffic parameters of their services, terminals have a variety of options available to them for requesting bandwidth. SAP (Service Access Point) is entities located in between the sub layers in order to convert the SDU to PDU. Basically, when PDUs of an upper layer are passed through the SAP to a lower layer, they are considered as SDU for that particular lower layer.

In TDD mode, a WiMAX MAC frame consists of two sub frames, DL-sub frame for downlink transmission and UL-sub frame for uplink transmission, as shown in figure 4. The DL-sub frame comprises a BP (Burst Preamble) and a FCH (Frame Control Header), followed by DL-MAP, UL-MAP, and a number of downlink payload bursts (DL-PL1, ..., DL-PLn). As will be discussed later, the UL-MAP contains the uplink scheduling results, i.e., the uplink bandwidth granted to each SS. The UL-sub frame starts with initial ranging contention slots and bandwidth request contention slots, followed by a number of uplink payload bursts (UL-PL1, ..., UL-PLn)(Figure 4). The uplink and downlink bursts are not necessarily equal and their length can be adjusted dynamically in order to adapt to the traffic variation. [29]

![WiMAX MAC frame structure](image_url)

**Fig. 4. WiMAX MAC frame structure [29]**

### 3. Problem statement

A WiMAX network is designed to incorporate different types of data streams, and it aims at providing QoS guarantee for all the data streams being served by WiMAX. The WiMAX protocol covers physical layer and MAC layer, and there are several challenges for QoS guarantee in WiMAX.
The IEEE 802.16 standard provides specification for the MAC and PHY layers for WiMAX and there are several challenges for QoS guarantee in WiMAX.

In the physical layer, one challenge is the uncertainty of the wireless channel, which makes the guarantee of broadband wireless data service difficult and renders the static resource allocation scheme unsuitable. In the MAC layer, one challenge is the diversified service types, which requires the WiMAX scheduling scheme to be adaptive to the various QoS parameters of different service types. There have been some studies of the WiMAX MAC scheduling problem [3], [4], [5], and [6].

The key components in WiMAX QoS guarantee are the admission control and the bandwidth allocation in BS. WiMAX standard defines adequate signalling schemes to support admission control and bandwidth allocation, but does not define the algorithms for them. This absence of definition allows more flexibility in the implementation of admission control and bandwidth allocation.

The research problem being investigated here is, after connections are admitted into the WiMAX network, how to allocate bandwidth resources and perform scheduling services, so that the QoS requirements of the connections can be satisfied.

### 3.1 What is QoS?

QoS refers to the ability of a network to provide improved service to selected network traffic over various underlying technologies including wired-based technologies (Frame Relay, ATM, Ethernet and 802.1 networks, SONET, and IP-routed networks) and wireless-based technologies (802.11, 802.15, 802.16, 802.20, 3G, IMS, etc). In particular, QoS features provide improved and more predictable network service by providing the following services:

- Supporting dedicated bandwidth
- Improving loss characteristics
- Avoiding and managing network congestion
- Shaping network traffic
- Setting traffic priorities across the network

Due to the differences in the wired-based and wireless-based access technologies, the detailed QoS implementations for both tend to be different, however they share common roots. What follows next are the common elements shared between wired-based and wireless-based access methods.

### 3.2 QoS and scheduling in WiMAX

A high level of QoS and scheduling support is one of the interesting features of the WiMAX standard. These service-provider features are especially valuable because of their ability to maximize air-link utilization and system throughput, as well as ensuring that SLAs (Service-Level Agreements) are met (Figure 5). The infrastructure to support various classes of services comes from the MAC implementation. QoS is enabled by the bandwidth request and grant mechanism between various subscriber stations and base stations. Primarily there are four buckets for the QoS (UGS, rtVR, nrtVR, and BE) to provide the service-class classification for video, audio, and data services, as they all require various levels of QoS.
requirements. The packet scheduler provides scheduling for different classes of services for a single user. This would mean meeting SLA requirements at the user level. Users can be classified into various priority levels, such as standard and premium.

Fig. 5. Packet scheduling, as specified by 802.16 [6]

3.3 Scheduling algorithm and their characteristic

In some cases, separate scheduling algorithms are implemented for the uplink and downlink traffic. Typically, a CAC (Call Admission Control) procedure is also implemented at the BS that ensures the load supplied by the SSs can be handled by the network. A CAC algorithm will admit a SS into the network if it can ensure that the minimum QoS requirements of the SS can be satisfied and the QoS of existing SSs will not deteriorate. The performance of the scheduling algorithm for the uplink traffic strongly depends on the CAC algorithm.

Scheduling has also been studied intensively in many disciplines, such as CPU task scheduling in operating systems, service scheduling in a client-server model, and events scheduling in communication and computer networks. Thus a lot of scheduling algorithms have been developed. However, compared with the traditional scheduling problems, the WiMAX MAC layer scheduling problem is unique and worth study for the following reasons.

First, the total bandwidth in a WiMAX network is adaptive since AMC (Adaptive Modelling and Coding) is deployed in the physical layer and the number of bytes each time slot can carry depends on the coding and modulation scheme. Second, multiple service types have been defined and their QoS requirements need to be satisfied at the same time. How to satisfy various QoS requirements of different service types simultaneously has not been addressed by any other wireless access standard before. Third, the time complexity of the WiMAX scheduling algorithm must be simple since real-time service demands a fast response from the central controller in BS.
Fourth, the frame boundary in the WiMAX MAC layer also serves as the scheduling boundary, which makes the WiMAX scheduling problem different from the continuous time scheduling problem. The above four characteristics make the resource allocation in the WiMAX MAC layer a challenging problem.

While some similarities to the wired world can be drawn, there are certain characteristics of the wireless environment that make scheduling particularly challenging. Five major issues in wireless scheduling are identified in [9]:

- **Wireless link variability:** Due to characteristics of the channel as well as location of the mobile subscribers.
- **Fairness:** Refers to optimizing the channel capacity by giving preference to spectrally efficient modulations while still allowing transmissions with more robust modulations (and hence, consuming a major amount of spectrum) to get their traffic through.
- **QoS:** Particularly for WiMAX, QoS support should be built into the scheduling algorithm to guarantee that QoS commitments are meet under normal conditions as well as under network degradation scenarios.
- **Data throughput and channel utilization:** Refers to optimizing the channel utilization while at the same time avoiding waste of bandwidth by transmitting over high loss links.
- **Power constrain and simplicity:** Be considerate of the terminals’ battery capacity as well as computational limitations both at the BS and MS.

### 3.4 Classification scheduling algorithms

Packet scheduling algorithms are implemented at both the BS and SSs. A scheduling algorithm at the SS is required to distribute the bandwidth allocation from the BS among its connections.

The scheduling algorithm at the SS needs to decide on the allocation of bandwidth among its connections. The scheduling algorithm implemented at the SS can be different than that at the BS.

The focus of our work is on scheduling algorithms executed at the BS for the uplink traffic in WiMAX i.e. traffic from the SSs to the BS. A scheduling algorithm for the uplink traffic is faced with challenges not faced by an algorithm for the downlink traffic. An uplink scheduling algorithm does not have all the information about the SSs such as the queue size. An uplink algorithm at the BS has to coordinate its decision with all the SSs where as a downlink algorithm is only concerned in communicating the decision locally to the BS.

In general, the scheduling algorithms can be classified as frame-based scheduling and sorted-based scheduling. Frame-based scheduling algorithms include WRR (Weighted Round Robin)[7], DRR (Deficit Round Robin)[8], etc. Sorted-based scheduling algorithms include WFQ (Weighted Fair Queue)[9], also known as PGPS (Packet-based Generalized Processor Sharing)[10], and a number of variations of WFQ such as WF2Q (Worst Case Fair Queuing)[11], SCFQ (Self-Clock Faire Queuing)[12].

The advantage of frame-based scheduling algorithms is their low computing complexity, while the disadvantage is the significant worst case delay. On the contrary, scheduling
algorithm in the WFQ family has better performance in worst case delay, but the algorithm complexity is much higher than that of the frame-based scheduling algorithms.

### 3.5 Uplink scheduling algorithms

In the coming subsections the fundamental scheduling algorithms will be briefly described.

#### 3.5.1 Round Robin

Round Robin as a scheduling algorithm is the most basic and least complex scheduling algorithm. It has a complexity value of $O(1)$ [13].

Basically the algorithm services the backlogged queues in a round robin fashion. Each time the scheduler pointer stop at a particular queue, one packet is dequeued from that queue and then the scheduler pointer goes to the next queue. This is shown in Figure 6.

![Fig. 6. RR Scheduler](image)

It distributes channel resources to all the SSs without any priority. The RR scheduler is simple and easy to implement. However, this technique is not suitable for systems with different levels of priority and systems with strongly varying sizes of traffic.

#### 3.5.2 Weighted Round Robin

An extension of the RR scheduler, the WRR scheduler, based on static weights. WRR [14] was designed to differentiate flows or queues to enable various service rates. It operates on the same bases of RR scheduling. However, unlike RR, WRR assigns a weight to each queue. The weight of an individual queue is equal to the relative share of the available system bandwidth. This means that, the number of packets dequeued from a queue varies according to the weight assigned to that queue. Consequently, this differentiation enables prioritization among the queues, and thus the SSes. [15]
3.5.3 Earliest deadline first

It is a work conserving algorithm originally proposed for real-time applications in wide area networks. The algorithm assigns deadline to each packet and allocates bandwidth to the SS that has the packet with the earliest deadline. Deadlines can be assigned to packets of a SS based on the SS’s maximum delay requirement. The EDF algorithm is suitable for SSs belonging to the UGS and rtVR scheduling services, since SSs in this class have stringent delay requirements. Since SSs belonging to the nrtVR service do not have a delay requirement, the EDF algorithm will schedule packets from these SSs only if there are no packets from SSs of UGS or rtVR class. [16]

3.5.4 Weighted fair queue

It is a packet-based approximation of the Generalized Processor Sharing (GPS) algorithm. GPS is an idealized algorithm that assumes a packet can be divided into bits and each bit can be scheduled separately. The WFQ algorithm results in superior performance compared to the WRR algorithm in the presence of variable size packets. The finish time of a packet is essentially the time the packet would have finished service under the GPS algorithm. The disadvantage of the WFQ algorithm is that it will service packets even if they wouldn’t have started service under the GPS algorithm. This is because the WFQ algorithm does not consider the start time of a packet.

3.5.5 Temporary removed packet

The TRS (Temporary Removal Scheduler) involves identifying the packet call power, depending on radio conditions, and then temporarily removing them from a scheduling list for a certain adjustable time period TR. The scheduling list contains all the SSs that can be served at the next frame. When TR expires, the temporarily removed packet is checked again. If an improvement is observed in the radio channel, the packet can be topped up in the scheduling list again, otherwise the process is repeated for TR duration. In poor radio conditions, the whole process can be repeated up to L times at the end of which, the removed packet is added to the scheduling list, independently of the current radio channel condition [18].

The temporary TRS can be combined with the RR scheduler.

The combined scheduler is called TRS+RR. For example, if there are k packet calls and only one of them is temporary removed, each packet call has a portion, equal to $\frac{1}{k-1}$, of the whole channel resources.

3.5.6 Maximum Signal to Interference Ration

The scheduler mSIR (Maximum Signal to Interference Ration) is based on the allocation of radio resources to subscriber stations which have the highest SIR. This scheduler allows a highly efficient utilization of radio resources. However, with the mSIR scheduler, the users with a SIR (Signal to Interference Ratio) that is always small may never be served.[18]
The TRS can be combined with the mSIR scheduler. The combined scheduler is called TRS + mSIR. This scheduler assigns the whole channel resources to the packet call that has the maximum value of the SNR (Signal to Noise Ratio). The station to be served has to belong to the scheduling list.

### 3.5.7 Reinforcement Learning

The scheduler RL (Reinforcement Learning) is based on the model of packet scheduling described by Hall and Mars [23]. The aim is to use different scheduling policies depending on which queues are not meeting their delay requirements. The state of the system represented by a set of $N - 1$ binary variables $\{s_1: s_{n-1}\}$, where each variable $s_i$ indicates whether traffic in the corresponding queue $q_i$ [24].

There is not variable corresponding to the best-effort queue $q_N$, since there is no mean delay requirement for that queue. For example, the state $\{0; 0; \ldots; 0\}$ represents that all queues have satisfied their mean delay constraint, while $\{1; 0; \ldots; 0\}$ represents that the mean delay requirements are being satisfied for all queues except $q_1$. Thus, if there are $N$ queues in the system including one best-effort queue, then there are $2^{N-1}$ possible states. In practice, the number of traffic classes is normally small, e.g., four classes in Cisco routers with priority queuing, in which case the number of states is acceptable.

At each timeslot, the scheduler must select an action $a \in \{a_1: a_N\}$, where $a_i$ is the action of choosing to service the packet at the head of queue $q_{i}$. The scheduler makes this selection by using a scheduling policy $\Pi$, which is a function that maps the current state of the system $s$ onto an action $a$. If the set of possible actions is denoted by $A$, and the set of possible system states is denoted by $S$, then $\Pi: S \rightarrow A$.

### 3.5.8 Hierarchical/hybrid algorithms

Hierarchical/hybrid algorithms build on the fact that scheduling services have different and sometimes conflicting requirements. UGS services must always have their delay and bandwidth commitment met, so simply reserving enough bandwidth for those services and controlling for oversubscription would be enough; rtVR services have little tolerance for delay and jitter, so an algorithm guaranteeing delay commitments would be more suitable; and finally, BE and nrtVR will always be hungry for bandwidth with no considerations for delay, so a throughput maximizing algorithm might be preferred.

While hierarchical refers to two or more levels of decisions to determine what packets to be scheduled, hybrid refers to the combination of several scheduling techniques (EDF for delay sensitive scheduling services such as rtVR and UGS, and WRR for nrtVR and BE for example). There could be hierarchical solutions that are not necessarily hybrid, but hybrid algorithms usually distribute the resources among different service classes, and then different scheduling techniques are used to schedule packets within each scheduling service, making them hierarchical in nature.

A two-tier hierarchical architecture is proposed in [24] for WiMAX uplink scheduling. In the higher hierarchy, strict prioritization is used to direct the traffic into the four queues, according to its type. Then, each queue is scheduled according to a particular algorithm, i.e.,
fixed allocation for UGS, EDF for rtVR, WFQ for nrtVR, and equal division of remaining bandwidth for BE. Although EDF takes care of the delay requirement of the rtVR, grouping multiple rtVR connections into one queue fails to guarantee the minimum bandwidth requirement of each individual rtVR connection. For example, one rtVR connection with tight delay budget may dominate the bandwidth allocation, resulting in starvation of other rtVR connections.

In [27], the authors use a first level of strict priority to allocate bandwidth to UGS, rtVR, nrtVR and BE services in that order; and then on a second level in the hierarchy, different scheduling techniques are used depending on the scheduling service: UGS, as the highest priority, has pre-allocated bandwidth, EDF is used for rtVR, WFQ for nrtVR, and FIFO for BE. Similarly, explains an algorithm that uses EDF for nrtVR and rtVR classes, and WFQ for nrtVR and BE classes.

In [27], the authors implement a two-level hierarchical scheme for the downlink in which an ARA (Aggregate Resource Allocation) component first estimates the amount of bandwidth required per scheduler class (rtVR, nrtVR, BE and UGS) and distributes it accordingly.

In [28], a SC (Service Criticality) based scheduling is proposed for the WiMAX network, where an SC index is calculated in every SS for each connection and then sent to BS, and BS sorts the SC of all the connections and assigns bandwidth according to the descending order of SC. SC is derived according to the buffer occupancy and waiting time of each connection. If a malicious connection always reports a high SC, or a connection is generating excessive traffic to occupy its sending buffer, this connection will dominate the available bandwidth and affect other connections.

4. Evaluation

This section presents the simulation results for the algorithms scheduling. For testing performance of algorithms, the introduced algorithms are implemented in the NS-2 (Network Simulator) [20] and WiMAX module [21] that is based on the WiMAX NIST module [20]. The MAC implementation contains the main features of the 802.16 standard, such as downlink and uplink transmission. We have also implemented the most important MAC signalling messages, such as UL-MAP and DL-MAP, authentication (PKM), capabilities (SBC), REG (Registration), DSA (Dynamic Service Addition) and DSC (Dynamic Service Change). The implemented PHY is OFDM.

| Lot size(byte) | Channel coding | modulation |
|----------------|----------------|------------|
| 108            | 3/4            | 64-QAM     |
| 96             | 2/3            | 64-QAM     |
| 36             | 3/4            | QPSK       |
| 24             | 1/2            | QPSK       |

Table 1. Slot size for OFDM PHY

The current implementation also supports differencing MCSs (Moulding Code Scheme). Table 1 shows present slot size for different modulations and channel coding types.
We present a simulation scenario to study thoroughly the proposed scheduling solution. The scenario will present a multi-service case, in which a provider has to support connections with different 802.16 classes and traffic characteristics.

The purpose of this scenario is to ensure that the scheduler at the BS takes the service class into account and allocates slots based on the QoS requirements and the request sizes sent by SSs. Another purpose is to test that the scheduler at the BS takes the MAC overhead into account. Table 1 presents information about which applications are active at scenario.

Regardless of the simulation scenario, the general parameters of the 802.16 network are the same (see Table 2). There is one BS that controls the traffic of the 802.16 network. The physical layer is OFDM. The BS uses the dynamic uplink/downlink slot assignment for the TDD mode. Both the BS and all SSs use packing and fragmentation in all simulation scenarios. The MAC level uses the largest possible PDU size. ARQ is turned off; neither the BS nor SSs use the CRC field while sending packets.

| Value | Parameter |
|-------|-----------|
| OFDM  | PHY       |
| 7MHz  | Bandwidth |
| 400   | Frame per Second |
| TDD   | Duplexing mode |
| OFF   | ARQ/CRC   |

Table 2. WiMAX parameter

We consider a general scenario, where rtVR and/or nrtVR connections are established. Connection i has an arrival rate of $\lambda_i$, a delay budget of $i$, and a minimum reserved bandwidth of MRR$_i$. For the sake of analytic tractability, we assume that the data arrival forms a Poisson process and all queues have infinite size. Other types of traffic (such as the more practical bursty traffic) are studied through simulations.

The main parameters of the simulation are represented in Table 3. Effects of these parameters are similar over results of all scheduling algorithms. Moreover, producers of this WiMAX module have used these values for testing performance of their simulator.

| Parameter            | Value       |
|----------------------|-------------|
| Frequency band       | 5 MHz       |
| Propagation model    | Two Ray Ground |
| Antenna model        | Omni antenna |
| Antenna height       | 1.5 m       |
| Transmit power       | 0.25        |
| Receive power threshold | 205e-12   |
| Frame duration       | 20 ms       |
| Cyclic prefix (CP)   | 0.25        |
| Simulation duration  | 100 s       |

Table 3. Main parameters of the simulation
In particular, we consider several comparable scheduling algorithms, including WRR, EDF, and TRS which is a representative WiMAX scheduling algorithm and has been patented and well received.

Besides packet drop rate and throughput that have been studied in analysis, we are also interested in the fairness performance, which is measured by Jain’s Fairness Index [22] defined as follows:

\[
f(x_1, x_2, \ldots, x_n) = \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n \sum_{i=1}^{n} x_i^2}
\]

(1)

Where \(x_i\) is the normalized throughput of connection \(i\), and \(n\) is the total number of connections. Each SS establishes a number of connections to the BS in our simulation. We consider ten rtVR connections and ten nrtVR connections. Each type of connection is associated with an MRR and a delay budget.

\[
X_i = \frac{TH_i}{MRR_i}
\]

(2)

ie, with \(TH_i\) and \(MRR_i\) stand for the connection \(i\)’s actual data rate and reserved data rate, respectively. The Jain’s Fairness Index ranges between 0 and 1. The higher the index, the better the fairness. If \(TH_i = MRR_i\) for all \(i\), or in other words, every connection obtains its reserved data rate, then \(x_i = 1\) for all \(i\), and Jain’s Fairness Index equals 1. All simulations and analytic calculations are done using NS2 simulator.

Fig. 7. Latency versus traffic

Figures 7, 8 show delay packets as a function of the traffic load submitted to the network. The data packets are generated by a streaming multimedia application. The diagram of UGS scheduling algorithm by considering delay is linear where its throughput is increasing. As mentioned above, the UGS traffic request is the highest priority. If a packet is available in this type of traffic it will be sent in no time. For accurate performance evaluation, we adopt the WiMAX physical layer standard OFDM_BPSK_1_2 in our simulations. [24]
The fairness of the scheduling algorithms under bursty traffic is shown in figure 9. As we can see, WRR always maintains almost high fairness, while the fairness of EDF algorithm is the worst among the four algorithms. This is due to the fact that some real time packets rtVR connections are dropped under high burstiness, and thus the throughput of rtVR decreases. [30], [31]
Figure 10 shows the latency as a function of rtVR+nrtVR traffic load. We verify that the TRS scheduler provides a decrease in the latency.

![Fig. 10. Latency versus Simulation Time](image)

Figure 11 shows the latency as a function of rtVR traffic load. We verify that the mSIR scheduler provides a decrease in the latency.

![Fig. 11. Latency versus Simulation Time](image)

In figure 12, the protocols have been compared on the base of throughput. As you see, TRS+RR throughput is greater than all.

![In figure 12, the protocols have been compared on the base of throughput. As you see, TRS+RR throughput is greater than all.](image)
5. References

[1] IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems, 2004, IEEE802.16. Available from: http://www.ieee802.org/16/.

[2] IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems—Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, 2005, IEEE802.16e.. Available from: http://www.ieee802.org/16/.

[3] Aura Ganz, Zvi Ganz, Kitti Wongthavarawat(18 September 2003). Multimedia Wireless Networks: Technologies, Standards, and QoS, Prentice Hall Publisher.

[4] Overcoming Barriers to High-Quality Voice over IP Deployments (2003), Intel Whitepaper.

[5] DiffServ-The Scalable End-to-End Quality of Service Model (August 2005), Cisco Whitepaper.

[6] WiMAX – Delivering on the Promise of Wireless Broadband (Second Quarter 2006), Xcell Journal - Issue 57.

[7] M. Katavenis, S. Sidiropoulos, and C. Courcoubetis. Weighted Round-Robin Cell Multiplexing in A General-Purpose ATM Switch Chip, IEEE J. Sel. Areas Commun., vol. 9, no. 8, pp. 1265–1279, Jan. 1991.

[8] M. Shreedhar and G. Varghese (1995), Efficient Fair Queuing Using Deficit Round Robin, in Proc. IEEE SIGCOMM, pp. 231–242. 135.

[9] A. Demeres, S. Keshav, and S. Shenker(1989). Analysis and Simulation of A Fair Queueing Algorithm, in Proc. IEEE SIGCOMM, pp. 1–12.
[10] A. Parekh and R. Gallager(1992). A Generalized Processor Sharing Approach to Flow Control: The Single Node Case, in Proc. IEEE INFOCOM, pp. 915–924.
[11] J. Bennet and H. Zhang(1996). WF2Q: Worst-Case Fair Weighted Fair Queueing, Proceedings of IEEE INFOCOM, 1996, pp. 120–128.
[12] S. Golestani (1994). A Self-Clocked Fair Queueing Scheme for Broadband Applications, Proceeding of IEEE INFOCOM, pp. 636–646.
[13] R. Jain, lecture notes (2007). A Survey of Scheduling Methods, University of Ohio.
[14] M. Katevenis, S. Sidiropoulos and C. Courcoubetis(1991). Weighted round robin cell multiplexing in a general purpose ATM switch chip, Selected Areas in Communications, IEEE Journal on 9(8), pp. 1265-1279.
[15] S. Belenki (2000). Traffic management in QoS networks: Overview and suggested improvements, Tech. Rep.
[16] M. Shreedhar and G. Varghese(June 1996). Efficient Fair Queuing using Deficit Round Robin, IEEE/ACM Transactions on Networking, vol. 1, pp. 375-385.
[17] T. Al_Khasib, H. Alnuweiri, H. Fattah and V. C. M. Leung(2005). Mini round robin: enhanced frame-based scheduling algorithm for multimedia networks, IEEE Communications, IEEE International Conference on ICC, pp. 363-368 Vol. 1.
[18] Nortel Networks, Introduction to quality of service (QoS)(September 2008), Nortel Networks Website, 2003. [Online]. Accessed on 1st of September 2008.
[19] C.F. Ball, F. Treml, X. Gaube, and A. Klein(September 2005). Performance Analysis of Temporary Removal Scheduling applied to mobile WiMAX Scenarios in Tight Frequency Reuse, the 16th Annual IEEE International Symposium On Personal Indoor and Mobile Radio Communications, PIMRC 2005, Berlin, 11 – 14.
[20] QoS-included WiMAX Module for NS-2 Simulator. First International Conference on Simulation Tools and Techniques for Communications Networks and Systems, SIMUTools 2008, Marseille, France, March 3-7, 2008.
[21] The network simulator ns-2(September 2007). Available from: http://www.isi.edu/nsnam/ns/.
[22] D. M. C. R. Jain and W. Hawe(1984). A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Systems, dEC Research Report, TR-301.
[23] J. Hall, P. Mars(December 1998). Satisfying QoS with a Learning Based Scheduling Algorithm, School of Engineering, University of Durham,.
[24] M. Taghipoor, G. Tavassoli and V. Hosseiini(April 2010). Guarantee QoS in WiMAX Networks with learning automata, ITNG 2010 Las Vegas, Nevada, USA. 12-14
[25] Ajay Chandra V. Gummalla, John o. Limb.Wireless Medium Access Control Protocols, IEEE Communications Surveys, 2000.
[26] Q. Liu, X. Wang, and G. Giannakis(May 2006). A Cross-Layer Scheduling Algorithm with QoS Support in Wireless Networks, IEEE Trans. Veh. Tech., vol. 55, no. 3, pp. 839–847.
[27] D. Niyato and E. Hossain (Dec. 2006). Queue-aware Uplink Bandwidth Allocation and Rate Control for Polling Service in IEEE 802.16 Broadband Wireless Networks, IEEE Trans. Mobile Comp., vol. 5, no. 8, pp. 668–679.
[28] A. Shejwal and A. Parhar(2007). Service Criticality Based Scheduling for IEEE 802.16 WirelessMAN, in Proc. 2nd IEEE Int. Conf. AusWreless, pp. 12-18.
[29] H. Chen, thesis (spring 2008). Scheduling and Resource Optimization in Next Generation Heterogeneous Wireless Networks, University of Louisiana.

[30] Jafari, Saeid. M., Taghipour, M. and Meybodi, M. R. (2011). Bandwidth Allocation in Wimax Networks Using Reinforcement Learning, World Applied Sciences Journal Vol. 15, No. 4, pp. 525-531.

[31] Jafari, Saeid. M., Taghipour, M. and Meybodi, M. R. (2011). Bandwidth Allocation in Wimax Networks Using Learning Automata, World Applied Sciences Journal Vol. 15, No. 4, pp. 576-583.
This book has been prepared to present state of the art on WiMAX Technology. It has been constructed with the support of many researchers around the world, working on resource allocation, quality of service and WiMAX applications. Such many different works on WiMAX, show the great worldwide importance of WiMAX as a wireless broadband access technology. This book is intended for readers interested in resource allocation and quality of service in wireless environments, which is known to be a complex problem. All chapters include both theoretical and technical information, which provides an in depth review of the most recent advances in the field for engineers and researchers, and other readers interested in WiMAX.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Majid Taghipoor, Saeid MJafari and Vahid Hosseini (2012). Scheduling Algorithm and Bandwidth Allocation in WiMAX, Quality of Service and Resource Allocation in WiMAX, Dr. Roberto Hincapie (Ed.), ISBN: 978-953-307-956-1, InTech, Available from: http://www.intechopen.com/books/quality-of-service-and-resource-allocation-in-wimax/scheduling-algorithm-and-bandwidth-allocation-in-wimax