ADAPTIVE RESOURCE ALLOCATION SCHEME FOR UPLINKS IN IEEE 802.16M SYSTEMS

NURUL HUDA MAT TAHIR

FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR
2013

ADAPTIVE RESOURCE ALLOCATION SCHEME FOR UPLINKS IN IEEE 802.16M SYSTEMS

NURUL HUDA MAT TAHIR

DISSERTATION SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGINEERING SCIENCE

FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR
2013

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **Nurul Huda Mat Tahir**

Registration/Matric No: **KGA 100076**

Name of Degree: **Master of Engineering Science (MEngSc)**

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work’**): Adaptive Resource Allocation Scheme for Uplinks in IEEE 802.16m Systems**

Field of Study: **Wireless Communications**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any except or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledge in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise. I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature ____________________ Date ____________________

Abstract

IEEE802.16m is a new standard for mobile WiMAX. It is expected to be the next 4G technology of wireless LAN offering high speed data rate and larger coverage area. Therefore to meet the requirements specified by IMT-advance (International Mobile Telecommunication Advance) a few amendments have been made from the earlier version of the standard. This thesis proposes a scheduler for uplink service and adapting the requirements outlined by the standard which includes the new advance air interface and an additional service class called aGPS. Dealing with scheduling Quality of Service (QoS) is still a scarce matter. To meet the QoS requirements, our scheduler uses a minimum group allocation where services that have QoS restriction will be put in this group. Only the remaining bandwidth will be distributed fairly using our enhanced deficit round robin (EDRR). EDRR gives variable effects on different types of user channel condition. To observe the performance, the proposed method is compared against deficit round robin (DRR). If the users are experiencing different channel conditions, DRR seems to be the best way of distributing the resources. However, when the users are having an identical channel condition, the method used by DRR in distributing the same amount of quantum number to all users will result in more packet drop for users with larger queues. Therefore DRR needs to be modified to not just considering the physical condition but also the state of user request. Thus the idea of EDRR is that the users with bulk data will have more chances to be mapped into an OFDMA frame. Since we do not want to violate the right of having more bandwidth for the users with the better channel quality, this method will only apply to those that having the same channel condition. The results show EDRR performs well than DRR by reducing the amount of packet drop.

Abstrak

IEEE802.16m adalah satu standard baru untuk WiMAX mudah alih. Ia dijangka menjadi teknologi 4G yang menawarkan kadar data berkelajuan tinggi. Oleh itu untuk memenuhi standard yang digariskan oleh IMT-advance (International Mobile Telecommunication Advance) beberapa pindaan telah dibuat. Tesis ini mencadangkan penjadual untuk perkhidmatan 'uplink' yang sesuai dengan standard yang digariskan oleh piawaian termasuk mengunakan 'new advance air interface' dan menambah kelas perkhidmatan yang dipanggil aGPS. Apabila berurusan dengan penjadualan, kualiti perkhidmatan (QoS) adalah satu perkara yang penting. Oleh itu bagi memenuhi keperluan QoS, penjadual kami menggunakan peruntukan kumpulan minimum di mana perkhidmatan yang mempunyai sekatan QoS akan mengambil bahagian dalam kumpulan ini. Hanya jalur lebar selebihnya akan diagihkan menggunakan defisit pusingan robin dipertingkat (EDRR). Untuk menilai prestasi system kaedah yang dicadangkan dibandingkan dengan defisit pusingan robin (DRR). Apabila pengguna yang mengalami keadaan saluran fizikal yang berbeza, DRR mampu menjadi cara terbaik mengagihkan sumber-sumber di kalangan mereka. Walau bagaimanapun, apabila pengguna mempunyai satu keadaan saluran yang sama, kaedah yang digunakan oleh DRR dalam mengagihkan nombor kuantum yang sama kepada semua pengguna akan mengakibatkan kehilangan data kepada pengguna yang mempunyai lebih data di dalam barisan. Oleh itu DRR perlu diubahsuai untuk bukan sahaja mengambil kita keadaan fizikal tetapi juga keadaan permintaan pengguna. Oleh itu, idea EDRR adalah pengguna dengan data yang besar akan mempunyai lebih banyak peluang untuk dipetakan ke dalam bingkai OFDMA. Oleh kerana kita tidak mahu melanggar hak mempunyai jalur lebar lebih bagi pengguna yang mempunyai kualiti saluran yang lebih baik, kaedah ini hanya akan dikenakan kepada mereka yang mempunyai keadaan

saluran yang sama. Keputusan menunjukkan EDRR menonjolkan prestasi yang lebih baik berbanding DRR di bawah senario yang sama. EDRR menyediakan jalur lebar lebih untuk pengguna yang mempunyai permintaan yang lebih supaya kemungkinan penurunan paket akan menjadi lebih rendah dalam bingkai seterusnya.

Acknowledgement

Thankfully with the help of others I am able to finish my dissertation. Special appreciation goes to my supervisor, Dr. Kamarul Ariffin Noordin, for his supports, caring, patience, and giving me opportunity on doing research in the wireless communication field. I would like to thank him for guiding me throughout the time and always kindly providing me with good suggestions. I would also like to express my appreciation to my lab mate for standing by me through the good nor bad and give good response to my research progress. Not to forget all my friends which are involved in my research life directly or indirectly for the friendship, memories and encouragement. Sincere thanks to my family which is always cheering me up and being beside me whenever in need. My acknowledgement also goes to all the technicians and office staffs of Electrical Engineering Department for their co-operations. Finally I would like to express gratitude to the University of Malaya which financially supported my research via GRA scheme.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2G	-	Second Generation
3G	-	Third Generation
3GPP	-	Third Generation Partnership Project
4G	-	Forth Generation
aGPS	-	Adaptive Grant Polling Service
BE	-	Best Effort
BS	-	Base Station
CDRR	-	Customized Deficit Round Robin
CID	-	Connection Identifier
CRC	-	Cyclic Redundancy Check
CQI	-	Channel Quality Information
CQICH	-	Channel Quality Indicator Channel
DC	-	Deficit Counter
DL	-	Downlink
DPRA	-	Dynamic Priority Resource Allocation
DRR	-	Deficit Round Robin
DRU	-	Distributive Resource Unit
DSL	-	Digital Subscriber Line
EA	-	Evolutionary Algorithm
EDRR	-	Enhanced Deficit Round Robin
EFS	-	Efficient and Fair Scheduling
ertpS	-	Extended Real Time Polling Service
FDD	-	Frequency Division Duplex
FEC	-	Forward Error Correction
FTTH	-	Fiber to the Home

FUSC	-	Fully Used Sub-Carrier
GA	-	Genetic Algorithm
GPI	-	Grant and Polling Interval
GSM	-	Global System Mobile Communications
HARQ	-	Hybrid Automatic Repeat Request
HOL	-	Head of Line
HSDPA	-	High Speed Downlink Packet Access
IEEE	-	Institute of Electrical and Electronics Engineers
IMT	-	International Mobile Telecommunication
ISI	-	Intersymbol Interference
ITU	-	International Telecommunication Union
LAN	-	Local Area Network
LOS	-	Line of Sight
LTE	-	Long Term Evolution
MAC	-	Media Access Control
MCK	-	Multiple Choice Knapsacks
MCS	-	Modulation and Coding Scheme
MIMO	-	Multiple Input Multiple Output
MLWDF	-	Maximum Largest Weighted Delay First
NLOS	-	Non Line of Sight
nrtPS	-	Non Real Time Polling Service
ODRR	-	Opportunistic Deficit Round Robin
OFDM	-	Orthogonal Frequency Division Multiplexing
OFDMA	-	Orthogonal Frequency Division Multiple Access
PC	-	Personal Computer
PHY	-	Physical

PMP	-	Point-to-multipoint
PRU	-	Physical Resource Unit
PUSC	-	Partially Used Sub-Carrier
QAM	-	Quadrature Amplitude Multiplexing
QoS	-	Quality of Service
QPSK	-	Quadrature Phase Shift Keying
RAS	-	Rate Assignment Strategy
RR	-	Round Robin
rtPS	-	Real Time Polling Service
SFID	-	Service Flow Identifier
SS	-	Subscriber
TDD	-	Time Division Duplex
TV	-	Television
UGS	-	Unsolicited Grant Service
UL	-	Uplink
USB	-	Universal Serial Bus
VoIP	-	Video over IP Protocol
WAN	-	Wide Area Network
WiMAX	-	Worldwide Interoperability for Microwave Access
WiFi	-	Wireless Fidelity

CHAPTER 1

INTRODUCTION

1.1 Overview and Motivation

Based on the current situation, Internet demand had increased tremendously. It is proportional to the invention of high technology communication tools that allow a wide range of application running at the same time. Furthermore the size of the tools is getting smaller and more convenient for us to mobilize and used it anywhere and everywhere. So with such demanding condition, hotspot which is created to cover a small range of area would be insufficient. The hotspot is developed under a standard known as WiFi which belong to the class of IEEE802.11. WiFi wirelessly connects users to the Internet backbone via an access point.

As mentioned earlier, limited coverage plus limited number of users that can be supported at one time are unfavorable. Users can only stay connected within the limited range and if the users are far from the access point, they will experience degradation of signal. Increase of demand requires more bandwidth capacity, larger coverage and support for moving users.

Thus IEEE802.16 which is also known as WiMAX was introduced. It is not just covering larger area but also supporting more user with higher speed. Now WiMAX is expanding to 4G system which corresponds to IEEE802.16m standard. But to realize the system is not as easy as one thinks. WiMAX is different from WiFi. So it cannot use the same system as WiFi does. Therefore new antenna installation needed for the network to run. WiMAX antenna is usually mounted on the highest building roof top to achieve non line of sight. Instead of the antenna the system is enhanced in order to support more users. The system must distribute the resources like the bandwidth intelligently since it is just not dealing with higher data transfer but also mobility. This is why we are investigating this topic specifically on how we could develop a scheduler which is responsible to distribute the available resources.

We are thinking about how to reduce the packet drop for congested area like in urban areas. But scheduler would face a problem in maintaining the connection of users with various applications and in various conditions. The main idea is that the scheduler should define priority so that the system would work efficiently in fairly manner. Notice that different service would demand different treatment in distributing the bandwidth. The treatment required for each service is defined by Quality of service (QoS). For example a user that wishes to make a phone call would need bandwidth allocation urgently than a user that is downloading data.

So the issue is that the scheduler must perform bandwidth aggregation based on priorities that is satisfying QoS while ensuring fairness towards other services. Fairness is a term that prevents any services from starvation. One of the cases is surfing Internet activity which is using best effort (BE) connection that does not have QoS requirement. We might just totally neglect the access of this kind of service since we are too concerned on the services that are more sensitive in delay. Eventually the BE data will be accumulated and the overwhelming data inside the buffer would likely result in starvation. Therefore a good scheduler must know how to treat each service class to obtain optimum throughput and being fair as well. In a simple way of explanation, the scheduler is important to help in keeping all services in demand still in connection. But scheduling is left as an open issue to allowing the vendors in differentiating their service type according to how they manage their networks.

There are two types of scheduling approach which are called downlink and uplink. The uplink connection poses more difficulties since BS does not have full information regarding data queue inside SS. BS only knows the bandwidth request from the polling session. Apart from that the scheduler also needs to be aware of the physical condition of each user in which moving user might posses poorer channel quality that results in lower bandwidth allocation.

Nowadays the realization of IEEE802.16m is still in progress. With the expectation that WiMAX could replace the 3G wireless WAN services like HSDPA, it is not surprising that even the portable device like the laptop or hand phone is equipped with WiMAX adapters. This proved that the world is already prepared for the upcoming emerging technology. There are three categories falls in wireless system which are:

i. Stationary

ii. Pedestrian

iii. Vehicular

Fixed WiMAX includes stationary and pedestrian users. Only mobile WiMAX can support all types of users including moving users. Unlike WiFi that is connected from access point via router or LAN card creating a limited coverage of hotspot, WiMAX consist only 2 main parts which is a WiMAX tower (BS) and WiMAX receiver (eg: embedded in mobile phone or laptop). Another advantage, WiMAX deploys flexible channel bandwidth selection which 3G uses a fixed channel bandwidth. Therefore allowing WiMAX to choose the best channel (frequency band) for the respective user to boost the system throughput. The flexibility is due to MIMO implementation in OFDMA physical layer. Further, the frequency diversity could exploit the system capacity and data rates.

One of the challenges of wireless broadband like WiMAX is in providing a competitive data rates compared to wired network service. It is known that wired network service relay the data through wired which is less vulnerable to environmental condition. But wired installation is more expensive and needs higher maintenance. Meanwhile mobile WiMAX uses air as the transmission medium. It more likely seems an ideal solution towards a cheaper and wider coverage area but wireless signal is exposed to unknown surrounding that may vary regularly. The situation would be fussier for the NLOS environment. As in urban environments users may be obstructed by the tall buildings. In order to overcome the matter the antenna is mounted on the highest building to get the most cleared signal path. Same goes to rural area where the transmission line is limited by mountains and trees. However wireless system could still experience the degradation of the received signal. The signal may affected by the multipath propagation and fading. Apart from that another factor like distance of the subscriber from the BS also contributes to signal loss. Therefore wireless technology would experience packet drop due to limited in bandwidth, packet jitter and longer delay for end to end user that gets affected by the surroundings.

1.2 Scope of project

In recent years, broadband has been in constant growth in telecommunication fields. This is proved by rapid increment of broadband subscription to over 200 million in less than 10 years. Outstanding features of the broadband is that it not only can provide higher data rates but also supporting other multimedia application such as real time audio, video conferencing, video streaming or online gaming attracting the consumer preference. It increases the quality of connection which then enhancing consumer lifestyle in seeking information, entertainment and dealing in business. Broadband wireless is about delivering the broadband signal via wireless means. Fundamentally there are two types of broadband wireless which are fixed wireless and mobile wireless broadband. Spurred by the arousing user demand various technologies like GSM, 2G, 3G, and WiMAX are evolved. WiMAX is first evolved from a fixed wireless broadband to mobile wireless. Is is first started in the year 1998 under the IEEE802.16 standard. The expansion of telecommunication field motivated the 3GPP and IEEE to find alternative solutions for a competitive wireless broadband services which are then called as advance LTE and mobile WiMAX. Both organizations strive to fulfill the 4G standard which is offering a very high data rate to support ruthless user demand. According to the standard outline by the International Telecommunication Union (ITU) 4G is expected to support fixed user at 100Mbps and 1Gbs for mobile user (Rakesh et al., 2010).

In this project we proposed how to allocate bandwidth to users with different service requested involved in uplink transmission. The method is developed especially for IEEE802.16m standard. The proposed method is called enhanced deficit round robin (EDRR) which is a modified version of DRR. The designed scheduler emphasized on aggregating bandwidth available to meet the QoS requirement first after which the left over bandwidth will be distributed in EDRR manner. In such a way we predicted that the scheduler will have lower packet drop probability and work efficiently in congested areas.

Since BS does not know the total information of bandwidth request for respective users, the scheduler is designed to work in both BS and SS. The scheduler at SS has access knowledge about the type and amount of data. Therefore once the scheduler at BS grants the bandwidth, SS will distribute them among the queued data inside the SS according to the QoS urgency level. For example user A is downloading a file from Internet while watching online movie. BS will grant appropriate amount of bandwidth to user A depending on the channel condition and the amount of bandwidth polling of all users. Then based on the granted bandwidth scheduler inside user A will divide the bandwidth and allocated them to the available services inside the data queue.

The process initiates from link budget analysis where the channel condition is reported in term of channel quality information (CQI). The amount of bandwidth polled is evaluated for bandwidth granted in BS. Lastly we map the data onto the Orthogonal Frequency Division Multiple Access (OFDMA) frame accordingly based on QoS restriction and EDRR. Simulation is performed in Matlab software where EDRR is compared against DRR as shown by the graph presented in Chapter 4.

1.3 Problem Definition and problem analysis

Voice call, video streaming, online gaming or checking email are examples of several services that are provided to end users. To keep the constant connectivity among them, QoS is set for each service. It is like a benchmark in describing the minimum requirement needs of each service to prevent the connection from dropping off. We can see that a service like the voice call is really popular nowadays. But transmission over wireless medium caused a significant problem on packet delay or latency. Jitter also degrades the reception quality of the signal. Thus to maintain a good voice conversation, delay and jitter should be considered. Same goes to other types of service like video streaming that is sensitive to delay but could tolerate latency. Therefore the idea is that in order to maintain the connection which eventually increasing the throughput, the QoS requirements should be fulfilled. All resources like power and bandwidth need to be distributed according to QoS characteristics.

When it comes to scheduling, it is well known that one of the crucial parts is meeting the QoS requirements. Each service demands a different way of treatment according to types of data supported. If the QoS requirement of the services is not fulfilled, the data will be dropped and retransmission is required. This does not just involve in losing of connection but also initiate retransmission that would make the network congested with the new incoming data along with dropped data.

Furthermore, apart from meeting QoS satisfaction, an efficient scheduler needs to be alert to the condition of the physical layer. Due to the distance and mobility of the users, the designed scheduler should adapt with the environmental effect such as multipath fading, Doppler effect and also shadowing. The fluctuations of environmental condition will surely affect the channel quality. Thus the scheduler needs to allocate a suitable amount of bandwidth to the respected users according to the bandwidth availability and channel quality so that it will be fairly distributed among users. Fairness is to abide the services that do not have any QoS requirement like Best Effort (BE) service class from starving. But the scheduling process is made differently between vendors to differentiate their priority of service.

In a communication system, scheduling is a crucial part. It is a process of sharing the bandwidth. For the uplink channel, request is sent from SS to BS. The grant from BS is embedded in the DL map sent from BS to SS along with the data intended for the respective users. While the UL map contains the information of the SS in demand namely the physical condition, modulation type and coding rate. The challenge is for the scheduler in SS to figure out the most suitable method to distribute the bandwidth fairly and optimized way without knowing the type of data inside the subscriber queue. Then SS will divide the granted bandwidth among the queued data according to QoS requirement.

The initial version of IEEE802.16 standard employs Orthogonal Frequency Division Multiplexing (OFDM), but nowadays it is preferable to use (OFDMA). OFDMA air interface improved the multipath propagation and provide the flexibility in allocating resources to user with multi data rate. It also enables prioritizing power allocation for the subscriber and matched them the based on channel quality. Inside OFDMA there are a number of slots. From the bandwidth granted via signaling process, backlogged data inside the queue will be organized according to the physical condition of the users and QoS. In WiMAX, slots will make up the frame. Each frame consists of uplink subframe and downlink subframe form in timeshared or known as Time Division Duplexing (TDD).

1.4 Objective of thesis

The objectives of the work are:

1. To propose a scheduler that would adapt to new standard outlined by IEEE802.16m while meeting QoS requirements for uplink transmission.
2. To study and evaluate the existing resource allocation methods in the current IEEE802.16 standard and enhance them.

1.5 Thesis contribution

The thesis contribution is:

1. Proposed a practical adaptive resource allocation scheme in the uplink direction which provides QoS satisfactions among connections for the IEEE802.16m standard.

1.6 Thesis Organization

This thesis is organized into 5 major parts:

Chapter 1: Introduction. In this part, an overview of the problem that will be solved in this thesis is presented. This part also includes objectives, motivation and contribution of the work.

Chapter 2: Literature survey. Here, the important subject of this thesis is defined (i.e. WiMAX evolution, resource allocation and IEEE802.16m standard).

Chapter 3: Proposed methodology. This chapter contains the proposed method used in this thesis and also the general view on the flow of the algorithm.

Chapter 4: Results and discussion. This chapter discusses the results of the proposed solution implemented and demonstrated through Matlab computer simulation.

Chapter 5: Conclusion. Chapter 5 is divided into two sections that are:

1. Conclusion for overall results. Clarifying overall flow of the project and presenting improvement that had been made.
2. Future work. Discuss about the future work of this thesis.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

WiMAX is an abbreviation of worldwide interoperability for microwave access which is first standardize under IEEE802.16 group on 1998. WiMAX is a rapid developing technology. With the added features of handover capability now WiMAX can be applied on mobile user. Despite that scheduling is still left as an open issue to help on vendor differentiates their service priority. This thesis will be discussing only on uplink connection. Mainly there are two types of WiMAX which are fixed WiMAX and mobile WiMAX. Figure 2.1 shows the evolution of IEEE802.16 standard.

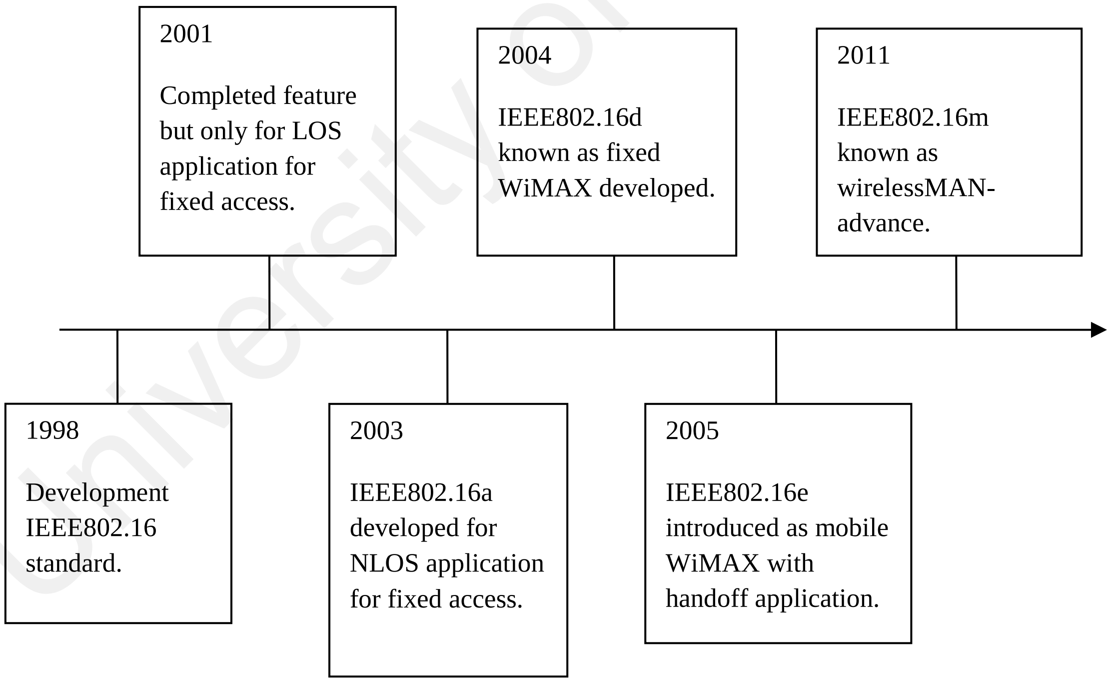

Figure 2.1: Development of IEEE802.16 standard (Ahmadi, 2011)

WiMAX is a desired network for the future because it offers a wide range of access. With WiMAX, terms like the 'wifi hotspot' could be no longer exist since the coverage can reach up to 3.5 miles in non line of sight (NLOS) condition and 2 miles in line of sight (LOS) condition. It usually supports OFDMA multiplexing technique which offers great advantages towards the establishes connection. This multiplexing type could diminish the Intersymbol Interference (ISI) as well as utilizing the multiuser diversity to achieve higher throughput. Besides it also mitigate multipath propagation and fading (Yen et al., 2009). In OFDMA one subchannel is divided into several orthogonal subcarriers. It is such like splitting a wideband channel into a group of narrowband channel while the smallest data allocation is called as a slot. A slot can contain different amount of symbol depending on the type of permutation scheme. Generally there are two types of permutation scheme for uplink which differ in term of frame format available (Papapanagiotou et al., 2009):

i. Partially Used Subcarrier (PUSC): Each slot consists of 2 OFDMA symbols and 24 subcarriers in one subchannel.

ii. Fully Used Subcarrier (FUSC): One OFDMA symbol in one slot and 48 data subcarriers in one subchannel.

The duplexing method deployed in WiMAX is the FDD and TDD. But this thesis only considers the TDD duplexing technique. Figure 2.2 is an example of the frame structure for uplink connection with the TDD duplexing technique.

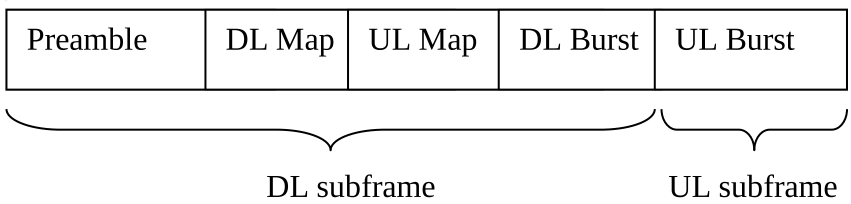

Figure 2.2: Frame structure

There are two types of connection in a network which called uplink (UL) and downlink (DL).

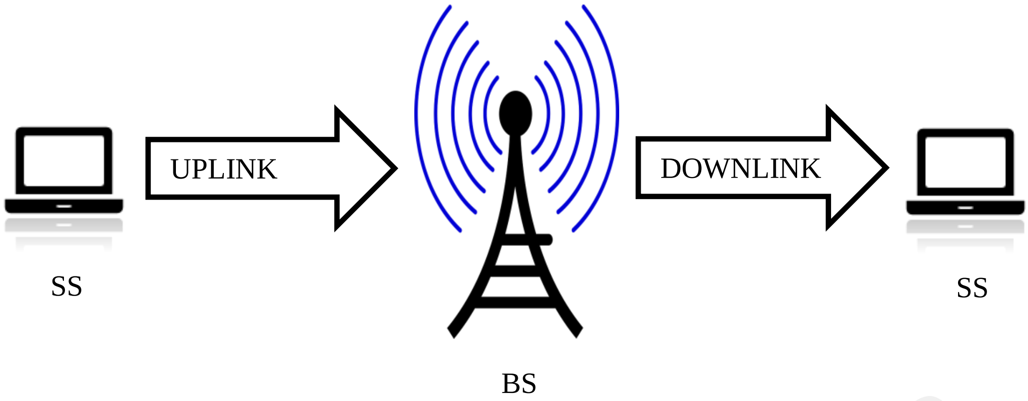

Figure 2.3: Network connection

UL is involving transmission from the subscriber (SS) to BS while DL is vice versa. DL transmission is easier to establish than UL because the BS already has knowledge of the packet size, type and length of data queued (So-in et al., 2001). Mainly there are two types of bandwidth request that categorized into implicit and explicit (Wang et al., 2008). Based on request bandwidth, BS will distribute the slots according to:

i. Quality of service (QoS).
ii. Resources: Bandwidth and power.
iii. Fairness: Equal allocation of network resources among various users operating either in good or bad channel state (Chowdhury & Misra, 2010).

2.2 IEEE802.16m Standard

IEEE802.16m is expected to be the next 4G technology. Therefore it has to meet the standard specified by International Mobile Telecommunication Advance (IMT-advance) which has speed of 100Mbit/s for mobile user and 1Gbit/s for low mobility user. To attain such achievement, IEEE802.16m is different from WiMAX 2009 standard in term of both PHY layer and MAC layer. There is also a new service class introduced which called as adaptive grant polling service (aGPS) (IEEE C802.16m, 2010). Ongoing progress will adapt new physical layer and MAC layer interface specified by IEEE 802.16 m standard. Since IEEE802.16m support mobile user, the mobility effect can be simulated by small scale fading, large scale fading and distance attenuation (Girici et al., (2008). To cope with the required high speed demand by 4G network, Physical (PHY) layer and Media Access Control (MAC) layer from IEEE802.16e is altered for adapting advance air interface. Overall, there will be few amendments in term of frame topology, new added service class, and altered modulation and coding techniques.

2.2.1 PHY Layer

New PHY layer design is employing advance air interface design to

i. Increase throughput by decreasing number of pilot carrier
ii. Reduce the access latency
iii. Support higher user mobility
iv. Minimize intracell and intercell interference
v. Improve reliability of control and data channel coverage especially at the cell edge.
vi. Decrease complexity and signaling overhead

The 4G technology extends the ability of communication further by increasing the speed and converging the communication range of mobile users. In legacy standard, the resource unit is a two dimensional slot consists of a group of OFDM symbols and subcarriers. The slot will be partitioned into pilot subcarrier and data subcarrier. Due to the high density of pilot subcarrier available bandwidth is diminished. Thus decreasing the throughput and resulting in a larger overhead. IEEE802.16m advance air interface tries to eliminate this drawback. The new frame structure uses PRU where each UL frame is divided into frequency partition while each frequency partition is divided into physical resource unit (PRU). PRU contain data and pilot subcarriers across the OFDM symbol.

IEEE802.16m is also designed to shorten the transmission time and accelerate HARQ transmission. Thus only 5ms radio frame is used for bandwidth of 5 MHz, 10 MHz or 20 MHz. To decrease the preamble length, 4 consecutive frames will form a superframe where a frame consists of n subframe. The number of subframe depends on the frame configuration which could be 8, 7, 6 or 5. Each subframe contains a number of symbols that varies upon the type of the subframe:

i. Type 1: Subframe consist of 6 OFDM symbols
ii. Type 2: Subframe consist of 7 OFDM symbols
iii. Type 3: Subframe consist of 5 OFDM symbols
iv. Type 4: Subframe consist of 9 OFDM symbols

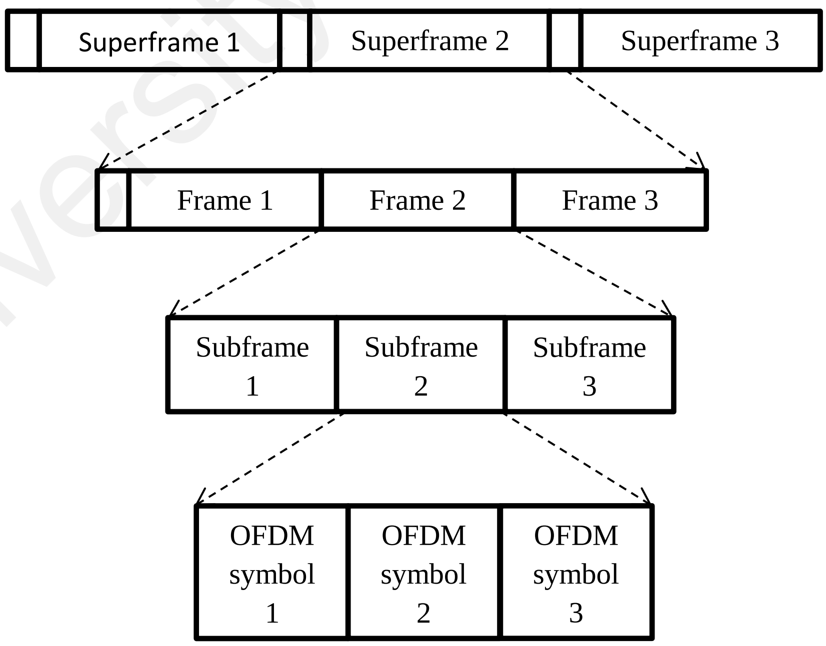

Figure 2.4: Uplink subchannelization and permutation for IEEE802.16m

According to standard, each UL frame is divided into frequency partition where each frequency partition is divided into physical resource unit (PRU). PRU is a basic resource allocation unit for resource allocation having the same size as DRU (including data and pilot subcarrier) which contains P_{sc} subcarrier across N_{sym} OFDM symbol. Number of P_{sc} is 18. While logical resource unit is a basic logical unit of resource allocation consists of $P_{sc}xN_{sym}$ subcarrier inclusive of pilot subcarrier that embedded in PRU. To achieve frequency diversity gain in multipath fading channel, a distributive resource unit (DRU) is introduced. DRU is comprised of a group of subcarrier across frequency partition.

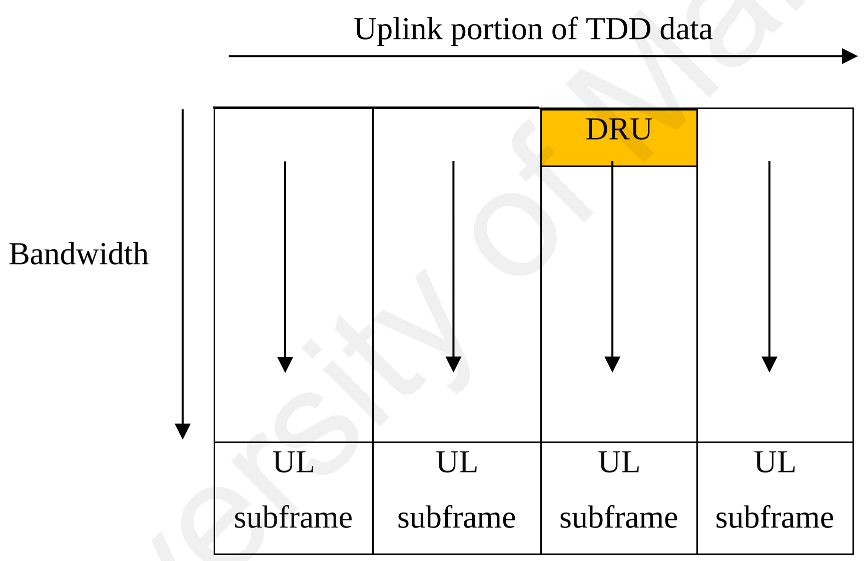

Figure 2.5: Frame topology

Acknowledged that IEEE802.16m is an amendments of IEEE802.16e in mobile WiMAX class. According to IEEE802.16e standard, BS chooses a suitable modulation and coding scheme from a set of supported modulation and coding schemes based on the channel quality. So there are a limited number of combinations of code rate and modulation order that can be selected. With fixed MCS, data burst is forced to fill in a number of slots even though not all slots will be fully occupied. Padding will fill up the unused space in the slot sacrificing available bandwidth. To overcome such situation, IEEE802.16m standard does not rely on MCS selection. The resource unit is divided into equally size FEC block with appropriate chosen level of modulation. Initially 16 bits of cyclic redundancy check (CRC) are added in a burst. If the addition of CRC and burst size exceeded the maximum FEC block size (4800bit), the burst will be partitioned into KFB FEC blocks. Thus all FEC blocks are identical in size and all space is occupied since it is partitioned according to data burst and not modulation level.

2.2.2 Mobile User

To support mobile networks more consideration is required rather than transmission via wired. This is due to variation of environmental surroundings and also position of end user leads to impairment of the communication channel (So-in et al., (2011). In a mobile communication system, transmitted signal may lose energy during transmission from transmitter to receiver. To simulate mobile user fast fading and slow fading will be considered in physical layer. Multipath fading and fading effect should be at attention in longer distance.

To get the channel gain $h_i(t)$ three types of loss need to be addressed

i. Distance attenuation

   Loss due to distance is a most basic loss in link adaptation analysis (Dai et al., 2006). It is due to decreasing of transmitting power that is spreading of radio waves in any direction throughout the transmission haul.

ii. Fast fading

   In urban area where the signal may obstructed by building or moving objects generating fast fading effects. It occurs when the signal component reflected, diffracted, delayed and scattered propagating in many paths towards receiver causing the Doppler shift where each multipath has own power gain and delay (Cai, 2003). A channel having only one fading path is called as frequency flat fading channel while frequency selective fading channel if fading path is more than one. Fast fading can be modeled by Weibull distribution, Nakagami distribution, Rayleigh distribution or Ricean distribution. This thesis will be considered the Rayleigh channel since it is a reasonable model for heavily built-up urban environments on radio signals and able to describe well on signal propagation. There are two main signal disturbance caused by fast fading:

- Multipath fading: Multipath fading channel occur when the signal travels through different path having different reflection from a transmitter with different arriving times. Thus causing fluctuation of the signal due to distortion and fading. Frequently used to model the multipath fading with no direct line of sight (LOS) path is Rayleigh distribution.
- Fading due to delay spread: Delay spread define at the time when signal first detects at receiver and the subsequent copies of it arriving at a delayed time. If multipath delay spreads relatively small to symbol period, the signal will experience flat fading and channel that signals propagate is in flat fading channel.

iii. Slow fading

Slow fading usually occurs when signal passing through a building and often represents by log normal distribution. Slow fading is usually associated with moving away from the transmitter and experiencing the expected reduction in signal strength.

In order to develop a good scheduler to distribute the resources it is essential to have the knowledge of the wireless channel characteristic. Apart from helping the researcher to create an efficient scheduler but it also critical for means of analyzing the performance of the system.

2.2.3 Quality of Service (QoS)

WiMAX is expected to deliver high speed data proficiently either in point-to-multipoint (PMP) or mesh operation (Kumar & Priyameenal, 2011). In PMP mode, a single BS is served by multiple users. Therefore for UL connection, all subscribers will share the same bandwidth in means of transmitting data to the BS. That allows us to get the insight of Quality of service (QoS).

QoS refers to an outline created to guarantee the performance of networks working within desired level. QoS includes in guiding how to treat the services according to their types. Thus in a complex traffic condition scheduling will have the significant effect on QoS satisfaction. It is known that IEEE802.16e only support five types of traffic services which is UGS, ertPS, rtPS, nrtPS and BE. IEEE802.16m further adding a new service class known as aGPS (Ahmadi, 2011). Therefore there will be total of 6 service classes to be scheduled. QoS concept is still the same where to associate packet transported through the MAC interface that are identified by transport CID to a service flow. Service flow is a unidirectional flow of packets either to UL/DL that is characterized by latency, jitter and throughput. All service flow identified by a 32bit service flow identifier (SFID) and each admitted/ active service flow has a 16 bit CID. The characteristics are:

i. The BS may grant or poll an MS periodically.
ii. Allow adaptation of grant and polling interval (GPI) and grant size according to traffic condition. For example the silence suppression enabled VoIP alternate between talk spurt and silence period.
iii. Mandatory QoS parameter are:

a. Maximum sustains traffic rate
b. Request/ transmission policy
c. Primary GPI
d. Primary grant size

Basically the services are divided into two groups which are real time (UGS, ertPS, aGPS and rtPS) and non real time (nrtPS and BE). UGS generates a real time data with a fixed data rate. So VoIP is not suitable for UGS application since it would wasting the bandwidth once the service is in its off state period of the voice call. Therefore VoIP is more compatible for ertPS application where combining UGS and rtPS characteristics allowing ertPS to deal with both real time and variable data rate (Adhicandra, 2010). aGPS is similar to ertPS except that during active period the data may vary in sizes between the primary and secondary parameter value (IEEE C802.16m, 2010). While nrtPS poses the same criteria as rtPS except it is interested in non real time data. Whereas data that is not sensitive to delay or jitter such as FTP or email would choose BE application. Table 2.1 below shows brief information regarding to the services available in mobile WiMAX.

Table 2.1: Characteristic of service flow defined in IEEE802.16m

Service flow	Characteristic	Activity
UGS	Real time service flow Fixed size data packet Delay sensitive	VoIP
rtPS	Real time service flow Variable sized data packet Delay sensitive	Streaming audio and video
ertPS	Real time service flow Fixed size data packet with silence suppression Delay sensitive	Voice with silence suppression
nrtPS	Delay tolerant data Variable size data packet Required minimum sustains data rate	File transfer
BE	Nonreal time service flow Variable data rate Do not have QoS requirement	Web browsing and email
aGPS	Real time service flow Variable size data packet with silence suppression Delay sensitive	Online gaming

Table 2.1 clearly shows that each service flow has its own limitation to be focused on such as delay, minimum throughput and minimum data rate. Since the services may have different satisfaction degree they should be addressed accordingly. Notice that only BE does not have any QoS constraint. Since BE is tolerable with delay and data rate, allocation for BE is usually granted after all the other services had been allocated the resources (Yen et al., 2009). It is observed that there are few types of QoS requirement need to be address on:

i. Throughput: Data rate transfer by the network usually in bit per second unit.

ii. Delay: Delay may be defined as the distance between packet and destinations. Some of the services just really sensitive to delay especially on voice application. Each service has different durability on delay except for BE which is does not have delay requirement.

iii. Jitter: Jitter has hardly been discussed by researcher. Jitter is not desired because it will lengthen the preamble length. It can be described by the distance between two consecutive slots of the same type of service from same subscriber type.

2.2.4 Modulation and Coding

Previously in IEEE802.16-2009 standard, BS chooses suitable modulation and coding scheme from a set of supported modulation and coding scheme based on channel quality. So there are a limited number of combinations of code rate R and modulation order M that can be selected. With fixed MCS, burst size would be different over unequal subframesize. To overcome such situation, IEEE802.16m standard does not rely on MCS selection. A cyclic redundancy check (CRC) of 16 bit size is added into a burst. If $CRC + Burstsize > 4800bit$ (exceed maximum FEC block size), the burst partition into KFB FEC blocks. If the burst partition into more than 1 FEC block, a 16 bit FEC block CRC is added to each FEC block. IEEE802.16m using convolutional turbo code with a minimum code rate of 1/3 (Ahmadi, 2011).

Modulation level will determine how much bit of data can be fitted inside DRU. Whereas physical layer only support burst size (N_{DB}) listed in Table 5. The more robust modulation level will allow more bits to be placed inside the block meanwhile lower modulation level causing data to be spread across the frequency increasing the bandwidth consumption thus decreasing the capacity. So basically each of data blocks will be in the same size, the only different is how much data can be stored inside a block. That is determined by the modulation level obtain based on channel conditions.

2.3 Scheduling

The important parts in ensuring QoS satisfaction are scheduling. The scheduler will act as a distributor to control the access to available resources such as power and bandwidth. It has to be acknowledged that an efficient scheduler should not only criticize on resource utilization in order to achieve the QoS goal but maximizing the throughput and fairness as well. Scheduling actually perform in the Media Access Control (MAC) layer. As explained in the figure below there are three sub layers constructed in the MAC layer.

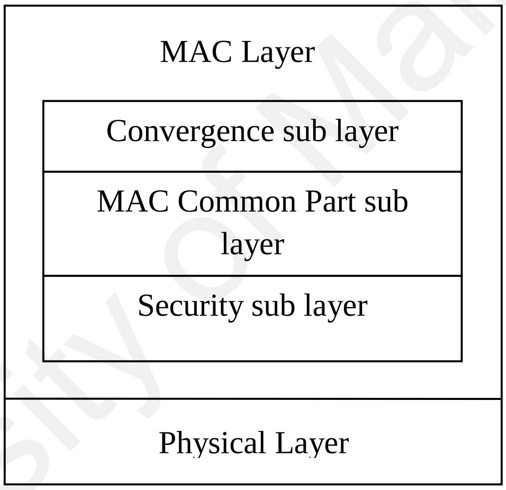

Figure 2.6: IEEE 802.16m architecture

The first layer which called as service specific convergence sub layer (CS) classifies all data received from the upper layer in form of MAC service data unit (SDUs) into a service flow using service flow identifier (SFID). Then the connection identifier (CID) that containing the temporary address for the transmitted data packet is added to the service flow. The second layer is known as MAC common part sub layer (MAC CPS). The allocation of bandwidth (request and grants), scheduling and QoS control is performed here. Initially SDUs is formatted into the MAC protocol data unit

(PDUs) to be passed to the physical layer. PDUs may in the form of a burst having same modulation and coding. PDU containing:

i. Service flow ID (SFID): Define QoS parameter of the service flow associated with the connection.

ii. Connection ID (CID): Identifying the transport connection over which information is delivered.

The last sub layer is security sub layer where the encryption of data and security key is managed here.

Scheduling mechanism is started once the SS polled the bandwidth to BS. The request is sent in the form of CID format (Wang et al., 2008). It includes the size of bandwidth and channel quality information (CQI). CQI is indicating the environmental condition that is generated from a feedback source in order to synchronize power allocation in SS (Reddy et al., 2007). Based on the reported physical layer information from CQI, it helps the BS to determine the modulation level and coding scheme. Once BS had decided the bandwidth aggregation to the respective user, it will send UL-MAP in DL subframe. DL subframe contains the information of grants bandwidth and also the downlink data. But BS only grants the bandwidth. Thus, the scheduling process is still not finished yet. The SS then plays its role by distributing the resources among the queued data.

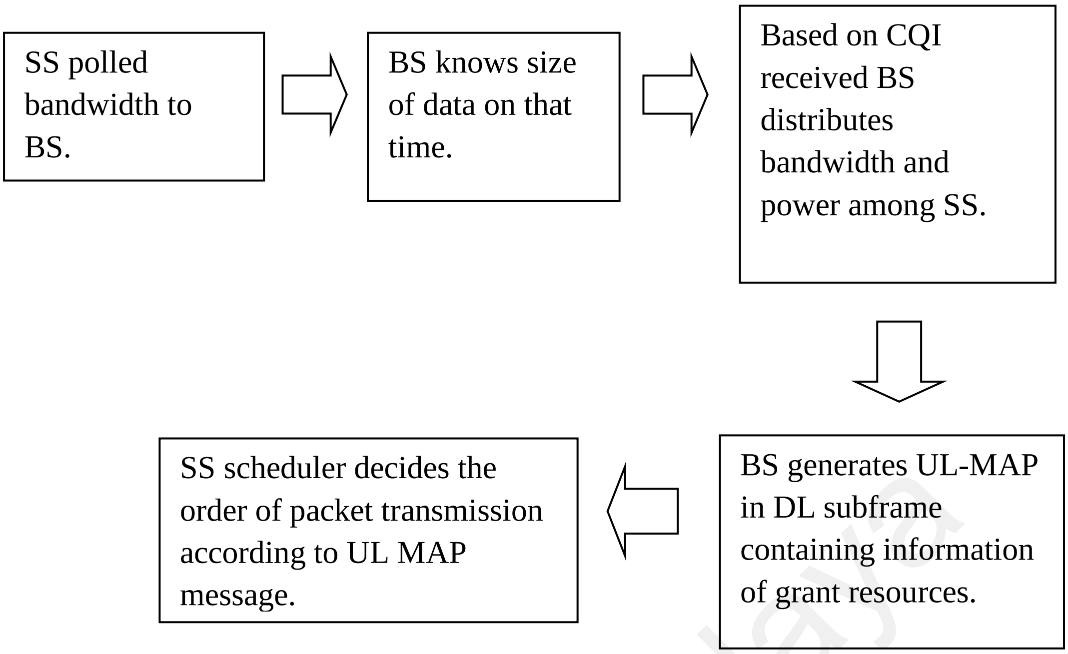

Figure 2.7: Flow for Uplink scheduling

2.4 Scheduling method

Rapid development of WiMAX showing there are so many interests pours in this technology to make it better. But for the scheduling problem regarding IEEE 802.16m standard still not widely discuss yet. It is looking forward to see upcoming research on this standard. Some of them did not consider power allocation or discuss on jitter. But the attempt to solve scheduling problems using evolutionary method is rising. Since the evolutionary method could handle the vast number of parameters while searching for the optimum data aggregation it is seen that there is a possibility they will be excelled in distributing the bandwidth. Despite that, effort of the researchers that trying to combine the optimization method is also interesting. With the rapid findings in the research field WiMAX may become a useful technology in years future. Scheduling approach can be classified into two major groups which are optimization method and Evolutionary algorithm (EA) method.

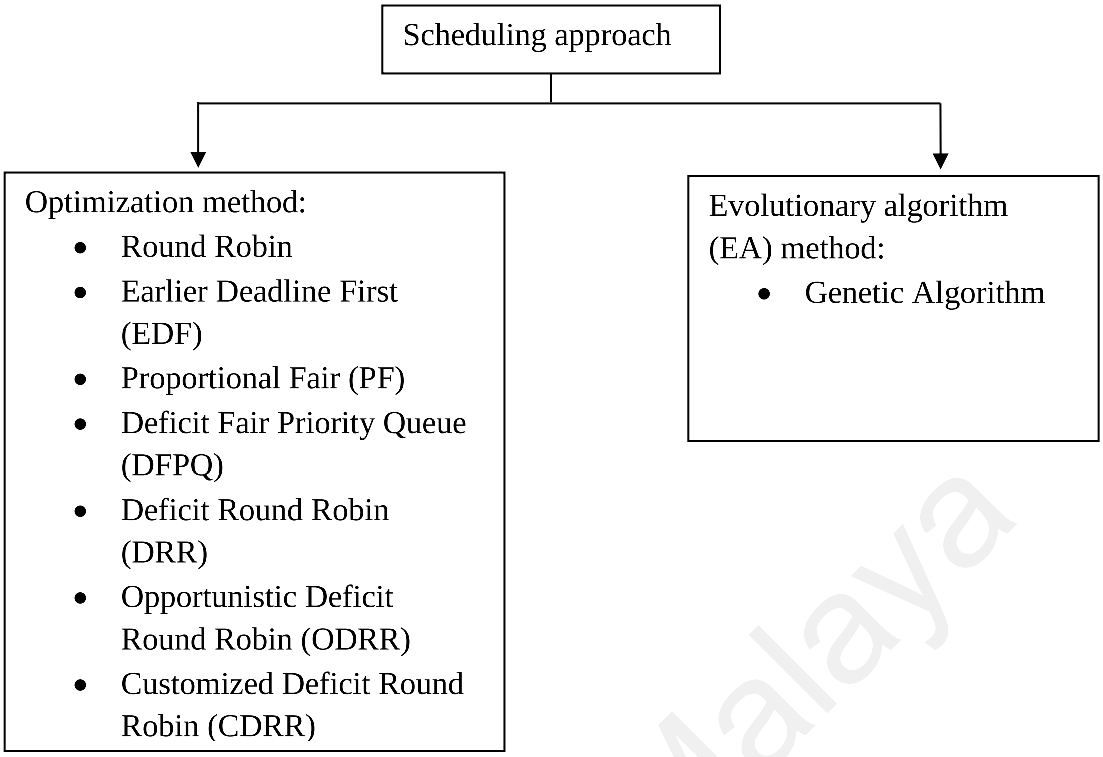

Figure 2.8: Scheduling approach cluster

Nowadays EA is widely used in various research fields to solve optimization problem. Scheduling problem also not excluded. Researchers like Chiu *et al.* (2010), Yen *et al.* (2009) and Reddy *et al.* (2007) are making an effort to adapt the scheduling process into biological resolution like Genetic Algorithm (GA). But due to its simplicity optimization approach is still being commonly used. Instead of distributing available bandwidth, scheduling problem also facing all the listed constraints:

i. Large overhead that will diminish the capacity of available bandwidth.

ii. The physical layer condition that is may varies and unpredictably. The environment of transmission path through the air may trigger the noise, interference and multipath propagation that would cause delay and also packet error.

iii. Packet drop due to unsatisfied QoS requirement.

iv. A scheduler that would provide fairness towards all user and services.

Despite that not all researchers consider those limiting factors.

2.4.1 Optimization method

The optimization method is usually easy to be implemented yet robust. But most of them are combined with another method to make it reliable in handling both real time and non real time application. Some researcher like Chowdry *et al.* (2010) combined two methods to satisfy QoS requirement of real time and non real time separately. While Wang *et al.* (2008) proposed a new combination of two optimization methods to enhance overall system.

Round Robin (RR) performing distribution of a set of task in circular action. But it does not consider on priority of services involved. Therefore to overcome the situation, Wang *et al.* (2008) combining RR with Proportional Fair (PF) algorithm. PF will act as controller to provide balance between the throughputs of the overall system and data rate served for each user. By this way overall algorithm is improved. They only consider for rtPS, nrtPS and BE traffic allocation since UGS need only once bandwidth polled (unsolicited granting) therefore same bandwidth size will be given to SS. The main server will organize all services available onto the frame according to priority with condition only one packet is allowed to transmit at a time. But if there are many requests of the same level of priority they will be served in round robin fashion. Since this algorithm is applied for scheduling user data that the bandwidth request has been granted by BS, they can assume SS has sufficient bandwidth for their queued traffic. Thus the algorithm responsible only in deciding which data packet transfers first. It is quite simple yet can be implemented for both real time and non real time. Distribute in RR fashion would not take packet size into consideration. RR will transmit one packet of each source repeatedly. Since services like nrtPS, rtPS and BE have variable data rate, this will create unfairness towards the users. Despite that this algorithm only responsible in distributing available grants bandwidth which is suited only for subscriber scheduler. They also did not consider the physical influence on transmission to make the algorithm more realistic in real life application.

Chowdry *et al.* (2010) using a different type of approach to separately solve different traffic requirements. The aim of the method proposed is to determine the delay of real time service while providing minimum sufficient data rate transmission for non real time application. The algorithm developed is combining Earliest deadline First (EDF) algorithm and Deficit Fair Priority Queue (DFPQ). EDF is for rtPS while DFPQ is for nrtPS and BE service. DFPQ is actually functioning just like ordinary deficit round robin method except for quantum size will decide either the packet can be transferred at that time or wait until next round robin round. Quantum size will help in distributing bandwidth fairly among the variable size data packet. Using both methods they find minimum bandwidth needed to satisfy QoS for each user. Then the remaining bandwidth is aggregated according the priority UGS>rtPS>nrtPS>BE.

Instead of combining the two methods, Rath *et al.* (2006) and Lee *et al.* (2006) use a single approach to deal with both real time and non real time communication service. Rath *et al.* (2006) proposing an enhanced version of Deficit Round Robin (DRR) which is known as Opportunistic Deficit Round Robin (ODRR). DRR is an improvement over RR and WRR (Ravichandran et al., 2010). A unit call quantum number will control the distribution of bandwidth. Each connection is assigned a state variable called DC (Deficit Counter). At the start of each round, DC_i of queue i is incremented by a specific service share (quantum). If the size of the packet, L_i, is less than or equal to DC_i, the scheduler allows the i^{th} queue to send a packet. Once the transmission is completed DC_i is decremented by L_i. The scheduler will repeatedly give bandwidth for each user in round robin manner until the quantum number of the user is deficit. As shown in Figure 2.9 the quantum number is decrement by the size of the data packet. The packet continued to be sent until the quantum number is obsolete (Kumar &

Priyameenal, 2011). Only then DRR will continue to serve another user (Kalyana & Reddy, 2009).

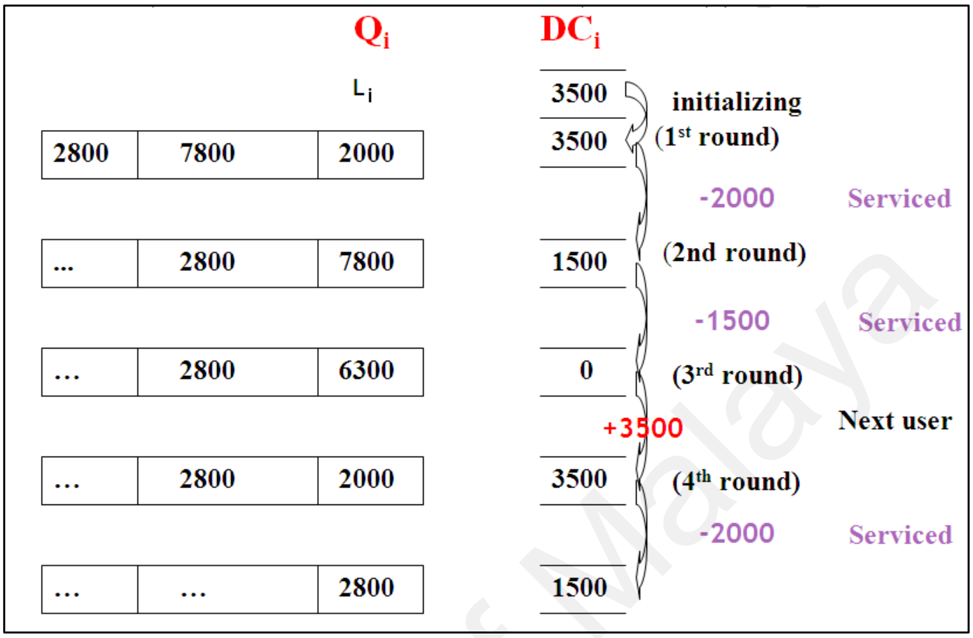

Figure 2.9: DRR assignment process

ODRR is responsible in distributing bandwidth of a set of the eligible user. The algorithm starts by selecting a group of user to be transmitted at that particular time. They decide it by using polling interval, *k* that represent the time chosen to give information about data queue by bandwidth polling done in BS. Thus the value of *k* will directly affect the fairness and delay of traffic service. In order to get the value of *k*, they make an assumption of two polling intervals. From that they calculate how much data rate can be transferred on that interval. But the value of *k* will be varied upon how the traffic conditions are. If the traffic is mixed up with various services minimum *k* will be used for polling mechanism otherwise vice versa. After that based on their weight which is in term of delay counter, they assign available bandwidth to active user group formed before. Assumption of delay is 50ms for rtPS, 200ms for nrtPS and 500ms for BE. The delay counter is the only parameter deciding which user transmits first to avoid

QoS violation. This algorithm is unique since not many researcher concerns about the polling interval in scheduling problem. Unfortunately it only focuses on the delay requirement whereas nrtPS QoS requirement is to have minimum data transfer rate. Therefore it is actually more suitable for real time application.

Another enhancement of DRR is the Customized Deficit Round Robin (CDRR) proposed by Laias *et al.* (2009). Preserving the simplicity of DRR method, CDRR is designed to improve in delay and fairness aspects. The services involved in this research are UGS, rtPS, nrtPS and BE. Each of the services has their own queue in the UL connection. CDRR started by emptying the UGS queue followed by rtPS data which nearing the deadline. Remaining bandwidth is then filled by rtPS and nrtPS data packet which served in DRR fashion. The last priority is going to BE service which not having any QoS boundaries. BE is served in FIFO manner.

Iyengar *et al.* (2009), Peng *et al.* (2007) and Lee *et al.* (2006) only consider real time traffic management. Iyengar *et al.* (2009) is proposing resource allocation for UGS and rtPS. In realizing the algorithm they assume that channel condition is same throughout time interval t and UGS transfer of a fixed amount of data. The main objectives are finding maximum throughput of system while satisfying the QoS of each subscriber. To satisfy both objectives, they divided the user allocation on subcarrier into two parts. First they assign user subcarrier to satisfy the demand of each user. If the chosen subcarrier able to satisfy the user needs, they try to allocate the residual slot to get the highest possible throughput. So the idea is to map user and subscriber that would result on maximum data rate. Hence QoS is satisfied and at the same time subcarrier is effectively utilized by suitable user. Since the value of SNR will affects power allocation and also number of bits can be allocated to the user, they try to solve this problem using bipartite graph approach. Bipartite graph is used to solve mapping problem between two disjoint sets. With assumption of one user can allocate to many subcarriers but one subcarrier cannot be shared by one user they mapped them using poly matching. With constraint demand of each user that must be satisfied, they distribute the power using waterfilling technique. But to minimize power consumption less number of subcarrier is allocated to one user. Compared to two other greedy algorithms that allocate user one by one to subcarrier, the proposed algorithm achieves the highest throughput that is always near to optimal throughput attainable.

Peng *et al.* (2007) only focus on scheduling rtPS communication service. It is well known that the real time application is sensitive to delay. But the BS only allocates bandwidth that is requested by SS. To get the bandwidth allocation SS need to wait until next frame and maybe at that time there are new rtPS data arrive. Therefore the queue is lengthening and to wait next granted bandwidth would generate unnecessary delay. They got the idea to solve this problem by making a prior estimation of bandwidth need by BS before the request is polled. The prediction is made by using the Lagrange interpolating function. The Lagrange interpolating function is commonly used to find optimizes solution between variable parameter without changing certain value. This method applies to estimating between how many data arrive and their time width to get the number of estimate time slot required for an SS. But since this only a prediction it is not always corrects and might leads to wastage. If there is another service type request for bandwidth, this would be violating the QoS requirement for others. Therefore prior estimation might be infeasible for the whole network.

Since ertPS have silence suppression and variable data rate, a specific scheduler is needed for this type of communication service. Therefore Lee *et al.* (2006) introduced a special technique to aggregate ertPS data in uplink connection. The condition of talk spurt and silence is known by the MAC sub layer called as Convergence sub layer (CS layer). Therefore BS will always know the condition of the voice they are handled with. When the talk activities are decreasing they change the bandwidth polling service to piggy back request type to reduce MAC overhead. Otherwise if the talk activities are increasing, they change back to bandwidth request mechanism to provide the larger bandwidth polled capability. So the concept is when the bandwidth request is expected to decrease, the bandwidth allocates or polling service will be small. Thus preventing from wastage of frame utilization. This method really designs specifically for ertPS characteristic therefore it is more likely to excel among other method.

2.4.2 Evolutionary Algorithm Method

Numerous researches exploiting Evolutionary Algorithm method considers on power and jitter. Jitter is solved by allocating one subscriber only to one user as proposed by Chiu *et al.* (2010). Besides they are also simulating noise in physical layer that make the scenario more realistic. For resource allocations that take only milliseconds it is infeasible to solve with computational complexity that affects the time consume. Therefore Chiu *et al.* (2010) adapt the resource distribution in UL into an intelligent method called Genetic Algorithm with SS grouping resource allocation (GGRA). The algorithm works based on GA operation which is a biological mimic's algorithm to find maximum throughput and meeting QoS needs while confronting the limitation in power and bandwidth. The solution search is based on randomization that iteratively enhance by selecting the best solution. GA also can be controlled by the number of iterations according to the time allowed. The method starts with grouping the SS according to their channel condition to cancel the effect of user interference. Based on selected group GA will distribute available bandwidth among them. The algorithm use rate assignment strategy (RAS) as a benchmark to decide the urgency level of all data services in the same group. RAS is based on residual lifetime. Residual lifetime is defined as the remaining frame need to be sent to particular service traffic head of line (HOL) packet, or else the packet will drop. The user that has minimum residual lifetime will be set to have the highest priority. To decrease overhead length, data in same user will consistently allocate onto the frame. After the respective user finishes the allocation, next user chosen from RAS will continue to fill in the frame. This process repeated for the next group until the frame full or all users data is obsolete. But as dealing with bad channel condition the length of the data burst need to be greater to avoid interference. Therefore the lower level of modulation is chosen. Lower level of modulation is good in term of power saving but bad in term of bandwidth consumption and throughput. Hence for respective user channel condition GA will try to find the most suit power to provide maximum throughput in the overall system. A set of available solution is generated by GA using mutation and crossover operation coming from parents chosen by fitness evaluation. The process will keep on circulating until the maximum iteration fulfills or maximum level of modulation which is 64-QAM is found. Eventually the algorithm would find the maximum throughput while guarantee QoS. GA is proven to be more effective in mapping data to whole frame instead of Efficient and Fair Scheduling (EFS) and Maximum Largest Weighted Delay First (MLWDF) which is slot by slot. This method greatly covers all of QoS requirements which are jitter, delay, power, throughput and packet loss.

Yen *et al.* (2009) proposing a Dynamic Priority Resource Allocation (DPRA) for UL connection. This method gives priority to all 4 multimedia applications regarding on their urgency level. The priority order will be UGS>rtPS >nrtPS>BE. They are also considering power allocation like Chiu *et al.* (2010). Despite that the algorithm allocated to respective SS on a subchannel according to priority not type of SS so jitter will be high thus increasing the preamble length. Initially they find the most urgent user to transmit their minimum required bit, then they find the optimal pair of subchannel for that user. After that they constantly allocate all user data inside the frame. They are proposing constant allocation where one subchannel that is selected by that user is constantly filled with all service types in order UGS, rtPS, nrtPS and BE. But if the bandwidth of respective subchannel is not enough next consequence subchannel is considered to be filled in with. But this time power is recheck back to match with that subchannel condition. There are possibilities that the number of slots required to fill in the remaining data might be decreasing. But the problem is if the next subchannel is already filled by another user, they need to be moved to other subchannel. This they called remapping. Allocation user specific for just one subchannel is good to diminish overhead length. But to allocate same user consistently would caused frame fully utilized before other urgent user to have the chance to transmit. So it is not fair for all users in the system.

Reddy *et al.* (2007) also use GA as a technique to compromise between user data rate and power. GA is desired because the parallel searching ability that minimize the time consume to find the solution even though when the number of subcarrier and user is increasing. The idea is by mapping subcarrier to a user in order to meet their data rate requirement while at the same time controlling the power consumption. The best match is done by confirming the fairness while distributing power using waterfilling technique. Waterfiling technique is a popular method of distributing power. This method is meant to fully utilize available power to achieve high throughput by aggregating lower power to poor channel condition and high power to better channel condition.

The interest pours in scheduling strategies had widened the idea in solving this matter. From all of the existing methods it is aware that even simpler method could give good performance. We realize that even though the method is seems easy to implement but if it could be enhanced, it will have potential to schedule the resources in a better way. Driven by Rath *et al.* (2006) and Laias *et al.* (2009) the noncomplex method like

DRR showed ignition of an improvement by enhancing it. Thus we come to the idea of improvising DRR with EDRR.

CHAPTER 3
METHODOLOGY

3.1 Research Methodology

The main task of a scheduler is to distribute bandwidth according to QoS necessity in a fairly manner. Failure of granting the bandwidth to the user which has strict QoS would be a catastrophe. We realized that BS would only know how much bandwidth are requested by each user from polling process where the SS sent information on how much bandwidth is needed. The request is sent in UL-MAP along with the uplink data. Due to this reason the scheduler cannot be solely placed at BS. Hence there will be one part of granting the bandwidth at the BS while another is distributing the bandwidth among the data queue inside the SS.

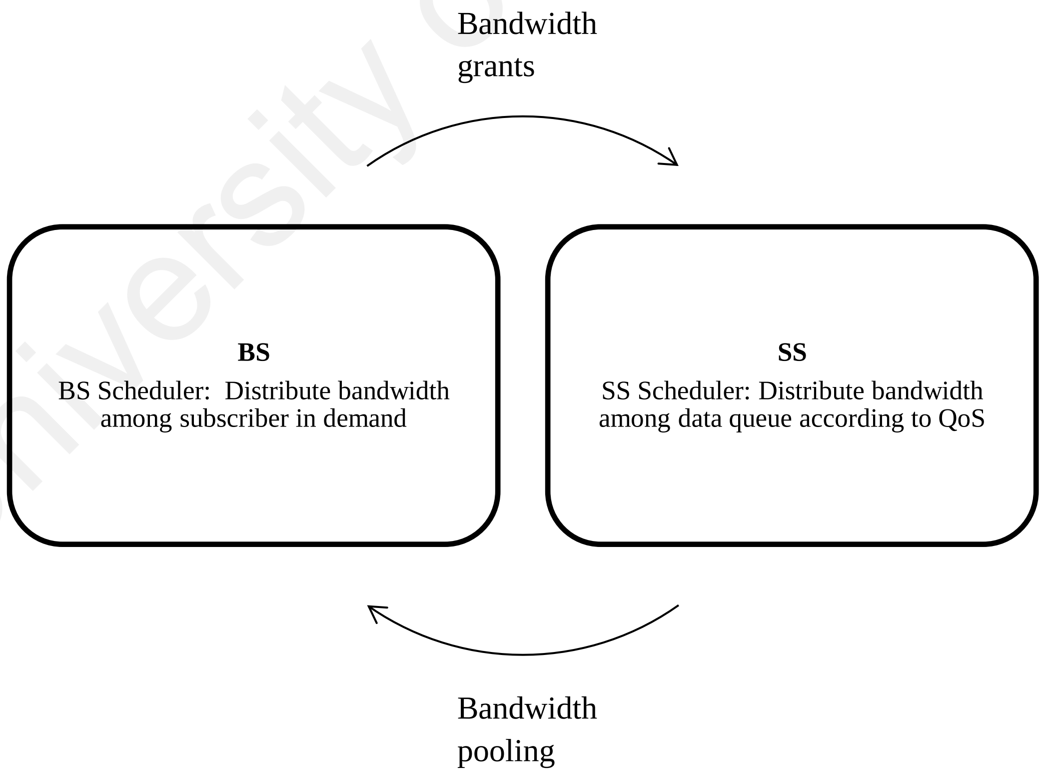

Figure 3.1: Scheduler

3.2 Scheduler

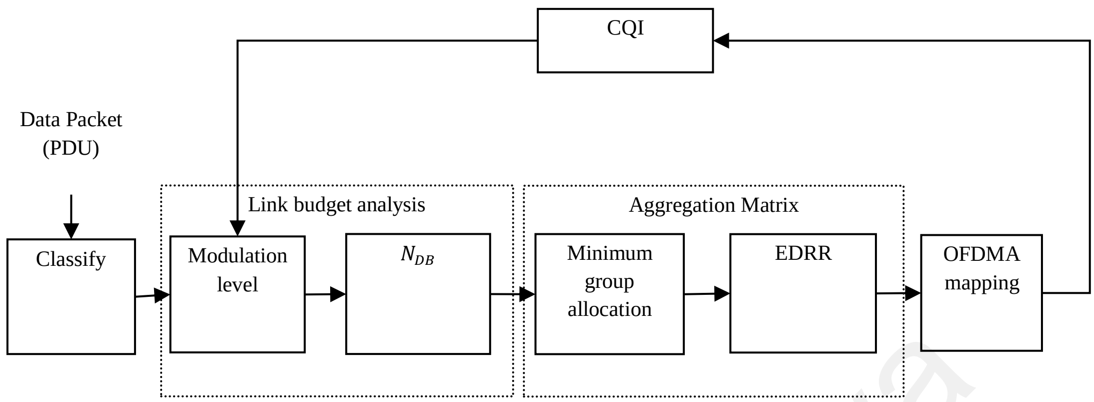

Figure 3.2: Proposed scheduler

Figure 3.2 illustrate overall operation of scheduler proposed in this thesis. Initially all data buffers on each subscriber at time t will be examined and grouped into six types. The data will stay inside the queue waiting to be served. At the same time based on the status reported by CQI, modulation level is determined for each user. Modulation level will affect the amount of bit that could be fitted in one DRU (N^{DB}). Then scheduler will calculate the bandwidth required for minimum allocation groups. If there is remaining bandwidth available, EDRR will commence. Once the bandwidth is fully used, all data will be mapped into the OFDMA frame. In conclusion the scheduler will guarantee the QoS requirement first before improving the throughput via EDRR.

3.2.1 Minimum Group Allocation

Minimum group allocation is a part in which to ensure the QoS restriction strictly followed. Refering to the characteristic of service classes mentioned in Chapter 2, it is known that UGS, ertPS and aGPS are very sensitive to delay. Therefore all data belonging to this type ($\delta_k^{UGS}, \delta_k^{ertPS}, \delta_k^{aGPS}$) will be straightly granted. rtPS service is also delay sensitive but it is more tolerant than the other two services discussed earlier. Hence packets queue of rtPS (δ_k^{rtPS}) for each user k is set to have delay counter(D_k^c)

which must be lower than the threshold delay (T_d) to prevent packet drop. The condition will be:

$$D_k^c < T_d \text{ where } D_k^c = nT_f \tag{3.1}$$

Where: n: Number of frame elapse since last transmission

T_f: Duration to transmit a frame

Unlike all services mention earlier, nrtPS is sensitive on minimum data rate but not delay. For nrtPS, each SS containing this type of data must send minimum of w_k packets to meet the minimum throughput required during every frame transmission. Since BE services do not have any QoS requirement, it will not participate in the minimum allocation group. After all minimum data have been allocated, they will be scheduled on the frame based on the urgency level. EDRR will start operating if there is remaining bandwidth left after serving the minimum allocation group. Since IEEE802.16m works for both static and mobile users, we want to emphasize more on the mobility effect. Therefore the distance between MS-BS is made equal for all users. Thus the physical condition will be mostly affected by the subscriber speed (v). The method is further illustrated by a flowchart in Figure 3.3.

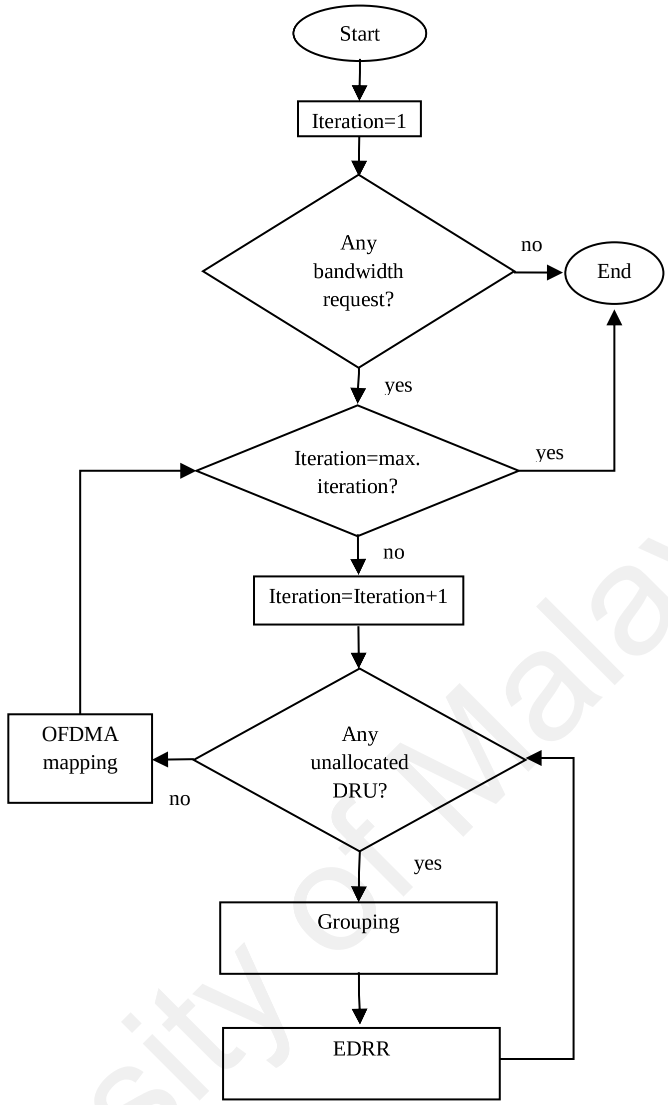

Figure 3.3: Flow chart of the proposed method

3.2.2 Enhanced Deficit Round Robin (EDRR)

EDRR is proposed to reduce the amount of packet drop. EDRR which an improvement over DRR method is designed to overcome the DRR impairment in determining the quantum number. EDRR also used the same assignment concept as DRR which is using the quantum number to organized the flow. Just how to calculate the quantum number would be different. We know that the quantum number indicates how much data can be sent to each user. In other word quantum number is the weight of bandwidth grants in the scheduling process. The main idea of EDRR is the quantum number will not only depend on the user channel state but also the amount of data queue. All subscribers in demand are first divided into four groups according to the speed of movement. According to ITU-R M.2134 (2008) the following classes of mobility are defined:

$$Group = \begin{cases} static\ user: v = 0km/h \\ pedestrian\ user: 0km/h > v > 10km/h \\ slow\ moving\ user: 10km/h > v > 120km/h \\ fast\ moving\ user: v > 120km/h \end{cases} \tag{3.2}$$

The quantum number would be differ according to the group type. User that has the worst channel condition would get a lower quantum number. For instance, a static user will always get lower bandwidth grants than a moving user due to the channel condition. But the user of equal physical condition does not necessarily have the same amount of quantum number. It means that within the same group if the user has less bandwidth request it will have a less allocation inside the frame. This will give the chance to user with more data to be sent and able to loosen up their queue so that it will have the space for incoming data packet. Eventually it will decrease packet drop for that particular user and thus enhancing the whole system. Therefore there will be two filter in calculating the quantum number. First the user is filtered into the respective group according outline by ITU, then the amount of data queued which belong to same type is analyzed. It is expected that the weight of the bandwidth grant is directly proportional to the traffic intensity. Thus two conditions is considered:

First condition: *If all users have a different physical condition (different group). Users which belong to different group will have a different amount of quantum number.*

$$Q_i = \frac{num\ bit\ DRU\ SSi}{num\ bit\ of\ FEC\ block} \tag{3.3}$$

Second condition: *If the users are in the same group*

If the users have been in the same group but the request bandwidth is not equal, EDRR will give more bandwidth to heavy loaded users.

$$Q_i = \frac{\text{Bandwidth request by } SSi}{max\ no\ of\ queue\ in\ the\ group}\, x\, \frac{num\ bit\ DRU\ SSi}{num\ bit\ of\ FEC\ block} \qquad (3.4)$$

Figure 3.4 further illustrate the pseudocode of the scheduler in detail.

```
Do

For frame=1:maximum frame

  Minimum DRU for UGS,
```

$$\beta_{(UGS)} = \frac{\delta_k^{UGS} x\ Num\ bit\ FEC\ block}{Num\ bit\ DRU\ SSi}$$

```
Minimum DRU for ertPS,
```

$$\beta_{(ertPS)} = \frac{\delta_k^{ertPS} x\ Num\ bit\ FEC\ block}{Num\ bit\ DRU\ SSi}$$

```
Minimum DRU for aGPS,
```

$$\beta_{(ertPS)} = \frac{\delta_k^{aGPS} x\ Num\ bit\ FEC\ block}{Num\ bit\ DRU\ SSi}$$

```
Minimum DRU for rtPS,
```

$$\beta_{(rtPS)} = \begin{cases} 0 & , D_k^c < T_d \\ \dfrac{\delta_k^{rtPS} x\ Num\ bit\ FEC\ block}{Num\ bit\ DRU\ SSi} & , D_k^c > T_d \end{cases}$$

```
Minimum DRU for nrtPS,
```

$$\beta_{(nrtPS)} = \frac{w_k\ x\ Num\ bit\ FEC\ block}{Num\ bit\ DRU\ SSi}$$

```
Total minimum DRU,
```

$$\beta_{(\min)} = \sum(\beta_{(UGS)} + \beta_{(ertPS)} + \beta_{(aGPS)} + \beta_{(rtPS)} + \beta_{(nrtPS)}$$

Figure 3.4: Proposed algorithm

Each of service classes is simulated as shown in Table 3.1 below.

Table 3.1: Services simulation

Services	Characteristic (Matlab simulation)
UGS	Real time with fixed data rate
aGPS	Real time with on and off data state. When the state is 'on' the data rate is varies. The value of variable data rate is between aGPS max and aGPS min
ertPS	Real time with on and off state. When the state is 'on' the data rate is fixed.
rtPS	Real time with variable data rate using Pareto distribution
nrtPS	Non real time with variable data size
BE	Non real time with fixed data size

3.3 Physical Layer

Each user is assumed to have a distinctive channel condition which is reported by CQI to the BS. In completing the system level simulator, link budget analysis is performed. The data will be aggregated inside the DRU with different capacity according to the user respond towards the channel condition.

3.3.1 Link Budget Analysis

Since the signal is subjected to loss, the mobile user is simulated through a Rayleigh channel where the Doppler effect takes into account. Furthermore, loss due to distance attenuation and log normal shadowing will be also considered. Therefore the channel gain can be calculated as:

$$h_{i,k}(t) = PL_i x Gain_{i,k} x L_{i,k} \tag{3.5}$$

3.3.1.1 Path Loss

Path loss will trigger the degradation of receiving power signal. Observe that path loss (PL_i) is differs based on the variation of BS-SS distance (D_i). The higher the distance the poorer physical condition will be. PL_i for each user i is calculated according this formula (Yen et al., 2009):

$$PL_i = 128.1 + 37.6\ log\ D_i \quad (3.6)$$

3.3.1.2 Rayleigh Channel

Rayleigh fading can be simulated using Jakes model (Dent et al., 1993). Therefore the normalized low-pass fading process is (Ukil et al., 2009):

$$U(t) = Uc(t) + jUs(t) \quad (3.7)$$

While:

$$Uc(t) = \frac{2}{\sqrt{N}} \sum_{n=0}^{M} a_n \cos(w_n t + \emptyset_n) \quad (3.8)$$

$$Us(t) = \frac{2}{\sqrt{N}} \sum_{n=0}^{M} b_n \cos(w_n t + \emptyset_n) \quad (3.9)$$

$$N = 4M + 2 \quad (4.0)$$

$$\beta_n = \frac{\pi n}{M}\ ,\ \ n = 1,2,3, \dots, M \quad (4.1)$$

$$a_n = 2cos\beta_n,\ \ n = 1,2,3, \dots, M \quad (4.2)$$

$$b_n = 2sin\beta_n,\ \ n = 1,2,3, \dots, M \quad (4.3)$$

$$w_d = 2\pi f_c \frac{v}{c} \quad (4.4)$$

$$w_n = w_d cos \frac{2\pi n}{N},\ \ n = 1,2,3, \dots, M \quad (4.5)$$

Where:

$U(t)$: Channel output

$Uc(t)$: Channel input

$Us(t)$: AWGN random variable at time t

$\emptyset_n$: Random phase

β_n: Oscillator gain

a_n: Angle of incoming wave

Calculating Jakes PDF, the gain of Rayleigh channel is obtained as:

$$Gain_{i,k} = 20\ log \left|\sqrt{Uc(t)^2 + jUs(t)^2}\right| \tag{4.6}$$

3.3.1.3 Log normal shadowing

If the signal goes through any obstacle path loss for each user can be calculated using log normal shadowing (LNS) path model (Wang et al., 2008):

$$L_{i,k} = L_{d_0} + 10n_i log_{10} \frac{d_{x,y}}{d_0} + X_\sigma \tag{4.7}$$

Where:

L_{d_0}: Path loss at reference distance d_0

n_i: Path loss exponent for each user

$d_{x,y}$: Link length between two nodes

X_σ: Zero mean Gaussian random variable with standard deviation σ

d_0 is defined as point allocated in far field antenna, where the value can be 1km for large cell, 100m for microcell and 1m for indoor cell.

Since each subcarrier will experience different signal degradation, the SINR for user i on subcarrier k (Hwang & Han, 2007) is:

$$SINR_{i,k} = \frac{p_i(t)\left|h_{i,k}(t)\right|^2}{2\sigma_i{}^2} \tag{4.8}$$

Where:

$p_i(t)$: Power allocated to user i at time t

$h_{i,k}(t)$: Channel gain of user i on subcarrier k at time t (product of distance attenuation, fast and slow fading)

σ_i: Variance of the additive white Gaussian noise (AWGN)

3.3.2 Burst Size Calculation

A few mapping techniques have been invented over the years such as the exponential effective SINR mapping (EESM), quasi-static, convex, and Shannon methods. It is common to use average SINRS but there are four reasons why average SINR is not competent for OFDMA system evaluation, namely (Andrews et al., 2007):

i. The extents of forward error correction (FEC) block bits are over the subcarriers.
ii. Each subcarrier posses a different SINR due to the frequency selectivity.
iii. Variation of SINR between FEC block bits and the average SINR will determine the reaction of the decoder.
iv. Although the bursts may have the same average SINR but the bursts that remark different channel and interference behavior will present different results of bit error rate (BER) or block error rate (BLER).

After doing the analysis the WiMAX Forum AWG group has come to a decision of using EESM as a default recommended method. Where SINR of many subscarriers are mapped into a single effective SINR (Song et al., 2011). Refering to the link budget analysis section where each subscarrier will experience independent SINR, burst size suitable for the subjected subchannel is calculated. Unlike previous IEEE802.16e standard the burst size is designed based on the channel condition not the modulation level avoiding padding. Padding can be abide by portioning the burst into a number of possible FEC blocks with predetermine size of 4800 bits per block. There will be a few steps to calculate the burst size according to SINR reported from the CQI:

i. Calculate SINR effective
ii. Map SINR effective to suitable type of modulation according to $I_{size\text{-}offset}$
iii. Calculate $I_{min\text{-}size}$
iv. Calculate index and get burst size from Table 3.4.

In EESM modeling β or the scaling factor is needed to calibrate the EESM mapping to the particular modulation and coding rate. To get β value, SINR is mapped to BLER. BLER for equivalent effective SINR and AWGN channel is map for the same MCS level. Instantaneous SINR is mapped into a single effective SINR as calculated as follows:

$$SINR_i^{eff} = -\beta_i \ln\{\frac{1}{Ns}\sum_{k=1}^{Ns} \exp(\frac{SINR_k^{eff}}{\beta_i})\} \tag{4.9}$$

Where $SINR_i^{eff}$ is the SINR of the user i, β is an adaptation factor that relies on the FEC type and MCS (Jain et al., 2008). Calculation of SINR effective is correlated to β value. So we need to generate SNR effective for respective β which the value is from -5dB to 20dB. Then we generate graphs of BER versus SNR for AWGN channel. Based on target BER we mapped SINR value according to each type of the modulation techniques. Then compare the β values gained from the SINR EESM and SINR AWGN. The comparisons of the SINR pair values will yield a mean squared difference for a given beta value (Lopez Aguilar, 2009).

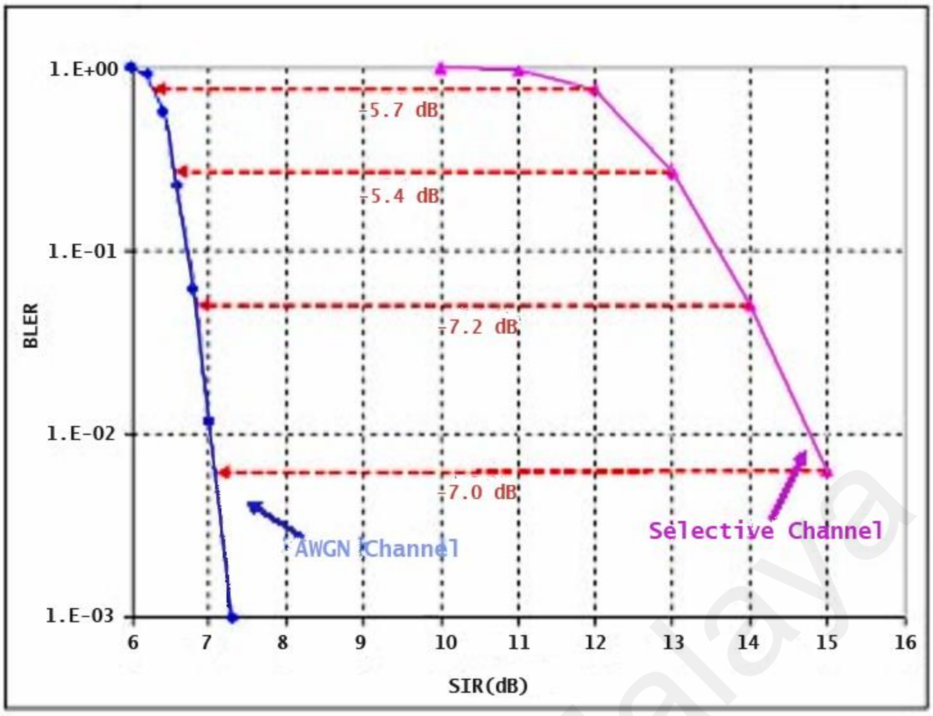

Figure 3.5: Channel model for the calibration of β value (Mumtaz et al., 2008)

To determine the type of modulation suitable for the user we choose the smallest β value. The β value is posses by finding the minimum value of the mean square error (r.m.s) between effective SINR and static SINR. By using equation (Mumtaz et al., 2008):

$$r.m.s = \frac{1}{Ns}\sum_{k=1}^{Ns} SINR_{AWGN} - SINR_{eff}(\beta)^2) \qquad (4.10)$$

$$\beta = \arg\min_{\beta}(r.m.s) \qquad (4.11)$$

$I_{size\text{-}offset}$ is defined as the offset used to compute the burst size index. It determined by number of spatial streams, modulation type and coding rate from the Table 3.1 where the spatial streams are the number of transmitting antennas. Using MIMO technology it is possible currently to run up to 4 spatial streams.

Table 3.2: MCS index

MCS index	Spatial stream	Modulation type	Coding rate
0	1	BPSK	1/2
1	1	QPSK	1/2
2	1	QPSK	3/4
3	1	16-QAM	1/2
4	1	16-QAM	3/4
5	1	64-QAM	2/3
6	1	64-QAM	3/4
7	1	64-QAM	5/6
8	2	BPSK	1/2
9	2	QPSK	1/2
10	2	QPSK	3/4
11	2	16-QAM	1/2
12	2	16-QAM	3/4
13	2	64-QAM	2/3
14	2	64-QAM	3/4
15	2	64-QAM	5/6
16	3	BPSK	1/2
17	3	QPSK	1/2
18	3	QPSK	3/4
19	3	16-QAM	1/2
20	3	16-QAM	3/4
21	3	64-QAM	2/3
22	3	64-QAM	3/4
23	3	64-QAM	5/6
24	4	BPSK	1/2
25	4	QPSK	1/2
26	4	QPSK	3/4
27	4	16-QAM	1/2
28	4	16-QAM	3/4
29	4	64-QAM	2/3
30	4	64-QAM	3/4
31	4	64-QAM	5/6

Allocation size can be calculated as the number of LRU multiplied by MIMO rank. Based on allocation size we can get $I_{minimum\text{-}size}$ from Table 3.3 (Ahmadi, 2011).

Table 3.3: Minimum size Index as a function of allocation size

Allocation size	$I_{minimum-size}$	Allocation size	$I_{minimum-size}$	Allocation size	$I_{minimum-size}$
1-3	1	16-18	15	58-64	26
4	2	19-20	16	65-72	27
5	4	21-22	17	73-82	28
6	6	23-25	18	83-90	29
7	8	26-28	19	91-102	30
8	9	29-32	20	103-116	31
9	10	33-35	21	117-131	32
10-11	11	36-40	22	132-145	33
12	12	41-45	23	146-164	34
13	13	46-50	24	165-184	35
14-15	14	51-57	25	192	36

Eventually since index=$I_{size-offset}$+$I_{min-size}$, based on Table 3.4 we now can get the size of FEC block that can be fit inside one DRU (N_{DB}) (Ahmadi, 2011).

Table 3.4: Supported Burst Sizes in IEEE802.16m

Index	N_{DB}(Byte)	K_{FB}	Index	N_{DB}(Byte)	KFB	Index	N_{DB}(Byte)	K_{FB}
1	6	1	23	90	1	45	1200	2
2	8	1	24	100	1	46	1416	3
3	9	1	25	114	1	47	1584	3
4	10	1	26	128	1	48	1800	3
5	11	1	27	145	1	49	1888	4
6	12	1	28	164	1	50	2112	4
7	13	1	29	181	1	51	2400	4
8	15	1	30	205	1	52	2640	5
9	17	1	31	233	1	53	3000	5
10	19	1	32	262	1	54	3600	6
11	22	1	33	291	1	55	4200	7
12	25	1	34	328	1	56	4800	8
13	27	1	35	368	1	57	5400	9
14	31	1	36	416	1	58	6000	10
15	36	1	37	472	1	59	6600	11
16	40	1	38	528	1	60	7200	12
17	44	1	39	600	1	61	7800	13
18	50	1	40	656	2	62	8400	14
19	57	1	41	736	2	63	9600	16
20	64	1	42	832	2	64	10800	18
21	71	1	43	944	2	65	12000	20
22	80	1	44	1056	2	66	14400	24

Furthermore, the data will be aggregated inside the DRU with different capacity according to user response towards the channel condition. N^{DB} represents the number of bytes supported in one DRU. If the burst exceeds 4800 bits, it will be partitioned into $n\, FEC\ blocks$ Assume that the bandwidth requested by user i is denoted by $\mathrm{B\_req_i}$. Then the number of DRU required to support the data burst is:

$$required\_DRU_i = \frac{\mathrm{B\_req_i}}{N_i^{DB} x 8} \tag{4.12}$$

Where one DRU is assumed could be shared by different user. DRU is spread across frequency inside the OFDMA frame network, thus the number of DRU inside a frame can be calculated as:

$$Num_{DRU} = \alpha\ X\ N_{sc},\ \alpha = \begin{cases} 8, subframe\ type\ 1 \\ 7, subframe\ type\ 2 \\ 6, subframe\ type\ 3 \\ 5, subframe\ type\ 4 \end{cases} \tag{4.13}$$

While the number of subchannel can be calculated as:

$$N_{sc} = \frac{B}{fs\ x\ P_{sc}} \tag{4.14}$$

Where:

B: Frame bandwidth

fs: Subcarrier spacing

P_{sc}: Number of subcarrier across a subchannel

Thus now we could obtain the suitable modulation level with respect to the channel condition and size of the burst. Now number of bits that could be fitted into one DRU is known. Depends on number of bits can be fit in one DRU, scheduler will calculate how many DRU needed for the current bandwidth polled. Then EDRR will perform its operation accordingly.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Simulation model

The proposed method is developed using MATLAB software. The data that were transmitted over the channel would experience the log normal shadowing, multipath fading and Doppler spread. In this simulation the radio resources are ignoring the physical layer synchronization, guards band and pilot. It is assumed that the available bandwidth is fully utilized. Our focus is the influence of the speed towards the channel condition. Therefore the distance from BS to all users have been kept identical to generate equal path loss. Other additional parameters that are used for link level simulatator are shown in Table 4.1. The values are solely to provide a consistent evaluation of the user.

Table 4.1: Simulation parameters

Parameter	Value
System bandwidth	10MHz
Number of iterations (number of frames)	10
Subcarrier frequency spacing	10940Hz
Duration of one superframe	20ms
Duration of one frame	5ms
Duration one frame elapse	10ms
Number of subframe	8 (Type 1)
Number of LRU per subframe	24
Receive noise level	-174dBm
Max delay for rtPS before packet drop	20ms
Min throughput (nrtPS)	192kbps
Max delay for UGS and ertPS before packet drop	5ms
Required BER	10^-5
Transmit power (PT)	23dBm

As the scheduler is meant to support IEEE802.16m, we want to analyze the effectiveness of the scheduler with respect to the mobility of the users. Therefore we simulated 3 different scenarios involving static users, pedestrian users and moving users. The simulation scenarios are as follows:

i. Scenario 1: Network comprises of 2 static users (user 1 and 2) and 2 pedestrian users (user 3 and 4).

ii. Scenario 2: Network comprises of 2 static users (user 1 and 2) and 2 moving users (user 3 and 4).

iii. Scenario 3: Network which number of user is kept increasing linearly to test the scalability of user.

The simulation model for scenario 1 and 2 is illustrated in Figure 4.1. The IEEE802.16m network consists of one BS covering 4 active subscribers which each user may come from any class of static, pedestrian,moving or fast moving user.

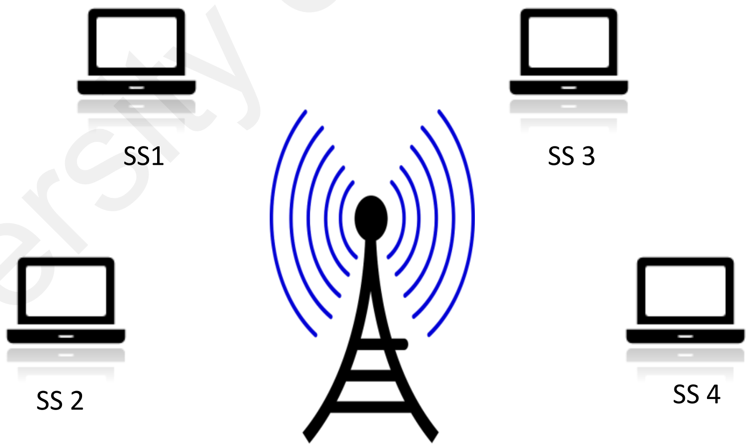

Figure 4.1: Simulation environment

Apart from the physical condition, the bandwidth requests from the subscribers are also important. Thus the load generated for all 4 users must be configured and known. In this regards we generated bulk data transfer to simulate the congested area. Thus we can see the effect of scheduler trying to optimize the utilization of the limited bandwidth. Apart from that we want to investigate what would be the result if the static users are in the same cell with the pedestrian or moving users, specifically how it would influence the throughput, packet drop and which services would be affected by the most. Therefore the data rate for each service for the corresponding user is set equal in each scenario so that we can analyze the pattern. We describe in more detail on the value of bandwidth polling as listed in Table 4.2 below:

Table 4.2: Simulation of loads for each user corresponding to type of service

	UGS (Mbps)	ertPS (Mbps)	aGPS (Mbps)	rtPS (Mbps)	nrtPS (Mbps)	BE (Mbps)
User 1	0.240	0.280	0.280	0.432	0.432	0.384
User 2	0.280	0.312	0.312	0.576	0.624	0.432
User 3	0.240	0.280	0.280	0.432	0.432	0.384
User 4	0.280	0.312	0.312	0.576	0.624	0.432

As we can see from the table, the 2nd and the 4th user are configured to have the identical bandwidth request that is higher than the 1st and the 3rd user. Notice that for each scenario even though the two users posses the same speed, the data inside the buffer are different. For instance in scenario 1 we have two static user which include user 1 and 2. User 1 and 2 will definitely have identical SINR because they belong to the same group but user 2 have more data in queue than user 1. We will further see the effects on this matter in the next section.

Simulation model for scenario 3 is made differently from scenario 1 and 2. In scenario 3 the scalability of the scheduler is tested. The simulation model is run 10 times with increasing number of the subscriber. The amount of data request and physical condition is randomly created but is kept identical for both EDRR and DRR for a fair comparison.

4.2 Performance Evaluation

To evaluate the robustness and the reliability of the proposed scheduler, EDRR is compared against the ordinary DRR. For scenario 1 and 2 there will be five types of figures collected and presented, namely:

i. Distribution of quantum number. As stated in methodology part the distribution of quantum number has divided into two conditions. Thus the calculation will be as stated in equation 3.3 and 3.4.

ii. Throughput for each user (in MAC layer not at receiver). The throughput is calculated on the transmitter side:

$$T_i = \frac{No\ of\ packet\ sent\ in\ kth\ iteration\ x\ No\ of\ bit\ per\ packet}{Time\ for\ one\ frame\ elapse} \quad (4.1)$$

iii. Percentage of packet drop for each user

$$PD_i = \frac{No\ of\ packet\ drop\ in\ kth\ iteration}{No\ of\ packet\ sent\ in\ kth\ iteration} x100 \quad (4.2)$$

iv. Total packet drop on each service

$$TD_i = \frac{No\ of\ packet\ drop\ in\ kth\ iterationx\ No\ of\ bit\ per\ packet}{Time\ for\ one\ frame\ elapse} \quad (4.3)$$

v. Comparison of EDRR and DRR performance in term of total number of packet drop.

Simulation 3 is intended for scalability test. Graph for the average amount of packet drop versus the number of user N is plotted. The calculation of the average number of packet drop is:

$$Q_{ave} = \frac{\sum\left(\frac{No\ of\ packet\ drop\ for\ N\ userx\ No\ of\ bit\ per\ packet}{Time\ for\ one\ frame\ elapse}\right)}{10} \quad (4.4)$$

4.3 Simulation Results

4.3.1 Scenario 1: Static Subscribers VS Pedestrian Subscribers

Scenario 1 will record the performance of both scheduler in dealing with static users and pedestrian users. It is known that the only difference between DRR and EDRR is the distribution of the quantum number. The quantum number is directly proportional to the amount of bandwidth grants to the respective user therefore it will surely affect the throughput and the amount of packet drop for the users on demand. For a detail comparison the figure of distribution of quantum number is generated to see how the aggregation was done.

The graph for quantum number is generated as shown in Figure 4.2. The pattern in Figure 4.2 (a) clearly imply that DRR gives almost equal quantum number to user which belong to the same group. That is why user 1 and user 2 have a nearly equal quantum number while user 3 has almost equal quantum number to user 4. Figure 4.2 (b) indicates that EDRR tried to tolerate between the physical layer condition and also the amount of bandwidth request. Even though user 1 and user 2 are in the same group but they have significant differences in the amount of the quantum number. It is proven that EDRR allocate more quantum number for user 2 than user 1 because user 2 has more data request. Therefore, larger quantum number was given so that user 2 has higher chance to transmit the data. The same thing occurs for user 3 and 4. Where user 4 which have more data inside the queue was given bigger chance to be placed inside OFDMA frame by granting larger quantum number. The effect of distribution of quantum number will be seen later on the next graph where the evaluation in term of throughput and packet drop was done.

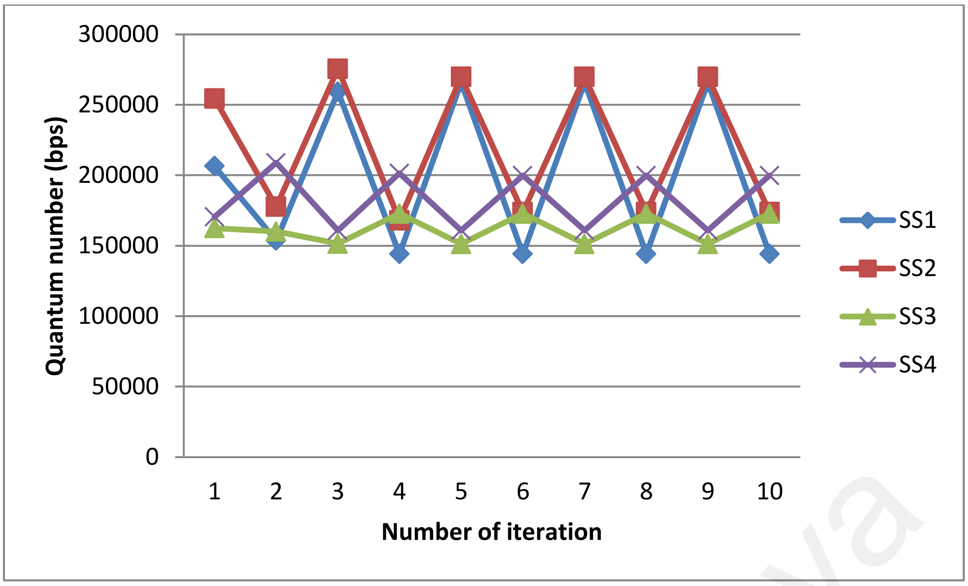

Figure 4.2 (a): Distribution of quantum number using DRR

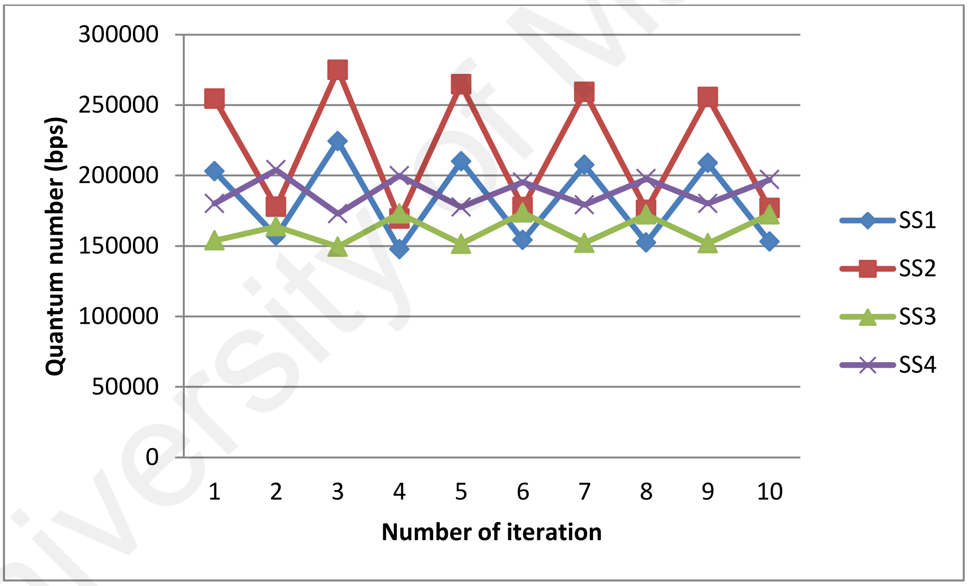

Figure 4.2 (b): Distribution of quantum number using EDRR

Figure 4.3 examines the throughput for each service provided in the system. The data are recorded each time the frame elapses in 10 ms. Figure 4.3 (a) depicts the DRR behavior while Figure 4.3 (b) represents EDRR. Figure 4.3 (a) clearly shows the characteristics of DRR which only considers the physical condition for the bandwidth aggregation. As shown in Figure 4.3 (a), DRR gives almost equal amount of quantum number to user which belong in the same group. Thus user 1 and 2 are having nearly equal throughput. This also applied for user 3 and 4. However, we observed that the throughput for user 3 and 4 is lower than user 1 and 2. This is because the user which belongs to the static group experiences more bandwidth grants than the pedestrian user. User 3 and 4 have lower bandwidth allocation because it faces turbulence from the environment factor which then leads to lower SINR performance. Poor SINR performance further initiates a lower quantum number which contributes to smaller bandwidth grants.

Figure 4.3 (b) generated by EDRR shows a quite a distinct difference between the throughput of user that belong to the same class. Both user 1 and 2 are static users but the throughput is not identical. User 3 and 4 also deal with the same situation where user 4 has more throughput than user 3. Recall that user 2 and 4 have more bandwidth request than user 1 and 3. Here is where the EDRR different from DRR in which apart from SINR consideration it allocates more bandwidth to the user which have more data queue or larger polled bandwidth. User 1 and 2 did belong to same group but since user 2 has more data in the buffer, the scheduler will give a bigger quantum number for user 2 as stated in Figure 4.2 (b) beforehand. This also happened to user 3 and 4. However, EDRR does not violate the right of getting more bandwidth allocation for the users with good channel circumstances. So we can see that static users always have higher throughput. Only the user within the same group will experience the difference. This would benefit in higher throughput for the sophisticated user which helps later in preventing the data in the queue from becoming overwhelmed.

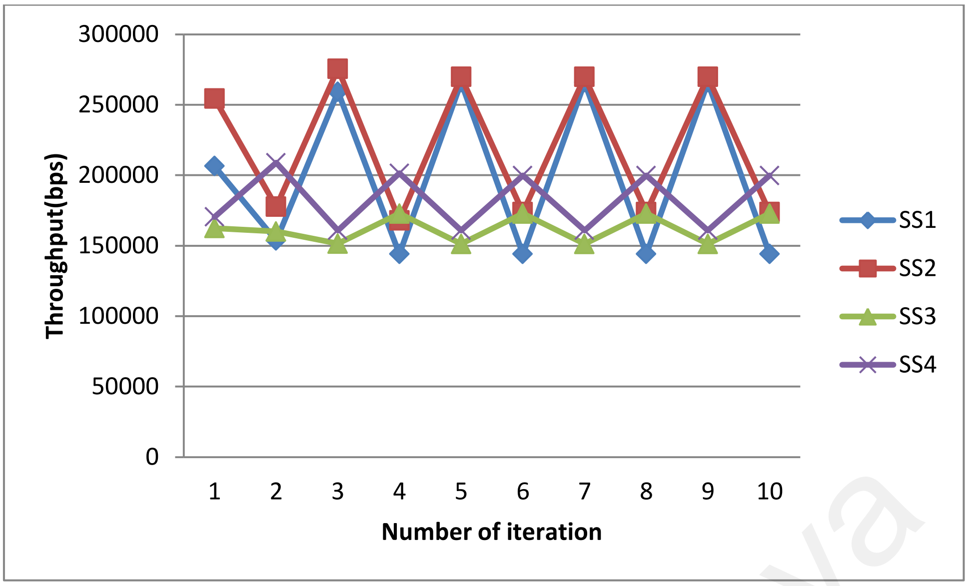

Figure 4.3 (a): Throughput for each user using DRR

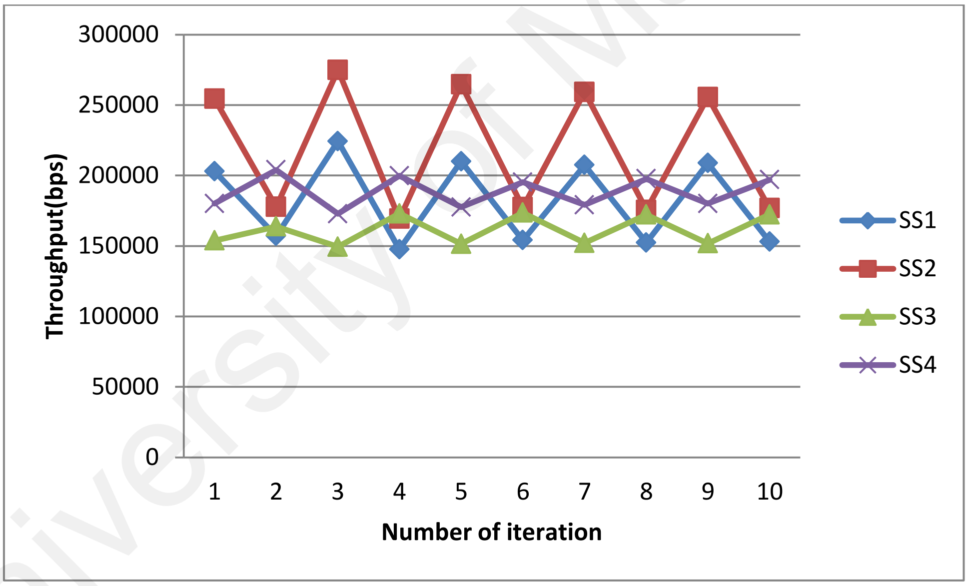

Figure 4.3 (b): Throughput for each user using EDRR

It has been set that the order of priority in bandwidth allocation inside the scheduler is UGS, ertPS, agPS, rtPS, nrtPS and BE. Figure 4.4 examines the packet drop for each type of services to see how the scheduler maintain the QOS requirements. As shown in Figure 4.4 (a) for DRR, all delayed sensitive data are transmitted before the deadline. Since the bandwidth is already full, nrtPS missed the chance to participate inside OFDMA frame causing the packet drop. BE service does not experience packet drop since BE does not have the QoS restriction. For EDRR in Figure 4.4 (b) it can be clearly seen that EDRR are facing the same circumstances of nrtPS packet drop as DRR but the amount is lower. This is due to the control of the bandwidth access according to the physical conditions and also bandwidth request. So the user with more bandwidth request can afford to loosen up its queue to give opportunity for the new incoming data on the next iteration. EDRR optimized the utilization of the system bandwidth which is not just at that particular iteration but also for the next frame of data transmission. This resulted in a lower packet drop in the long run.

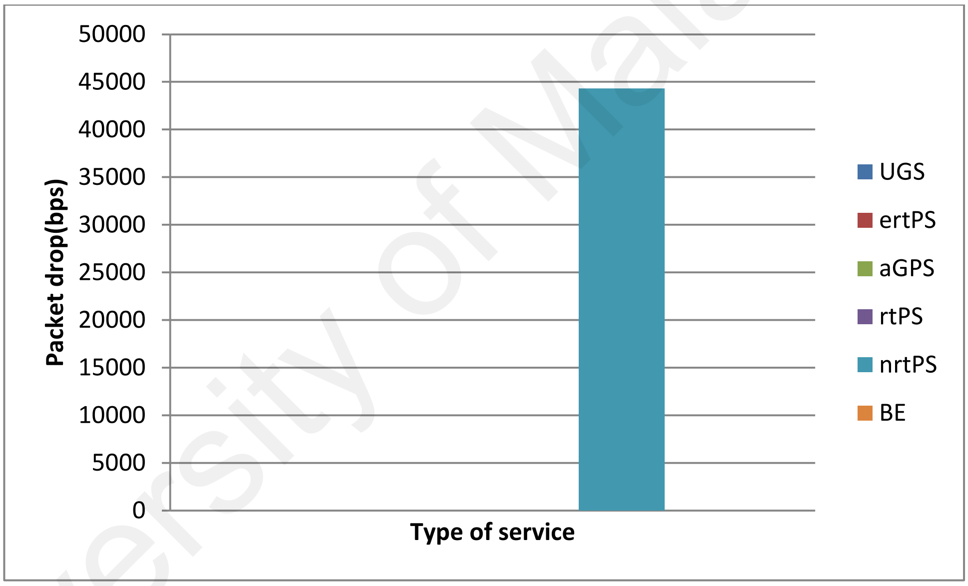

Figure 4.4 (a): Percentage of packet drop for each service using DRR

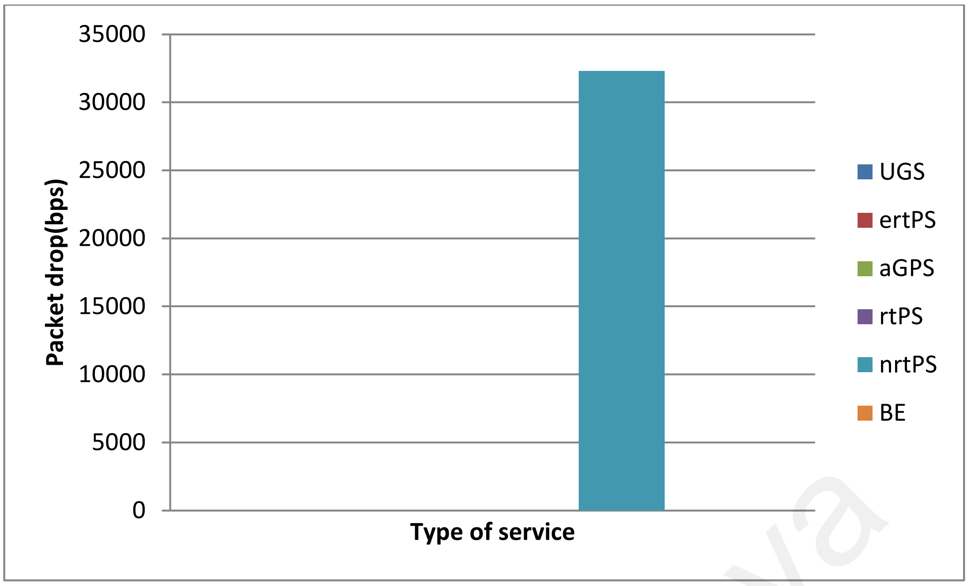

Figure 4.4 (b): Percentage of packet drop for each service using EDRR

Figure 4.5 further illustrates the total amount of packet drop throughout the simulation. Figure 4.5 (a) shows that user 4 in DRR does experience unsatisfied QoS requirement on the 4th, 6th ,8th and 10th iteration. As discussed before this is due to lower bandwidth grants since it has the worst channel condition than user 1 and 2. Even though user 3 is in the same physical condition with user 4 but it has the lowest amount of data inside the queue. So user 3 manages to at least serve the services which have the QoS restriction, saving it from any data loss.

Figure 4.5 (b) shows that EDRR is able to lower the amount of packet drop in each iteration. As discussed earlier, the ability of EDRR to distribute the resources carefully among the same group of users according to the bandwidth request gives great benefit in this situation. EDRR tolerate with the amount of bandwidth grants from user 3 and 4 where scheduler gives user 4 a higher amount of bandwidth access than user 3 so that it could support the larger request. Therefore the amount of packet drop is lower.

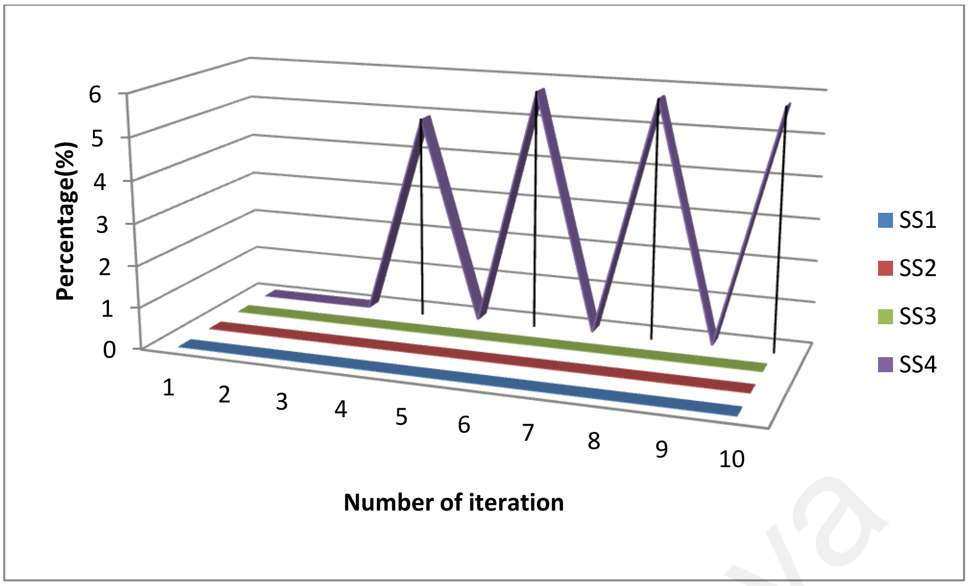

Figure 4.5 (a): Percentage of packet drop for each user using DRR

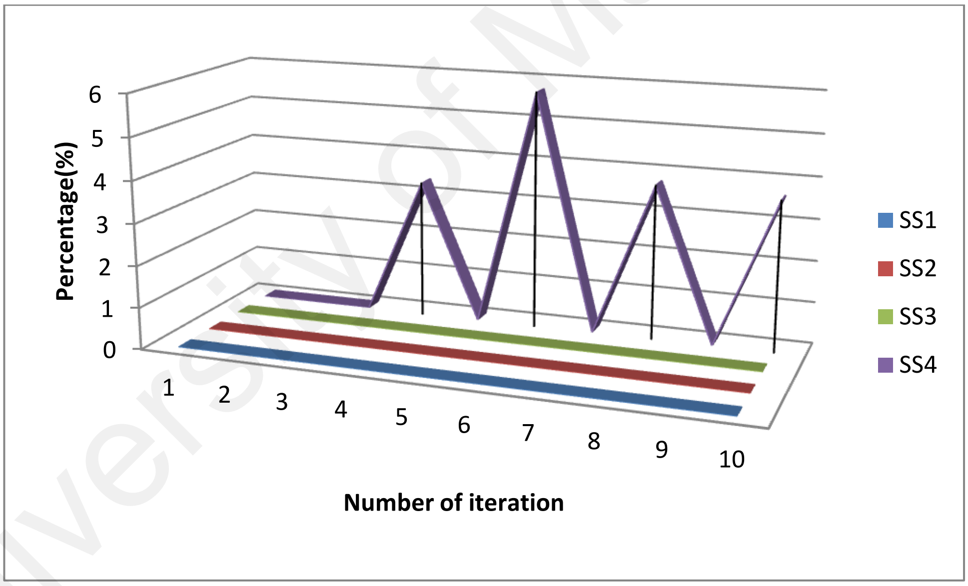

Figure 4.5 (b): Percentage of packet drop for each user using EDRR

We already analyzed the throughput of the users and packet drop for each service. Now we want to see how good is EDRR in comparison to DRR. The benchmark for this evaluation is the amount of packet drop for the entire system. Notice that DRR gives almost the same opportunity for each user belonging to the same group. Therefore even though user 4 requires more bandwidth at a particular time, the scheduler would not allow it since it distributes the bandwidth equally towards all users having an identical channel condition. Since user 4 is a heavily loaded user, the bandwidth grant would not be able to support the large number of packets waiting to be served. The data packet would be eventually dropped, contributing to the higher data loss in DRR case. Figure 4.6 shows the comparison of EDRR and DRR in term of the total packet drop in bits per second. The graph proves that EDRR posses better performance than DRR by reducing the data loss to 27.08%.

The percentage is gained from the calculation of:

$$Loss = \frac{Tot\ no\ of\ packet\ loss\ on\ DRR - Tot\ no\ of\ packet\ loss\ on\ EDRR}{Tot\ no\ of\ packet\ loss\ on\ DRR} x100 \qquad (4.4)$$

$$Loss = \frac{44266.67 - 32281.37}{44266.67} x100 = 27.08\% \qquad (4.5)$$

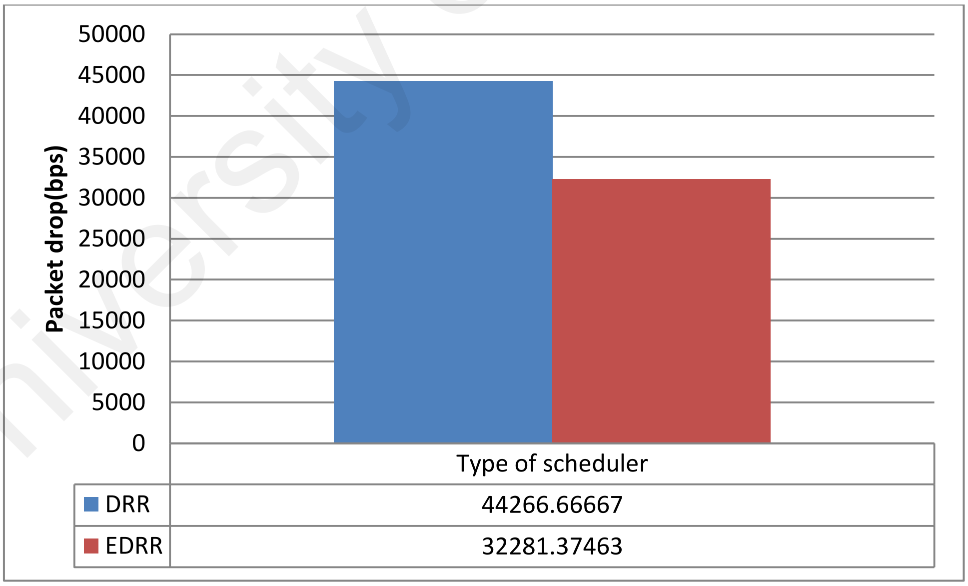

Figure 4.6: Total number of packet drop for each scheduler

4.3.2 Scenario 2: Static Subscribers VS Moving Subscribers

All simulation parameters for scenario 2 is identical to scenario 1 except that the pedestrian user is replaced by moving user. Figure 4.7 (a) shows the distribution of quantum number for DRR while Figure 4.7 (b) is generated for EDRR. Both figures shown insignificant difference since the simulation model for scenario 2 was replaced by moving user. So the physical condition is poorer than scenario 1. Therefore larger amount of bandwidth is required to transfer one block of data. Initially both schedulers will try to satisfy QoS. Then with the limited bandwidth left they struggle to distribute the quantum number among the user. Hence there is no clear difference between them. Despite there are still slight differences on the number. For a better view the average distribution for the respective user in scenario 2 is further generated in Table 4.3. It can be seen that for EDDR, user with larger data inside the queue like user 2 and user 4 had a slightly greater quantum number than DRR. While EDRR allocate the lower amount of quantum number than DRR , for user 1 and user 3 which has less amount of bandwidth request. EDRR is making an effort to transmit as much as they can for users with higher request. Even this little difference could lower the amount of packet drop that will be shown later.

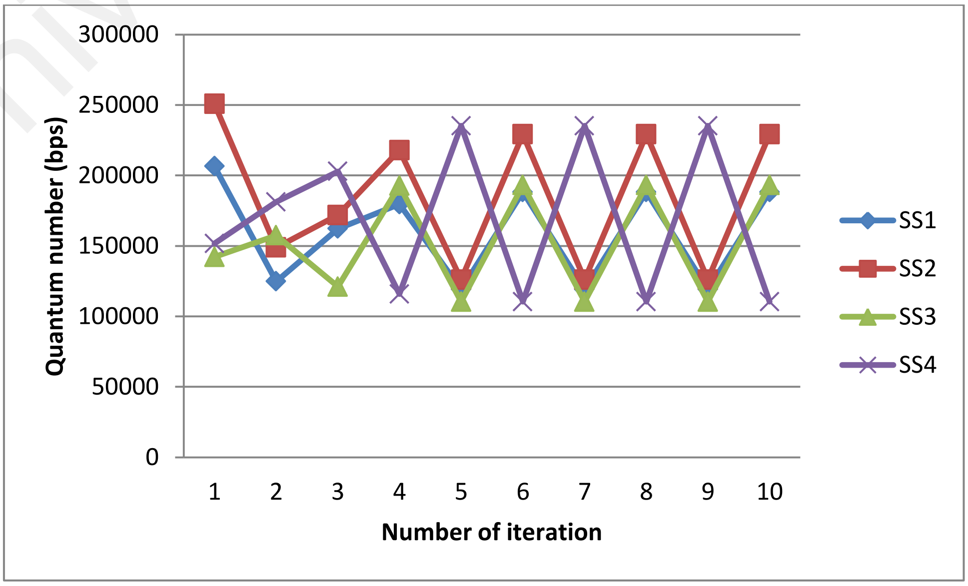

Figure 4.7 (a): Distribution of quantum number using DRR

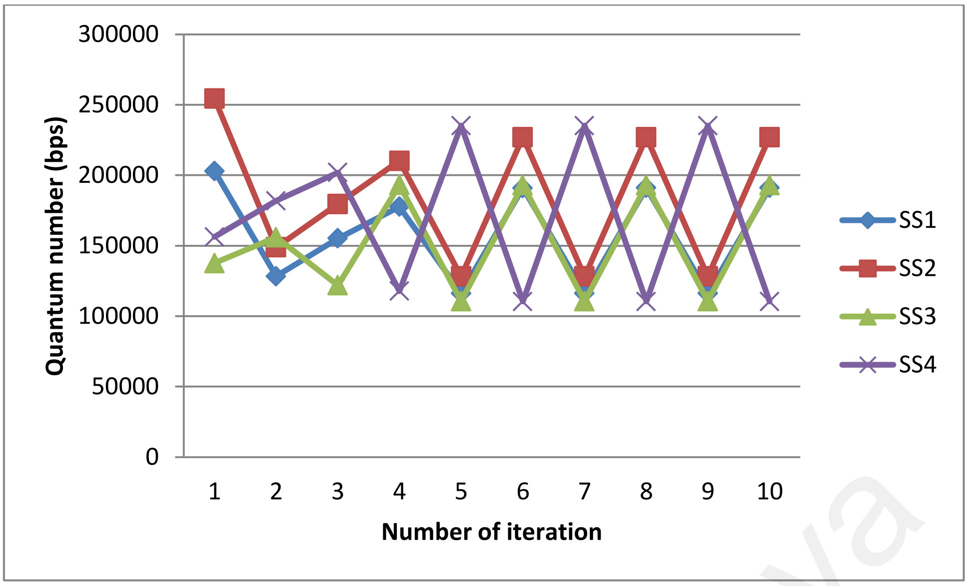

Figure 4.7 (b): Distribution of quantum number using EDRR

Table 4.3: Average of quantum number distribution

	Average of quantum number distribution	
	DRR	EDRR
SS1	1594560	1585574
SS2	1854895	1858560
SS3	1523584	1518729
SS4	1688256	1694412

The throughput of DRR and EDRR as plotted in figure 4.8 (a) and 4.8 (b) respectively does not show much difference. This is because the moving user with limited bandwidth encounters lower SINR than pedestrian user. Thus both scheduler is striving to fulfill the QoS restriction for all service classes in demand.

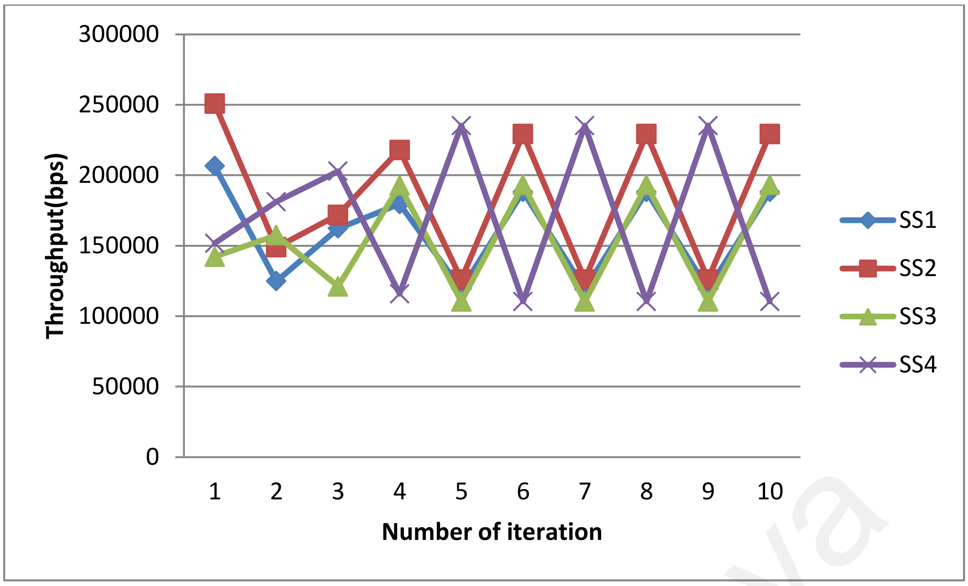

Figure 4.8 (a): Throughput for each user using DRR

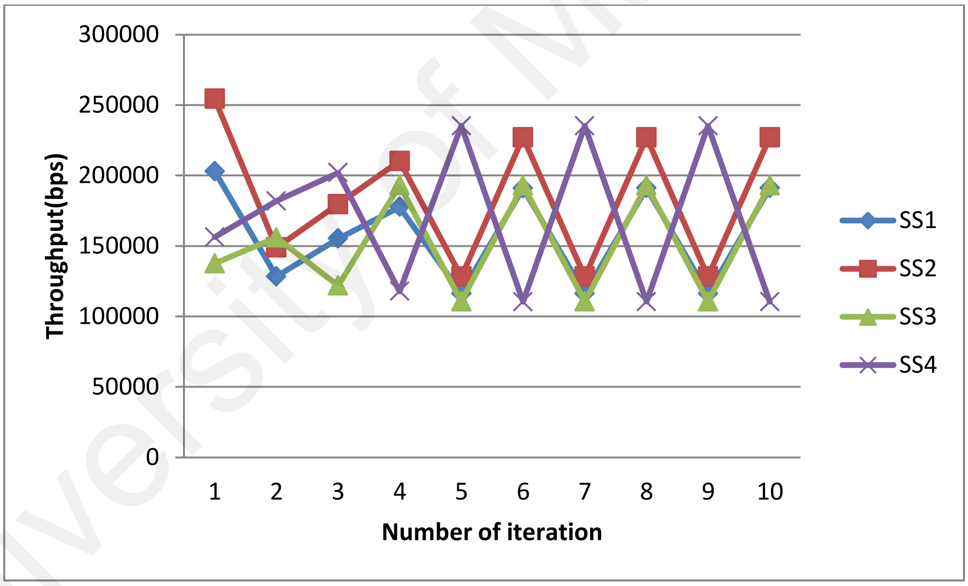

Figure 4.8 (b): Throughput for each user using EDRR

Like scenario 1, scenario 2 also faced loss of data. However, since scenario 2 have to lower the quantum number due to lower SINR than scenario 1, packet drop for DRR and EDRR scheduler has increased. As shown in Figure 4.9 (a) and (b), this time it is not only nrtps which has fallen into dissatisfaction but also rtPS. According to the minimum allocation group described earlier in the methodology section, to satisfy the delay sensitive services like UGS, ertPS and agPS, they will be fitted in first inside the

OFDMA frame. This could lead to meeting the deadline for service like rtPS. Since nrtPS is sent only after rtPS has been transmitted, any packet loss experience by rtPS will surely affect nrtPS as well. Furthermore, it will be even worst for nrtPS since all the minimum data is halted from transmission.

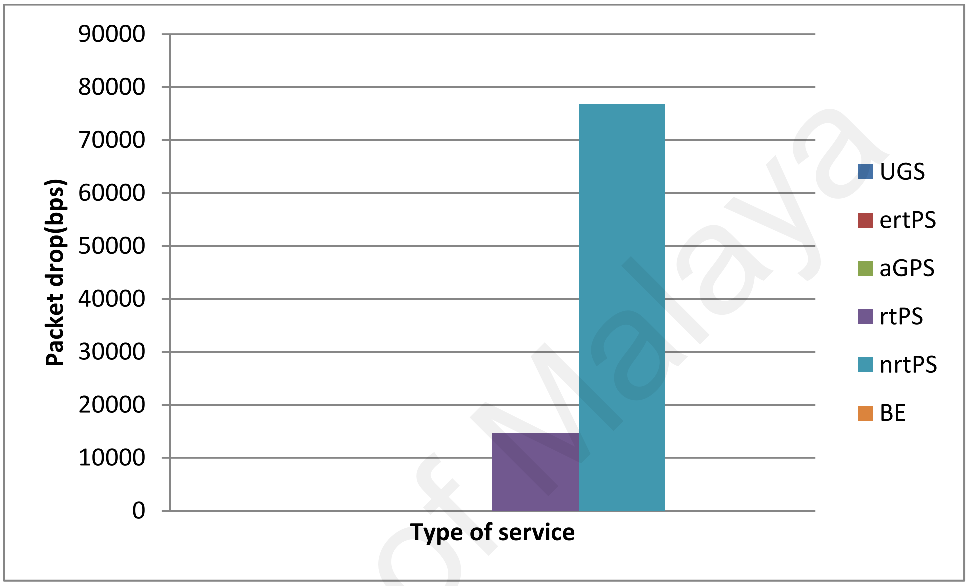

Figure 4.9 (a): Percentage of packet drop for each service using DRR

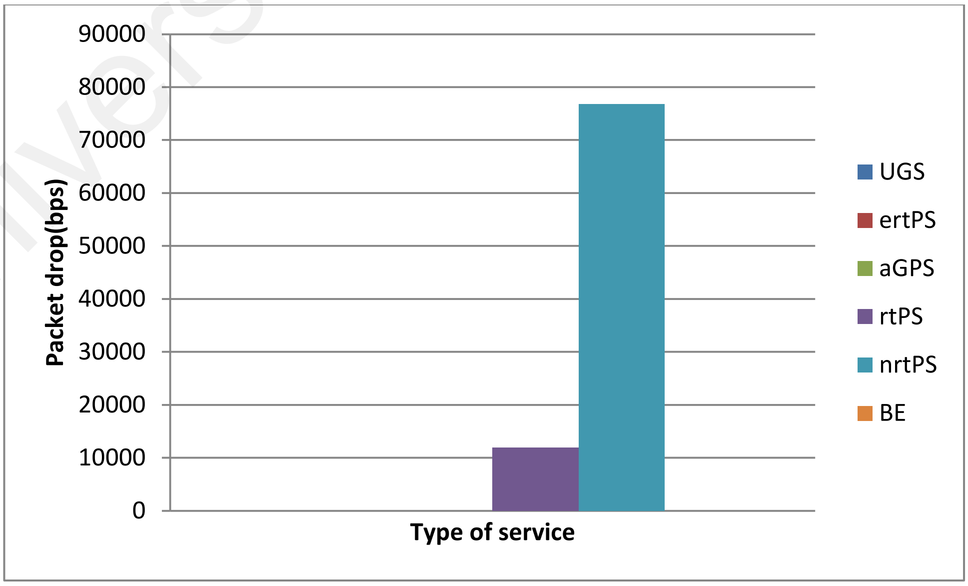

Figure 4.9 (b): Percentage of packet drop for each service using EDRR

Figure 4.10 depicts the effect of scheduler to the amount of packet drop in each iteration. In scenario 2 more users are experiencing packet drop compared to scenario 1. Now not only user 4 experienced the unsatisfaction of QoS but all users on the 2nd iteration as well. However, as expected EDRR shows a lower packet drop.

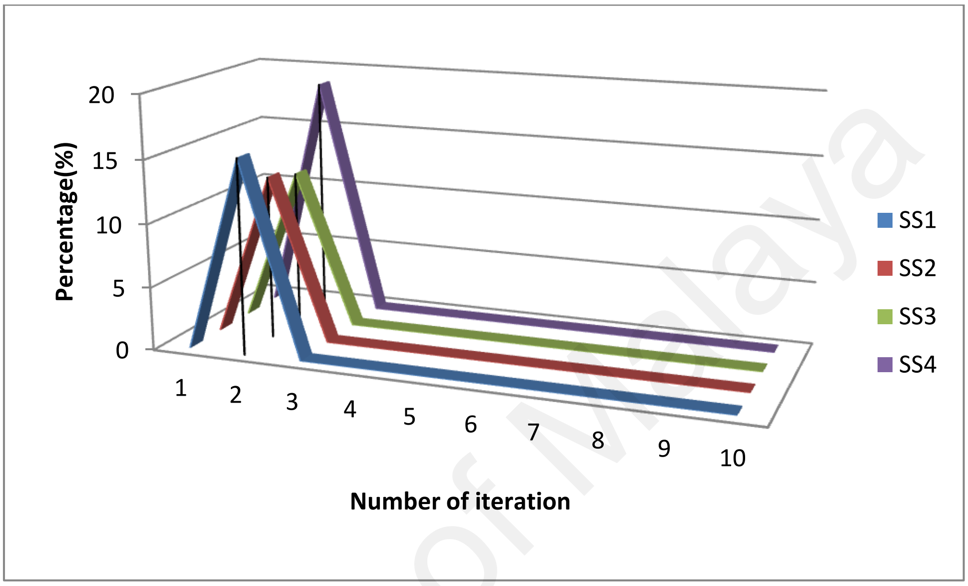

Figure 4.10 (a): Percentage of packet drop for each user using DRR

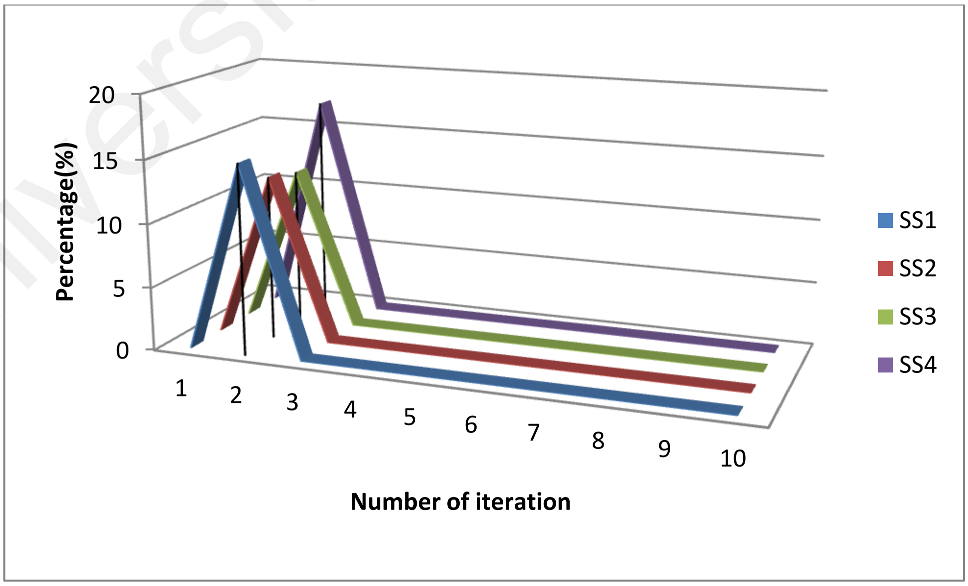

Figure 4.10 (b): Percentage of packet drop for each user using EDRR

The total number of packet drop in the system is presented in Figure 4.11 in which it is demonstrated that EDRR outperforms DRR by 3.04%. The graph shows that even though with limited bandwidth, EDRR would enhance the system performance by wisely arranging the allocation for each user. Using the same equation as (4.4):

$$Loss = \frac{91456 - 88673.52}{91456} x100 = 3.04\% \tag{4.6}$$

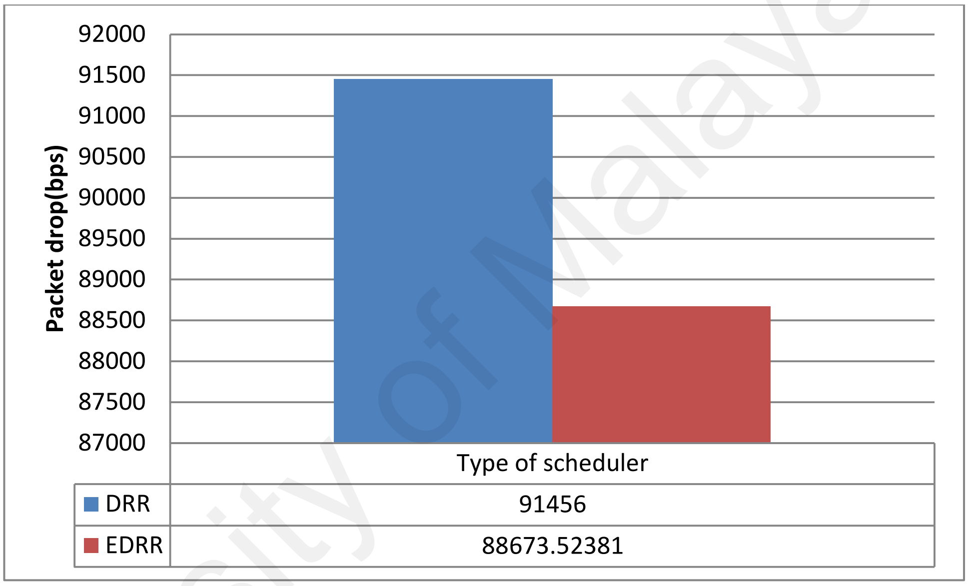

Figure 4.11: Total number of packet drop for each scheduler

We already developed the system model for both DRR and EDRR based on the IEEE802.16m in two simulative scenarios. As expected EDRR could help in minimizing the amount of packet drop even though the traffic intensity increased. Therefore both objectives are fulfilled.

4.3.3 Scenario 3: Scalability of Scheduler

Figure 4.12 is the graph for the average amount of packet drop versus N number of user. The pattern shows that EDRR always outperform DRR method as the amount of packet drop for EDRR always lower than DRR. This proves that even in a different environment with an increasing number of users with various amounts and demand and also speed of movement EDRR still have better performance than DRR.

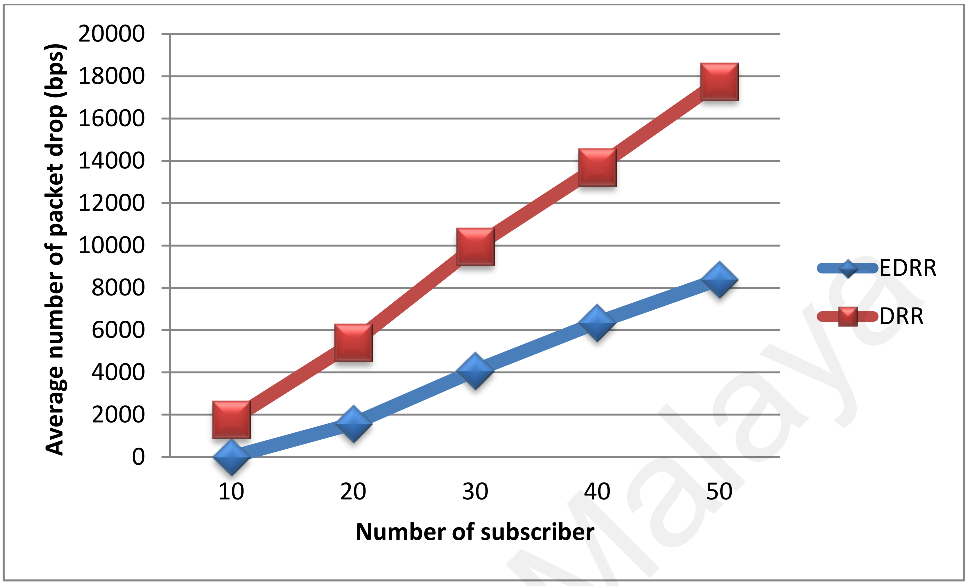

Figure 4.12: Amount of packet drop with increasing number of user

CHAPTER 5

CONCLUSION

5.1 Conclusion

A good scheduler should be able to reduce if not diminish the packet drop for all users, meaning it should meet the QoS level matter for the service classes. However, in a system, due to movement and distance from BS each user might experience a variety of impairments in channel condition. Generally user with bad channel quality will consume more bandwidth than the user with a good channel quality. It would be a waste if we provide more bandwidth for the user with poor environment circumstance. Therefore scheduler will play important role in electing which user is eligible to obtain more bandwidth based on the reported CQI. The wise distribution of available resources will result in higher performance.

In this thesis, a scheduler for IEEE802.16m is proposed. The scheduler is adapting the new physical layer interface, added service class and convolutional turbo code scheme. Our scheduler is mainly divided into 2 parts which are minimum allocation group and EDRR. The minimum allocation group emphasizes that important service will remain as top interest in data aggregation. Consequently the minimum packet which satisfies the QoS requirement is priory placed on the OFDMA frame. After that the left over bandwidth is managed by EDRR. EDRR is a modification from DRR method which is proven in enhancing the system performance as shown by the simulative results. EDRR uses the concept in which users in different group would face different treatment on quantum number size while users in the same group would have an unequal quantum number which is further depends on the bandwidth request. The reason we demonstrated three types of scenarios is to observe the influence of the mobility towards the scheduler. We want to see the EDRR reaction in balancing the mobility effect with the limited bandwidth.

Starting with simulation 1, the results show that EDRR outperformed DRR by decrementing the amount of packet drop. This is due to users that have more bandwidth request have higher chance to be placed in the OFDMA frame. It is done by granting larger quantum number for the user with higher demand initiating in higher transmission of data. Thus loosen up the data queue for the sophisticated user. Since less data accumulated inside the queue, the emptied space is ready for the next incoming bulk data minimizing the packet drop that leads to less retransmission which only makes the traffic much busier on the next iteration. As a result, EDRR would not just benefit in current iteration but also on the next round. Further boost the overall system performance by increasing the throughput of the users. But to be remembered user that has good channel quality is still getting the privilege to get more bandwidth allocation. That could be the reason why only user 4 has packet drop even though user 2 and 4 both have the same amount of data inside the queue. Notice that user 1 and 2 always has higher throughput than user 3 and 4.

Results in simulation 2 validate that even though dealing with faster moving users, EDRR manage to lower the amount of packet drop. But this time the percentage is less than scenario 1 due to limited bandwidth access because the transmission power is kept identical for both simulations. Hence a lower modulation level is chosen in scenario 2 to balance out between the power consumption and physical condition. Since the physical condition is directly proportional to quantum number, poor channel condition for user 3 and 4 in scenario 2 resulting on lower quantum number than scenario 1. Therefore bigger bandwidth required to allocate the data packet. Hence the number of data packets can be occupied in the frame is reduced, producing lower throughput and higher packet drop in simulation 2.

Scenario 3 proves that even though with gradually increasing number of user EDRR still manage to reduce the amount of packet drop better than DRR. Despite, in all scenarios, the designed scheduler which operates in IEEE802.16m standard proved to succeed in dealing with the mobility competently.

5.2 Future Works

For further research jitter will be in concern. Where the order of the data packet is organized in the frame is counted. To decrease the jitter, instead of distributing available bandwidth the data packet is positioned according to the type of subscriber. But to place the data packet one by one would surely consume more time. So the scheduler may be enhanced with intelligent method to increase the system robustness. Hence the data aggregation could be done faster.

Bibliography

Adhicandra, I. (2010). Measuring Data and VoIP Traffic in WiMAX Networks. *Journal of Telecommunications*, *2*(1), 1-6.

Ahmadi, S. (2011). *Mobile WiMAX: A System Approach to Understanding IEEE802.16m Radio Access Technology*. Elsvier.

Andrews, Jeffrey G., Arunabha Ghosh, R. M. (2007). *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*. Prentice Hall.

Cai, J. (2003). Doppler spread estimation for mobile OFDM systems in Rayleigh fading channels. *IEEE Transactions on Consumer Electronics*, *49*(4), 973-977. doi:10.1109/TCE.2003.1261183

Chiu, Y., Chang, C.-ju, Feng, K.-ten, & Ren, F.-ching. (2010). GGRA : A Feasible Resource-Allocation Scheme by Optimization Technique for IEEE 802 . 16 Uplink Systems. *IEEE Transactions on Vehicular Technology*, *59*(3), 1393-1401.

Chowdhury, P., & Misra, I. S. (2010). A Fair and Efficient Packet Scheduling Scheme for IEEE 802 . 16 Broadband Wireless Access Systems, *1*(3).

Dai, H., Cherniakov, M., & Hong, J. (2006). Link Budget Analysis in Mobile Communication System. *International Conference on Communication Technology, 2006. ICCT ’06.* (pp. 1-4). Ieee. doi:10.1109/ICCT.2006.341977

Dent, P., Bottomley, G. E., & Croft, T. (1993). Jakes fading model revisited. *Electronics Letters*, *29*(13), 1162 - 1163.

Girici, T., Zhu, C., Agre, J. R., & Ephremides, A. (2008). Practical Resource Allocation Algorithms for QoS in OFDMA-based Wireless Systems. *IEEE CCNC 2008 proceedings* (pp. 897-901).

Hwang, J.-Y., & Han, Y. (2007). An adaptive traffic allocation scheduling for mobile wimax. *IEEE International Symposium on Personal PIMRC*.

IEEE 802.16m-08/004r3. (2010).

Iyengar, R., Kar, K., Sikdar, B., & Luo, X. (2009). QoS Provisioning and Radio Resource Allocation in OFDMA based WiMAX Systems (Vol. 1, pp. 1-5). New York.

Jain, Raj, So-in, Chakchai, and A.-K. A. T. (2008). System Level Modeling of IEEE 802.16E Mobile WIMAX Networks: Key Issues, (October), 73-79.

Kalyana Chakravarthy, C., & Reddy, P. (2009). Modified Opportunistic Deficit Round Robin Scheduling for improved QOS in IEEE 802 . 16 WBA networks. *(IJCSIS) International Journal of Computer Science and Information Security*, *6*(2), 75-81.

Kumar, D., & Priyameenal, V. (2011). Adaptive Packet Scheduling Algorithm for Real-Time Services in Wi-MAX Networks. *IEEE-International Conference on Recent Trends in Information Technology, ICRTIT 2011*.

Laias, E., Awan, I., & Chan, P. M. L. (2009). Fair and Latency Aware Uplink Scheduler In IEEE 802 . 16 Using Customized Deficit Round Robin. *International Conference on Advanced Information Networking and Applications Workshops 2009* (pp. 425-432). doi:10.1109/WAINA.2009.101

Lee, H., Kwon, T., Cho, D.-ho, & Science, C. (2006). Extended-rtPS Algorithm for VoIP Services in IEEE 802 . 16 systems. *Proceedings of IEEE ICC*.

Lopez Aguilar, F. (2009). Effective SNR mapping algorithms for link prediction model in 802. *International Conference on Ultra Modern Telecommunications & Workshops, 2009. ICUMT '09.* (pp. 1-6).

Mumtaz, S., Gamerio, A., & Rodriguez, J. (2008). EESM for IEEE 802.16e: WiMAX. *Seventh IEEE/ACIS International Conference on Computer and Information Science* (pp. 361-366). doi:10.1109/ICIS.2008.33

Papapanagiotou, I., Toumpakaris, D., & Devetsikiotis, M. (2009). A survey on next generation mobile WiMAX networks: objectives, features and technical challenges. *IEEE Communications Surveys & Tutorials*, *11*(4), 3-18. doi:10.1109/SURV.2009.090402

Peng, Z., Guangxi, Z., Hongzhi, L., & Haibin, S. (2007). Adaptive Scheduling Strategy for WiMAX Real-time Communication. *Proceedings of International Symposium on Intelligent Signal Processing and Communication Systems* (pp. 718-721). China.

REPORT ITU-R M: Requirements related to technical performance for IMT-Advanced radio interface(s). (2008).

Rakesh, J., A, W. V., & Dalal, U. (2010). A Survey of Mobile WiMAX IEEE 802 . 16m. *(IJCSIS) International Journal of Computer Science and Information Security*, *8*(1), 125-131.

Rath, H. K., Bhorkar, A., & Sharma, V. (2006). An Opportunistic Uplink Scheduling Scheme to Achieve Bandwidth Fairness and Delay for Multiclass Traffic in Wi-Max (IEEE 802 . 16) Broadband Wireless Networks. *IEEE GLOBECOM* (pp. 0-4).

Ravichandiran, C., Raj, P., & Vaithiyanathan. (2010). Analysis and Modification of Scheduling Algorithm for IEEE 802 . 16e (Mobile WiMAX). *Second International Conference on Communication Software and Networks 2010*. doi:10.1109/ICCSN.2010.87

Reddy, Y. B., Gajendar, N., Taylor, P., & Madden, D. (2007). Computationally Efficient Resource Allocation in OFDM Systems: Genetic Algorithm Approach. *Fourth International Conference on Information Technology (ITNG'07)* (pp. 36-41). Ieee. doi:10.1109/ITNG.2007.57

So-in, C., Member, S., Jain, R., & Tamimi, A.-karim. (2001). Scheduling in IEEE 802 . 16e Mobile WiMAX. *IEEE J-SAC-BAN*.

Song, H., Kwan, R., & Zhang, J. (2011). Approximations of EESM Effective SNR Distribution. *IEEE Transactions on Communications*, *59*(2), 603-612. doi:10.1109/TCOMM.2011.011811.100056

Ukil, A., Sen, J., & Bera, D. (2009). A New Optimization Scheme for Resource Allocation in OFDMA Based WiMAX Systems. *International Conference on Computer Engineering and Technology 2009*, 145-149. Ieee. doi:10.1109/ICCET.2009.10

Wang, Y., Chan, S., Zukerman, M., & Harris, R. J. (2008). Priority-Based fair Scheduling for Multimedia WiMAX Uplink Traffic. *IEEE International Conference on Communications 2008*, 301-305. Ieee. doi:10.1109/ICC.2008.63

Yen, C.-ming, & Chang, C.-ju. (2009). Dynamic Priority Resource Allocation for Uplinks in. *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, *58*(8), 4587-4597.

Yoon, J.-pil, Kim, W.-jae, Baek, J.-young, & Suh, Y.-joo. (2008). Efficient Uplink Resource Allocation for Power Saving in IEEE 802 . 16 OFDMA Systems. *IEEE Vehicular Technology Conference, 2008. VTC Spring 2008.* (pp. 2167-2171).

Publications

1. Tahir, N. H. M., & Noordin, K. A. (2011, 14-16 Nov. 2011). Adaptive resource allocation scheme for uplinks in IEEE802.16 systems. Paper presented at the Information Technology and Multimedia (ICIM) on International Conference, 2011.