HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Electrical and Communications Engineering

Antti Mäkeläinen

Analysis of Handoff Performance in Mobile WiMAX Networks

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Engineering

3.12.2007
Espoo, Finland

Supervisor: Prof. Riku Jäntti

Instructor: Lic.Sc. (Tech.) Markus Peuhkuri

HELSINKI UNIVERSITY OF TECHNOLOGY

Abstract of the Master's Thesis

Author:	Antti Mäkeläinen	
Name of the Thesis:	Analysis of Handoff Performance in Mobile WiMAX Networks	
Date:	3.12.2007	Number of pages: 9 + 72 + 11
Department:	Department of Electrical and Communications Engineering	
Professorship:	Telecommunication Technology	
Supervisor:	Prof. Riku Jäntti	
Instructor:	Lic.Sc. (Tech.) Markus Peuhkuri	

This thesis introduces a new 802.16e-2005 amendment to the 802.16-2004 standard, generally known as WiMAX. The 802.16e-2005, or the Mobile WiMAX, introduces the most significant new feature, the support for handoffs, which can be considered as a basic requirement for mobile communication system. The mandatory handoff method in Mobile WiMAX is the Hard Handover and the two optional soft handoff methods are called Macro Diversity Handover and Fast Base Station Switching. To support mobility, the Mobile WiMAX introduces also the Scalable OFDMA, which is a multiplexing scheme that allows adjustments of bandwidth according to the physical conditions of the used channel at a certain moment. This makes possible the versatile deployment of various environments with different propagation characteristics. In addition, a mobile device has also the weakness of operating with limited power resources, which is addressed in Mobile WiMAX by introducing two power saving modes: the Sleep Mode and the Idle Mode.

The Mobile WiMAX technology is first presented as literature study and later on the handoff performance is tested with NS-2 simulations. The simulator did not initially include support for Mobile WiMAX so two add-on modules from NIST had to be installed. The handoff latency was measured and the parameters of the simulator adjusted in order to achieve best possible handoff times. The goal was to find out which parameters had the greatest impact on the handoff duration and to compare the results to the objectives set by the WiMAX Forum. The WiMAX Forum says that the Mobile WiMAX supports mobility up to 120 km/h and the handoff should take less than 50 ms.

The results showed that some of the parameters did not have an influence at all and some could be enhanced to achieve faster handoffs. The handoff times remained below the 50 ms limit up to 20 m/s (72 km/h). The promised higher speed handoffs are designed to use the soft handoff methods, which were not supported by the simulator. Additionally, the other lacks in the NIST add-on modules restricted feasible comparison of the simulation results to the standardized Mobile WiMAX.

Keywords: Mobile WiMAX, 802.16e-2005, mobility, handoff, handoff latency, HHO, MDHO, FBSS, simulation, NS-2

TEKNILLINEN KORKEAKOULU Diplomityön tiivistelmä

Tekijä:	Antti Mäkeläinen	
Työn nimi:	Yhteydenvaihdon suorituskykyanalyysi Mobiili WiMAX verkossa	
Päivämäärä:	3.12.2007	Sivumäärä: 9 + 72 + 11
Osasto:	Sähkö- ja tietoliikennetekniikan osasto	
Professuuri:	Tietoliikennetekniikka	
Työn valvoja:	Prof. Riku Jäntti	
Työn ohjaaja:	TkL Markus Peuhkuri	

Diplomityö esittelee uuden 802.16e-2005 lisäyksen 802.16-2004 standardiin, joka yleisesti tunnetaan nimellä WiMAX. 802.16e-2005, tai Mobiili WiMAX, tuo esiin tärkeimmän uuden ominaisuuden, tuen yhteydenvaihdolle, jota voidaan pitää perustavanlaatuisena ominaisuutena liikkuvalle tietoliikennejärjestelmälle. Pakollinen yhteydenvaihtotapa Mobiili WiMAX:ssa on nimeltään Hard Handover ja kaksi muuta valinnaista tapaa ovat Macro Diversity Handover ja Fast Base Station Switching, jotka perustuvat pehmeään yhteyden siirtoon. Liikkuvuuden mahdollistamiseksi Mobiili WiMAX esittelee myös skaalautuvan OFDMA:n, joka on hetkellisten olosuhteiden mukaisen taajuuskaistan säätämisen salliva kanavointimenetelmä. Tällöin voidaan monipuolisesti hyödyntää kukin toimintaympäristö erilaisilla yhteyslaaduilla. Lisäksi liikkuvan laitteen virrankäyttö on rajattua, joten Mobiili WiMAX tarjoaa kaksi uutta virransäästötilaa: Sleep Mode ja Idle Mode.

Mobiili WiMAX teknologia esitellään aluksi kirjallisuuskatsauksen muodossa, jonka jälkeen yhteydenvaihdon suorituskykyä tutkitaan NS-2 simulaatiossa. Simulaattori ei oletuksena sisältänyt tukea Mobiili WiMAX:lle, joten kaksi lisäosaa NIST:ltä piti asentaa. Yhteydenvaihdon latenssi mitattiin ja simulaation parametreja säädettiin nopeampien yhteydenvaihtojen saavuttamiseksi. Tavoitteena oli löytää parametrit, joiden muuttaminen vaikutti merkittävimmin yhteydenvaihdon kestoon sekä verrata tuloksia WiMAX Forum:n asettamiin tavoitteisiin. WiMAX Forum kertoo Mobiili WiMAX:n tukevan liikkuvuutta 120 km/h saakka ja yhteydenvaihdon keston jäävän alle 50 ms:n.

Tulosten mukaan joillakin parametreilla ei ollut lainkaan vaikutusta ja toisia pystyttiin parantamaan nopeampien yhteydenvaihtojen saavuttamiseksi. Yhteydenvaihtoihin kuluneet ajat jäivät 50 ms:n rajan alle aina 20 m/s (72 km/h) nopeuteen saakka. Luvatut korkeampien nopeuksien yhteydenvaihdot käyttävät mainittuja valinnaisia menetelmiä, joita simulaattori ei kuitenkaan tukenut. Lisäksi muut puutteet NIST:n lisäosissa rajoittivat simulaatiotulosten vertailua standardoituun Mobiili WiMAX:iin.

Avainsanat:	Mobiili WiMAX, 802.16e-2005, liikkuvuus, yhteydenvaihto, yhteydenvaihdon latenssi, HHO, MDHO, FBSS, simulaatio, NS-2

PREFACE

This Master's Thesis was carried out in the Networking Laboratory of the Helsinki University of Technology.

First, I would like to thank the funding organization of the project for making the thesis possible and providing an interesting topic to work on. I would also like to express my gratitude to my supervisor, Prof. Riku Jäntti, for advanced and fast feedback. My instructor, Lic.Sc. (Tech.) Markus Peuhkuri, has been an invaluable help and support during this process and I would like to thank him for all the comments and hints that guided my work forward. Additionally, I wish to thank Marko Repo, with whom I shared the office, for a pleasant atmosphere to work at.

Finally, I want to express my deepest gratitude to my wife, Laura, for the endless support and understanding throughout my studies and during this thesis process. I love you.

Espoo, December 3rd, 2007

Antti Mäkeläinen

TABLE OF CONTENTS

PREFACE .... **iv**

TABLE OF CONTENTS .... **v**

ABBREVIATIONS .... **vii**

1 INTRODUCTION **1**

2 MOBILE WIMAX RADIO NETWORKS **2**

2.1 Mobile WiMAX Physical (PHY) Layer ***2***

2.1.1 Orthogonal Frequency Division Multiple Access Basics 2
2.1.2 OFDMA Symbol Structure and Sub-Channelization 6
2.1.3 Time Division Duplex Frame Structure 7
2.1.4 Other Advanced PHY Layer Features 9

2.2 Mobile WiMAX Medium Access Control (MAC) Layer ***12***

2.2.1 Data/control plane 13
2.2.2 Quality of Service (QoS) Support 16
2.2.3 MAC Scheduling Service 18
2.2.4 Bandwidth Allocation and Request Mechanisms 19
2.2.5 Mobility Management 21
2.2.6 Security 21

2.3 Other Advanced Features of Mobile WiMAX ***23***

2.3.1 Smart Antenna Technologies 23
2.3.2 Fractional Frequency Reuse 25
2.3.3 Multicast and Broadcast Service 27

2.4 Performance ***28***

2.5 Other Candidates/Competitors ***28***

2.5.1 3G/3.XG/4G (Post-GSM Generations) 28
2.5.2 FuTURE Project 30
2.5.3 Mobile Wireless Broadband Access (802.20) 31
2.5.4 Media Independent Handover (802.21) 32
2.5.5 Wireless Regional Area Network (802.22) 33
2.5.6 Wireless Local Area Network (802.11x) 33

2.6 Summary ***35***

3 MOBILITY **36**

3.1 Network Architecture ***36***

3.1.1 Access Service Network 38
3.1.2 Connectivity Service Network 38
3.1.3 Network Reference Model 38
3.1.4 Interworking with Other Technologies 41

3.2 Handoff ***41***
3.2.1 Handoff Types 42
3.2.2 Handoff Process 46
3.2.3 Handoff Method Comparison 52
3.3 Power Management ***53***
3.3.1 Sleep Mode 54
3.3.2 Idle Mode 56
3.4 Summary ***59***
4 SIMULATIONS **60**
4.1 Scenario ***60***
4.2 Components ***61***
4.2.1 Neighbor Discovery -module 61
4.2.2 Media Independent Handover -module 62
4.3 Parameters ***63***
4.3.1 Constant Parameters 63
4.3.2 Adjusted Parameters 63
4.4 Velocity of MS ***64***
4.5 Summary ***66***
5 CONCLUSIONS **67**
REFERENCES **69**
APPENDIX **A.I**
A.1 Simulation Results ***A.I***
A.2 Simulation Code ***A.V***

ABBREVIATIONS

3GPP/2	3rd Generation Partnership Project/version 2
AAA	Authentication, Authorization and Accounting
AAS	Advanced Antenna Systems
AC	Access Concentrator
ACK	Acknowledgment
AES	Advanced Encryption Standard
AES-CCM	AES-CTR mode with CBC-MAC
AK	Authorization Key
AKA	Authentication and Key Agreement
AMC	Adaptive Modulation and Coding
AMS	Adaptive MIMO Switching
ASN	Access Service Network
ASN-GW	Access Service Network Gateway
BE	Best Effort
BPSK	Binary Phase Shift Keying
BS	Base Station
BSID	Base Station Identity
BSS	Basic Service Set
BTC	Block Turbo Code
CBC-MAC	Cipher Block Chaining Message Authentication Coder
CC (1)	Chase Combining
CC (2)	Convolutional Coding
CCI	Co-Channel Interference
CDMA	Code Division Multiple Access
CID	Connection Identifier
CINR	Carrier to Interference plus Noise Ratio
CMAC	Cipher based Message Authentication Code
CP	Cyclic Prefix
CQICH	Channel Quality Indicator Channel
CRC	Cyclic Redundancy Check
CSN	Connectivity Service Network
CTC	Convolutional Turbo Coding
CTR	Counter Mode Encryption
DC	Direct Current
DCD	DL Channel Descriptor
DL	Downlink
DoA	Direction of Arrival
DP	Decision Point
DSL	Digital Subscriber Line
DSx	DSA, DSC, or DSD; Dynamic Service Addition/Change/Deletion
EAP	Extensible Authentication Protocol
EP	Enforcement Point
ertPS	Extended Real-Time Polling Service
FBSS	Fast Base Station Switching
FCH	Frame Control Header
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FFT	Fast Fourier Transform
FSS	Frequency Selective Scheduling

FTP	File Transfer Protocol
FUSC	Full Usage of Sub-channels
FuTURE	Future Technologies for a Universal Radio Environment
GMC	Generalized Multi-Carrier
GPRS	General Packet Radio Service
GRD	Guard (interval)
GSM	Global System for Mobile communications
HA	Home Agent
HARQ	Hybrid Automatic Repeat Request
HHO	Hard Handoff
HMAC	keyed-Hash Message Authentication Code
HO	Handoff, or handover
HSDPA	High Speed Downlink Packet Access
HSOPA	High Speed OFDM Packet Access
HSPA	High Speed Packet Access
HSUPA	High Speed Uplink Packet Access
ID	Identifier
IE	Information Element
IEEE	Institute of Electrical and Electronics Engineers
IFFT	Inverse Fast Fourier Transform
IMT-A	International Mobile Telecommunications Advanced
IP (IPV4 or IPv6)	Internet Protocol (version 4 or 6)
IR	Incremental Redundancy
ISI	Inter Symbol Interference
ITU	International Telecommunication Union
KEK	Key Encryption Key
LDPC	Low Density Parity check Code
LSB	Least Significant Bit
LTE	Long Term Evolution
MAC	Medium Access Control
MAP	Map, mapping, definition
MBS	Multicast and Broadcast Service
MBWA	Mobile Broadband Wireless Access
MD5	Message-Digest algorithm 5
MDHO	Macro Diversity Handover
MIH	Media Independent Handover
MIMO	Multiple Input Multiple Output
MPEG	Moving Picture Experts Group
MS	Mobile Station
MSB	Most Significant Bit
MS-CHAPv2	Microsoft-Challenge Handshake Authentication Protocol
MSH-DSCH	Mesh Mode Schedule with Distributed Scheduling
NACK	Negative Acknowledgment
NAP	Network Access Provider
ND	Neighbor Discovery
NIST	National Institute of Standards and Technology
NRM	Network Reference Model
nrtPS	Non Real-Time Polling Service
NS-2	Network Simulator version 2
NSP	Network Service Provider
NWG	Network Working Group
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PDA	Personal Digital Assistant

PDU	Protocol Data Unit
PKMv1/2	Privacy Key Management version 1 or 2
PMP	Point-to-multipoint
PRBS	Pseudo-Random Binary Sequences
PSTN	Public Switched Telephone Network
PUSC	Partial Usage of Sub-channels
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RA	Router Advertisement
RoF	Radio over Fiber
RRA	Radio Resource Agent
RRC	Radio Resource Controller
RRM	Radio Resource Management
RS	Router Solicitation
RSA	public key cryptography method developed by Rivest, Shamir, and Adleman
RTG	Receive/Transmit Transition Gap
rtPS	Real-Time Polling Service
SA	Security Association
SAID	Security Association Identity
SAP	Service Access Point
SDMA	Space-Division Multiple Access
SDU	Service Data Unit
SFN	Single Frequency Network
SIM	Subscriber Identity Module
SIMO	Single Input Multiple Output
SM	Spatial Multiplexing
SNR	Signal-to-Noise Ratio
S-OFDMA	Scalable OFDMA (also SOFDMA)
SS	Subscriber Station
STBC	Space-Time Block Code
STC	Space-Time Coding
TCP/IP	Transmission Control Protocol/Internet Protocol
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TEK	Traffic Encryption Key
TLS	Transport Layer Security
TTG	Transmit/Receive Transition Gaps
TTLS	Tunneled TLS
TUSC	Tiled Use of Sub-channel
UCD	UL Channel Descriptor
UGS	Unsolicited Grant Service
UL	Uplink
UMTS	Universal Mobile Telecommunications System
VoIP	Voice over IP
WAVE	Wireless Access for the Vehicular Environment
WCDMA	Wideband Code Division Multiple Access
WiMAX	Worldwide Interoperability for Microwave Access
Wi-Fi	Wireless Fidelity
WLAN	Wireless Local Area Network
WPA(2)	Wi-Fi Protected Access (version 2)
WRAN	Wireless Regional Area Network
VR-(N)RT	Variable-Rate (Non-)Real-Time

1 INTRODUCTION

This thesis introduces the IEEE 802.16e-2005 standard, also known as Mobile WiMAX (Worldwide Interoperability for Microwave Access), which defines the physical (PHY) and the Medium Access Control (MAC) layers of it. The network architecture of Mobile WiMAX is defined by the WiMAX Forum [1]. The 802.16-2005 [2] is the new, mobile version of the older WiMAX specification known as 802.16-2004 [3], which is a wireless, but fixed, data transmission scheme for providing broadband connection to metropolitan areas. The traditional WiMAX has lacked the ability for the user to move during the transmission. Due to a moving user, the need for the change of serving base station, a handoff, creates demands for the Mobile WiMAX. The handoffs, or as in some sources handovers, should be fast enough so that the ongoing video call or Voice over IP (VoIP) conversation is not interrupted, at least for so long that the user notices it.

The communications industry is heading towards wireless data transfer with great speed and several competing technologies are emerging to replace the old ones. The traditional Wireless Local Area Network (WLAN) has gained a strong place in the market and is definitely the leader for short distance wireless networks. However, the coverage and mobility are adequate for indoor usage only. The Mobile WiMAX is planned to be independent or to extend the mobile access when a user exits the WLAN hotspot coverage area.

The thesis is roughly divided in three major parts: the literature review discussing mainly the PHY and MAC properties of Mobile WiMAX, the Mobile WiMAX mobility, and the simulations. First, some facts about Mobile WiMAX are presented in order to understand the technology at hand. The chapter 2 includes also a brief overview to main competing technologies of Mobile WiMAX. The issues characteristic to mobility are discussed in more detail in a chapter of their own. The simulations part introduces the methods used and results achieved while testing the performance of Mobile WiMAX handoffs. Finally, a summary and conclusions based on simulations performed are also presented. Even some suggestions for future work are given.

2 MOBILE WIMAX RADIO NETWORKS

This chapter is committed to the Mobile WiMAX (802.16e-2005) technology. First, physical and MAC layer properties are introduced. Later on, the security issues are discussed with only very basic mobility capabilities, since the mobility issues have been dedicated an own chapter after this one. Some of the matters discussed here have a key role making the Mobile WiMAX mobile, but they are nevertheless left in this chapter to keep them in the right context. The material in this chapter is based mostly on references [2], [3], [4], and [5].

Additionally, some possibly competing or co-existing technologies in wireless communications are introduced. The traditional WLAN (802.11-family) is mentioned with the post-GSM technologies as well as some other IEEE 802 standards.

2.1 Mobile WiMAX Physical (PHY) Layer

The 802.16-2004 specification and the 802.16e-2005 amendment define five PHY alternatives:

- WirelessMAN-SC (Wireless Metropolitan Area Network using Single Carrier Modulation for use in the 10-66 GHz bandwidth region)
- WirelessMAN-SCa (Wireless Metropolitan Area Network using Single Carrier Modulation for use on bandwidths below 11 GHz)
- WirelessMAN-OFDM (Wireless Metropolitan Area Network using OFDM)
- WirelessMAN-OFDMA (Wireless Metropolitan Area Network using OFDMA)
- WirelessHUMAN (Wireless High-Speed Unlicensed Metropolitan Area Network)

The OFDMA-version creates the basis for functionality of Mobile WiMAX and therefore discussed more in this chapter, while the others are left aside.

2.1.1 Orthogonal Frequency Division Multiple Access Basics

Orthogonal Frequency Division Multiplexing (OFDM) is a digital modulation scheme suited especially well for terrestrial broadcasting. It can handle multipath propagation and delays between received signals. The OFDM is sensitive to frequency changes as Doppler shift while the Mobile Station (MS) is moving. However, the delay spread is not a great problem in the OFDM because of the increased symbol duration.

The Orthogonal Frequency Division Multiple Access (OFDMA) is a version of OFDM and intended for several user mobile communications environments. It is the solution considered to be the modulation scheme in most future advanced wireless communications technologies, as examples can be mentioned the Long Term Evolution (LTE, discussed later in 2.5.1), the Mobile Broadband Wireless Access (MBWA, discussed later in 2.5.3), or the Mobile WiMAX. In [6], the possible success of Mobile WiMAX is doubted and they are presenting that by 2009 the number of global subscribers would be over 920 million for 3G and only

14.3 million for (Mobile) WiMAX. Another view is presented in [7], where by 2009 the (Mobile) WiMAX would have 40 percent share of wireless broadband markets. Nevertheless, these issues and figures can vary significantly depending on the announcing party, mainly whether their business is related or not to the WiMAX. The actual figures are probably something between the mentioned and this depends on the success of launching the new technology.

The OFDMA has several advantages over traditional Code Division Multiple Access (CDMA)-versions used in post-GSM 3G technologies. The spectral efficiency is higher and the fading can be tolerated better. In OFDMA data streams from different users are combined to sub-channels in both Downlink (DL) and Uplink (UL). However, there are some drawbacks as well. Since the manufacturing of OFDMA electronics is rather complex, the expenses rise at the same time. Additionally, the Co-Channel Interference (CCI) from neighboring cells is less disturbing in CDMA than in OFDM. The CCI can although be mitigated by using Fractional Frequency Reuse (described later in sub-section 2.3.2). [8]

OFDM Basic Principle

With OFDM the used bandwidth is divided into several frequency sub-carriers so that they are orthogonal to each other. The stream of input data is separated into multiple, parallel sub-streams with reduced data rate. Then the sub-streams are modulated individually and sent on separate sub-carriers. Consequence of this is the increase in symbol duration.

Since the long signal duration decreases Inter Symbol Interference (ISI) caused by multipath propagation, it is efficient to transmit the low-rate streams in parallel, instead of one high-rate stream. The signal duration is long, so by using a proper guard interval, the ISI can be avoided totally, assuming the guard interval is longer than the difference between the first and last multipath echo. The Figure 2.1 below illustrates the principle of several sub-streams combined at the transmitter and separated again at the receiver.

As seen in the Figure 2.1 the information is coded and modulated across the sub-carriers before performing an Inverse Fast Fourier Transform (IFFT). The IFFT takes advantage of the frequency diversity of the multipath channel. Finally, before transmitting the data, the streams are combined to a single signal and sent to the air interface. At the receiver end the procedure is the same, but in reversed order. The 802.16e specification defines the Fast Fourier Transform (FFT) size to be 128, 512, 1024, or 2048 with respective channel bandwidths 1.25, 5, 10, and 20 MHz. However, the Mobile WiMAX allows other bandwidth profiles to be used as well, but the sub-carrier frequency can not be kept constant anymore (more in the next sub-subsection).

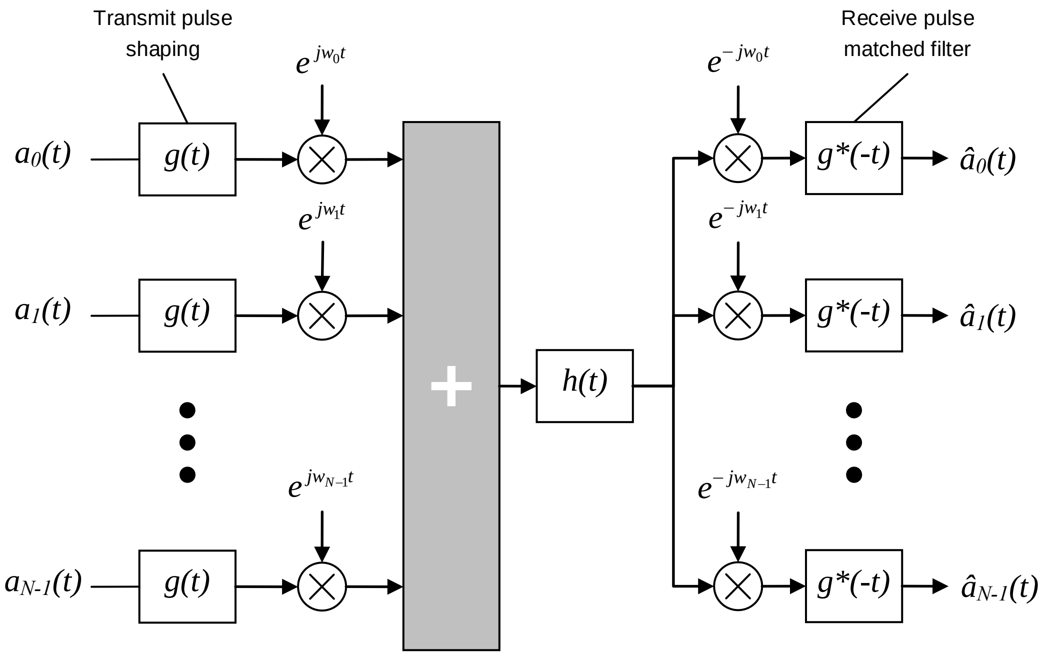

Figure 2.1 - Basic Architecture of an OFDM System [4]

The available resources of OFDM can be divided into time and frequency domains. In the time domain OFDM symbols can be used and frequency domain has sub-carriers. Both of these can be utilized for individual users by using sub-channels. [4]

Scalable OFDMA

Scalable OFDMA (S-OFDMA or SOFDMA) creates the basis for 802.16e-2005. Basically S-OFDMA means a possibility to adjust the used bandwidth and this way different environments with varying spectral requirements can be served. The bandwidth adjustment can be chosen between 1.25-20 MHz as described in Table 2.1 below. The scalability is realized with FFT size variations and the frequency spacing of sub-carriers is defined to be 10.94 kHz.

Table 2.1 - OFDMA Scalability Parameters [4]

Parameters	**Values**			
System Channel Bandwidth (MHz)	1.25	5	10	20
Sampling Frequency (MHz)	1.4	5.6	11.2	22.4
FFT Size	128	512	1024	2048
Number of Sub-Channels	2	8	16	32
Sub-Carrier Frequency Spacing	10.94 kHz			
Useful Symbol Time (Tb=1/f)	91.4 µs			
Guard Time (Tg=Tb/8)	11.4 µs			
OFDMA Symbol Duration (Ts=Tb+Tg)	102.9 µs			
Number of OFDMA Symbols (5 ms frame)	48			

Cyclic Prefix

During the guard interval a Cyclic Prefix (CP) is sent and usually the CP has the same length as the guard interval. The CP consists of the end of the symbol placed in the beginning of the new symbol, as can be seen in Figure 2.2.

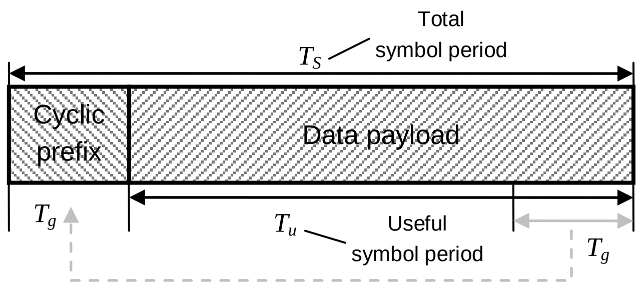

Figure 2.2 - Insertion of Cyclic Prefix [4]

The task of the CP is to settle the echoes from multipath propagation before the actual data can be processed. There are also other benefits while using the CP. For example, inter-block interference (hence, the interference between symbols *n* and *n+2*) is prevented and the channel seems circular. In addition, low-complexity frequency domain equalization is allowed. A negative aspect with the use of CP is the extra overhead needed and therefore the bandwidth efficiency is affected. However, the channel bandwidth can be used in an efficient way for data transmission since the OFDM spectrum fades fast outside the actual window containing the carriers, as can be seen in Figure 2.3. It is also important to keep the CP length defined by the Base Station (BS) during initialization, since the change of it would force all other MSs to resynchronize.

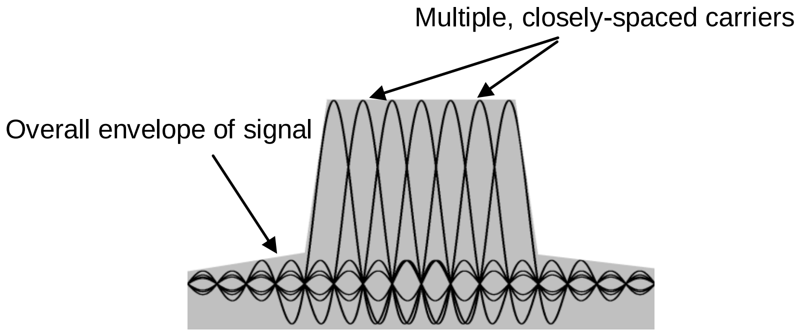

Figure 2.3 - OFDM Spectrum

The selection of FFT size is a matter of balance between problems caused by multipath propagation, Doppler shift, and cost/complexity. If the FFT size is increased, the sub-carrier spacing is decreased and the total symbol length is increased. Both of these make the system stronger against multipath delay spread. However, by narrowing the sub-carrier spacing, the Doppler shift causes intercarrier interference in mobile applications. Additionally, larger FFT sizes require more computing power as well, which again increases the costs. [9]

2.1.2 OFDMA Symbol Structure and Sub-Channelization

The OFDMA symbol structure is shown in Figure 2.4. As can be pointed, three types of sub-carriers are used in OFDMA symbols. Data sub-carriers handle the transmission of data, pilot sub-carriers are for the estimation and synchronization use, and null sub-carriers have no transmission, but they are intended for guard bands and Direct Current (DC) carriers.

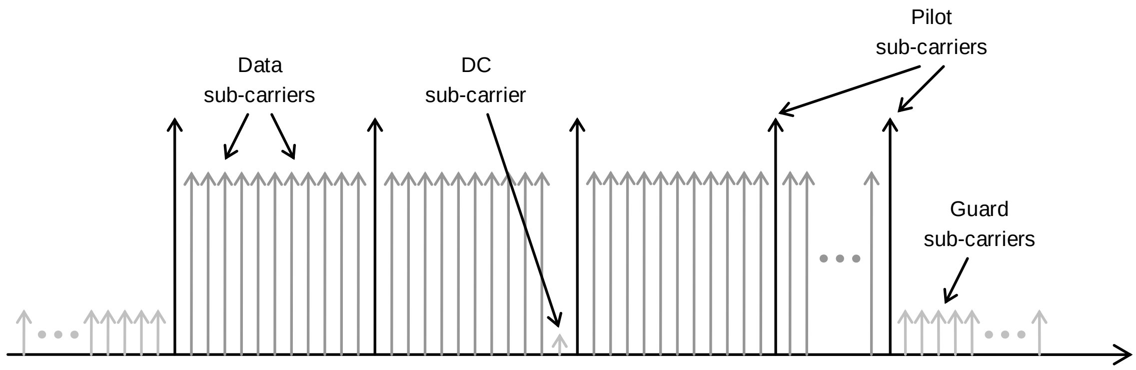

Figure 2.4 - OFDMA Sub-carrier Structure [4]

The definitions sub-carrier and sub-channel may be confusing, but they can be clarified in the following way: a subset of data or pilot sub-carriers is a sub-channel and the OFDM symbol consists of several sub-channels. Sub-channelization defines the smallest time-frequency resource unit to be a slot. One slot is the same as 48 data tones, in other words sub-carriers. Sub-channelization is supported in both link directions.

Sub-carrier permutation of sub-channelization can be done in two ways in Mobile WiMAX, with diversity or contiguous permutation. In the most cases diversity permutations are better suited for mobile environment while the contiguous permutations are more applicable for environment with fixed, portable, or low-mobility devices. The choice between these two can be made either to increase throughput or to give more flexibility considering the movement of the user.

Diversity Permutation

The diversity permutation, also known as distributed permutation [10], arranges sub-carriers in a pseudo-random way to create a sub-channel. This enables frequency diversity and averages the inter-cell interference. The available permutations are Downlink Full Usage of Sub-channels (DL FUSC), Downlink Partial Usage of Sub-channels (DL PUSC) and UL PUSC. The 802.16e-2005 defines also optional permutations, such as optional FUSC (OFUSC), Tiled Use of Sub-channel 1 and 2 (TUSC1/2).

Contiguous Permutation

The adjacent, or the contiguous permutation, creates a sub-channel by grouping a block of contiguous sub-carriers. The available modes for contiguous permutation are the Adaptive Modulation and Coding for both DL and UL (DL AMC and UL AMC). The optional AMC allows modulation and coding adjustments to be performed based on current channel conditions.

The Table 2.2 below describes the main differences between the two permutation types.

Table 2.2 - Comparison of Permutation Modes [11]

	Contiguous sub-carrier permutation (AMC)	**Diversity sub-carrier permutation (PUSC, FUSC)**
Benefits	Sub-channelization gain; Frequency selective loading gain	Sub-channelization gain; Frequency diversity; Inter-cell Interference averaging
Scheduling	Advanced frequency scheduler to explore frequency selectivity gain	Simple scheduler; Rely on frequency diversity to achieve robust transmission
Channel condition	Stationary channel	Fast-changing channel
Favorable smart antenna technology	Beamforming	MIMO

2.1.3 Time Division Duplex Frame Structure

The Mobile WiMAX used to support only Time Division Duplex (TDD) but recently full and half-duplex Frequency Division Duplex (FDD) support has been added too. This is mainly because of local restrictions in some areas. A drawback for TDD is that it needs to be synchronized over the whole system, but however, there are several reasons for preferring the use of TDD. The ratio of DL/UL data rates can be adjusted freely while with the FDD the ratio is always constant, and in most cases symmetric. The use of TDD assures channel reciprocity, which gives better support of link adaptation, MIMO (Multiple Input Multiple Output) and other closed loop advanced antenna technologies. On the other hand, FDD requires a pair channel while TDD can share one for both DL/UL traffic. Additionally, from the economical point of view, FDD transceivers are more complex and therefore more expensive to manufacture. [4]

An OFDMA frame begins with a preamble and continues with both DL and UL sub-frames, which can have the length ratio varying from 3:1 to 1:1. They are separated by Transmit/Receive and Receive/Transmit Transition Gaps (TTG and RTG). The TTG follows the DL sub-frame and RTG the UL. These are used for collision avoidance. Figure 2.5 demonstrates the structure of an OFDMA TDD frame. There are additional and optional fields as well that can be used in the sub-frames. The mandatory and optional fields are discussed in sub-subsections after the figure. [2] [9]

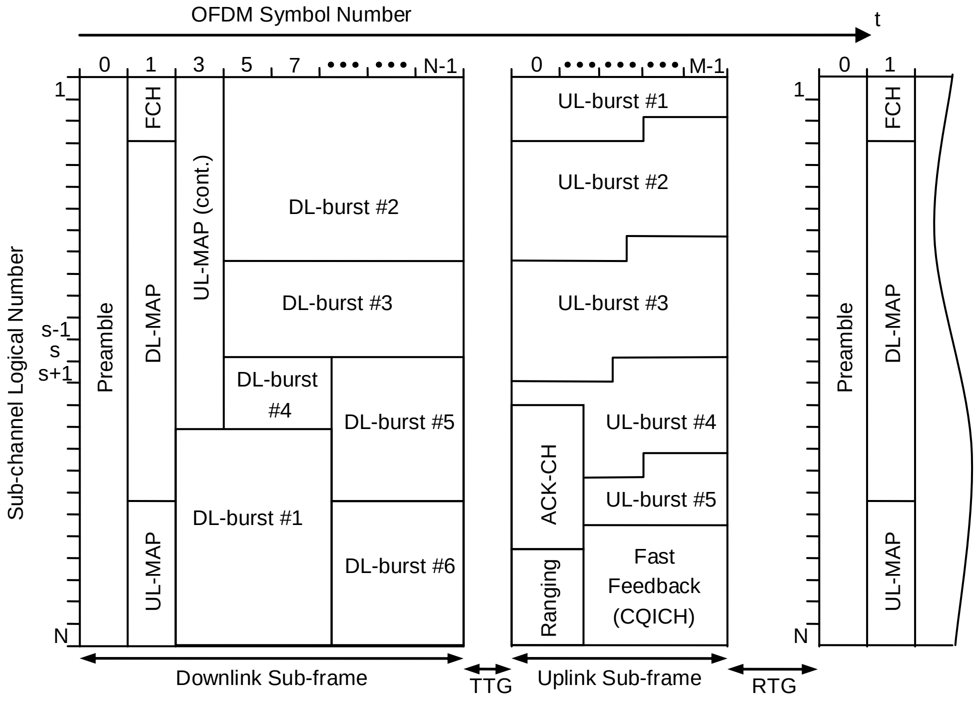

Figure 2.5 - OFDMA Frame Structure in TDD

Preamble

As mentioned, the preamble is the first symbol in an OFDMA (TDD) frame. It can consist of one or two symbol, depending on the type of preamble. There is a preamble before both the DL and UL sub-frames. The preamble before the DL sub-frame is longer with two symbols while the UL preamble consists of only one symbol. The task of preamble is the frequency synchronization and adjustments for amplitude and phase of the signal.

An optional scheme for preambles is supported as well, where short preambles, called midambles, can be used in UL after 8, 16, or 32 symbols, or in DL in the beginning of each burst. The use of more frequent preambles gives tolerance against rapid variations in time.

Frame Control Header (FCH)

The preamble is followed by a Frame Control Header (FCH), which contains frame configuration information about the length of MAP message following the FCH, modulation and coding scheme, and the available sub-carriers.

DL/UL Mapping (MAP) messages

After the FCH, the DL sub-frame contains broadcast DL-MAP and UL-MAP messages. They are intended for mapping the data region allocations of different users within the frame. The MAP messages describe the burst profiles (define the used modulation and coding scheme) for each user. The MAP messages are considered to contain very important information so they are usually sent over a robust link with low modulation, for example Binary Phase Shift Keying (BPSK). In a case of several users with small packets (VoIP), the overhead can increase significantly. However, in such cases the Mobile WiMAX has an optional solution to use several sub-MAP messages that have dedicated control messages to different users and that use higher rates determined individually by their respective channel properties. To achieve even more efficiency the MAP messages can even be compressed. The MAP messages can also be followed by DL/UL Channel Descriptor (DCD and UCD) messages that contain physical channel characteristics for DL and UL, respectively.

DL Bursts

The actual data payload from different users is carried within DL bursts that can have varying size or type depending on the application of a user. Additionally, the number of bursts is flexible and can change on a frame-by-frame basis, as the other metrics as well. The frame itself can also be between 2 to 20 ms.

UL Ranging

The UL sub-frame contains a field for ranging purposes. This field is reserved for the MSs to do periodic closed-loop time, frequency, and power adjustments with ranging requests. Additionally, requesting UL bandwidth can be realized with this field. The ranging process is described in more detail later in sub-section 3.2.2 when the handoff process is discussed. The ranging channel can additionally be used to transmit best effort data which would be too small to dedicate an actual UL allocation for it.

UL Bursts

The UL bursts have the same functionality as the bursts in DL direction; hence they provide the way to carry the different sized data from several users served by the same BS.

Optional fields

The UL sub-frame can have two optional fields that can be used to enhance the performance of Mobile WiMAX. The UL fast feedback channel (UL CQICH) is intended for the MS to give feedback about channel-state information that can be used at the BS scheduler. The UL acknowledgment channel (UL ACK) gives the MS a chance to give feedback to DL HARQ (Hybrid Automatic Repeat Request) acknowledgements in this field.

2.1.4 Other Advanced PHY Layer Features

To achieve performance improvements in coverage and capacity Mobile WiMAX has received some new features compared to the traditional WiMAX.

Adaptive Modulation and Coding (AMC)

Modulation techniques required for the DL direction in Mobile WiMAX are Quadrature Phase Shift Keying (QPSK), 16-point Quadrature Amplitude Modulation (16QAM), and 64QAM. The last one is optional in the UL direction. Coding is achieved with Convolutional Coding (CC) or Convolutional Turbo Coding (CTC) with variable code rate and repetition coding. Optional coding methods are Block Turbo Code (BTC) and Low Density Parity check Code (LDPC). Table 2.3 below describes the supported modulation and coding methods.

Table 2.3 - Supported Modulations and Codes (*optional*) [4]

		DL	**UL**
Modulation		QPSK, 16QAM, 64QAM	QPSK, 16QAM, *64QAM*
Code rate	**CC**	1/2, 2/3, 3/4, 5/6	1/2, 2/3, *5/6*
	CTC	1/2, 2/3, 3/4, 5/6	1/2, 2/3, *5/6*
	Repetition	x2, x4, x6	x2, x4, x6

The modulation and the coding have a direct impact on the achievable data rate, hence, one reaches higher rates with the 64QAM, but the channel is more vulnerable to disturbances, and vice versa for the QPSK. Since there are several possibilities to choose the modulation and the coding, the range of data rates in Mobile WiMAX is rather versatile. There are some achievable theoretical rates mentioned later in Table 2.6 while advanced antenna technologies are introduced.

Fast Feedback Channel (CQICH)

As already mentioned, the OFDMA UL sub-frame has an optional Channel Quality Indicator Channel (CQICH) used for delivering information about the channel conditions from MS to the scheduler of the BS. Some information, such as physical Carrier to Interference plus Noise Ratio (CINR), effective CINR, MIMO mode selection, and frequency selective sub-channel selection can also be delivered back to the BS. While using TDD, channel reciprocity offers an accurate channel condition measure for the link adaptation. [4]

Hybrid Automatic Repeat Request (HARQ)

Automatic repeat request methods are intended for situations when the sent packet has not been received properly, for example due to bit errors, and a retransmission is required. The HARQ is an optional part of the Mobile WiMAX MAC and is supported only by the OFDMA PHY. The use of HARQ and the needed parameters are negotiated during the process of network entry, or re-entry in case of a handoff. There is support for HARQ based on per-terminal or per-connection usage, hence, by all the active CIDs of a terminal or by a single CID, respectively. However, these two must not be used simultaneously in a single terminal. In [2], it is defined that a Subscriber Station (SS) should use the per-terminal implementation, while an MS should use the per-connection one. The SS is a general term of a station in WiMAX and MS is a special case supporting mobility, however the SS can also be referred to as Fixed WiMAX station.

While using HARQ, several MAC Protocol Data Units (PDU) can be combined to a HARQ packet. The HARQ packet is formed by adding a Cyclic Redundancy Check (CRC) field to the MAC PDUs. An example is shown in Figure 2.6 below. The parity field contains the information for possible error detection and correction.

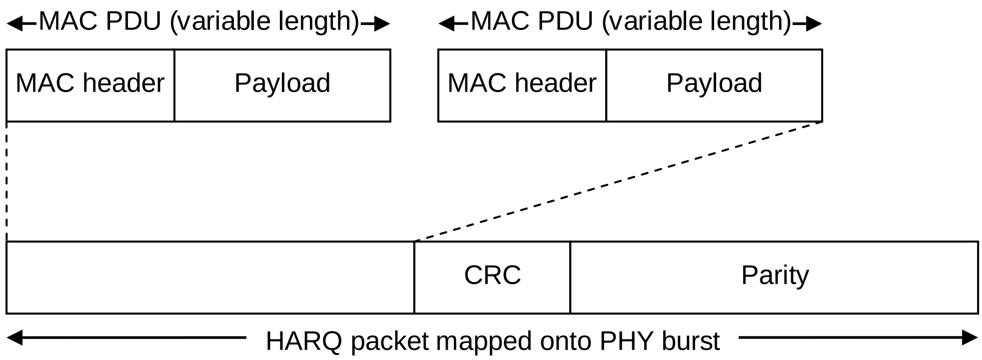

Figure 2.6 - Construction of HARQ Encoder Packet [2]

The HARQ is fundamentally a protocol with a stop-and-wait nature with support for several HARQ channels per connection. Each of these channels can have an encoder packet waiting simultaneously. This way the bandwidth resource of a waiting HARQ process can be used for data transfer of some other process. As a result, the total throughput can be improved. However, to avoid problems with memory usage, the number of channels should be kept low enough.

As stated earlier, the UL sub-frame includes an option for dedicated acknowledgment channel (UL ACK) to be used with HARQ ACK/NACK signaling from MS to the BS. The BS uses a HARQ Bitmap IE (Information Element) for the DL. Upon successful reception the MS sends an ACK and in the contrary situation a Negative Acknowledgment (NACK) is transmitted to the BS. After NACK or no ACK at all after a certain time is received, the BS retransmits the HARQ packet until an ACK is received. The receiver has stored the previous, failed packet and can try to combine the information from the first one with the retransmission. Hence, the decoding of the packet can succeed even though the retransmission also has faults, as long as the correct information from both of them is enough to create a complete packet. Here the HARQ functionality is described in the DL direction, but the scenario can be changed to the UL as well.

There are two main variants for HARQ operation in Mobile WiMAX, the Chase Combining (CC) and Incremental Redundancy (IR). Their operations vary in the procedure of retransmission described in the previous paragraph. The CC always retransmits the packet with the same encoding as the original packet and tries to combine the received data from retransmission(s) with the original data, as described above. On the other hand, the IR relies on changing the encoding to a version with better tolerance against faulty transmissions. The original message is sent using a high level coding with only detection and possibly simple correction capabilities. When an error is detected at the receiver, the packet is stored at buffer and receiver notifies the transmitter that error has happened. The transmitter sends back the original parity block with additional bits and a code to achieve a better error detection/correction.

The HARQ increases the reliability of transmission in general and additionally the performance of the connection on the cell edge can be improved. The signaling in WiMAX allows fully asynchronous operation, which makes a variable delay between retransmissions possible. This gives more flexibility to the scheduler, but additional overhead is required for each retransmission allocation. According to WiMAX Forum [4], when using HARQ with AMC and CQICH, Mobile WiMAX can create a link for a MS moving at vehicular speeds over 120 km/h. [2] [3] [4] [9]

2.2 Mobile WiMAX Medium Access Control (MAC) Layer

The development of Mobile WiMAX was considered to fulfill requirements for all major traffic types present today. This means that it will have to support voice, data and video, even simultaneously. The normal voice calls are very sensitive to latency and on the other hand video-streams demand capacity for transmission. Data traffic requires also a high-speed connection, but the bandwidth usage comes in bursts. The allocated resource can vary from a single time slot to the entire frame. Variations can be made frame-by-frame.

The WiMAX introduces two sharing mechanisms for the air interface. The Point-to-Multipoint (PMP) networking and Mesh networking are described in the following sub-sections followed by other available services related to the MAC layer in Mobile WiMAX.

The DL in WiMAX runs on a PMP networking basis, which means the principle of a single BS transmitting to one or several users. On a certain frequency channel and antenna sector all SSs, fixed or mobile, receive the same transmission from the BS, unless it is clearly defined in DL-MAP that a certain sub-frame is for a certain SS. The CIDs in received PDUs are inspected by the receiver SSs and only those addressed to them are held. The UL is shared on a demand basis and the resources are given according to the services needed. Naturally, the BS determines whether the SS has the right to access them or not.

The WiMAX offers also optional Mesh networking, which reminds the ad-hoc functionality used in WLAN (802.11) networks. Hence, the traffic can be routed through other SSs or communication between two SSs without a BS is possible. The Mesh network has a station, with a direct connection to the backhaul services outside the Mesh network, named as Mesh BS while all other stations are called Mesh SSs. The terminology in Mesh networks differs slightly from the used to as the stations are called nodes and the terms neighbor, neighborhood, and extended neighborhood are presented. Nodes with direct links (one hop distance) to each other are called neighbors and neighbors of a node create a neighborhood. The extended neighborhood includes the neighbors of neighborhood (two-hop distance).

Unlike in PMP, the Mesh BS can not transmit without agreeing with other nodes, hence, up to the nodes in the extended neighborhood. With distributed scheduling, all nodes broadcast periodically according agreed schedules their current schedules, with possible proposed changes to them, to all neighbors in a two-hop distance. The nodes must ensure that their transmission does not result in collisions with any other data or control traffic from members of normal or extended neighborhoods, in both DL and UL. Another Mesh networking scheduling method is called centralized scheduling, which is based on the Mesh BS that collects resource requests from neighboring SSs within determined distance of hops. It decides permissions for DL and UL traffic, and informs the Mesh SSs within the hop range about the granted resources. [2] [3]

2.2.1 Data/control plane

Addressing and Connections

In PMP, every air interface in a MS is given a unique MAC address; just like for example regular Ethernet network adapters. The address is used with initial ranging process and, additionally, as a part of authentication process. A connection between BS and MS is identified with a CID. When the MS contacts the BS, two management connection pairs (DL and UL) are created. Additionally, an optional third pair can be created as well. These connections form three different Quality of Service (QoS) levels available for management traffic. The first connection, called basic, is intended for exchanging short and time-urgent messages, while the other, called primary, can be used for longer and more delay-tolerant management traffic between the MS and BS MACs. The optional, third connection is called secondary management connection, which is used for delay-tolerant, standards-based messages. Additionally, a broadcast connection for delivery of some management messages is also available.

The CIDs for above management connections are defined in ranging or registration response messages, RNG-RSP and REG-RSP, respectively. Both members of a connection pair (DL and UL) share the same CID. The information sent on these connections shall never be sent on transmission connections intended for data traffic.

Transmission requests are based on management CIDs because the allowable bandwidth can vary for different connections even while still having the same service type. This can occur for example in an office, where a single MS is serving several users with different service limits. In such case the MS would make the requests for all the users. Additionally, traffic from several higher level sessions (such as TCP/IP) with common service requirement parameters can be combined to a single connection, since the addressing (sources and destinations) for local area network is encapsulated in the payload part of transmissions.

The Mesh networking also uses MAC addresses, just like the PMP does above. Though, here the authentication does not concern just the MS and BS as in PMP, but here the node and the network identify each other. After successful authentication the node receives a Node Identifier (Node ID), requested from the Mesh BS. Additionally, the nodes create Link Identifiers (Link ID) between every neighboring node they are linked to. The Link IDs are used with distributed scheduling in order to identify requests and grants of resources. In Mesh networking the traffic is broadcast to all nodes, which can determine the granted schedule by investigating the Node ID of the transmitter and Link ID in a Mesh Mode Schedule with Distributed Scheduling (MSH-DSCH) message. [2] [3]

MAC Protocol Data Unit (PDU) formats

The structure of a MAC PDU is demonstrated in Figure 2.7 below. The PDU begins with a generic MAC header field that has a fixed length. The second field is the payload which may also be empty. The payload can, but does not have to, consist of subheaders or MAC Service Data Units (SDU) and/or fragments thereof. The length of the payload can vary, so the MAC PDU can not be explicitly determined in bytes. The final field is used for a CRC, which is required for OFDM and OFDMA PHY layers, while for some 802.16-2004 PHY layers it is optional.

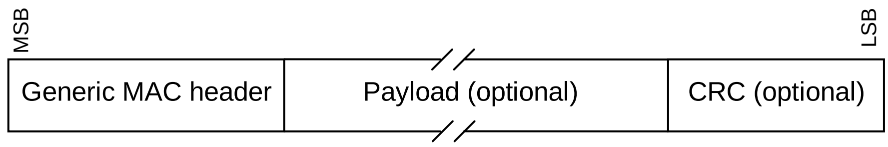

Figure 2.7 - MAC PDU Format [3]

MAC Management Messages

In this sub-subsection some important management messages for the MAC layer are presented. The messages are transmitted in the payload part of the PDU.

Downlink/Uplink Channel Descriptor (DCD/UCD) messages are sent periodically and the function for them is to provide the characteristics of the downlink and uplink physical channels.

Downlink/Uplink map (DL/UL-MAP) messages are intended for definition of access to the downlink/uplink information.

Ranging Request/Response (RNG-REQ/RSP) messages are a request-response pair during the initial network entry process. The RNG-REQ is sent by the MS during the initialization and later in a periodic way. The ranging process determines the delay in the network with request for power and/or downlink burst profile changes. RNG-RSP message is a response to the previous RNG-REQ message. The RNG-RSP can also be sent asynchronously in order to apply adjustments according to measured values from other received data or MAC messages. Hence, the MS may receive RNG-RSP message anytime, not just upon request.

Registration Request/Response (REG-REQ/RSP) messages are used during the initialization phase. The MS request registration by sending the REG-REQ message to the BS and the BS responses with REG-RSP. The messages include information about more detailed properties of the connection to be created.

Construction and Transmission of MAC PDUs

The flow process for MAC PDU construction is described in Figure 2.8 below and the different phases appearing are shortly explained in this sub-subsection.

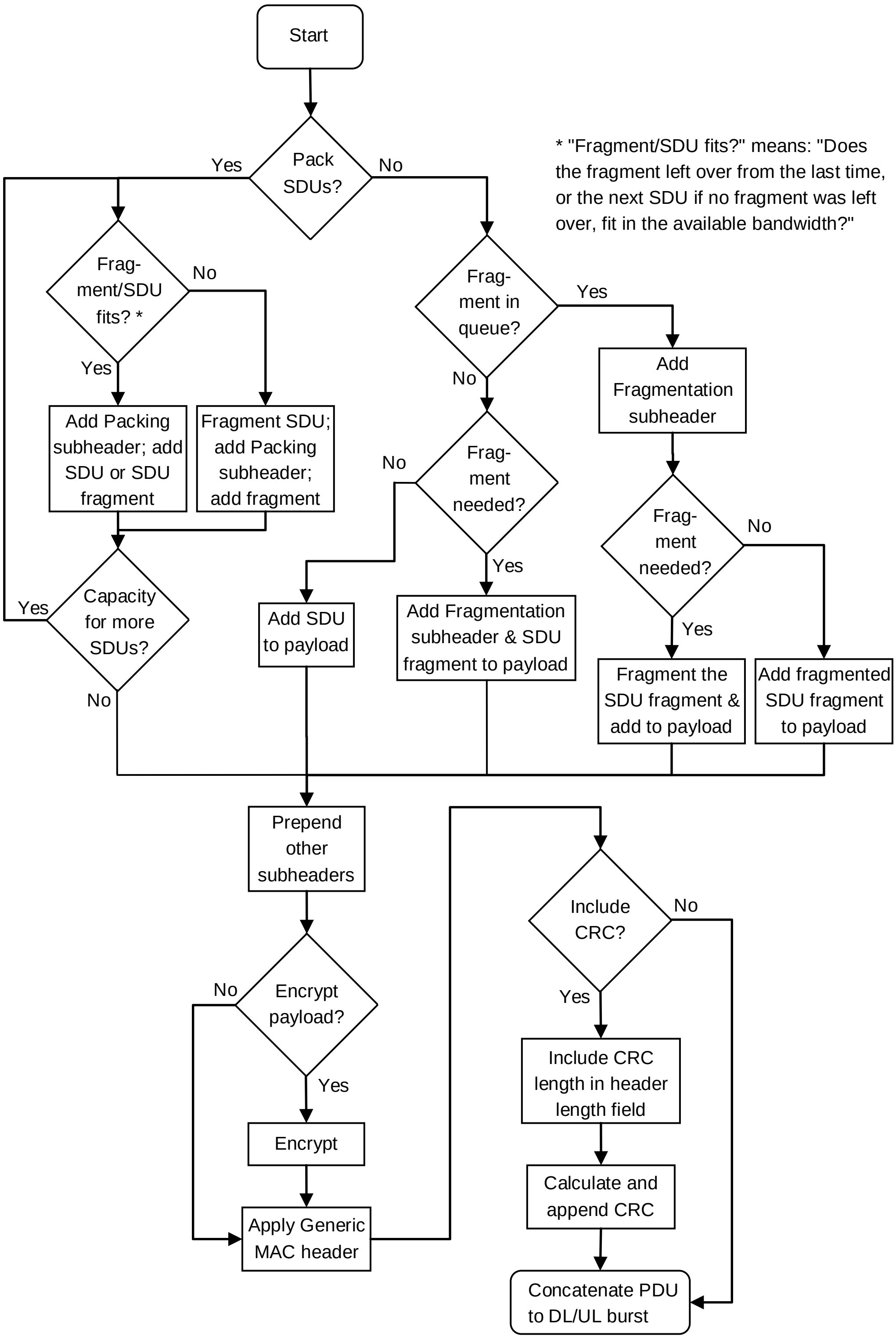

Figure 2.8 - Construction of MAC PDU [3]

Concatenation is a procedure for combining multiple MAC PDUs into a single transmission, in DL or UL. Every PDU has a unique CID that allows the receiving MAC to resolve the MAC SDU from one or more received PDUs and to deliver the SDU to a respective MAC Service Access Point (SAP). Additionally, the PDUs for MAC management messages, user data, and bandwidth requests can also be included in the same transmission. Figure 2.9 illustrates the situation with an example for UL burst transmission.

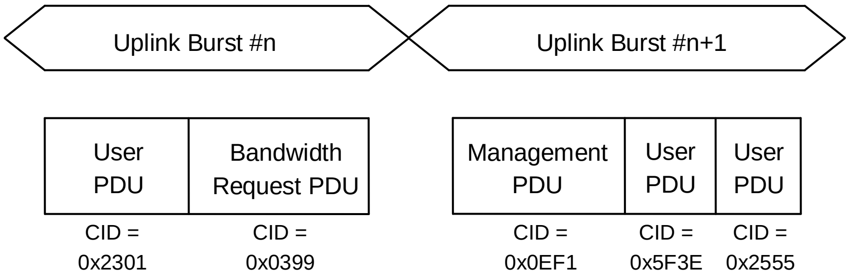

Figure 2.9 - MAC PDU Concatenation Showing Example CIDs [3]

Fragmentation can occur when MAC SDU or management message can not fit into a MAC PDU and have to be divided into several PDUs. The fragmentation and reassembly are needed to better utilize the available bandwidth and they must be supported by the equipment.

Packing allows several SDUs to be packed into a single PDU. The decision is made in the transmitting station and the unpacking is made obligatory for all stations.

CRC calculation can be required by some connections and in these cases the CRC field is added to respective PDUs containing data for those connections. The CRC will include the generic MAC header and the payload part of the PDU. In case encryption is used, the CRC calculation is performed after it.

Encryption is performed to the payload of MAC PDU if the connection is mapped to a Security Association (SA). The encryption and data authentication are done in the transmitting end according to the specifications of the SA and the reverse operation of decryption and data authentication at the receiver is based on that same SA. As stated, only the payload is encrypted while the generic MAC header consists of the unencrypted information required for decryption. If a PDU mapped to an SA requiring encryption is received unencrypted, it is simply rejected.

Padding is intended for filling up the unused parts of an allocated space within a data burst. The space must be in a known state, which can be achieved by setting every unused byte to a stuff byte value (0xFF). In case the size of the unused section is greater than (or exactly) a MAC header length, it can also be formatted as an MAC PDU. [2] [3]

2.2.2 Quality of Service (QoS) Support

Mobile WiMAX is suited for supplying various QoS methods for different types of data services and applications. This is achieved with the sufficient data rates, adjustable capacities in both DL and UL, the fine resource granularity, and flexible mechanism for resource allocation. To provide QoS in Mobile WiMAX, so called *service flows* are designed. These flows are unidirectional packets with certain QoS parameters and Figure 2.10 is demonstrating the principle.

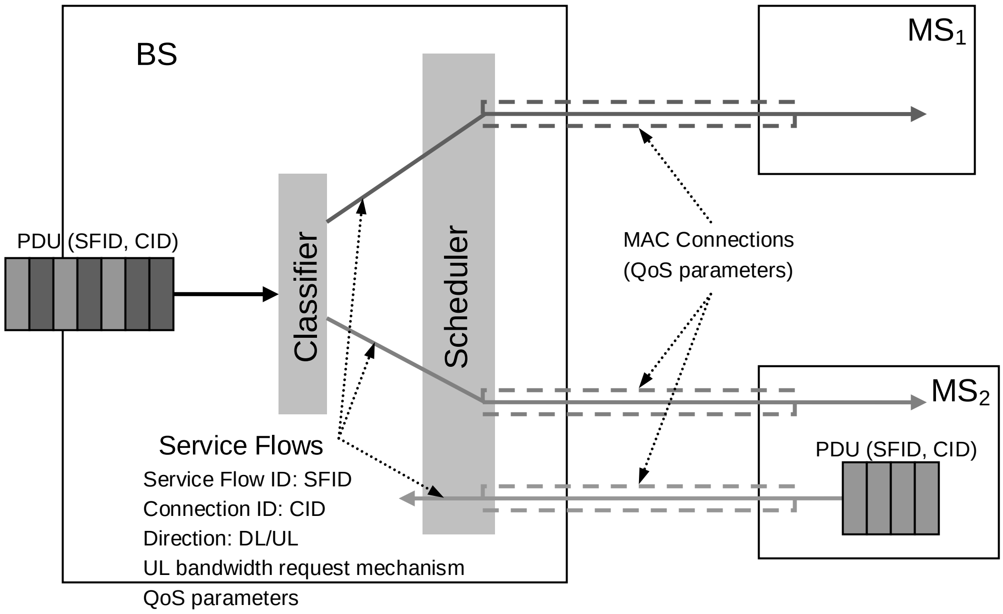

Figure 2.10 - Mobile WiMAX QoS Support [4]

When some type of data service (voice, data, etc.) is wanted to be offered, a connection has to be created between the BS and the MS. This is done by first building a unidirectional logical link between the peer MACs. The packets at the MAC interface are attached with information of a service flow to be delivered over the connection. The service flow has certain QoS parameters that give the scheduler a chance to do decisions for transmission priorities, even during the transmission the parameters can be changed according to the desired service type.

Since the QoS is connection-oriented, it can be effectively controlled during the transmission. Additionally, this enables an end-to-end QoS even over the air interface, which usually is the main problem in wireless communications. The service flow principle is supported in both DL and UL. The Table 2.4 demonstrates used specifications for different data services and applications.

Table 2.4 - Mobile WiMAX Applications and QoS [4]

QoS Category	Application	QoS Specification
UGS Unsolicited Grant Service	VoIP	• Maximum Sustained Rate • Maximum Latency Tolerance • Jitter Tolerance
rtPS Real-Time Polling Service	Streaming Audio or Video	• Minimum Reserved Rate • Maximum Sustained Rate • Maximum Latency Tolerance • Traffic Priority
ertPS Extended Real-Time Polling Service	Voice with Activity Detection (VoIP)	• Minimum Reserved Rate • Maximum Sustained Rate • Maximum Latency Tolerance • Jitter Tolerance • Traffic Priority
nrtPS Non Real-Time Polling Service	File Transfer Protocol (FTP)	• Minimum Reserved Rate • Maximum Sustained Rate • Traffic Priority
BE Best Effort Service	Data Transfer, Web Browsing, etc.	• Maximum Sustained Rate • Traffic Priority

2.2.3 MAC Scheduling Service

A mobile user creates special requirements for the transmission because of the continuous changes in the used transmission medium. To achieve efficient connections with data, audio, or video, an efficient scheduling service is needed. The Mobile WiMAX MAC scheduling service includes Fast Data Scheduler, Scheduling for both DL and UL, Dynamic Resource Allocation, QoS Oriented Scheduling, and Frequency Selective Scheduling. These are introduced in the following sub-subsections.

The scheduling of data for transmission over the connection between the BS and the MS is done at the MAC scheduler. A connection is associated with a single scheduling service, which is defined by a set of QoS specifications as listed in the Table 2.4 earlier. The BS schedules the requests and grants for UL transmissions by providing bandwidth or arranging bandwidth request opportunities for MSs. The scheduling type and related QoS specifications are needed for the BS to be able to predict the needs of the MS regarding throughput and latency of the UL traffic. In the following sub-subsection, the QoS categories are explained in more detail. [2] [3]

Uplink request/grant scheduling

Unsolicited Grant Service
The UGS focuses on UL traffic with real-time requirements with fixed packet length sent on a periodic basis. As mentioned in the Table 2.4, the VoIP (without activity detection) traffic falls into this category. The service provides fixed-size grants periodically and in real-time. The grants are predictable by the MS so the excess overhead and latency can be minimized. The size of the grants is defined based on the maximum sustained traffic rate of the service flow and the size has to be at least large enough to include the fixed-size data from the application with the needed header information.

Real-Time Polling Service
The rtPS is intended for similar use as UGS, but here the packet size may vary. An excellent example of this kind of traffic could be an MPEG compressed video stream. The MS has a possibility to request for a needed bandwidth based on the size of the transmission. Resulting from the variable need of bandwidth, the rtPS requires more overhead than the previous UGS. The BS provides request opportunities sent periodically on unicast messages.

Extended Real-Time Polling Service
The ertPS is a combination of the two previous services, UGS and rtPS. The BS provides unicast grants in an unsolicited manner with the possibility to change the requested grant size. The ertPS design allows traffic for VoIP with activity detection. Hence, when there are silence periods in the communication, the packets require less space and the bandwidth can be saved. The BS provides periodic allocations that can be used either for bandwidth requests or data transfer. Similar to UGS, there is a maximum sustained traffic rate, which defines the default size of the allocations, but, however, with ertPS this value can be changed by request.

Non Real-Time Polling Service
The nrtPS offers polls that are sent on unicast and on a regular basis. With nrtPS the request opportunities are available even on situations with congestion in the network. The intervals between polls are not constant but, anyway, they are kept short enough to provide sufficient service for the MS.

Best Effort Service
The BE service is intended for traffic without strict requirements for latency and/or QoS in general. Typical traffic type could be normal web browsing, which typically consists of data bursts on more or less random intervals. As the name of the service suggests, the aim is to just offer the capacity left from higher service types. [2] [3]

2.2.4 Bandwidth Allocation and Request Mechanisms

In order to transmit anything the MS must request bandwidth for the transmission from the BS. Depending on the QoS category in use, there are several possibilities, which are described in the following sub-subsections.

Requests

A request is a message to the BS which tells that the MS has something to transmit and needs some UL bandwidth for it. The requests can be transmitted through a separate bandwidth request header or by an optional piggyback request. The requested size consists of the needed amount of bytes for the MAC header and payload, but not the PHY overhead. The requests can be transmitted with any UL allocation, except during initial ranging interval. Additionally, the MS is not allowed to request bandwidth for a connection without any PDUs to be transmitted on the respective connection. The request can be either incremental or aggregate, in other words, when an incremental request arrives at the BS, the existing value is adjusted accordingly while with the aggregate alternative the old one is replaced. The optional piggyback requests are always incremental, since they do not include the type field indicating the type of the request.

Grants

The BS offers grants to the MS for bandwidth allocation according the requests sent earlier. The requests are associated in the MS with certain connections by their CIDs, but the grants are associated with the basic CID of the MS. This means that the BS just grants bandwidth for the MS, but the actual use is determined at the MS. The situation might also be that the granted bandwidth is less than requested initially. In such case, the MS must decide the connection(s), if any, that are allowed to use the allocation. The MS can simply settle to the decision of BS and drop the SDU or to request again.

Polling

As told in the previous paragraph, the bandwidth requests are always linked to CIDs and the grants are addressed to certain MSs, which is the case for polling as well. The process of polling occurs when BS allocates bandwidth for MS(s) specially for making bandwidth requests. The allocation can be pointed to single MS (unicast) or to a group of MSs (multi-/broadcast). While polling MS(s), no specific message is sent, but the UL-MAP contains bandwidth allocation enough for a MS(s) to reply with a bandwidth request. In single BS polling the basic CID of the MS is used and with multi-/broadcast polling the UL-MAP includes a special CID dedicated to multi-/broadcast. [2] [3]

Properties of MAC Scheduler

The Mobile WiMAX includes a **Fast Data Scheduler** to arrange the usage of the available capacity in a way that meets the needs of the current application. The packets have been given certain QoS requirements as described in the previous sub-section 2.2.3 and based on these "tags" the BS gives priorities to certain traffic classes. The scheduler supports **Scheduling for Both DL and UL**. The UL scheduling sets constraints to the delivery of correct information about connection quality and the needed QoS to the BS. The **Dynamic Resource Allocation** is also supported for both DL and UL traffic. The resources can be allocated both in time and frequency. The allocation is done by inserting MAP messages in the beginning of each frame. Hence, changes on a frame-by-frame basis are possible, which allows quick reaction for variations in the connection. The scheduling is **QoS Oriented**, which means that the data transport scheduling in MAC is taken care of separately for each connection and these connections are put into specific data service classes with QoS requirements and the scheduling is performed accordingly. With **Frequency Selective Scheduling, t**he MAC scheduler is designed to exploit sub-channels of different kind. When sub-carriers in the sub-channels are pseudo-randomly distributed across the bandwidth (e.g. PUSC permutation), the sub-channels are of similar quality. If contiguous permutation (e.g. AMC) is used, the attenuation between sub-channels can vary. The advantage with FSS is that the weak sub-channels of the mobile users can be avoided and select the better ones for transmission. The FSS creates some overhead in the CQI in the UL direction but on the other hand the total capacity is increased in case of low mobility. [4]

2.2.5 Mobility Management

The main issues for mobile application in general are the inadequate power resources with challenges in handoffs. The Mobile WiMAX has two modes to allow power saving when possible: the Sleep Mode and the Idle Mode. The Mobile WiMAX is also said [4] to support seamless (connection not broken while changing the BS) handoffs, up to the speed of 100 km/h. Mobility issues are dealt in more detail later in chapter 3 and are therefore omitted here.

2.2.6 Security

Mobile and wireless applications demand more security compared to traditional fixed and wired connections. When using the air as transmission interface, it is open for all, which results in the fact that the traffic and the signaling have to be encrypted in order to keep the communication confidential between MS and BS. In addition, building up the connection with a new MS/BS requires authentication in both ends. The Mobile WiMAX supports mutual device/user authentication with a flexible key management protocol. Also support for strong traffic encryption and signaling message protection are included. The fast handovers need enhancements to security protocols as well. These are described more in the following sub-subsections. [4]

Key Management Protocol

The security issues of Mobile WiMAX are defined as a sublayer in MAC layer. The used protocol, and the basis for WiMAX security, is Privacy and Key Management protocol with versions 1 and 2 (PKMv1/2). The (Fixed) WiMAX supports only the first version while the Mobile WiMAX supports both.

Device/User Authentication and Authorization

When a Mobile WiMAX device is turned on it tries to connect to a BS. However, a reliable connection between the user and the BS has to be ensured. This can be performed by mutual or one-sided (the BS authenticates the MS) authentication provided by the PKM. The PKM also allows periodic reauthentication/reauthorization and key refresh. It may use the Extensible Authentication Protocol (EAP) [12] or X.509 digital certificates [13] with RSA (see Abbreviations) public-key encryption algorithm [14]. Yet another solution is to use a sequence starting with RSA authentication and followed by EAP authentication.

The PKM creates an Authorization Key (AK), which has the length of 160 bits and is a shared secret key between the MS and the BS. After the establishment of AK, the Key Encryption Key (KEK) is created from the AK. Again, the KEK is used to encrypt later PKM exchanges of Traffic Encryption Key (TEK).

The RSA based solution is based on a unique X.509 certificate which is issued by the MS manufacturer. The X.509 certificate includes the Public Key (PK) of the MS and its MAC address. When the MS is requesting an AK, it sends the digital certificate to the BS which confirms the certificate and creates the AK based on the received PK.

The EAP based solution uses an exclusive credential provided by the operator. This can be a Subscriber Identity Module (SIM) or an X.509 certificate mentioned above. The operator chooses the most suitable method and the EAP type accordingly. With SIM the used type is EAP- Authentication and Key Agreement (AKA) or with X.509 EAP-TLS (Transport Layer Security). There is also a possibility to use EAP-TTLS (Tunneled TLS) for Microsoft-Challenge Handshake Authentication Protocol (MS-CHAPv2).

After the authentication the MS still needs authorization and a Security Association Identity (SAID) from the BS. Upon authorization request the BS contacts the Authentication, Authorization and Accounting (AAA) server in the network and after successful communication with the AAA, sends back to the MS the AK, a lifetime key and the SAID. The SA is defined as security information set shared between the BS and the MS(s) connected to the BS. The SA can be primary, static, or dynamic. The primary SA is created during the initialization phase of the MS and static SAs are provided within the BS. Additionally, the dynamic SAs are generated in real time based on the creation of service flows. Since there can be several simultaneous service flows for one MS, it is also possible to have several dynamic SAs. However, these must be compatible with the service type(s) the MS is allowed to access. [4] [15]

Traffic Encryption

The Mobile WiMAX MAC layer uses AES-CCM (Advanced Encryption Standard-CTR mode with CBC-MAC, where CTR comes from Counter Mode Encryption and CBC-MAC from Cipher Block Chaining Message Authentication Coder) ciphering for assuring the privacy of user data. The basis for created keys comes from the EAP authentication described above. A KEK is created directly from the AK and is used for creating a TEK. The TEK is generated in the BS and is a random number created with the TEK encryption algorithm with the KEK as the encryption key. The key used for data traffic encryption is this TEK. The keys are also changed periodically.

Control Message Protection

The traffic for controlling the connection needs also to be protected. The solution for Mobile WiMAX control data protection is a scheme with AES based CMAC (Cipher based Message Authentication Code) or MD5 (Message-Digest algorithm 5) based HMAC (keyed-Hash Message Authentication Code).

Fast Handoff Support

The fast handoffs supported by the Mobile WiMAX set requirements also for the privacy of the information exchanged during the change of BSs. The handoff needs to be done in a way that so called man-in-the-middle -attacks are not possible, or at least are made more difficult. Additionally, there must be a balance between latency and reliability, since more reliability equals more latency, and vice versa.

The Mobile WiMAX offers a possibility to use a pre-authentication, which is simply authentication performed before handoff with a target BS in order to make the handoff faster. However, the exact mechanism for pre-authentication is left outside the scope of the 802.16e-2005 standard.

Vulnerabilities

According to [16] the Mobile WiMAX security faces also some vulnerability issues. The first case deals with the initial network entry confidentiality. During the network entry several basic factors such as physical parameters, performance factors, and security contexts between the MS and the BS are agreed. However, this information is not protected. As stated earlier the Mobile WiMAX has control message protection with CMAC/HMAC, but these are used only during the normal operation. Hence, it could be possible for malicious users or networks to interfere the important handshake scheme.

Another presented weakness in Mobile WiMAX security is related to the access network. The authors of [16] want to point out that the 802.16e-2005 (or 802.16-2004) specification does not provide any security measures within the access network, but the specification only assumes the access service network (see sub-section 3.1.1 about network architecture) to be trusted.

2.3 Other Advanced Features of Mobile WiMAX

IEEE has also introduced several technologies to improve the performance of data transmission in Mobile WiMAX. Here are presented some of them.

2.3.1 Smart Antenna Technologies

Traditionally, multiple-antenna-systems have required complex computing on signals and therefore consumed too many resources (battery, processor, etc.). The OFDMA makes the computing much simpler by using vector-flat sub-carriers. It also does not need complex equalizers to compensate the losses on certain frequencies. Furthermore, OFDMA in general is very well suitable for smart antenna technologies. Actually, most of the next generation broadband communication systems are designed to be based on MIMO-OFDM/OFDMA. There are several smart antenna technologies supported in Mobile WiMAX. In the following sub-subsections we give short descriptions of three supported alternatives. [4]

Beamforming

Beamforming is a method where several antennas with weighted signals are used to create stronger beams. The technique improves capacity of transmission and also the distance between MS and BS can be increased. In [17], three types of beamforming realization are presented. The first method is called Switched Beam, which allows switching between several narrow beam antennas or between different beams in an antenna array. The second option is called Dynamically Phased Array, which uses a Direction of Arrival (DoA) algorithm. The DoA gives the user a chance to dynamically adjust the beam. Both of these methods increase the signal strength, but are suffering from scattering and multipath propagation, especially in urban areas. The third method is known as Adaptive Array or Adaptive Beamforming. This method determines the beamforming parameters adaptively based on information from both channel and interference conditions. Hence, the influence of signal reflections in an urban environment can be fought against.

Space-Time Code

In Space-Time Coding (STC), also known as MIMO Matrix A, two identical data streams are sent in the DL providing space and time diversity. With the STC the Signal-to-Noise Ratio (SNR) can be increased and as a consequence the coverage area and/or the capacity of the system can be enhanced. The Mobile WiMAX supports transmit diversity such as Alamouti code in order to supply spatial diversity and decrease fade margin. [4] [17]

Spatial Multiplexing

Spatial Multiplexing (SM), also known as MIMO Matrix B, is similar to STC but here the transmitted streams are not the same, hence they have the data scheduled for transmission divided on each stream. SM increases both peak data rates as well as total throughput of the connection. However, the receiver can increase the throughput only if it has several receiving antennas too. The throughput is increased because the multiple streams transmitted from the MS can be received simultaneously. Each user has only one transmit antenna in the UL direction, but two separate users can transmit collaboratively in the same slot. This appears as if transmitting is done by spatially multiplexing the two streams from two antennas of the same user. This kind of functionality is called UL collaborative SM. [4] [17]

The alternatives in the Table 2.5 below can be adjusted and changed adaptively in order to meet the current connection requirements. Adaptive MIMO Switching (AMS) is used to change between different MIMO modes. This method focuses on improving the spectral efficiency while still keeping the coverage area unchanged. On the other hand, the SM can increase the (peak) throughput, but the coverage area may suffer in conditions with disturbances in the connection.

Table 2.5 - Advanced Antenna Options (Nt/r=number of transmit/receive antennas) [4]

Link	Beamforming	Space time coding	Spatial multiplexing
DL	Nt≥2, Nr≥1	Nt=2, Nr≥1 Matrix A	Nt=2, Nr≥2 Matrix B, vertical encoding
UL	Nt≥1, Nr≥2	N/A	Nt=1, Nr≥2 Two-user collaborative SM

The Table 2.6 below describes the (theoretical) data rates for different antenna configurations using Single/Multiple Input Multiple Output (SIMO/MIMO).

Table 2.6 - Data Rates for SIMO/MIMO Configurations (10 MHz channel, 5 ms frame, PUSC sub-channel, 44 data OFDM symbols) [4]

DL/UL Ratio			1:0	3:1	2:1	3:2	1:1	0:1
User peak rate (Mbit/s)	SIMO (1X2)	DL	31.68	23.04	20.16	18.72	15.84	0
		UL	0	4.03	5.04	6.05	7.06	14.11
	MIMO (2X2)	DL	63.36	46.08	40.32	37.44	31.68	0
		UL	0	4.03	5.04	6.05	7.06	14.11
Sector peak rate (Mbit/s)	SIMO (1X2)	DL	31.68	23.04	20.16	18.72	15.84	0
		UL	0	4.03	5.04	6.05	7.06	14.11
	MIMO (2X2)	DL	63.36	46.08	40.32	37.44	31.68	0
		UL	0	8.06	10.08	12.10	14.12	28.22

2.3.2 Fractional Frequency Reuse

To take advantage of the most of channel bandwidth the Mobile WiMAX supports fractional frequency reuse. The used reuse pattern is "one", or 1x1. Hence, all BSs use the same frequency channel. This method introduces significant Co-Channel Interference (CCI), especially on cell borders, and the users suffer from poor connection quality. A solution lies in the properties of OFDMA and in the fact that Mobile WiMAX gives users sub-channels that are only a small part of the whole bandwidth. The CCI can be mitigated by just arranging the used sub-channels properly. Therefore there is no need to design complicated frequency tables for neighboring BSs.

The sub-channel arrangement is made possible by segmentation and a permutation zone. In the segmentation some, or even all, available sub-channels are in use for deploying single instance of MAC. The permutation zone is described as a number of contiguous OFDMA symbols in DL or UL that use the same permutation. Additionally, there may be more than one permutation zone in the DL/UL sub-frame. The Figure 2.11 below shows the frame structure with multiple zones.

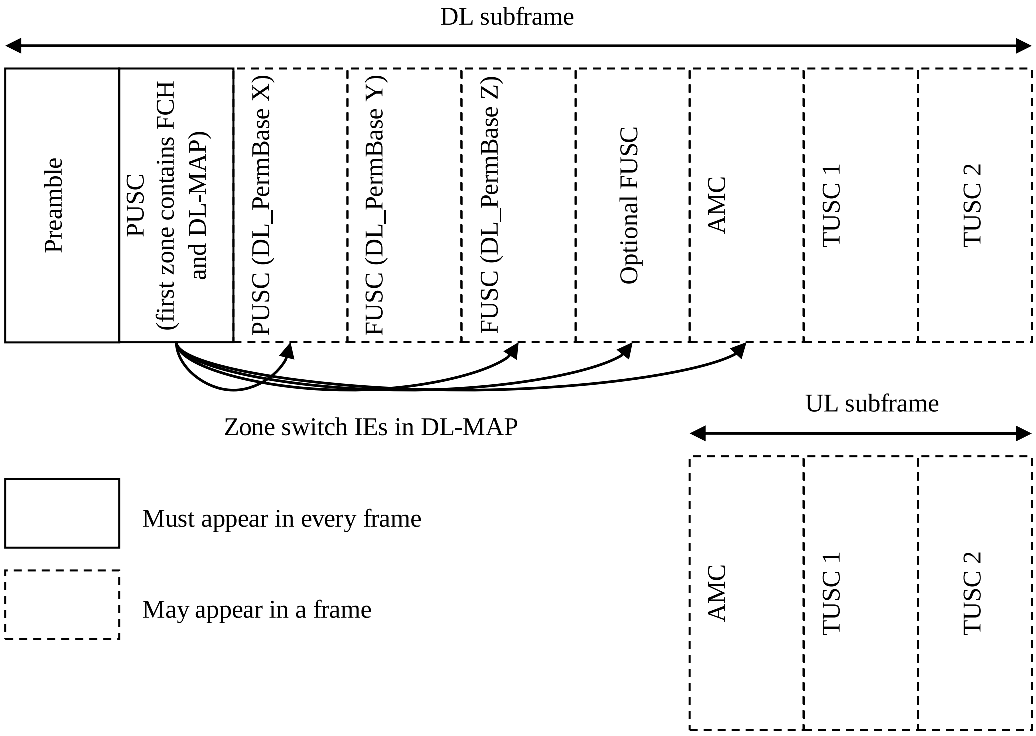

Figure 2.11 - Multi-Zone Frame Structure [4]

The reuse of sub-channels can be arranged so that when the MS is close to the BS, it can use all the sub-channels available and when approaching the borders of a cell, only a part of possible sub-channels are used. The partly used sub-channels must be different than the ones used in the neighboring (and interfering) BS. Figure 2.12 illustrates the principle of fractional frequency reuse, hence near the BS all sub-channel groups F1, F2, and F3 can be used simultaneously while they are used separately at the cell edges. Although this method gives users at cell edge better chances to maintain the connection with BS, the accessible bandwidth is limited and therefore the data rates are affected.

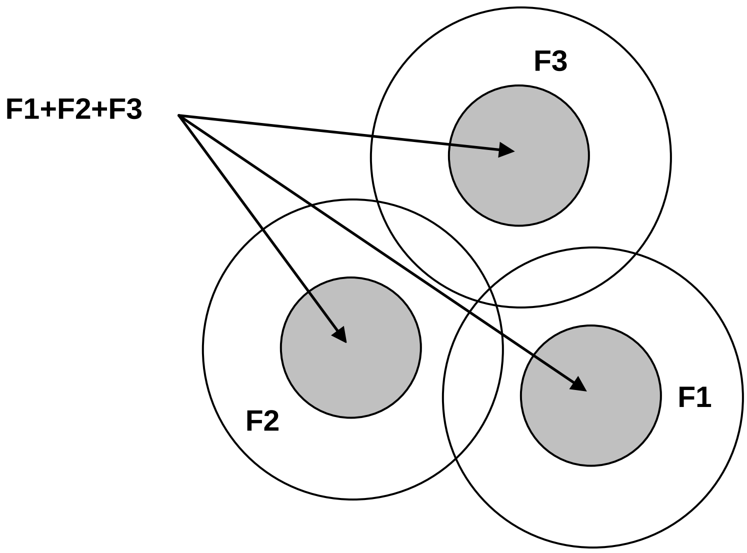

Figure 2.12 - Fractional Frequency Reuse [4]

2.3.3 Multicast and Broadcast Service

There can be service types that require the data to be transmitted to several MSs simultaneously either by broadcast or multicast. These service flows have defined QoS requirements with possibly encryption of the messages. The situation can be that there are some MSs registered to the BS and some MSs are not registered to the BS since saving power in Idle Mode (see sub-section 3.3.2). However, the multi-/broadcast traffic needs to reach these as well.

The Mobile WiMAX supports Multicast and Broadcast Service (MBS) and offers two types of access to be used: the single-BS and multi-BS access. With single-BS access the multi-/broadcast connections are established within one BS, while the multi-BS access type is using multiple BSs to transmit data from service flows. The MS can support both access types, but HARQ can not be combined with either of them. The MS has to be registered in order to initialize the MBS and to create a multicast connection for MBS data, with a certain service flow ID regarding the chosen service. In multi-BS mode, the BSs capable for MBS are attached to a MBS zone that is defined as a set of BSs where the same CID and SA is in use for certain service flows. [2]

Single-BS access

The single-BS access connections are created between the BS and each MS separately. However, the CID used for the multicast connection is the same for all MSs on the same channel. The data for the connection is received and processed at each MSs MAC, but the BS needs to send the multicast MAC SDU only once per BS channel. In a case that the DL multicast traffic needs to be encrypted, the MSs will have an extra SA, which allows the encryption with certain keys, that are not related to the encryption keys used in normal MS-BS transmissions. [2]

Multi-BS Access

With multi-BS MBS the MSs registered to a certain MBS connection can receive (encrypted) MAC PDUs from multiple BSs at a certain time period. The BSs transmit multi-/broadcast traffic and the MSs can receive the traffic regardless of their location within the MBS zone. The BSs participating in the same multi-BS MBS need to be synchronized since they are transmitting common multi-/broadcast data. The multi-BS MBS connection uses a shared CID for both MSs and BSs on the same channel involved in the connection. The synchronization of transmissions allows MSs to receive the same data from several BSs, which can significantly increase reliability of reception, especially at cell borders.

Unlike in previous single-BS version, the multi-BS MBS does not require the MS to be registered to any BS, but the association to the MBS zone is enough for receiving the CID and SA to be able to receive the multi-/broadcast traffic. The usage of multi-BS MBS mode is allowed with Single Frequency Network (SFN) operation, based on the synchronization requirement. Additionally, the duration of MBS zones is adjustable and therefore the bandwidth reservation for MBS traffic can be controlled according the needs at a specific moment. Multiple MBS zones can also be used. Power savings are also possible because the MBS can be accessed when MS is in Idle Mode (discussed later in 3.3.2). [2] [4]

2.4 Performance

The actual performance of Mobile WiMAX is strongly dependent on the selections made by the network operator regarding for example the used equipment or the level of provided service. It is not meaningful to describe all the possible solutions and the capabilities of them. In this section, some theoretical performance parameters announced are discussed.

Even the theoretical capacity is depended on possible usage of various features provided by the Mobile WiMAX standard. Some data rates were presented already, while discussing the optional MIMO technology, where the maximum rates were several and few tens of megabits per second in DL and UL, respectively. The figures, however, describe a single channel that is often shared by several users. The user data rates will remain in a few megabits per second.

The MS distance from the BS in Mobile WiMAX is usually few kilometers on urban and suburban scenarios, but can be extended to even a few tens of kilometers for rural users. It has to be remembered that the extension of distance forces lower modulation and coding to be used and therefore the data rates are affected in a negative way.

The WiMAX Forum [1] has published documents that discuss the performance of the Mobile WiMAX. In [18], a comparison between the Mobile WiMAX and 3G technologies is presented while the other document [17] portrays arguments for different deployment scenarios of Mobile WiMAX in the access network.

2.5 Other Candidates/Competitors

The following sub-sections present some candidates that are designed to fulfill similar needs than the Mobile WiMAX. Some technologies are not direct competitors but more like involved in the subject. The technologies are only introduced very superficially since the weight in the thesis is on Mobile WiMAX. At the end of each sub-section, the main differences or benefits/drawbacks compared to Mobile WiMAX are shortly described.

2.5.1 3G/3.XG/4G (Post-GSM Generations)

Universal Mobile Telecommunications System (UMTS)

The UMTS is the European standard for the follower of GSM. The purpose has been to deliver also data services to mobile users, but it still carries the burden from 1G and 2G, systems designed for mobile *voice* transmission. This results in not so excellent data capabilities. The UMTS reaches the data rate of only a few hundreds of kilobits in a second. However, it has a higher range in terms of supported mobile speeds.

High-Speed Packet Access (HSPA)

The HSPA is designed to extend and improve the performance of the UMTS. It has been divided into two standards for DL and UL, HSDPA and HSUPA. The HSDPA is currently entering the markets [19], but on the other hand HSUPA is still under development and the world is waiting for commercial availability of products supporting the standard [20]. The future of HSPA lies in High-Speed OFDM Packet Access (HSOPA), which could be a competitive rival to the Mobile WiMAX.

The HSDPA is based on a Wideband Code Division Multiple Access (WCDMA) and the main task is to improve the existing UMTS performance. The UMTS offers a maximum data rate of 384 kbit/s while HSDPA is promised to reach a maximum of 14.4 Mbit/s. However the maximum rate is not yet available for commercial use and the common rate for HSDPA devices is below 2 Mbit/s.

The HSDPA mobility is handled with Fast Cell Selection instead of soft handoff. The Fast Cell Selection is a similar technique to the Fast Base Station Switching (FBSS) in Mobile WiMAX, which is discussed later in sub-section 3.2.1. Hence, the terminal selects the most suitable cell for downlink from an active set of nearby BSs (or Node-Bs, as they are called in HSPA environment) using uplink signaling. [21]

The HSUPA is the pair for HSDPA, dedicated to the traffic in UL direction. It is promised to reach data rate of 5.8 Mbit/s. Together HSDPA and HSUPA are simply called HSPA.

Long Term Evolution (LTE)

The LTE, or also known as the earlier mentioned HSOPA, is the name of the follower of HSPA. It is projected to reach data rates of 100 Mbit/s and 50 Mbit/s for DL and UL direction, respectively. This is assumed with a 20 MHz spectrum allocation.

The LTE mobility is mainly intended for pedestrian speeds (0-15 km/h), however higher vehicular speeds (15-120 km/h) should also be supported with high performance. In higher speeds up to 350 km/h, or even 500 km/h depending on the frequency band, the connection should be kept alive. The planned coverage of 5 km should meet all the expectations for mobility stated above and additionally with slight degradation within 30 km cells. [22]

Comparison to Mobile WiMAX

The recent technologies based on the GSM technology have a strong foothold on the market and the network infrastructure already exists. Obviously, it is easier for operators to gradually update the current equipment than starting from scratch. With GSM this is also a drawback, since it was initially not designed for data traffic, but merely to provide voice services. The LTE, which would be the most even competitor for the Mobile WiMAX, is few years behind in standardization and the predecessor HSPA can not compete in terms of data transmit capabilities. On the other hand, the Mobile WiMAX falls behind in performance when discussing the highest supported speeds of the mobile station.

2.5.2 FuTURE Project

Chinese researchers have developed an own project heading beyond 3G. Their system is called the FuTURE project (Future Technologies for a Universal Radio Environment) and according to news spread worldwide at the end of January 2007, the first test cases are built in Shanghai. Exact facts about the technology are not available, but something can be evaluated.

The frequency range is between 3-5 GHz in 20-30 MHz bands. The peak data rate in high mobility cases should be between 40-100 Mbit/s and for lower mobility the rate should exceed 100 Mbit/s. QoS should meet basically the same requirements as in wired systems. The transmission system will be based on OFDM or Generalized Multi-Carrier (GMC) techniques, additionally with TDMA/FDMA/CDMA functionalities to allow even more flexibility for supporting dynamic transmission rate. The wireless technology in FuTURE takes advantage of MIMO, link adaptivity (for example AMC), environmental adaptation, and turbo iterative receivers. In addition, new channel coding/decoding techniques have been developed.

The project has also demonstrated a concept of generalized cell structure for the network. It is based on several antenna sensors distributed across the cell and these sensors are connected to the BS with Radio over Fiber (RoF). The RoF cells form a generalized cell, as shown in Figure 2.13. [23]

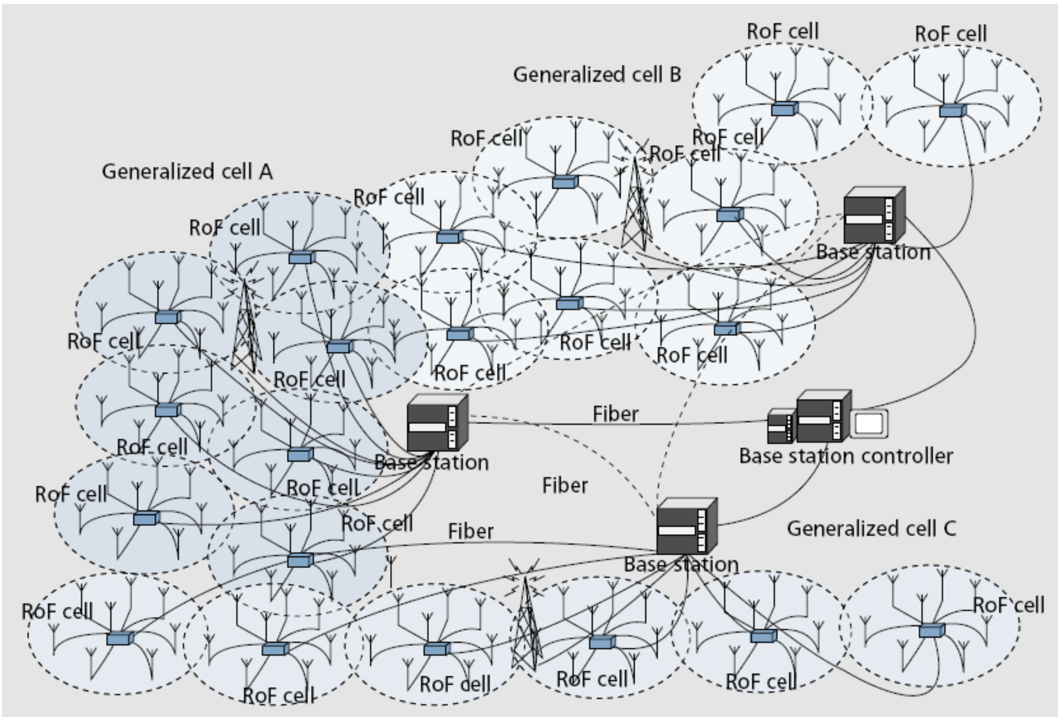

Figure 2.13 - Structure of the RoF Network [23]

Comparison to Mobile WiMAX

The information available about FuTURE does not really allow very thorough comparison, but based on the issues above, the FuTURE would have slightly better data rate performance than the Mobile WiMAX. However, the Mobile WiMAX is already standardized while the FuTURE is still very much in progress. It can also be assumed that the project is more intended on domestic markets of China rather than to become a wireless standard used worldwide.

2.5.3 Mobile Wireless Broadband Access (802.20)

The IEEE 802.20 workgroup is defining a standard, known as Mobile Broadband Wireless Access (MBWA), which is rather similar to the Mobile WiMAX. The greatest difference is the goal to achieve mobility at high vehicular speeds, even up to 250 km/h. The corresponding maximum speed, according to Mobile WiMAX developers, was about 100 km/h.

There has been strong discussion about the sensibility of developing two very similar wireless communication technologies. The 802.20 working group was even suspended during the summer of 2006, though the reason was suspicions of too heavy influence of certain parties (instances behind Flash-OFDM technology [24] tried to force it to be the standardized 802.20) in the workgroup. The result was a suspension for three months and changes in the standardizing organization. [25]

The MBWA has the advantage of being designed from the scratch instead of the burden from previous WiMAX versions which the Mobile WiMAX suffers from. It has been designed to be fully mobile with high speeds and providing fast enough connection.

The PHY layer of 802.20 gets its basis from the WiMAX standards, but since the goal is set to create a standard for a very mobile user, different techniques are needed to fulfill these expectations.

The standardization of MBWA is still in progress and therefore final decisions are not yet made. The 802.20 is designed to function on licensed bands below 3.5 GHz. The channel bandwidth can vary from 1.25 to 40 MHz with both TDD and FDD duplexing. The targeted data rates are megabits in DL and hundreds of kilobits in UL.

The data transmission is planned to be handled with OFDMA in very similar way that is in use in Mobile WiMAX. This would ease and speed up the development since the technology is already there. Nevertheless, some suggestions are made that the DL traffic would use the OFDMA, but the UL direction would be CDMA-based.

Modulation and coding methods are practically identical to the ones used in 802.16. Modulation possibilities vary from BPSK to 64QAM with convolutional and turbo coding. As in Mobile WiMAX, there is also full scalability according the channel conditions. Hence, high-speed mobility can not take advantage of the higher transmission rates with for example the 64QAM-modulation.

Advanced transmission mechanisms designed to be used with MBWA include basically the same as with the family of 802.16 standards. These are for example Advanced Antenna Systems (AAS), Space-Time Block Code (STBC), spatial multiplexing and MIMO. There have also been suggestions that Space-Division Multiple Access (SDMA) could be used.

The MBWA MAC layer functionalities take also advantage of the work already done while designing the 802.16(e). Since the Mobile WiMAX uses also OFDMA, it can be figured that the connection mechanisms are very similar with MBWA. The difference may arise if, as suggested, CDMA is used for the UL traffic. This requires a separate reverse-link access channel for users to be allowed to request transmission slots. The hybrid OFDMA/CDMA DL/UL scheme would also make the transmission and framing more complicated since the UL signaling would have to be sent with CDMA and the feedback from the BS is received in OFDMA. However, the frame format will most likely stay very similar to the one used in 802.16. [26]

Comparison to Mobile WiMAX

As already stated, the greatest differences between the MBWA and the Mobile WiMAX are the support for higher mobility and the drive for it from the beginning of standardization. However, there is a price to pay for high mobility and the presumed data rates are going to be much less than the pronounced values for Mobile WiMAX.

2.5.4 Media Independent Handover (802.21)

The 802.21 is IEEE's rather new workgroup heading for co-operation of different wired and wireless networks. The standard could include 802.3 (basic Ethernet), 802.11, 802.16(e), and different cellular xG networks with other 802-based technologies, making the handoffs between them possible. The main focus is concentrated on vertical handoffs but horizontal ones are also possible. The vertical handoff is a handoff between different networks while the horizontal handoff occurs within a single network.

Usually there exist multiple different types of networks in the reach of a MS. The devices designed for multiple networks, like the latest mobile phones, that operate with GSM (GPRS, WCDMA, etc.) and WLAN, additionally even with Bluetooth, could change the currently serving network automatically with the 802.21. The situation so far has been that the user has to do a manual change of connection type. For ongoing video call, for example, it would be advantageous to choose the network with best performance or on the other hand, the greater expenses of the data transmission in a GSM-based network could be avoided if a vacant 802.11 hotspot is available. Another critical issue with mobile devices is the power consumption, hence switching to a technology requiring less power would result in a longer battery life.

IEEE has defined the scope of the 802.21 to be in the initiation and preparation of handoff rather than the execution of it. The initiation includes discovering and selecting the available networks (with compatible technology) in the reach of the MS. The handoff preparation handles the layer 2 and IP connectivity. The actual transfer of the connection (handoff signaling, context transfer, and packet reception) is left outside the scope of 802.21. [27]

Comparison to Mobile WiMAX

Since the MIH is not a data transport technology it can not be compared with Mobile WiMAX. If it is successfully standardized, it becomes a significant companion for Mobile WiMAX products equipped with other transmission technologies as well.

2.5.5 Wireless Regional Area Network (802.22)

IEEE has also the workgroup 802.22 with the technology called Wireless Regional Area Network (WRAN). The key idea behind WRAN is the usage of cognitive radios on unused, licensed bands such as the TV broadcast frequencies. The markets of the WRAN are designed to be the rural and remote areas that do not possess the access to traditional broadband technologies (xDSL, cable, etc.). The available frequencies for TV broadcasting are especially suitable for this kind of market groups because of the very favorable propagation characteristics. The service requirements are the same as for the Mobile WiMAX. Hence, WRAN has to support data, voice, and even audio/video traffic. [28]

Comparison to Mobile WiMAX

Although the Mobile WiMAX supports also stationary use, the WRAN can rather be considered to be a competitor to the 802.16-2004. WRAN has a longer serving range from the BS, resulting in a greater coverage area and less BSs. Another significant difference is already mentioned above when the cognitive radios are mentioned. The initial requirement of WRAN, when applying the use of TV broadcasting frequencies, was the condition that the unlicensed data traffic would not interfere with the existing TV broadcasts in the area. This kind of functionality requires cognitive radios that detect the unavailable parts of the frequency band. The issue is also considered in the 802.16h task group, but the detection of competing frequencies is only considering ones with WiMAX functionalities.

2.5.6 Wireless Local Area Network (802.11x)

Currently the WLAN, usually denoting the IEEE 802.11-family [29], is the solution to be used when creating a wireless network inside companies, homes, or other public/private buildings. The 802.11g is the market version today, while the aging 802.11b is retiring, mostly because of lacks in the connection data rates. The 802.11a is still in use on some locations where the backhaul connection to the access point (the name for a BS within 802.11x-standards) can not be created with wired connection and the clients subsequently use

802.11b or 802.11g for communication with the access point. The 802.11n task group is currently standardizing the latest version and the first products are appearing on the markets.

The g-amendment gives the maximum data rate of 54 Mbit/s and rather good security with Wi-Fi Protected Access (WPA). The older WEP can also be used, but since it is rather easy to break, the WPA is preferred. The WPA, or the more recent WPA2, security protocols have been developed in a working group of their own, the 802.11i. The emerging 802.11n uses MIMO increasing the peak rate by a decade compared to the 802.11g. However, as with the 802.11g, the typical rate remains much lower; in 802.11n in the region of 200 Mbit/s.

The 802.11-technologies are well suited for communications with short distance to the access point. The coverage indoors is only a couple of tens of meters and in perfect conditions outdoors the signal can travel a few hundred meters. The 802.11 alone is not competing the same markets with Mobile WiMAX, but it can be more or less considered as a good companion to a device working with Mobile WiMAX.

There are also other amendments in the 802.11-family that are focused on providing a certain service with some of the transmission amendments. The 802.11u working group is based on similar thoughts as the earlier mentioned 802.21, hence interoperability of 802-technologies with other networks such as cellular. [30]

The 802.11p, or Wireless Access for the Vehicular Environment (WAVE), is also a standard under way. The purpose with it is to give vehicles a way to communicate with the roadside or other vehicles. The 802.11p is based on functioning on a licensed band around 5.9 GHz, offering an average coverage of a few hundred meters and a data rate of 6 Mbit/s. As the name suggests, this protocol is intended to be used in vehicles and the proposed usage scenarios could be in toll collection, vehicle safety services, and/or commerce transaction via cars. [31]

There have also been attempts to increase the mobile use of 802.11 technologies. The 802.11r, or also know as fast roaming or fast Basic Service Set (BSS) transition, is focused on providing handoffs for a moving MS.

However, as stated earlier, the 802.11 is not a direct competitor to the Mobile WiMAX and therefore not described in more detail in this work.

Comparison to Mobile WiMAX

The WLAN devices can already today support data rates equal and more than the ones announced to be reached with the Mobile WiMAX, but the lacks in allowed distance from the access point and the handoff capabilities make it rather to be a good companion than a direct competitor. The examples already exist, where WLAN and UMTS/HSPA is combined in a cell phone. This is a very likely scenario with WLAN and Mobile WiMAX as well.

2.6 Summary

The second chapter introduced the fundamental properties of PHY and MAC layers in Mobile WiMAX. The S-OFDMA was described as the basis for Mobile WiMAX PHY with the capability to adjust the channel bandwidth by changing the FFT size.

The Mobile WiMAX frame structure was also explained with overview to some other PHY layer features. The MAC layer description included the addressing and connection properties of MAC, the PDU structure and construction, introduction to service flows with certain QoS class, the MAC scheduling service, and the security methods used with Mobile WiMAX. The MAC layer contained also the presented supports for advanced antenna systems and fractional frequency reuse. As a new feature to WiMAX, the MBS was also introduced.

The performance figures are usually used to compare a certain technology with other similar ones. The section describing performance issues gave some sight to the scales of data rates or MS-BS distances. Finally, some competitors of Mobile WiMAX were shortly introduced, though especially the WLAN can be considered more as a companion in a common device.

In addition, the features supporting mobility would have been in right context in this chapter. However, in order to emphasize their meaning in Mobile WiMAX, they have been dedicated a chapter of their own, which is discussed next.

3 MOBILITY

The mobile use of devices is a rapidly increasing trend in communications today, mostly due to new technologies providing sufficient data rates able to compete with traditional wired connections used at homes. The users want to have access to the same services as they are used to no matter where they happen to be. This creates a fact that the emails, instant messaging, multimedia streaming, and web browsing are coming to mobile phones, PDAs, and laptops. For some users the mobile wireless network is the only choice available, for example in areas without communications network architecture. If there is only need for a network for a short period of time on a certain location, the network can be set up quickly with movable BSs. This could be a usage scenario for a military or United Nations peace keeping group.

However, the development of networks supporting mobility sets several requirements for a device planned to be used "on the go". The first, and maybe the most important, requirement is the ability of a device to change the serving BS according to the movements of the user. Additionally, this has to be performed without disturbances in the connection and maintaining the confidentiality between the MS and both, old and new, BSs. A very likely usage scenario could be a user in a public transportation vehicle or in a normal car which means that the handoffs and communicating in general need to be supported even in a vehicle moving with a rather high speed.

The second obstacle a mobile device has to face is the limited power resources. The batteries can carry only a certain amount of charge and they have to be recharged on a regular basis. The capacity and the physical dimensions of the battery create a challenge to manufacturers trying to balance between the two.

As described earlier, the mobility of a user has been the major problem with traditional WiMAX functionality, since it supported only nomadic access. By nomadic access the user was able to change the location of the subscriber station, but without the support for handoffs. The new 802.16e is told to solve these problems and truly make WiMAX go mobile. In the following sections we go through the building blocks of mobility in WiMAX by starting with description of the network architecture behind Mobile WiMAX.

3.1 Network Architecture

Previously there has been only discussion about the connection between the MS and the BS. This section introduces the architecture behind the BS (from the MS point of view). The key property of Mobile WiMAX network is the all-IP (both IPv4 and IPv6) platform which leaves out the traditional circuit switched alternatives. This allows financial savings since there is no need to maintain both types of core networks.

The 802.16e-2005 standard defines only the air interface while the implementation of the network connecting the BSs and providing the access to the Internet is left to service providers. However, the WiMAX Forum has established a Network Working Group (NWG) that defines a Network Reference Model (NRM) [32], shown in Figure 3.2, to be used when constructing a Mobile WiMAX network. Before the model is investigated in more detail, the main elements of a Mobile WiMAX network are introduced.

There are several solutions existing but Figure 3.1 below describes the basic structure of the Mobile WiMAX network. The MSs are connected via air to BSs. The BSs connect via routers to an access gateway which again is connected to a Connectivity Service Network (CSN). The CSN has the functionalities of a home agent and authentication with the access to the Internet.

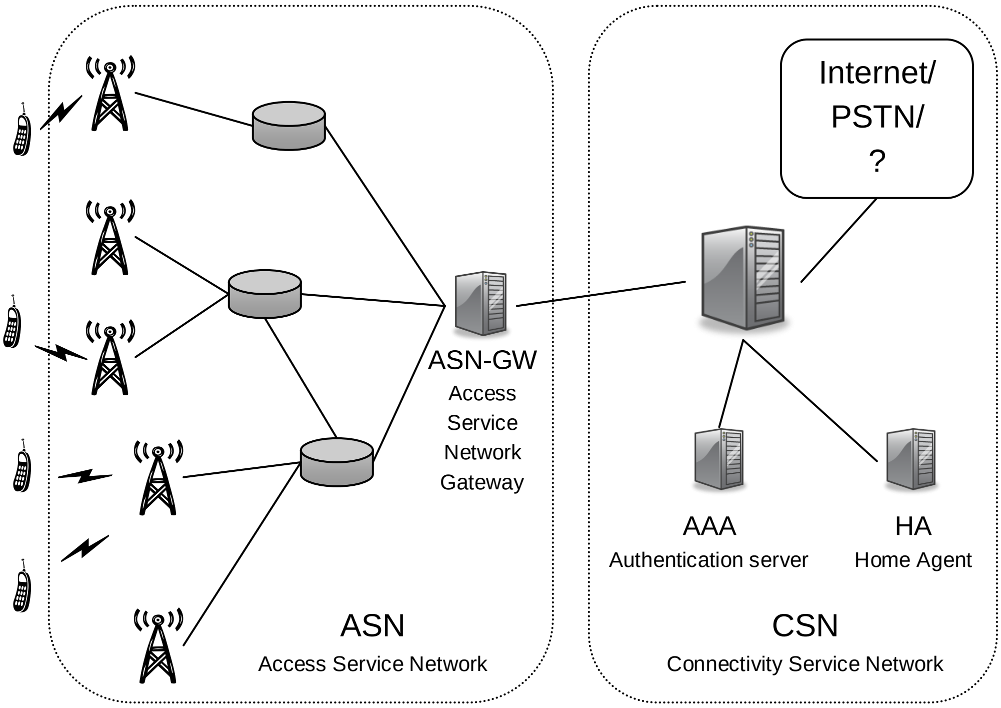

Figure 3.1 - Mobile WiMAX Network Structure

Mobile WiMAX network manufacturers use different names for their equipment, although their tasks are the same. For instance, in some sources Network Access Provider (NAP) equals the Access Service Network (ASN) or Access Service Network Gateway (ASN-GW) is known as Access Concentrator (AC).

The connections from the BSs are usually wired and optical cables, but in some cases other solutions can be used too. These decisions are dependent on local infrastructure since building a new optical connection to a remote BS site would be financially unprofitable. In such cases a fixed radio link can be a sensible solution.

3.1.1 Access Service Network

The ASN is defined in [32] to provide all the needed network components to offer radio access to a WiMAX subscriber. It has the task of providing the WiMAX-L2 connectivity to a WiMAX-MS. It also transfers the authentication, authorization, and user accounting information to the home network service provider. Network discovery and selection of a preferred Network Service Provider (NSP) to the user are taken care of in the ASN. Additionally, the ASN has to deal with relay functionality to enable L3 connectivity to a WiMAX MS, hence providing an IP address. The effective usage of radio resources are also handled with the ASN.

The requirements and tasks above are intended for WiMAX in general, but the mobile applications have yet other demands. Hence, ASN/CSN anchored mobility must be supported. The ASN anchored mobility is described as the functionalities needed to allow an MS to change a BS, hence handoffs, as long as the foreign agent is not changed. On the other hand, the CSN anchored mobility allows the change of anchor point within the ASN, but the CSN anchor point is the same. The Mobile WiMAX ASN must also support paging and tunneling between the ASN and the CSN.

3.1.2 Connectivity Service Network

The general task of the CSN is to provide IP connectivity services to the WiMAX subscriber(s). In [32] the CSN is given some other functions, which are not defined as required properties, but however, they usually are present.

The CSN can allocate the MS IP address and other endpoint parameters. As mentioned earlier, the CSN usually is the one providing the internet access for the users and the AAA proxy or server may be located in the CSN too. It should also implement the policy and admission control (allow/deny certain services) according to the subscription profiles of the users. The support for tunneling between the ASN and the CSN with the billing and inter-operator settlement may be available. Inter-CSN tunneling will allow roaming in the networks of other service providers, while inter-ASN mobility allows MS handoffs between different ASN, but within the same CSN. The CSN provides also WiMAX services such as different IP multimedia services or navigation/location services.

3.1.3 Network Reference Model

The NRM defines operational entities and reference points over witch the interoperability is achieved between network elements (Figure 3.2). The previously described elements, MS, ASN, and CSN create the basis for the NRM. These are connected by reference points R1-R5. The network entities mentioned above form a group of functional entities that can be realized in a single physical functional entity or may be divided to several different physical entities. Hence, the functionalities of the ASN can be constructed within one device, or divided to several. The choice belongs to the operator, but the NWG has defined profiles A, B, and C, in order to assist in the network implementation. These profiles are discussed later in this sub-section. [32]

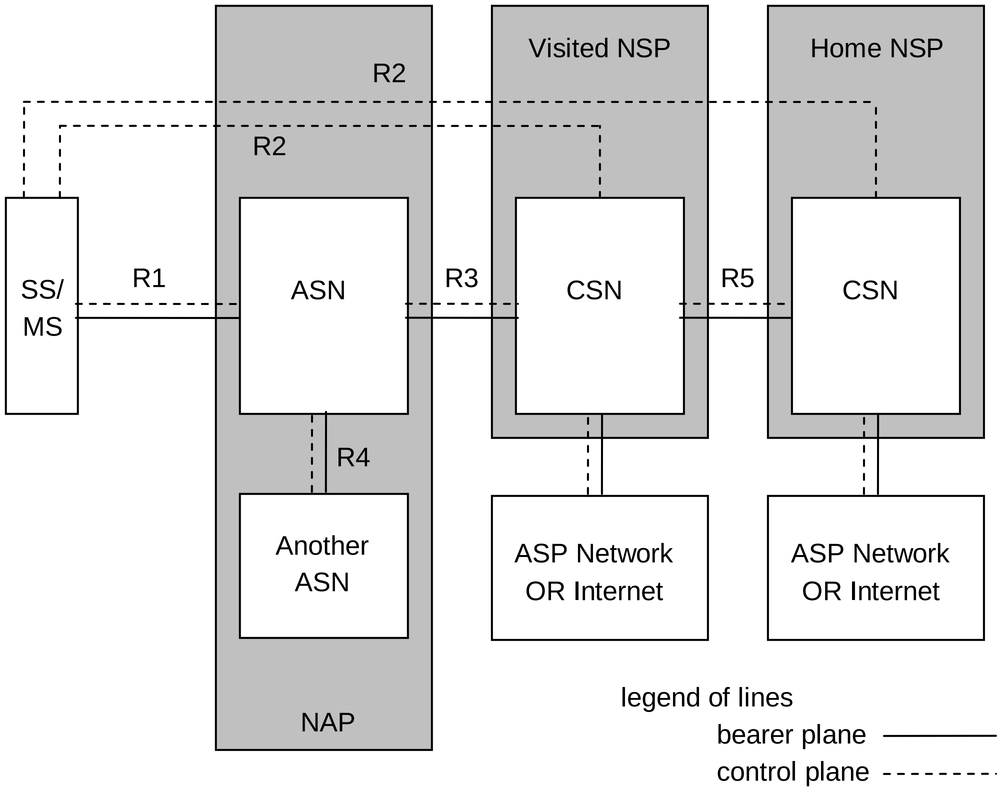

Figure 3.2 - Network Reference Model [32]

Reference points

The reference points R1-R5 in Figure 3.2 are defined as follows. The **R1** is indicating the protocols and procedures used between the MS and the ASN. The **R2** is intended for protocols and procedures between the MS and the CSN, which is associated with authentication, authorization of services, and management of IP host configuration. Since the actual physical connection chain to CSN goes through ASN, the R2 is considered as logical link. Additionally, the authentication in R2 is operated with the home NSP CSN, while the other can also be operated in the visited NSP CSN. The **R3** includes the control plane protocols between the ASN and the CSN. It supports AAA, policy enforcement, and mobility management capabilities. The **R4** consists of protocols used within the ASN. It can be used to allow MS mobility management between different ASNs and ASN-GWs. The **R5** includes the protocols needed for communication between the visited and home CSN.

ASN Reference Model

The Figure 3.3 describes the reference model for ASN. As mentioned in the previous sub-subsection the ASN shares the R1 reference point with the MS, the R3 with the CSN, and R4 with other ASNs. The ASN must have at least one BS and one ASN-GW. The BS(s) and ASN-GW(s) are connected with logical links referenced as **R6**. The reference point **R8** is intended for communication between BSs to assure fast and seamless handoffs.

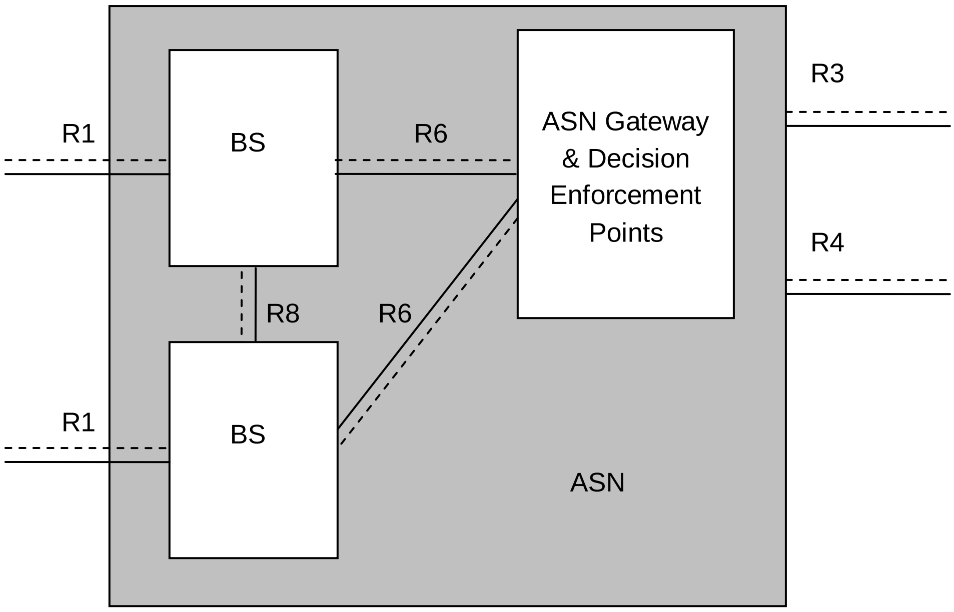

Figure 3.3 - ASN Reference Model Containing a Single ASN-GW [32]

The ASN-GW can optionally be decomposed to two groups of functions, the Decision Point (DP) and Enforcement Point (EP), which are connected via reference point **R7**, as shown in Figure 3.4.

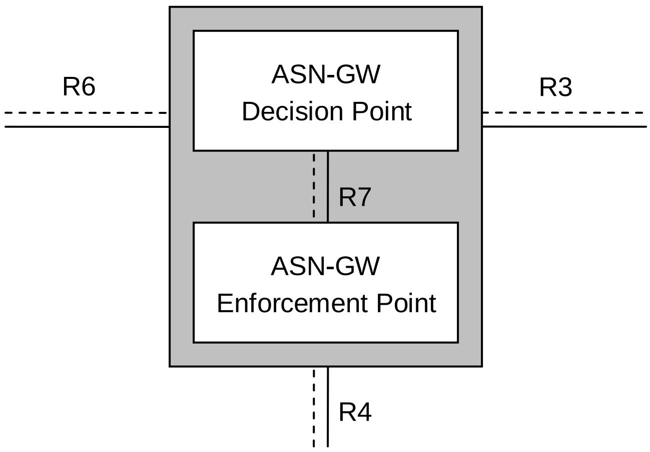

Figure 3.4 - ASN-GW Decomposition Reference Diagram [32]

ASN Profiles

As already mentioned, the NWG has created three profiles for manufacturers and operators to guide them in creating a working Mobile WiMAX network. Here only basic functions are described and the more detailed view can be found at [33]. The **Profile A** has the handoff control at the ASN-GW. Also the Radio Resource Controller (RRC) is located at the ASN-GW, which allows Radio Resource Management (RRM) with multiple BSs. The RRM can be understood as methods trying to increase reliability and efficiency of wireless transmission. A BS contains a Radio Resource Agents (RRA) that maintains a database of radio resource indicators. Finally, the R4 and R6 physical connections are used for ASN anchored mobility within the BSs. The **Profile B** does not make any assumptions on the physical location of different ASN entities and they can be freely distributed or grouped within the ASN. The requirement, however, is that the Profile B ASN can still communicate through R3 and/or R4 with other ASNs regardless of their profile types. The last, the **Profile C,** is similar to the profile A, but here the handoff control and the RRC are located at the BS. RRC and RRA in the BS allow RRM within that BS and the ASN-GW is equipped with an RRC relay for delivering RRM messages from BS to BS via R6

3.1.4 Interworking with Other Technologies

Since the access network in Mobile WiMAX scenarios is based on all-IP, it is rather simple to attach other access network with different technologies to the Mobile WiMAX. The NWG specifies scenarios for 3GPP, 3GPP2, and DSL interworking with Mobile WiMAX. The intention is to combine different technologies to a working solution with for example common AAA servers. This way the operators with several networks can achieve financial savings since certain services can be shared. The customers benefit as well since the needed services can be obtained from a single operator with common accounting.

3.2 Handoff

A special requirement for a mobile device is the ability to change the serving BS if there exists another BS with, for example, better link quality in the reach of the MS. The handoff, in some sources referred as handover, is a procedure with an intention to switch the network connection access point of the MS without data loss or disturbing the existing connection(s).

First, for a handoff to be even possible, one needs to have at least two BSs, the currently serving and the handoff target(s), and an MS within reach of both BSs. The handoff usually is understood as a change of serving BS, but it does not necessarily mean that the BS must be changed. In some cases the handoff can occur also within the same BS, though within different channels. This handoff type is called intra-cell handoff, while the other option is called inter-cell handoff. Handoffs between different technologies are also possible, as already mentioned while discussing the MIH standard. The horizontal handoff was defined to be a handoff within a single technology network, while the vertical handoff changes the network.

The reasons for handoff can be various and here are listed only some of them:

- signal strength is not enough for maintaining proper connection at the edge of the cell
- BS capacity is full and more traffic is pending
- disturbing co-channel interference from neighboring cell
- behavior of MS changes, for example in a case of fast-moving MS suddenly stopping, a large cell size can be adjusted to a smaller one with better capacity
- faster or cheaper network is available (if vertical handoffs are supported)

The handoff has roughly two major types, a hard and a soft handoff, with different variants of these depending on the used technology. The hard handoff is performed, when the connection to the serving BS is broken before creating the new connection with the target BS. With soft handoff the connection is transferred to the new BS and after successfully continuing communications the old BS can be released. The hard handoff can be very efficient regarding the channel usage, since only one channel is occupied simultaneously. This makes the equipment also cheaper because it does not have to support two or more channels in parallel. However, it can cause unrecoverable damage to the connection in case the handoff fails. The benefit of soft handoff is the reliability since the connection is broken only after finding a working connection. The drawback of soft handoffs is the required computational capacity in the equipment, which consumes money and power. Additionally, the use of several channels per user decreases the overall capacity of the BS.

Usually, the handoff process follows a common pattern. The BS maintains a list of neighbors that can be used in a case a served MS needing to perform handoff. The connection quality is constantly monitored and at some point the decision for a handoff is made. The criteria for the decision may be for example something listed in handoff reasons above. Before performing handoff an appropriate candidate must be chosen and then the handoff procedure is continued based on the current application and technology. The exact procedures vary depending on used technology and usually within the technology several alternatives are available as well.

In WiMAX scenarios the technology has to be 802.16e-2005 since the 802.16-2004 does not support handoffs at all. Additionally, there must be way to measure connection quality, since the transmission medium is constantly in change. To be able to perform handoffs, the technology must define a scheme for decision making to initiate them. A procedure for discovering competing BSs is also needed.

The handoff should also be as fast as possible, at least fast enough to keep current IP connections alive. Data traffic is not so sensitive to larger delays but real-time voice or video (or both simultaneously) requires a swift change of the serving BS.

3.2.1 Handoff Types

The 802.16e specification has a support for three handoff methods: the Hard Handoff (HHO), the Fast Base Station Switching (FBSS), and the Macro Diversity Handover (MDHO). The first one is required while the others are optional ones. The WiMAX forum [1] has been working on the HHO designing enhanced techniques to achieve handoffs (layer 2) in less than 50 milliseconds.

The Table 3.1 presents the greatest difference between the traditional WiMAX and the new mobile version. As can be seen, the traditional WiMAX does not support handoffs at all.

Table 3.1 - Comparison of Mobility in 802.16-2004 and 802.16e-2005 [34]

Access	Location / speed	Handoff	802.16-2004	802.16e-2005
Fixed access	single / stationary	no	yes	yes
Nomadic access	multiple / stationary	no	yes	yes
Portability	multiple / walking speed	hard handoff	no	yes
Simple mobility	multiple / low vehicular speed	hard handoff	no	yes
Full mobility	multiple / high vehicular speed	soft handoff	no	yes

Hard Handoff (HHO)

The hard handoff is a procedure to change the serving BS using a "brake-before-make" -way, in other words the connection to the old BS is broken before a new BS is connected. This way the excess signaling traffic can be avoided during the handoff, but the time before the connection is again in normal operation can be longer. [35]

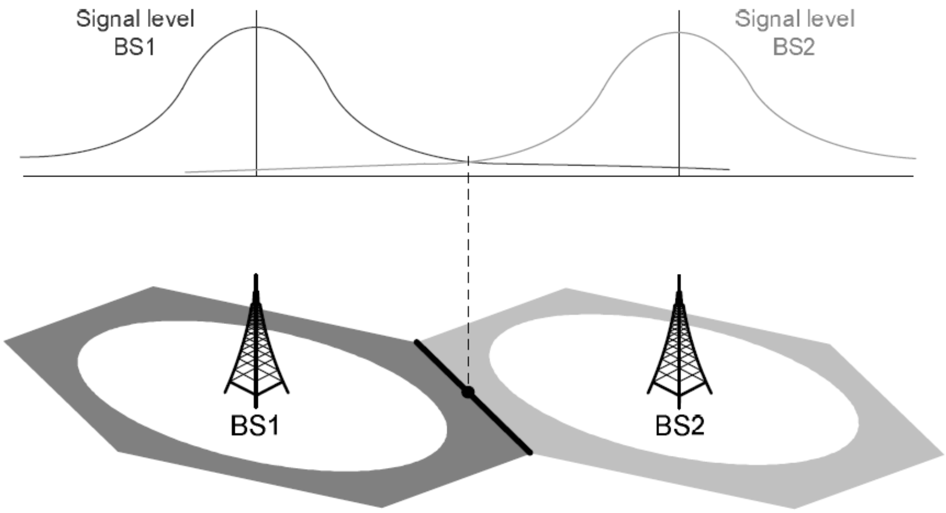

Figure 3.5 - Hard Handoff Realization [36]

While connected to a BS, the MS listens to the link-layer messages in case a new BS's periodically broadcasted neighbor advertisement message (MOB_NBR-ADV) is received. These messages are used for identification of networks and distributing the properties they have. The information received can give, for example, facts about the signal quality from a neighboring BS. If a better BS is not found, the MS can store the information for possible future handoffs. The HHO is described later in sub-section 3.2.2 while introducing the handoff process. Figure 3.5 above demonstrates the situation when a moving user reaches a point where the signal level is better with another BS. Additionally, a decision criteria hysteresis needs to be included to avoid constant handoffs back and forth between BSs.

Macro Diversity Handover (MDHO)

The MDHO is an optional handoff scheme for the Mobile WiMAX and therefore needs to be supported by both the MS and the BS. The MS keeps a list of BSs capable to the MDHO on its coverage area (as can be seen in Figure 3.6). This group is called a diversity set, or in some sources an active set. There is always one BS in the diversity set that is defined as an anchor BS. The normal functionality is a special case of MDHO when there is only one BS in the diversity set.

There might be also BSs that can be reached with the MS, but the signal is too weak for real traffic. These BSs are kept outside the diversity set and named as neighbor BSs. Naturally, while moving towards a neighbor BS, at some moment the signal is strong enough and the BS can be included in the diversity set, or the other way round. The measured factor is long-term CINR which is compared to the defined limits for adding/dropping a BS from the diversity set.

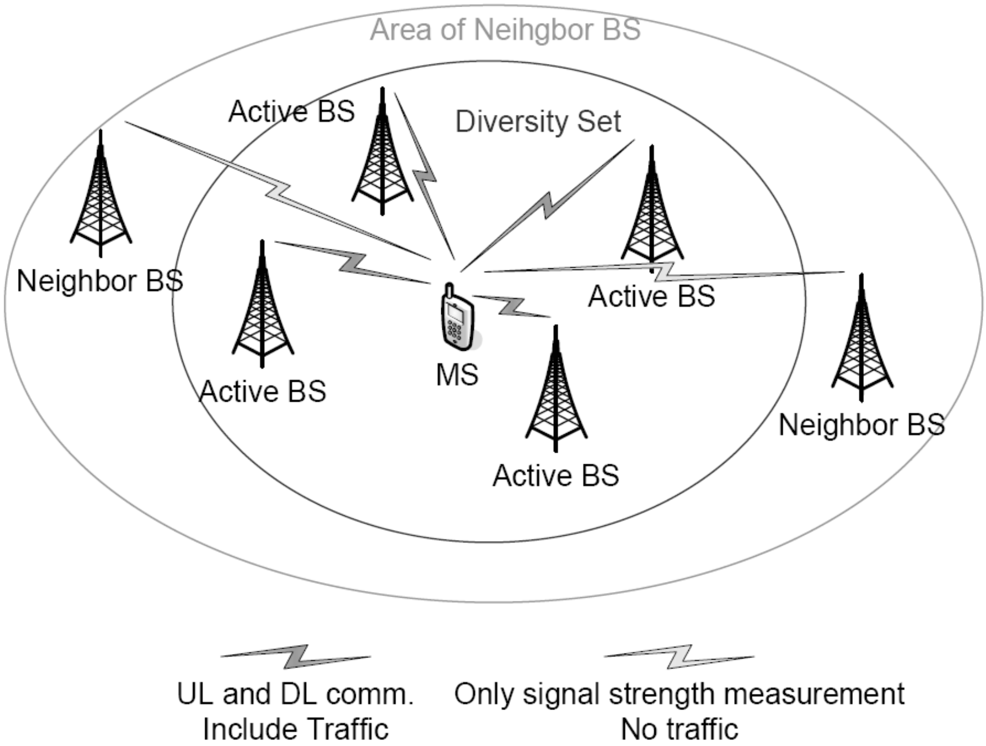

Figure 3.6 - Macro Diversity Handover [36]

The MS has two ways to monitor DL control information and broadcast messages. Either it listens to only the anchor BS for burst allocation information of other (non-anchor) BSs in the diversity set or it listens to all the BSs in the diversity set. While monitoring all the diversity set BSs, a DL/UL-MAP message from any BS may include information for the other BSs.

The procedure of MDHO is started by the MS when it decides to receive and/or transmit from multiple BSs at the same time interval. For DL traffic, two or more BSs transmit the data to the MS and the diversity combining is performed in the MS. For the UL traffic, the transmission from the MS is received by the diversity set BSs and selection diversity is performed.

The MDHO requires several terms to be fulfilled before it can be used. First of all, the involving BSs communicate through the RRAs at each station and they are synchronized on a common time source, since the frames sent by the BSs at a certain time frame have to be received at the MS within the prefix interval. The BSs frame structures have to be synchronized and the frequency assignment has to be the same. Additionally, the same set of CIDs has to be used by all the BSs that form connections with the MS. Furthermore, all the BSs should send the same MAC/PHY PDUs to the MS. Finally, the BSs involved in MDHO must share MAC context. By MAC context is meant everything a BS and an MS usually share from encryption information to information exchanged during network entry. [2]

Fast Base Station Switching (FBSS)

The FBSS is based on a similar principle as the MDHO above. Again both the MS and the BSs have to support the FBSS. A diversity set is kept in the MS and the BS but the MS communicates only with one BS in the diversity set (see Figure 3.7 below). The currently serving BS is named as an anchor BS. In FBSS the communication, including the signaling traffic focuses on only one BS at a moment but the anchor BS can be changed for every frame separately. Naturally, the changing is possible only if there are multiple BSs in the diversity set. The adding/dropping of members of the diversity set is similar to the one with MDHO above.

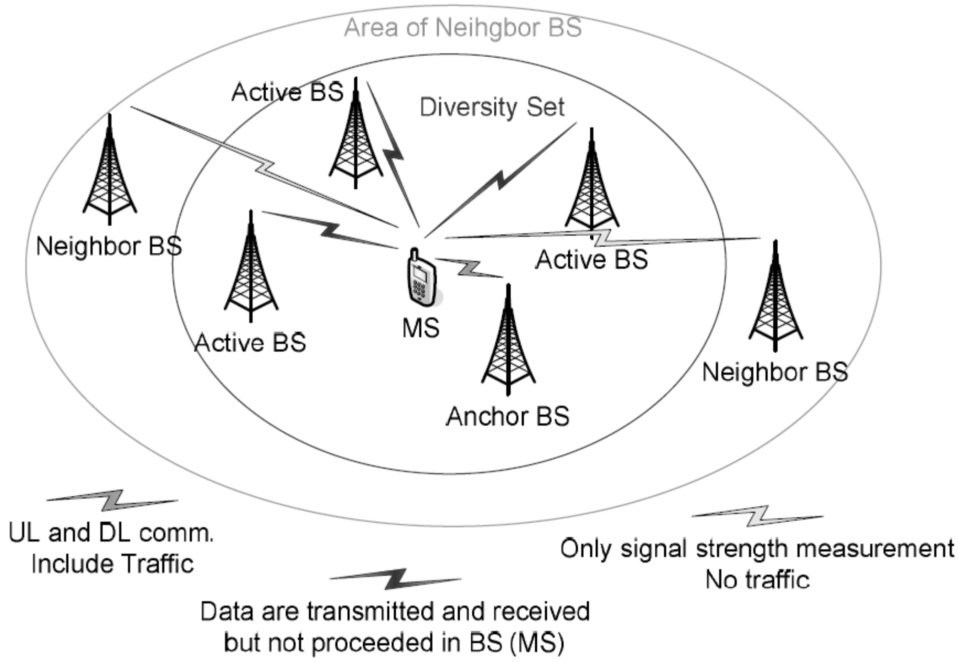

Figure 3.7 - Fast Base Station Switching [36]

In fact, all the BSs in the diversity set receive the data addressed to the MS, but only one of them transmits the data over the air interface while the others eventually drop the received packets. The operation of FBSS is based on the decisions of MS regarding the used (anchor) BS and these decisions are transmitted on the CQICH channel or by MS/BS initiated request message. Again, the decision of MS overrules the ones of BS.

The requirements of FBSS are the same as earlier with MDHO without the demand for same set of CIDs and MAC/PHY PDUs. [2]

3.2.2 Handoff Process

The handoff process in Mobile WiMAX is described in the following sub-subsections. The Mobile WiMAX specification [2] defines the procedures during the handoff, but the making of handoff decision is left outside the scope of it.

Generally, the decision for a handoff can be determined based on various properties and values. As described in [37], the decision attribute is a combination of network conditions, system performance, application types, power requirements, MS conditions, user preferences, security, and cost. The network conditions and system performance can be improved by balancing the load of heavily occupied BSs to less active BSs, assuming possible within other requirements. Different applications in the mobile device can set requirements to the currently serving BS and it might be that it does not support all the needed technologies. Additionally, if a new BS can provide sufficient service with better power saving or security properties than the currently serving BS, it can be useful for the MS to perform a handoff to the new one. The costs and user preference can define that the network of the own service provider is used from several available networks.

The MS conditions are measured constantly and, if a certain level of degradation is noticed in some of the defined parameters, the handoff decision can be initiated. These parameters may include signal strength, BS coverage area, data rate, service cost, reliability, security, battery power, and network latency. [37]

In Figure 3.8, a combination of network entry and handoff processes is presented. It can be seen that the two procedures are very similar to each other.

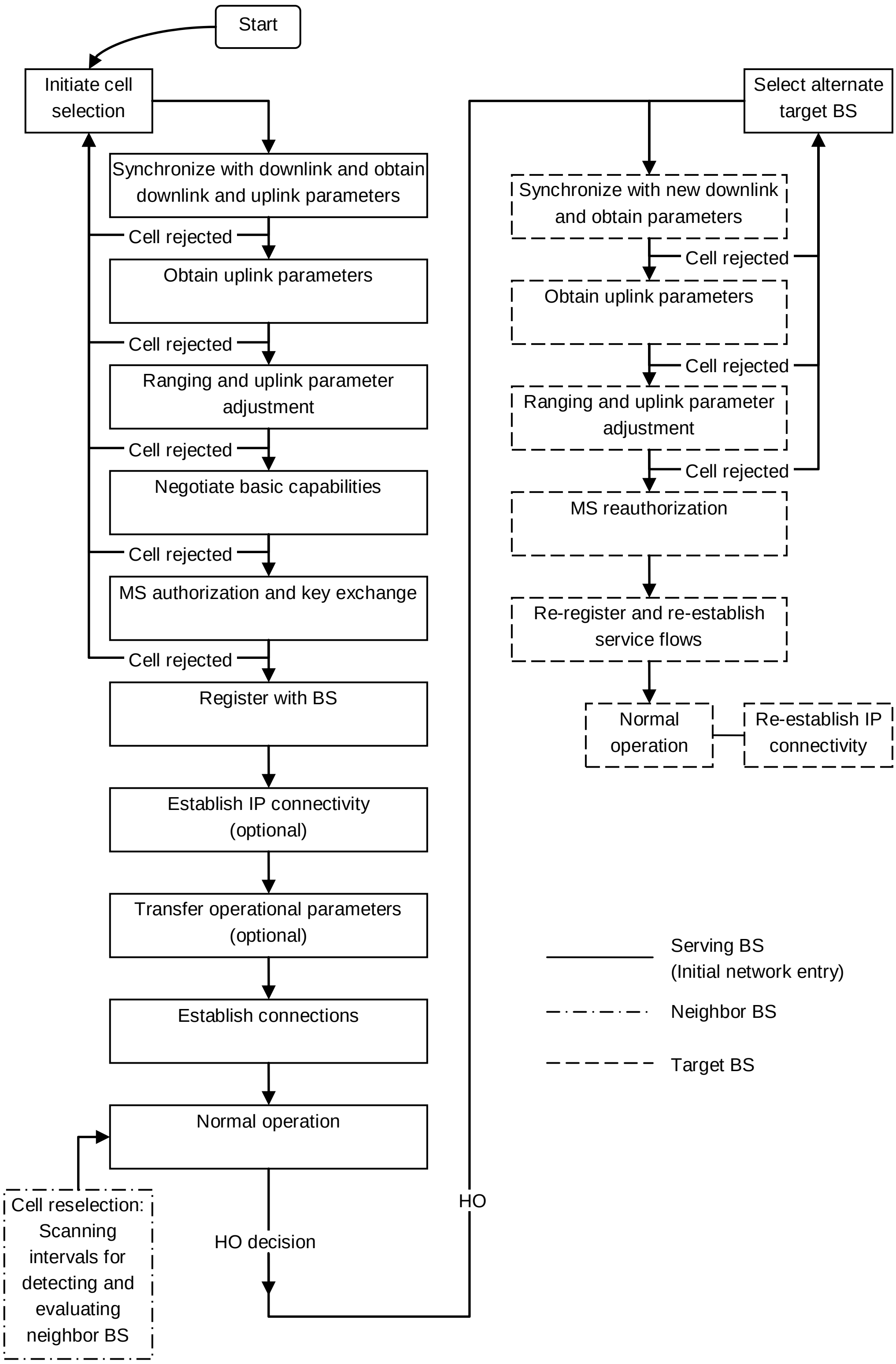

Figure 3.8 - Initial Network Entry and Handoff [2]

Cell Reselection

The cell reselection is a process with intention to find a potential BS for handoff. The MS has several possibilities to use while evaluating the possible change of the serving BS. It can exploit the information in neighbor advertisement messages (MOB_NBR-ADV). The

MOB_NBR-ADV message is sent periodically by the BS and the intention is to identify the network and to give the MS information about neighbor BS(s) for possible handoff or initial network entry. The BS stores the MAC addresses and indexes of neighbor BS(s) as mapping-tables and transmits them in the MOB_NBR-ADV message. The message includes also several other fields and is described in greater detail in [2]. Additionally, the MS can send a request for scanning interval(s) or sleep-intervals to be used for scanning and/or ranging the neighboring BS(s). This process is just a survey about handoff alternatives and the connection is not yet broken with the serving BS.

The Figure 3.9 describes the performed procedures during cell (re)selection, including ranging. The process begins with synchronization to the first BS and DL/UL parameters (DL- and UL-MAP, DCD, and UCD messages) are acquired. When the air interface parameters are received the channel measurements can be launched by sending a ranging request message (RNG-REQ). The BS responds with a ranging response message (RNG-RSP). These steps are described in more detail in the following sub-subsections.

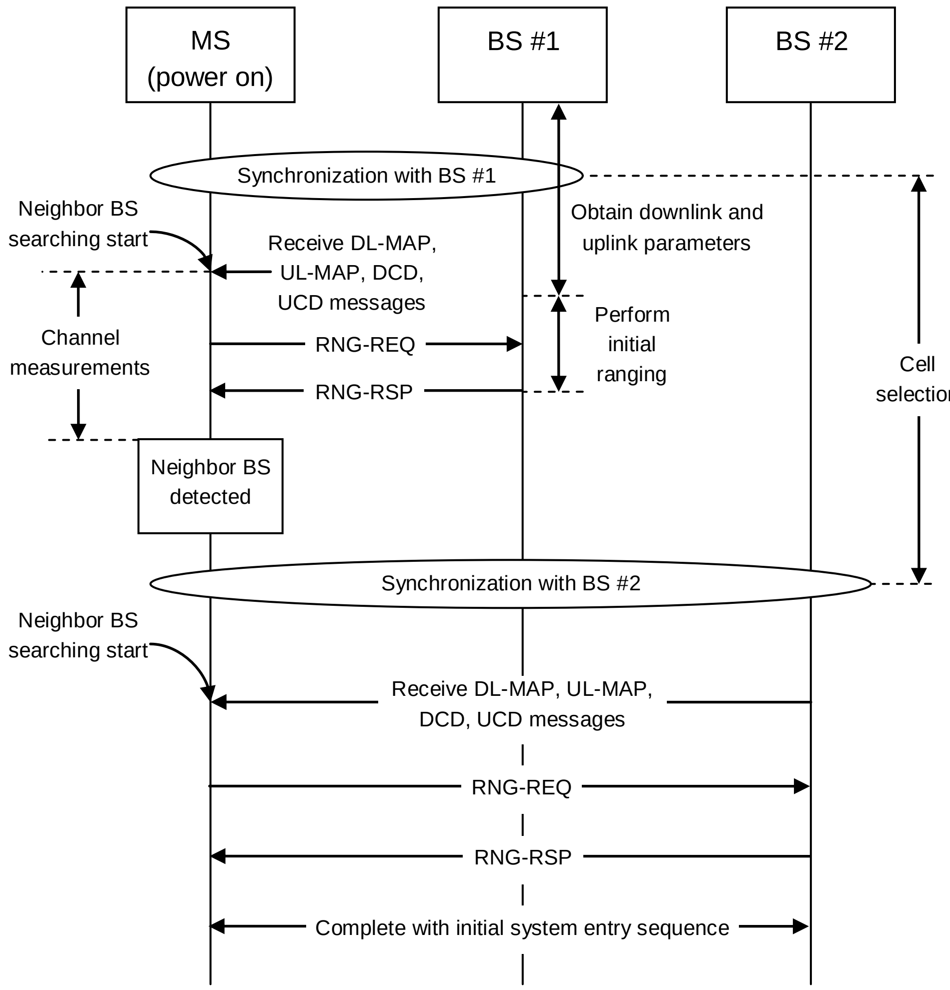

Figure 3.9 - Cell Selection with Ranging [2]

The messaging during the handoff is shown in Figure 3.10, as a case of MS initiated handoff. Shortly described, the process begins with a MOB_MSHO-REQ message sent by the MS to the serving BS. The BS gathers information about neighboring BSs and informs MS with MOB_HO-RSP. The MS confirms the target BS and initiates the actual process of handoff with MOB_HO-IND sent to the serving BS. Then the MS synchronizes with the target BS and retrieves the basic connection parameters (DL_MAP, UL-MAP, DCD, and UCD messages). After successfully resolving the mentioned messages, the MS requests ranging (RNG-REQ) and target BS responds with RNG-RSP. With successful response the network re-entry is performed and the old serving BS released. As stated earlier, these steps are discussed in more detail in the following sub-subsections.

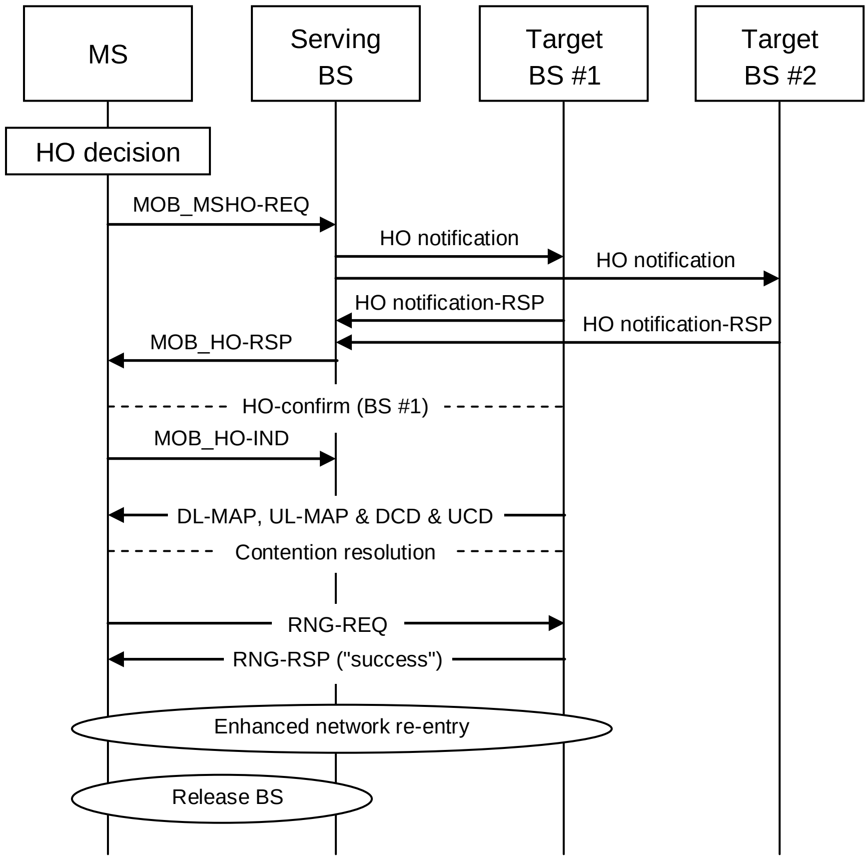

Figure 3.10 - Messaging during a MS Initiated Handoff

Handoff Decision and Initiation

The actual handoff begins when a decision is made that the MS changes the serving BS. The decision can be made at the MS, the BS, or on the network. The following step is sending a notification message, not obligatory but recommended, by either the MS (MOB_MSHO-REQ) or the BS (MOB_BSHO-REQ). However, if the notification message is sent, a response (MOB_MSHO-RSP or MOB_MSHO-RSP) is required. In a case when both send notification messages, the one sent by the MS has a priority over the one sent by the BS. Both notifications may include one or more possible target BS(s), which have been for example scanned earlier. There is also a possibility for the serving BS to communicate through the backbone with the possible target BS(s).

The serving BS has only a possibility to force the MS to handoff, not to define the target BS. The MS can choose or neglect recommended options for the target BS without restrictions.

The handoff decision is confirmed with a MOB_HO-IND message. The MOB_HO-IND is sent by the MS and it tells the BS whether the MS is really proceeding with the handoff or not. The message can include also other information related to BS selection:

- 0b00: HO (serving BS release, HO cancel, or HO reject)
- 0b01: MDHO/FBSS anchor update (confirm, cancel, or reject)
- 0b10: MDHO/FBSS diversity set update (confirm, cancel, or reject)
- 0b11: Reserved

Synchronization to Target BS DL

After the handoff is initialized the MS synchronizes with target BS's DL and UL transmissions by obtaining the required parameters. If the MS has received a neighbor advertisement earlier, the synchronization procedure can be faster. The advertisement needs to include target BS Identity (BSID), physical frequency, DCD, and UCD. If a handoff notification was sent by the serving BS and received by the target BS (via backbone connection), non-contention-based initial ranging opportunities can be assigned.

Ranging

After the synchronization of the DL/UL parameters the MS starts the ranging phase. The two possibilities available are initial or handoff ranging. The ranging is a phase that consists of several processes between the MS and the target BS in order to communicate the properties of the transmission link.

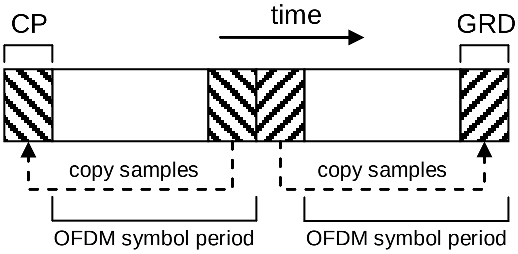

Figure 3.11 - Initial/Handoff Ranging Transmission for OFDMA [2]

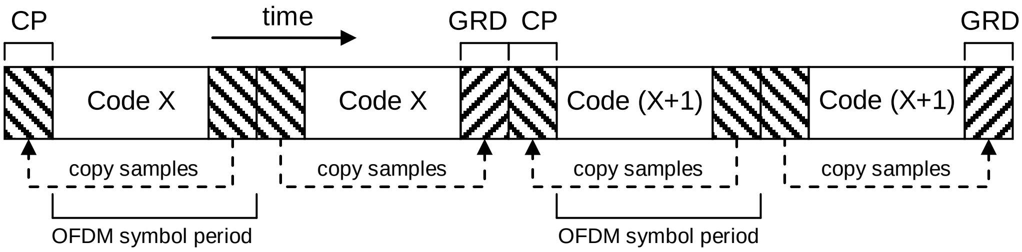

Figure 3.12 - Initial/handoff Ranging Transmission for OFDMA(using two consecutive initial ranging codes) [2]

Figures 3.11 and 3.12 demonstrate the structure of ranging code to be used during initial or handoff ranging, with single and two consecutive ranging codes. Periodic ranging and bandwidth requests, used during normal operation, use similar structure, but they last only for one or two OFDM symbol periods while initial/handoff ranging has the codes in pairs two or four periods.

The initial ranging begins by the MS choosing a ranging slot for a ranging code (using CDMA) to be sent to the target BS. The ranging codes are created as Pseudo-Random Binary Sequences (PRBS) with the generator described below in Figure 3.13, which is implementing the polynomial $1+X^1+X^4+X^7+X^{15}$. The seed used is b14...b0 = 0,0,1,0,1,0,1,1,s0,s1,s2, s3,s4,s5,s6, where s6 is the LSB of the seed and s0...s6 = UL_PermBase (an integer value of 7 LSBs of the Permutation Base parameter transmitted within the UCD) with s6 as the MSB of it.

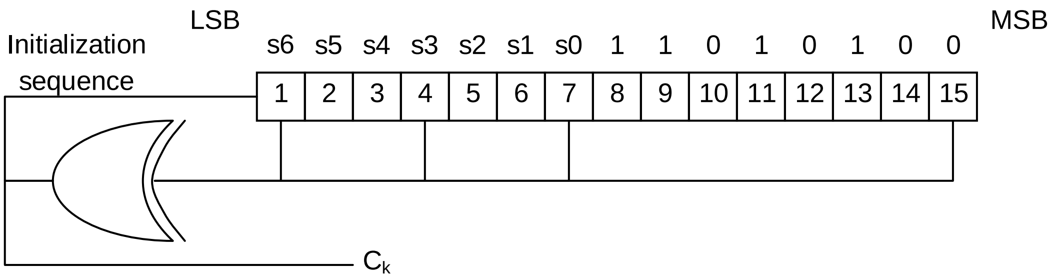

Figure 3.13 - PRBS Generator for Ranging Code Generation [2]

As a reaction to the MS ranging the BS broadcasts a ranging response message marked with the sent slot and code for the MS to identify the correct response. The response is broadcast since the BS can not know which MS sent the ranging code. The response message includes also all the needed adjustments (such as power, time, or frequency corrections) with status notification. If the status is "continue", the MS will repeat sending the ranging code, generated according to the description above. With "success" status the BS allocates bandwidth for the MS and the ranging is over.

The difference with handoff ranging is that the sent CDMA code is selected from a special handoff-ranging domain. Additionally, if the target BS is informed in advance about the coming handoff, it can directly allocate the needed bandwidth. Another possibility for shortening the handoff duration is to use target BS information acquired earlier during scanning interval. This information can decrease the amount of needed RNG-REQ/RSP interactions, but it has to be recent enough to qualify accurate.

Network Re-entry

The network re-entry is performed in a similar way as the initial network entry, when a MS is turned on. The process of re-entry during handoff can, however, be enhanced and therefore made faster. Figure 3.8 (initial network entry and handoff) already described the phases of network (re-)entry and the handoff process so far has included the ranging phase. The next step is to negotiate the basic capabilities regarding for example modulation/demodulation. The reauthorization of MS and key exchange is performed and the MS registers with the target BS, which is intended for agreement of for example ARQ or CRC capabilities. Now, the MS has re-entered the network of the target BS and the service flows can be re-established with proceeding to the normal operation. Finally, the old serving BS can be released.

The target BS can acquire information from the serving BS via backbone connection, or even from other network entities. With this information, the basic capabilities negotiation, registration, privacy key management, authentication and/or encryption key establishment phases can be skipped for enhancement of the network re-entry and therefore the entire handoff process. [2] [6]

Handoff Cancellation

The MS can cancel the handoff process anytime after the sending of MOB_MSHO/ BSHO_REQ message, as long as the above mentioned Resource_Retain_Timer has not expired. The cancellation is done by sending a message (MOB_HO-IND) containing a handoff cancel option.

Termination of MS Context

After the handshake with the target BS is completed the connection to the serving BS has to be broken. The termination message (MOB_HO-IND) with a code indicating BS release is sent to the serving BS. Upon receiving the message, the serving BS starts a Resource_Retain_Timer. This timer defines when all context (information in queues, counters, timers, etc.) related to the MS is retained. However, in a case when the target BS sends a backbone message of successful MS network attachment with it, the timer can be bypassed and the MAC context and PDUs related to the MS removed from the old serving BS.

Drops during Handoff

There can be a situation during the handoff process when the MS has stopped communicating with the serving BS before the normal cell selection or termination of MS context have been completed. This situation is called a drop and the MS can detect it by failed demodulation of DL, or by exceeding the limit for consecutive RNG-REQ retries. On the other hand, the BS can notice a drop when the limit for inviting ranging request messages is exceeded.

If the MS detects a drop while trying to establish a connection with a target BS, it can attempt network re-entry with its preferred target BS as through cell reselection. Additionally, it can resume communicating with the serving BS by sending a handoff cancellation message

3.2.3 Handoff Method Comparison

The handoff methods provided by the 802.16e-2005 standard offer the same basic service intended for slightly different use. The HHO is best suitable for low mobility scenarios while the MDHO/FBSS can provide better performance for users moving with greater speeds. However, this does not mean that the MDHO/FBSS would be less suitable for low mobility.

The differences of hard and soft handoffs in general were already discussed at the beginning of section 3.2. The HHO is an efficient method that does not require multiple channels during the handoff, unlike the two other soft handoff methods. However, in a case of failed handoff the connection can be completely broken. This can be avoided with MDHO/FBSS, which on the other hand demand more capacity in terms of bandwidth efficiency (at least two parallel channels reserved during the handoff). Additionally, complex methods require also more computing power, which again makes the equipment more expensive and tend to consume more battery power as well.

The soft handoff methods can improve link quality while the MS is operating with a weak signal to the BS(s). Especially the MDHO can perform well because of several combined transmissions from different BSs. The probability that all the diversity set BSs would have an unusable signal is rather low. The FBSS can not take advantage of several simultaneous transmissions, but it can choose the best BS on a frame-by-frame basis. For HHO the situation can be very demanding and it could be forced to handoff back and forth between two (or more) BSs.

3.3 Power Management

Every device suffers from the fact that they need power to operate. Fixed devices usually have a power outlet nearby, but the situation is a lot different when the device becomes mobile. They have to manage without a recharge for days, even for a week. Therefore, it is necessary to design ways to do the required tasks with as low as possible power consumption. The size of the mobile device also sets restrictions to the physical size, and though, for the capacity of the battery. The supposed power consumption of the device must be known so that the battery capacity and size can be designed accordingly.

The main procedures consuming power in a wireless device are the transmitting and receiving the data to and from the air interface, and, in some devices, the processor and the display can also take a significant part of the available power capacity. Therefore, it is advantageous to minimize the unnecessary use of different components while they are not needed. For example, the antenna(s) of a Mobile WiMAX device would be sensible to turn off while there is no data to send or receive. However, in order to keep the connection active, the antenna(s) can not just be totally turned off, but the BS must be contacted periodically. It is apparent that a Mobile WiMAX device needs a method, on the one hand, to save power resources, and on the other hand, to maintain link quality. For this, the Mobile WiMAX introduces two modes for handling the issues with power usage, the Sleep Mode and the Idle Mode.

The battery must also be able to maintain the user interface of the mobile device. This depends on the device type, but the various input and output methods can be significant power consumers. A user having a video call requires the power used by a display, a camera, a speaker, a microphone and an antenna for transmitting and receiving data, with the computation of all of these. However, from these, the scope of Mobile WiMAX specification can only address to the usage of the antenna(s), while the development of the other components is left to their respective manufacturers. In the following sub-sections, the more detailed description of the mentioned methods for issuing the power saving in Mobile WiMAX is presented.

3.3.1 Sleep Mode

The sleep mode gives the MS a possibility to go to a state where the serving BS can not communicate with it. The timing is scheduled before-hand between the MS and the BS. This kind of functionality can improve the battery life of the MS and also give the BS more possibilities to exploit a crowded air interface. The sleep mode has to be supported by the BS but it is optional for the MS.

The MS does not necessarily have to go to "sleep" while switched to a sleep mode. It can also use the time for example by collecting information about surrounding BSs (besides the one serving). This information can be used to make handoffs faster.

The sleep mode can be used even with optional handoff mechanisms MDHO and FBSS. The assumption for this is that the timeout value for diversity set updates must not expire during the sleep mode.

Sleep Classes

While a Mobile WiMAX device is on sleep mode it can have either listening or sleeping window on. As the names suggest, the listening window is used for communicating periods with the BS and the sleeping window makes the MS disappear, at least from the BS's point of view. The traffic affected by the sleep mode is only in DL direction, since the MS always knows when it has UL traffic and can interrupt the sleep mode. Additionally, there are three types of classes to be used during sleep mode activity. [2]

Type I class is intended for Best Effort (BE) or Variable-Rate Non-Real-Time (VR-NRT) type of traffic. With this class a variable duration of sleep duration is used. The terminal enters the sleep mode in case there is no traffic with the type described above. The procedure starts with a sleep window with a certain initial length. Then a short listening window is introduced and in case there is no traffic for the MS another sleep window is implemented. However, the second sleep window is (usually) twice as long as the initial one. Again, short listening interval appears and with no traffic the terminal goes to sleep. The doubling of sleeping window can be restricted to a preferred value, otherwise, without traffic, the window would increase infinitely. In case the MS detects incoming traffic during the listening window, it resumes to normal operation. All the parameters (window lengths and execution timestamps) are negotiated with the BS before making any changes to the state of the terminal.

Type II class has similar ways to perform as type I above, but here the sleeping window length is constant. The suitable traffic types are Variable-Rate Real-Time (VR-RT) or Unsolicited Grant Service (UGS). This class is well suited for VoIP traffic since there an amount of data is sent periodically and the interval between transmissions can be spent in sleep. The difference here, compared to the type I, is that the terminal sends/receives data during the listening window and returns back to sleep. With type I, in case of incoming traffic during the listening window, the MS returns to normal operation. Later it can resume the sleep mode by starting with the initial sleeping window as described in the previous paragraph.

Type III class is intended to enable BS initiated sleep mode activation. With type III the MS goes to sleep for only one period and then returns to normal operation. Type III can be used with multicast traffic or with some management operations such as periodic ranging. The length of the sleeping period can be adjusted, for example, according to the needed time for the ranging.

Naturally, there may be several connection types active simultaneously. The user may be browsing the internet during a VoIP call. The data transfer would use type I sleep mode and the VoIP call would take the type II. In these hybrid cases the terminal can truly save power if the sleep windows of various classes are overlapping. Figure 3.14 below is demonstrating the situation.

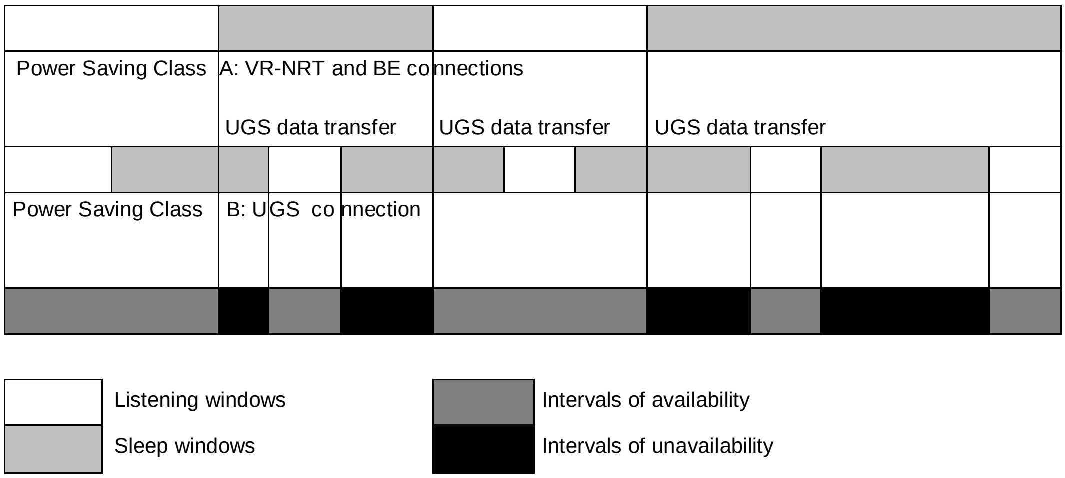

Figure 3.14 - Example of Sleep Mode Operations with Two Power Saving Classes [2]

The following Figure 3.15 gives an example of initiation of the sleep mode (type I). In case the MS is the one initiating the procedure, the first message is sent to the BS (highlighted in gray in figure). Otherwise, in the BS initiated case, the procedure is exactly the same, just without the first MOB_SLP-REQ -message

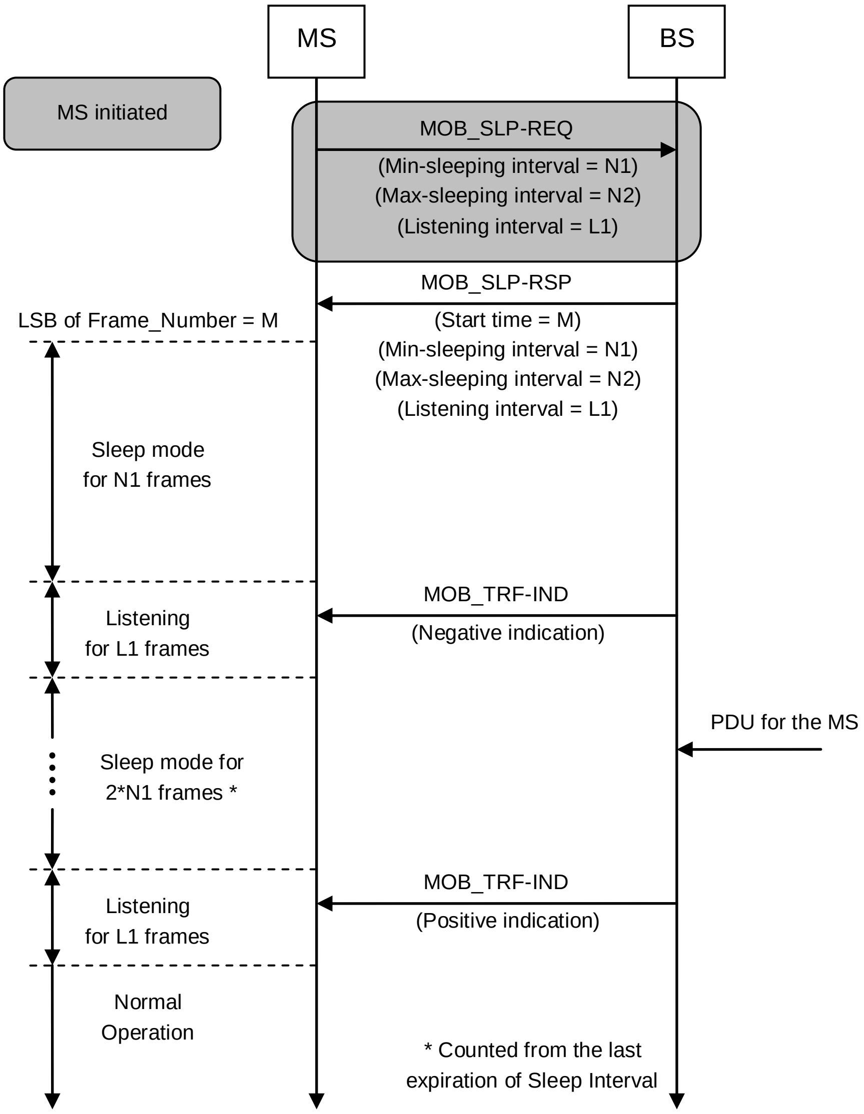

Figure 3.15 - Sleep Mode Example: MS/BS Initiated [2]

3.3.2 Idle Mode

The idle mode is designed for situations where the MS can, on certain time periods, receive broadcast traffic. The mode does not require registration to a specific BS. Presumed scenario is for example a motorway with multiple BSs along it and the MS is in a car traveling on the road. With the idle mode no handoffs or other signaling are needed between the MS and the BS. While the signaling traffic is dropped, the usable air interface resources are better available for those needing them. If needed, the MS can be alerted about incoming traffic in a simple and timely manner.

In a network supporting the idle mode, the BSs form logical groups (paging groups). The purpose is to create large enough coverage area in a way that the MS would not have to communicate in UL, yet be reachable by DL paging. On the other hand, the area has to be kept sufficiently small in order to avoid excess paging overhead. The BS groups are usually overlapping and some of the BSs may belong to several groups.

During the idle mode the terminal can have several different states of operation. These are briefly introduced in the following sub-subsections. [2]

MS Idle Mode Initiation

The initiation of idle mode can be executed from either the MS or the BS. In either case, a response is required before the MS can be unregistered from the BS. As mentioned, MS operates unregistered in the idle mode. After successful agreement of MS entering the idle mode the MS terminates the normal operation with the serving BS and a paging controller is defined. The paging controller is the serving BS or some other network entity, which is controlling the idle mode activity for the MS. Additionally, timers for idle mode are set in the MS and the paging controller. The paging controller timer (Idle Mode System Timer) and the MS timer (Idle Mode Timer) define the intervals for sending MS location updates to the paging controller. In the case of expiration of these timers the idle mode is considered to be ended and the MS returns to normal operation.

Cell Selection

The MS can also select a new BS during the previously described MS idle mode initiation. The MS evaluates the available BSs and chooses the one with best air interface DL properties. This BS is named as preferred BS. However, the terminal can remain with the serving BS without changes.

MS Broadcast Paging Message Time Synchronization

The MS evaluates a BS paging interval from DCD and DL-MAP of the selected BS for the idle mode. The contained information about frame size and number give the time till next BS paging opportunity. Nevertheless, the MS has to calculate the required time for DL scanning, decoding of the DCD and DL-MAP, and synchronization. The needed time is taken from the interval between paging opportunities and the remaining time gives the MS paging unavailability interval.

MS Paging Unavailable Interval

During the unavailability time the MS can carry out several tasks. It can turn some properties off for power saving, scan neighbor BSs, re-select the preferred BS, perform ranging, or any other tasks not requiring DL traffic. If the preferred BS is changed the terminal will return to the previous time synchronization state.

MS Paging Listening Interval

Before able to receive a BS broadcast paging message the MS has to scan, decode the DCD and DL-MAP, and synchronize with the preferred BS. After the listening interval the MS can return to unavailable state, unless otherwise decided. In case something being received during the listening interval, the MS has to react according the information acquired.

BS Paging Interval

The BS paging interval is intended for the BS to be able to send one or several broadcast paging messages. However, it is not necessary to send a message at all if there is no need for it. The interval is configured by information retrieved from the backbone network.

BS Broadcast Paging Message

The task of a BS broadcast paging message is to notify a certain MS about incoming traffic or to request location update, without the need for full network registration. The message is sent during the BS paging interval on the broadcast CID or the idle mode multicast CID. The MSs are identified in the message by their MAC address hash and a single message can include several MAC addresses. The message has a certain action code upon which the MS reacts. The available codes are:

- 0b00: no action required
- 0b01: perform ranging to establish location and acknowledge message
- 0b10: perform initial network entry
- 0b11: reserved

After sending the code for ranging or network entry the BS waits for a response from the paged MS. If no response is received until the next MS listening interval, the message is sent again. There is a counter for unsuccessful broadcasts and when it expires the MS is considered unavailable. Additionally, a message to update the list of MSs in idle mode is sent via backbone to all the BSs belonging to the same paging group.

Paging Availability Mode Termination

The paging availability mode termination can be initiated by the MS anytime simply by re-entering to the network. Otherwise, the BS can determine the MS to be unavailable (as described above in the previous paragraph) if a response is not received to a BS broadcast paging message. Additionally, the expiration of the Idle Mode System Timer (see the earlier MS Idle Mode Initiation sub-subsection) terminates the idle mode operation.

Location Update

The process of location update consists of evaluating the condition for update and the actual processing of the update. There are four different conditions upon which the location update process can be performed. **The paging group update** is carried out when the MS detects a change in the paging group. The information can be obtained from the preferred BS in DCD or broadcast paging message. **The timer update** is launched by the MS before the idle mode timer expires. **The power down update** is intended to be used when the terminal is turned off and a notification about this is sent to the paging controller. The paging controller can update the current status of the MS and remove all information concerning it. Finally, **the MAC hash skip threshold update** is executed by the MS when the MAC hash skip threshold is successively exceeded by the MS MAC hash skip counter. After the update, the BS and the MS re-initialize their respective MAC hash skip counters, assuming the update process has been successful. The updates mentioned above are performed with the preferred BS using either secure or unsecure process (the details can be found from [2]).

Network Re-entry

In order to return to normal operation from the idle mode the MS has perform a network re-entry. After successfully performing the required operations (ranging, authentication, etc.) the target BS informs the paging controller about the network re-entry. The paging controller can, but do not have to, send a backbone message to the BS which started the idle mode with the MS. The message informs the BS that the MS has resumed normal operation at the new serving BS. [2] [3]

3.4 Summary

The third chapter discussed the Mobile WiMAX with emphasis to the issues providing the mobility. The network architecture was described first in order to better understand the actions during the handoff, which was introduced next. Additionally, the supported power saving features were presented.

The Mobile WiMAX network constructed from two main parts, the ASN and the CSN. The ASN consisted of the all-IP access network between the BSs and ASN-GW. The CSN connected the ASN-GW to a network, which could be the internet, PSTN, or something else. The CSN also included other administrative functionalities, for example an AAA server.

The handoffs make the mobility possible and three different methods were presented for the Mobile WiMAX. The mandatory version was the HHO, which was the basic way to change the BS by first breaking the connection with the old one and then registering to a new one. The two other optional methods, the MDHO and the FBSS, used several BSs during communications and therefore the concept of handoff was not as straightforward as with the HHO.

The power saving is an important feature for a mobile device and the Mobile WiMAX introduced two methods. The Sleep Mode and the Idle Mode were demonstrated to be used when the MS had nothing to send or receive. This chapter ends the literature study part of the thesis and the next chapter describes the performed simulations.

4 SIMULATIONS

The goal for simulations was to test the properties of Mobile WiMAX in practice, or at least, less theoretically. To keep the simulations simple enough, a very basic scheme was planned. At the time of designing the simulations only the Network Simulator [38] (version 2.29 with additional WiMAX and Mobility packets from a NIST project [39]), later in text NS-2, was the only simulator available with support for Mobile WiMAX. However, the additional packets did not have full functionality of the 802.16e-2005 standard and the simulation measurements with result interpretation had to be performed accordingly.

The handoffs between BSs were the special target of interest. They were illustrating the HHO method, though it was not possible to perform according to the standard because of lacks in the simulator, or to be more precise the lacks of the WiMAX add-on module. The add-on description [40] defines as the main supported features the PHY layer of WirelessMAN-OFDM with only TDD, messages for network entry management without support for authentication, 802.16e extensions for scanning and handoff, and fragmentation and reassembly of frames. Important Mobile WiMAX related features left out of scope are the mentioned authentication, WirelessMAN-OFDMA, FDD, ARQ, service flow and QoS scheduling and periodic ranging with power adjustment. As can be seen, some very basic Mobile WiMAX features are not supported and this has to be remembered when interpreting the simulation results. The lack of certain features (for example authentication) can speed up the handoff times, but on the other hand, some other features would make the handoffs take longer time since the missing elements are not supported.

The scenario consisted of three BSs evenly aligned on a line in a way that the coverage areas of two neighboring BSs had some overlap. Certain constant values (for example the cell size, the transmit power of BSs, the route of MS) were selected for the simulation and the handoff times were tried to make faster by adjusting the properties of the WiMAX module in the NS-2. Finally, the tests were also run with speeds 1-40 m/s (3.6-144 km/h) with 1 m/s steps. The assumed traffic was constant bit rate with data rate of 1.2 Mbit/s.

4.1 Scenario

The all-in-one package of the NS-2 did not include support for Mobile WiMAX and therefore additional components were required. Two packets, the WiMAX and the Mobility module, from NIST were installed to achieve simulations of mobile scenarios.

Since one of the example scenarios in the modules was very similar to the designed simulation scheme, it was taken as a basis for our simulations. The basic idea is shown below in Figure 4.1. There is a MS traveling through the coverage areas of three 802.16e BSs (BS0, BS1, and BS2).

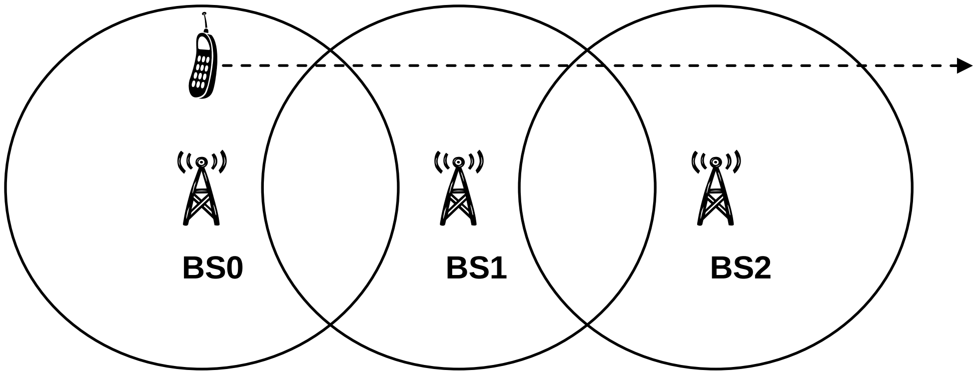

Figure 4.1 - Simulation Scenario

The BSs are aligned on a straight line so that they have 750 meters between each other and their coverage areas have a radius of 500m. At the start of the simulation the distance between the MS and the BS0 is 310 meters. The MS begins moving, as shown in Figure 4.1, while keeping the shortest distance to the BS in mentioned 310 meters. This value was selected in random and, naturally, it could be anything else too. However, the MS has to be in the reach of the BS coverage. The reasonable maximum distance is the crossing point of the edges of the cells, approximately 330 meters from the center line.

4.2 Components

The simulation code (see Appendix A.2) was based on the mentioned Mobility and WiMAX packets from NIST, especially the Neighbor Discovery (ND) and Media Independent Handover (MIH) modules were the key elements used in the simulation code.

4.2.1 Neighbor Discovery -module

The Neighbor Discovery (ND) module was designed to provide movement detection for layer 3. Its task is to create IP addresses when a network is changed. The module is a part of the MIH packet (described in the next subsection) and is intended to support multiple interface types, such as Ethernet, WLAN, UMTS, and, in this case, Mobile WiMAX. The ND agent uses broadcast or unicast messages according the technology in use.

The ND agent is located in all nodes, but the configuration in NS-2 has to be done according the type of the node in the network. For example, Ethernet or UMTS networks do not have a capability to send broadcast messages in NS-2 whereas WLAN has. The ND agent can be configured to send unicast messages according a pre-configured list of targets.

The functionality of the ND agent depends on the role of the node in the network, in other words, whether the node is a router or a host. The router functionality consists of sending unsolicited Router Advertisements (RA) periodically to the hosts. The possible sending period is defined with parameters *minRtrAdvInterval* and *maxRtrAdvInterval*. In case a router receives a Router Solicitation (RS) from a host, it sends an RA, assuming the time from previous sending is between the values of parameters described above. If a router receives an RA, it is discarded.

The hosts can ask for an RA with RS messages. When an RA is received the included prefix information is compared to the existing tables and possibly new values are added. Additionally, an expiration timer is attached to an RA message, which tells when to abandon the prefix information in case a new one is not received. [41] [42]

4.2.2 Media Independent Handover -module

The Media Independent Handover (MIH) module is a part of the NIST Seamless Mobility project and was developed to control handoffs with various technologies. The functionality is based on MIH Function (MIHF). It works on layer 3 and can communicate between local and remote interfaces. The remote interfaces can be contacted via another MIHF. This is illustrated in Figure 4.2 below.

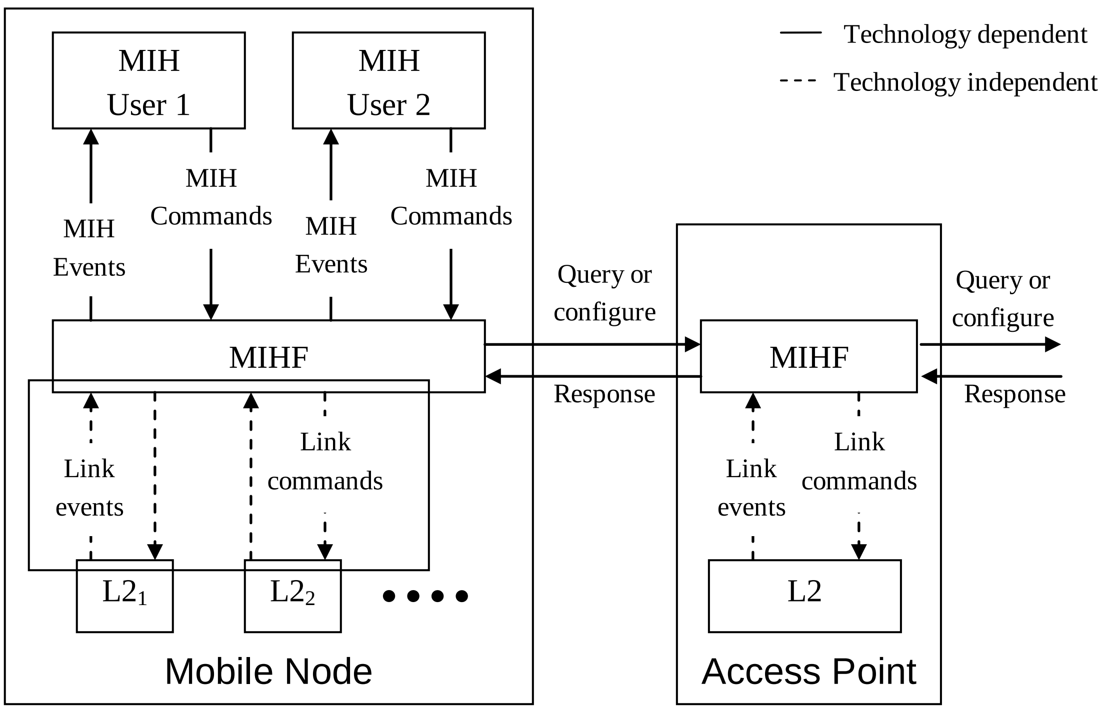

Figure 4.2 - MIH Design Overview [41]

The NS-2 required enhancements since the handoffs are not supported by default. The additions included the support for multiple technologies and modification of default implementation intended for 802.11. Additionally, a special node suitable for multiple interfaces had to be designed with support for subnet discovery and change of address (ND module). The solution was called a *MultiFaceNode* which is a virtual node controlling the different technologies and interfaces. [41]

4.3 Parameters

The intention with simulations was to find parameters that affect the most in the handoff performance. The values measured were the number of sent, received, and dropped packets with times for both handoffs. The time of handoff was determined to be the time difference between the last received packet from the old BS and the first received packet from the new BS. The two handoffs had never exactly the same length in time and usually the difference varied from a few milliseconds to a few tens of milliseconds.

4.3.1 Constant Parameters

Some of the parameters were set as constant in order to keep the simulation simple. The BS coverage area and transmit power as well as operating frequency were unchanged, and they were the same for all three BSs. Additionally, the velocity was kept in 10 m/s and only after all the adjustments different speeds were tried (more in section 4.4). With lower speeds the length of simulation time had to be adjusted to allow two handoffs.

The topology of simulation was the same for all simulations. This included the locations of BSs and the route of the MS. The details were already described earlier in section 4.1.

The data was sent with constant bit rate so that the packet size was 1500 bytes and a packet was sent with 10 ms interval. These result in the data rate of 1.2 Mbit/s, which is rather moderate compared to the announced theoretical maximum rates (Table 2.6). However, the selected rate is already nearly sufficient to provide a MPEG-1 video stream, which needs a data rate up to 1.5 Mbit/s. [43] Since the goal was not to test the capacity, the low data rate was reasonable.

4.3.2 Adjusted Parameters

Since the simulator was deterministic (the results with unchanged parameters are always the same), it was enough to run the simulation only once per change of a parameter. Certain pre-set values from the example file were used as a basis and were changed while trying to find the shortest handoff times. Some of the parameters were defined in the original example file and the rest are additional parameters offered by the Mac/802_16 in NS-2. Nevertheless, the additional parameters have an impact on the simulation, although they are not written in the simulation code. If a parameter is not separately defined in the code, they use default values described in the WiMAX module files. The Appendix A.1 presents the adjusted parameters one-by-one with brief description of the parameter function and the possible influence to the handoff.

The adjustments displayed that some of the parameters did not have an influence at all to the handoff latency, but there were others that had an impact on the latency. The Link Going Down -factor was one of the parameters with rather significant influence. It determined the detection sensitivity of a failing link. It was important to detect the link failure, on one hand, early enough before the connection is broken, or on the other hand, late enough, to avoid unnecessary handoffs. Several timer and timeout parameters had also their impact on the handoff latency. This was quite understandable since they usually define some time to wait before some function is performed. If the function is somehow related to the handoff process, it can delay the process even significantly. One of these was the *t21_timeout_*, which defined the time for a MS to search for the DL-MAP message on a certain channel. As stated above, the Appendix A.1 explains all the parameters adjusted and the results accomplished during the simulations.

When the adjustment of parameters was completed, the amount of sent packets during one simulation was 17897 and 17893 of them was also received. These are the combined values of all three BSs. The reported packet drop was five packets. As can be calculated, the presented figures did not add up (17893+5 = 17898 ≠ 17897), hence there was one additional packet drop. This was explained by examining the simulation data, which showed, that the first drop comes already before any packets are sent. The timestamp of the first drop was at 24.14 ms, while the first data packet was sent at 38.23 ms and received at 39.16 ms.

Another interesting point was found out, when the time difference between transmission and reception time stamps were compared. When the MS communicated with the first BS, the difference was just below 1 ms, while with the second BS it is was already 21 ms and with the last BS it doubled to 41 ms. It seemed that the simulator is circulating the packets through the first BS although each BS is defined to have direct link of their own to the presumed access network. Nevertheless, the handoff times were not influenced by this and the simulator was known to have weak points. Therefore, the problem was left for future study.

With the adjusted parameters the handoff times were 32.06 ms for the first handoff and 33.75 ms for the second one. During the adjustments, either handoff latency value was occasionally even below 30 ms, but in such cases the other was significantly larger.

There might be even better values for the parameters, since some of them had a rather unpredictable influence on the handoff times. As mentioned earlier, the main goal was to find the parameters that are the most sensitive in influencing the handoff duration.

4.4 Velocity of MS

When the adjustments of NS-2 and WiMAX-module parameters were performed, the influence of velocity of the MS was also investigated. In the previous section the speed of the MS was set to constant 10 m/s (36 km/h). In these simulations the simulation parameters were untouched and only the speed of the MS was changed.

The simulations were done with MS speeds between 1 and 40 m/s with 1 m/s step. The 40 m/s equals to 144 km/h, which is well above the mentioned (subsection 2.2.5) 100 km/h "limit" for a seamless handoff. For the slowest speeds the overall length of the simulation time had to be increased in order to allow two handoffs. Additionally, higher speeds 50, 60, 75, and 100 m/s were also simulated. Although these are not within the standard, the intention was to try the sensibility of the simulation.

The handoff times varied in the region of 40 ms and stayed nicely below the 50 ms limit until the MS reached the velocity of 20 m/s, apart from few exceptions that exceeded the limit by only few milliseconds. After this, the times show a more or less steady growth up to 150 ms region with the 40 m/s MS speed. The handoff times (for 1-40 m/s) are drawn in Figures 4.3 and 4.4 below.

The higher speeds (50, 60, 75, and 100 m/s) showed also a steady increase while the velocity grew. The handoff time with 100 m/s, according to the simulations, was just below 0.2 seconds. The simulator seemed to handle the speed and there were no traffic problems between the MS and the BS, but in real life the situation would hardly be the same, or even possible to achieve.

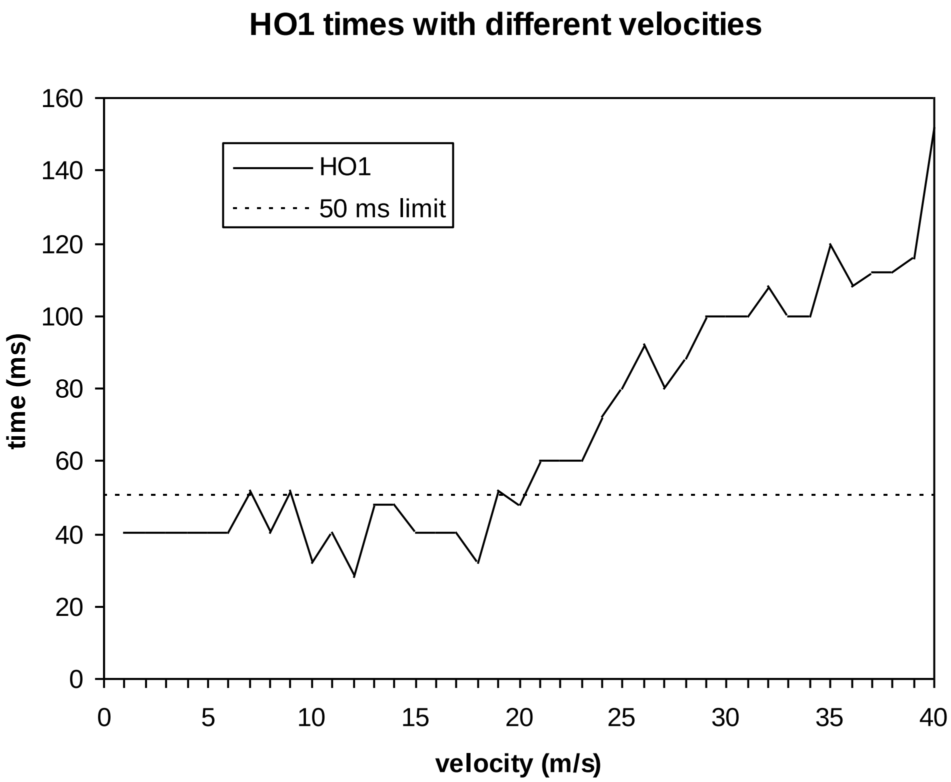

Figure 4.3 - Handoff Times for the First Handoff

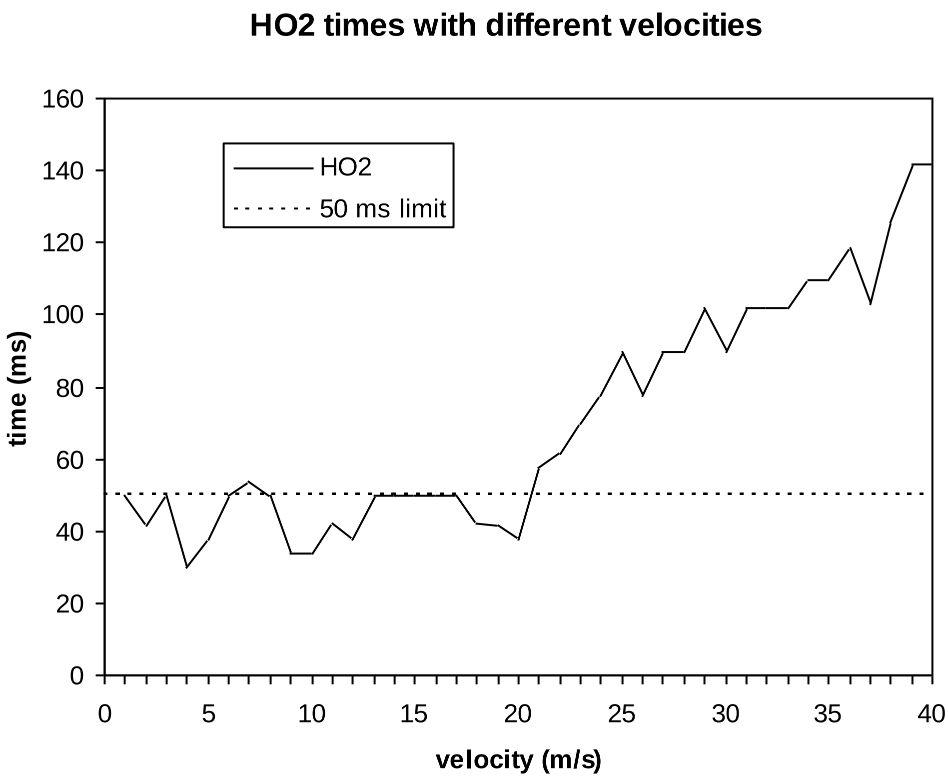

Figure 4.4 - Handoff Times for the Second Handoff

4.5 Summary

The fourth chapter discussed the simulation part of the thesis. The designed scenario was very simple with three BSs placed equally in a row and the MS moving through the coverage areas of all of them. The used software was NS-2 with two add-on modules from NIST. The handoff latency was measured during the handoffs and the parameters of the simulator were adjusted one-by-one in order to achieve faster handoff times. The goal was to find out the parameters with the most influence on the latency and comparison of the values to the 50 ms limit set by the WiMAX Forum. After the adjustments were performed and the best possible values for this scenario solved, the simulations were additionally run with different velocities between 1-40 m/s. Though, the simulation add-ons had some lacks and were not functioning according to the standard, the results seemed to follow the values promised by the WiMAX Forum.

5 CONCLUSIONS

This thesis has introduced a new mobile and wireless transmission technology standardized by IEEE 802.16 working group. The Mobile WiMAX (based on the 802.16e-2005) is an amendment to the earlier 802.16-2004, known also as Fixed WiMAX or just WiMAX. The main advance from the 802.16-2004 is the scalable OFDMA, which allows the adjustment of bandwidth depending on the spectral requirements at certain moment. This allows different environments to be served with high performance. Additionally, several other advanced methods are introduced for Mobile WiMAX in order to support the true mobility of it. As examples can be mentioned the support for multicast and broadcast with MBS, two power saving modes (Idle and Sleep Mode), and finally the most important feature: the handoff support.

The Mobile WiMAX offers three handoff methods. The mandatory HHO and the optional soft handoff types MDHO and FBSS are discussed and the HHO is used as the technology for the change of BS in the simulation part of the thesis.

The simulations were performed with the Network Simulator version 2 (NS-2). The simulator itself did not support the Mobile WiMAX and additional modules had to be obtained. The (only) add-ons available for this purpose were the WiMAX and Mobility packets from NIST. Their solution was incomplete, but anyway it was able to perform Mobile WiMAX handoffs.

The simulation scenario consisted of one MS traveling through the coverage areas of three BSs. The total of two handoffs was performed when moving from coverage area of one BS to the other. The MS was moving with constant speed of 10 m/s and sending constant bit rate traffic with data rate 1.2 Mbit/s. The latency during both handoffs and the number of sent/received/dropped packets was examined. The simulator parameters were adjusted when trying to find out the fastest ones considering the handoff duration. Additionally, after the adjustments, the parameters were kept constant and only the speed of MS was changed between 1-40 m/s, with 1 m/s step.

The main task with simulations was to determine the parameters that had the greatest impact on the handoff latency when adjusted. As a result, some of the parameters did not influence the handoff times at all, but changing some of them even slightly had direct consequences. Significant impact was met for example with Link Going Down -factor, which determines the sensitivity of detecting a failing link, or different timeout and timer values such as the *t21_timeout*, which defines the time used for searching DL-MAP.

The obtained results give a preliminary view to Mobile WiMAX technology, because of the lacks in the NIST add-ons. In order to achieve more reliable results the simulations should be repeated when the software is developed enough to be in line with the standardization, or even with actual devices certified to comply with the standard.

The Mobile WiMAX discussed in this thesis is still in early stages of its life cycle and offers still much research topics related to the handoffs or the technology in general. Straight follow-ups relating to this thesis could be for example further experiments with the used simulation scenario. One could add several MSs moving simultaneously through the coverage areas of the BSs or try similar parameter adjustment with several different velocities of the MS. It could be interesting to see the possible changes in parameter values whether the MS is moving with low or high speed. The different velocities were already simulated in this thesis, but the parameters were untouched during that simulation. It has to be noted, however, that the simulator, as it was during the simulations, still needs heavy development towards the standard in order to be convincing.

The first products supporting 802.16e-2005 standard are currently appearing or has just appeared on the markets and at some point in the future testing of those could give more information about capabilities of Mobile WiMAX in practice. Several companies developing simulators have already published their equipment, which are also in accordance with the standard. The simulations with NS-2 or similar software-only based solutions are always in some way handicapped when compared to the actual physical devices.

The IEEE 802.16 working group has approved a new task group, which is denoted with letter "m". The 802.16m [44] is designed to be the next generation of WiMAX providing data rates even up to 1 Gbit/s. The standardization of 802.16m is planned to be ready by the end of 2009 and the new standard should be backwards compatible with the previous WiMAX version, which was not the case with 802.16-2004 and 802.16e-2005. Additionally, the IEEE has set a goal to fulfill International Mobile Telecommunications Advanced (IMT-A) requirements set by the ITU. The IMT-A can be understood as a next generation (4G) wireless in a similar way as the IMT-2000 [45] was 3G. [46]

REFERENCES

Note: All links confirmed working on November 28th, 2007.

[1] The WiMAX Forum, Internet home page, http://www.wimaxforum.org/ (November 11, 2007)

[2] IEEE 802.16e-2005, IEEE Standard for Local and metropolitan area networks - Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems - Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1, http://standards.ieee.org/getieee802/download/802.16e-2005.pdf (February 28, 2006)

[3] IEEE 802.16-2004, IEEE Standard for Local and metropolitan area networks - Part 16: Air Interface for Broadband Wireless Access Systems, http://standards.ieee.org/getieee802/download/802.16-2004.pdf (October 1, 2004)

[4] WiMAX Forum: Mobile WiMAX – Part I: A Technical Overview and Performance Evaluation, http://www.wimaxforum.org/news/downloads/Mobile_WiMAX_Part1_Overview_and_Performance.pdf (August, 2006)

[5] J.H. Stott, The how and why of COFDM, http://www.ebu.ch/en/technical/trev/trev_278-stott.pdf (January 22, 1999)

[6] Qualcomm, Mobile WiMAX: Hype and Realities, http://www.wimaxtrends.com/docs/reports/QUALCOMM%20WiMAX%20for%20Yankee%20Group%20Ver%205.pdf (May 3, 2007)

[7] C. Taylor, WiMax to steal 3G and DSL market share, http://www.enn.ie/article/9569371.html (December 9, 2004)

[8] J. Shandle, All about OFDMA, http://www.wirelessnetdesignline.com/howto/broadband/198000601 (March 13, 2007)

[9] J. Andrews, A. Ghosh, and R. Muhamed; Nuts and Bolts of WiMAX - Parts I-VIII, http://www.mobilehandsetdesignline.com/howto/201302525 [Part I], http://www.mobilehandsetdesignline.com/howto/201310829 [Part II], http://www.mobilehandsetdesignline.com/howto/201500260 [Part III], http://www.mobilehandsetdesignline.com/howto/201800379 [Part IV], http://www.mobilehandsetdesignline.com/howto/201801097 [Part V], http://www.mobilehandsetdesignline.com/howto/201801617 [Part VI], http://www.mobilehandsetdesignline.com/howto/201801953 [Part VII], http://www.mobilehandsetdesignline.com/howto/201802483 [Part VIII] (August 8-27, 2007)

[10] H. Yaghoobi, Scalable OFDMA Physical Layer in IEEE 802.16 WirelessMAN, ftp://download.intel.com/technology/itj/2004/volume08issue03/art03_scalableofdma/vol8_art03.pdf (August 20, 2004)

[11] K. Li, IEEE 802.16e-2005 Air Interface Overview, http://santos.ee.ntu.edu.tw/mobile/Speech/IEEE%20802.16e%20OFDMA%20AirInterface%20Overview.pdf (June 05, 2006)

[12] B. Aboba, L. Blunk, J. Vollbrecht, J. Carlson, and H. Levkowetz, Extensible Authentication Protocol (EAP), RFC3748, http://www.ietf.org/rfc/rfc3748.txt (June, 2004)

[13] R. Housley, W. Polk, W. Ford, and D. Solo, Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile, RFC3280, http://www.ietf.org/rfc/rfc3280.txt (April, 2002)

[14] RSA Laboratories, PKCS #1 v2.1: RSA Cryptography Standard, ftp://ftp.rsasecurity.com/pub/pkcs/pkcs-1/pkcs-1v2-1.pdf (June 14, 2002)

[15] Airspan, White Paper: Mobile WiMAX Security, http://www.airspan.com/pdfs/WP_Mobile_WiMAX_Security.pdf (2007)

[16] T. Shon and W. Choi, An Analysis of Mobile WiMAX Security: Vulnerabilities and Solutions, http://www.springerlink.com/content/d03p14w7720x842l/fulltext.pdf (August 24, 2007)

[17] WiMAX Forum, A Comparative Analysis of Mobile WiMAX Deployment Alternatives in the Access Network, http://www.wimaxforum.org/technology/downloads/mobile_wimax_deployment_alternatives.pdf (May, 2007)

[18] WiMAX Forum, Mobile WiMAX – Part II: A Comparative Analysis, http://www.wimaxforum.org/news/downloads/Mobile_WiMAX_Part2_Comparative_Analysis.pdf (May, 2006)

[19] T. Smith, Samsung ships first HSDPA phone you can buy, http://www.reghardware.co.uk/2006/05/17/samsung_ships_first_hspda_phone/ (May 17, 2006)

[20] 3G.co.uk, First HSUPA Data Call with Nokia Flexi WCDMA Base Station, http://www.3g.co.uk/PR/December06/4071.htm (December 18, 2006)

[21] Nortel, HSDPA and beyond, http://faculty.stut.edu.tw/~cwycwy/951/N9490006.pdf (December 2, 2006)

[22] 3GPP, UTRA-UTRAN Long Term Evolution (LTE) and 3GPP System Architecture Evolution (SAE), http://www.3gpp.org/Highlights/LTE/LTE.htm (October 4, 2006)

[23] X. Yu, G. Chen, M. Chen, and X. Gao, Toward Beyond 3G: The FuTURE Project in China, http://ieeexplore.ieee.org/iel5/35/30130/01381877.pdf (January, 2005)

[24] Qualcomm, Flash-OFDM, Internet home page, http://www.qualcomm.com/technology/flash-ofdm/

[25] S. Mills (IEEE), Re: Status of 802.20, http://grouper.ieee.org/groups/802/mbwa/email/pdf00015.pdf (June 15, 2006)

[26] J. Thiel, Metropolitan and Regional Wireless Networking: 802.16, 802.20 and 802.22, http://www.cs.wustl.edu/~jain/cse574-06/ftp/wimax.pdf (May 9, 2006)

[27] V. Gupta, IEEE P802.21 Tutorial, http://www.ieee802.org/21/Tutorials/802%2021-IEEE-Tutorial.ppt (July 17, 2006)

[28] C. Cordeiro, K. Challapali, D. Birru, and N. Sai Shankar, IEEE 802.22: The First Worldwide Wireless Standard based on Cognitive Radios, http://www.ececs.uc.edu/~cordeicm/papers/dyspan05_802-22.pdf (November 8, 2005)

[29] IEEE 802.11 Working Group, Internet home page, http://www.ieee802.org/11/

[30] IEEE 802.11u Task Group, Internet home page, http://grouper.ieee.org/groups/802/11/Reports/tgu_update.htm

[31] S. Churchill, Intelligent Transportation Gets 802.11p, http://www.dailywireless.org/2004/07/15/intelligent-transportation-gets-80211p/ (July 15, 2004)

[32] WiMAX Forum, WiMAX Forum Network Architecture (Stage 2: Architecture Tenets, Reference Model and Reference Points) [Part 1], http://www.wimaxforum.org/technology/documents/WiMAX_End-to-End_Network_Systems_Architecture_Stage_2-3_Release_1.1.0.zip (July 11, 2007)

[33] WiMAX Forum, WiMAX Forum Network Architecture (Stage 2: Architecture Tenets, Reference Model and Reference Points) [Part 2], http://www.wimaxforum.org/technology/documents/WiMAX_End-to-End_Network_Systems_Architecture_Stage_2-3_Release_1.1.0.zip (July 11, 2007)

[34] Z. Becvar, J. Zelenka, and R. Bestak, COMPARISON OF HANDOVERS IN UMTS AND WIMAX, http://fireworks.intranet.gr/fireworks_docspublic/Fireworks_6CTUPB005a.pdf (September 26, 2006)

[35] Conniq.com, Handoff mechanism in Mobile WiMAX, http://www.conniq.com/WiMAX/handoff.htm

[36] Z. Becvar and J. Zelenka, Handovers in the Mobile WiMAX, http://fireworks.intranet.gr/fireworks_docspublic/Fireworks_6CTUPB008a.pdf (December 4, 2006)

[37] Y. Nkansah-Gyekye and J. Agbinya, Vertical Handoff Decision Algorithm for UMTS-WLAN, http://ieeexplore.ieee.org/iel5/4299640/4299641/04299686.pdf (August 27-30, 2007)

[38] The Network Simulator - NS-2, Internet home page, http://www.isi.edu/nsnam/ns/

[39] NIST Seamless and Secure Mobility -project, Internet home page, http://w3.antd.nist.gov/seamlessandsecure.shtml

[40] R. Rouil, The Network Simulator NS-2 NIST add-on - IEEE 802.16 model (MAC+PHY) - Draft 1.1, http://w3.antd.nist.gov/seamlessandsecure/files/80216/doc/wimax_module.pdf (April, 2007) [The link addresses to Draft 1.2]

[41] NIST, The Network Simulator NS-2 NIST add-on - IEEE 802.21 model (based on IEEE P802.21/D03.00) - Draft 1.0, http://w3.antd.nist.gov/seamlessandsecure/files/mobility/doc/MIH_module.pdf (January, 2007)

[42] NIST, The Network Simulator NS-2 NIST add-on - Neighbor Discovery - Draft 1.0, http://w3.antd.nist.gov/seamlessandsecure/files/mobility/doc/Neighbor_discovery.pdf (January, 2007)

[43] L. Chiariglione, Short MPEG-1 description, http://www.chiariglione.org/mpeg/standards/mpeg-1/mpeg-1.htm (June, 1996)

[44] IEEE 802.16 Task Group m (TGm), Internet home page, http://www.ieee802.org/16/tgm/

[45] ITU, International Mobile Telecommunications-2000 (IMT-2000), Internet home page, http://www.itu.int/home/imt.html

[46] S. Churchill, WiMAX 802.16m: 1Gbps, http://www.dailywireless.org/2007/02/20/wimax-80216m-100-mbps/ (February 20, 2007)

APPENDIX

A.1 Simulation Results

The following paragraphs present the adjusted values and give a short description about the influence of adjustments, if any, on the handoff latency. The parameters provided by the Mac/802_16 of NS-2 are presented first, following the parameters of the ND module and finally, some other adjusted parameters are discussed as well.

Mac/802_16

The following parameters are a part of the WiMAX module.

scan_iteration_

As the name suggests, the number of scan iterations was selected with *scan_iteration_*. The influence is easy to understand, since more iteration times equal longer durations for handoffs. The best value found out was two iteration times.

lgd_factor_

The Link Going Down -factor determines the sensitivity of detecting a failing link. If the value is set to 1, there is no detection. There seemed to be a rather significant influence for selection of *lgd_factor_*. The fastest handoff times were achieved when the value was set to 1.1. Between values 1.2-1.6 the results were rather similar, but above 1.7 the handoff times increased considerably.

scan_duration_

The duration of scan interval, in frames, defines the length of the scanning period. The value should be kept long enough to ensure successful scanning, but on the other hand, short enough to keep the elapsed time moderate. The simulations gave two alternatives for the best value, four and six frames. These values were used in the next measurements until a few parameters later the smaller value was selected.

interleaving_interval_

The interleaving interval is a time of normal operation between scanning periods in the MS. The best value for *interleaving_interval_* was six frames. Changes in the value below 20 frames kept the variations of handoff times in rather small region, but the mentioned value was the best.

t21_timeout_

This is a timeout value for MS to search for a DL-MAP message. In other words, on a certain channel a DL-MAP message have to be found within this time. This is rather similar to the *lost_dlmap_interval_*, but *t21_timeout_* relates to finding the DL-MAP, not receiving an updating message. Values between 6 and 28 ms gave the same results. However, even an increase of 1 ms from 28 ms resulted in handoff times of several seconds. The value to be used in the simulation was selected as 20 ms.

client_timeout_

The *client_timeout_* defines a timer value for detecting out of range MS. There did not seem to be a great influence on selecting the timeout, but the best results were obtained with 7 ms.

queue_length_

The *queue_length_* describes the size of the sending buffer of the MS. Since the used data rate is not very demanding, there is not a need for large buffer. With the selected data rate already size of 2 packets was enough to keep simulation undisturbed. If the buffer size is only 1 packet, there is some degeneration in the handoff times. The smallest possible value is the mentioned 1 packet, since otherwise there would not be any traffic.

frame_duration_

As the name trivially suggests, this parameters defines the size, or length, of a frame. The simulation refused to accept any other values but 4 or 5 ms. The shorter time gave better results.

lost_dlmap_interval_

This parameter is a timeout value for last reception of a DL-MAP message. If a new DL-MAP does not arrive within the interval, the MS loses synchronization with the BS. The simulations indicate that the value should be 9 ms or longer. Already with 8 ms the handoff times increase to multiple seconds and the amount of dropped packets becomes significant. The default value (0.6 s) was chosen.

lost_ulmap_interval_

This is the same as above but for the UL-MAP message. Here the critical limit was 5 ms, which still resulted in good results. With 4 ms the NS-2 reported "Segmentation fault" and shorter times meant there was no traffic. The default value (0.6 s) was again chosen for the simulation.

rng_backoff_start_

This parameter defines the initial backoff window size for ranging. The window size should be 2, 3, or 4 slots. These values gave the same results and larger ones increased the handoff times, even rather significantly. The size of 1 resulted in slightly longer times for the handoffs. The chosen value was 2 slots.

t44_timeout_

The *t44_timeout_* is a timeout value for scan requests. The results indicated that the timeout value should be at least 5 ms, or larger. The chosen value was the default, 10 ms.

The following parameters of the Mac/802_16 did not have any influence on the handoff times:

dcd_interval_

DCD is a message providing information about the physical conditions of the DL channel and the *dce_interval_* defines the broadcast interval of the DCD.

ucd_interval_

This is basically the same as the previous one, but for the UL channel.

contention_rng_retry_

Number of retries on ranging requests (contention mode).

invited_rng_retry_

Number of retries on ranging requests (invited mode).

request_retry_
Number of retries on bandwidth allocation requests.
reg_req_retry_
Number of retries on registration requests.
dsx_req_retry_
Number of retries on DSx requests.
dsx_rsp_retry_
Number of retries on DSx responses.
rng_backoff_stop_
Maximal backoff window size for ranging.
bw_backoff_start_
Initial backoff window size for bandwidth.
bw_backoff_stop_
Maximal backoff window size for bandwidth.
scan_req_retry_
Number of retries on scan requests.
max_dir_scan_time_
Maximum scanning time for each neighbor BSs.
nbr_adv_interval_
Interval between two MOB-NBR_ADV (Neighbor Advertisement) messages. There was no influence, at least when kept below 10 seconds.
t3_timeout_
*T3_timeout_* is the timeout value for receiving a ranging response message. There was no influence on the simulation results. The only exception was the zero-timeout, which increased the handoff times slightly.
t6_timeout_
Registration response timeout is defined with *t6_timeout_*. No influence, except with 1 ms the NS-2 gave "Segmentation fault".
t16_timeout_
The bandwidth request timeout is set with *t16_timeout_*. This parameter did not have an impact on the simulation results, at least while kept below 5 seconds. If the value was set to 0, the handoff times experienced a slight increase.

Agent/ND

The following four parameters are a part of the Neighbor Discovery module. [42]
minRtrAdvInterval_
The minimum interval between consecutive RAs is defined with *minRtrAdvInterval_*. The best result was achieved when the interval was set to zero.
maxRtrAdvInterval_
This is the opposite for above minimum RA interval, hence, the maximum interval. The selected value was 10 seconds, which also resulted in best results.

minDelayBetweenRA_

This parameter identifies the minimum time between two consecutive RAs by overriding the default value. The results of the simulation were rather varying. With some values the difference between the first and the second handoff was over 1 s. As long as the delay was kept below 2.96 s, the results were below the seamless handoff recommendation. The best handoff latency value was reached with 30 ms.

maxRADelay_

The *maxRADelay_* defines the maximum delay for replying to an RS. This parameter can be recommended to keep very low. The handoff times increased already with 50 ms delay. The most efficient handoff was accomplished when the delay was chosen at 5 ms, or below. The delay of 5 ms was selected to be used in the simulation.

Others

seed

Seed is used for creation of random numbers in a NS-2 simulation. Different values of seed did not have any influence on the handoff performance.

default_modulation

The AMC offers different modulation techniques for Mobile WiMAX to compensate different scenario requirements, for example lower modulation can be chosen with poor connection. However, the used NS-2 WiMAX model did not provide support for AMC and it was only possible to select the preferred modulation used in the simulation. The available modulations were BPSK(1/2), QPSK(1/2, 3/4), 16QAM(1/2, 3/4), and 64QAM(2/3, 3/4). The BPSK and QPSK alternatives resulted in greater amounts of dropped packets and longer times for the handoff. The options to be considered for this simulation were the QAM-modulations. The differences between these were minimal and the 16QAM 3/4 was chosen.

contention_size

This parameter is used for definition of the number of contention slots allocated for initial ranging and bandwidth requests in each frame. It did not have an influence on the handoff latency.

A.2 Simulation Code

```
#                       CN  0.0.0
#                       |
#                       R1  1.0.0
#                      /|\
#                     / | \
#                    /  |  \
#                   R2  R3  R4
#
#                2.0.0 3.0.0 4.0.0
#
#              MN1-------------->
#               2.0.1
#

#read arguments
set seed                             5555
Mac/802_16 set scan_iteration_       2
set use_going_down                   1

if {$use_going_down == 1} {
    Mac/802_16 set lgd_factor_        1.1
} else {
    Mac/802_16 set lgd_factor_        1.0
}

Mac/802_16 set scan_duration_         4
Mac/802_16 set interleaving_interval_ 6
Mac/802_16 set dcd_interval_         5 ;#max 10s
Mac/802_16 set ucd_interval_         1 ;#max 10s
set default_modulation               OFDM_16QAM_3_4 ;#OFDM_BPSK_1_2
set contention_size                  5 ;#for initial ranging and bw
Mac/802_16 set t21_timeout_          0.02 ;#max 10s, to replace the timer for looking at
preamble
Mac/802_16 set client_timeout_       0.007

Mac/802_16 set lost_ulmap_interval_ 0.76

#define frequency of RA at base station
Agent/ND set minRtrAdvInterval_ 0
Agent/ND set maxRtrAdvInterval_ 10
Agent/ND set router_lifetime_   250
Agent/ND set minDelayBetweenRA_ 0.03
Agent/ND set maxRADelay_        0.005

#define debug values
Agent/ND set debug_ 1
Agent/ND set send-RS 1
Agent/MIH set debug_ 1
Mac/802_16 set debug_ 0
Agent/MIHUser/IFMNGMT/MIPV6 set debug_ 1
Agent/MIHUser/IFMNGMT/MIPV6/Handover/Handover1 set debug_ 1

# Handover
Agent/MIHUser/IFMNGMT/MIPV6/Handover/Handover1 set case_ 3

#define coverage area for base station:
#default frequency 3.5e+9 hz
Phy/WirelessPhy set Pt_ 15
Phy/WirelessPhy set RXThresh_  1.215e-09           ;#500m coverage
Phy/WirelessPhy set CSThresh_ [expr 0.8 *[Phy/WirelessPhy set RXThresh_]]
```

```
# Parameter for wireless nodes
set opt(chan)           Channel/WirelessChannel    ;# channel type
set opt(prop)           Propagation/TwoRayGround   ;# radio-propagation model
set opt(netif)          Phy/WirelessPhy/OFDM       ;# network interface type
set opt(mac)            Mac/802_16                 ;# MAC type
set opt(ifq)            Queue/DropTail/PriQueue    ;# interface queue type
set opt(ll)             LL                         ;# link layer type
set opt(ant)            Antenna/OmniAntenna        ;# antenna model
set opt(ifqlen)         50                       ;# max packet in ifq
set opt(adhocRouting)   NOAH                       ;# routing protocol
set opt(nbMN)           1                          ;# number of Mobile Nodes
set opt(nbBS)           3                          ;# number of base stations
set opt(x)             3000                      ;# X dimension of the topography
set opt(y)             3000                      ;# Y dimension of the topography
set opt(mnSender)       1

#Rate at which the nodes start moving
set moveStart 1.0
set moveStop 179.0
#Speed of the mobile nodes (m/sec)
set moveSpeed 10

#destination of the MN
set X_dst 2999.0
set Y_dst 1310

#defines function for flushing and closing files
proc finish {} {
    global ns tf f0 f1 f2
        $ns flush-trace
        close $f0
        close $f1
        close $f2
        close $tf
#        exec xgraph out0.tr -geometry 800x400 &
#       exec xgraph out1.tr -geometry 800x400 &
       # exec xgraph out2.tr -geometry 800x400 &
       exit 0
}

#create the simulator
set ns [new Simulator]
$ns use-newtrace

#create the topography
set topo [new Topography]
$topo load_flatgrid $opt(x) $opt(y)

#open file for trace
set tf [open out.res w]
set namtrace [open out.nam w]
$ns trace-all $tf
$ns namtrace-all-wireless $namtrace $opt(x) $opt(y)

#**********
set f0 [open out0.tr w] ;#for the 1st BS
set f1 [open out1.tr w] ;#for the 2nd BS
set f2 [open out2.tr w] ;#for the 3rd BS
#**********

# set up for hierarchical routing (needed for routing over a basestation)
$ns node-config -addressType hierarchical
AddrParams set domain_num_  9                      ;# domain number
```

```
AddrParams set cluster_num_ {1 1 1 1 1 1 1 1 1}    ;# cluster number for each domain
lappend tmp 1                                      ;# router CN
lappend tmp 1                                      ;# router 1
lappend tmp 3                                      ;# 802.16 MNs+BS
lappend tmp 3                                      ;# 802.16 MNs+BS
lappend tmp 3                                      ;# 802.16 MNs+BS
lappend tmp 1                                      ;# MultifaceNode
lappend tmp 1                                      ;# MultifaceNode
lappend tmp 1                                      ;# MultifaceNode
lappend tmp 1                                      ;# MultifaceNode
AddrParams set nodes_num_ $tmp

# Node address for router0 and router1 are 4 and 5, respectively.
set CN [$ns node 0.0.0]
$CN install-default-ifmanager
set router1 [$ns node 1.0.0]

# connect links
$ns duplex-link $router1 $CN 100MBit 30ms DropTail 1000

# Create God
create-god [expr ($opt(nbMN) + $opt(nbBS))]         ;

# Create multi-interface node

$ns node-config  -multiIf ON
for {set i 0} {$i < $opt(nbBS)} {incr i} {
    set multiFaceNode($i) [$ns node [expr 5+$i].0.0]
}

set multiFaceNode_wl [$ns node [expr 5+$opt(nbBS)].0.0]

$ns node-config  -multiIf OFF

#creates the Access Point (Base station)
$ns node-config -adhocRouting $opt(adhocRouting) \
                 -llType $opt(ll) \
                 -macType $opt(mac) \
                 -ifqType $opt(ifq) \
                 -ifqLen $opt(ifqlen) \
                 -antType $opt(ant) \
                 -propType $opt(prop)    \
                 -phyType $opt(netif) \
                 -channel [new $opt(chan)] \
                 -topoInstance $topo \
                 -wiredRouting ON \
                 -agentTrace ON \
                 -routerTrace ON \
                 -macTrace ON  \
                 -movementTrace OFF

# configure each Base station 802.16
for {set i 0} {$i < $opt(nbBS) } {incr i} {
    Mac/802_16 set debug_ 1
    set bstation($i) [$ns node [expr 2+$i].0.0]
    $bstation($i) random-motion 0

#    if {$i == 1} {
#    $bstation($i) set X_ 1091.515
#    $bstation($i) set Y_ 600
#}

#if {$i == 2} {
#    $bstation($i) set X_ 1250
```

```
#    $bstation($i) set Y_ 1200
#}

#if {$i == 0} {
#    $bstation($i) set X_ 1250
#    $bstation($i) set Y_ 800
#}

    $bstation($i) set X_ [expr 500 + $i*750]
    $bstation($i) set Y_ 1000.0

    $bstation($i) set Z_ 0.0

    set clas($i) [new SDUClassifier/Dest]
    [$bstation($i) set mac_(0)] add-classifier $clas($i)
    #set the scheduler for the node. Must be changed to -shed [new $opt(sched)]
    set bs_sched($i) [new WimaxScheduler/BS]
    $bs_sched($i) set-default-modulation $default_modulation
    $bs_sched($i) set-contention-size 5
    [$bstation($i) set mac_(0)] set-scheduler $bs_sched($i)
    [$bstation($i) set mac_(0)] set-channel [expr $i*2]

    #add MOB_SCN handler
    set wimaxctrl($i) [new Agent/WimaxCtrl]
    $wimaxctrl($i) set-mac [$bstation($i) set mac_(0)]
    $ns attach-agent $bstation($i) $wimaxctrl($i)

    puts "Bstation: tcl=$bstation($i); id=[$bstation($i) id]; addr=[$bstation($i) node-addr]
X=[expr 500.0 + $i*750.0] Y=1000.0"
}

$wimaxctrl(0) add-neighbor [$bstation(1) set mac_(0)] $bstation(1)
$wimaxctrl(1) add-neighbor [$bstation(2) set mac_(0)] $bstation(2)
$wimaxctrl(1) add-neighbor [$bstation(0) set mac_(0)] $bstation(0)
$wimaxctrl(2) add-neighbor [$bstation(1) set mac_(0)] $bstation(1)

# creation of the mobile nodes
$ns node-config -wiredRouting OFF \
                -macTrace ON                                                     ;# Mobile
nodes cannot do routing.

for {set i 0} {$i < $opt(nbBS)} {incr i} {
    #create 1 node in each cell to init cells
    Mac/802_16 set debug_ 1
    set m_node_($i) [$ns node [expr 2+$i].0.1]
    $m_node_($i) random-motion 0                                    ;# disable random
motion
    $m_node_($i) base-station [AddrParams addr2id [$bstation($i) node-addr]] ;#attach mn to
basestation

#    if {$i == 1} {
#    $m_node_($i) set X_ 1091.515
#    $m_node_($i) set Y_ 600
#    }

#    if {$i == 2} {
#    $m_node_($i) set X_ 1250
#    $m_node_($i) set Y_ 1200
#    }

#    if {$i == 0} {
#    $m_node_($i) set X_ 1250
#    $m_node_($i) set Y_ 800
#    }
```

```
    $m_node_($i) set X_ [expr 500.0 + $i*750.0]
    $m_node_($i) set Y_ 1000.0
    $m_node_($i) set Z_ 0.0

    set clas_mn($i) [new SDUClassifier/Dest]
    [$m_node_($i) set mac_(0)] add-classifier $clas_mn($i)
    #set the scheduler for the node. Must be changed to -shed [new $opt(sched)]
    set ss_sched($i) [new WimaxScheduler/SS]
    [$m_node_($i) set mac_(0)] set-scheduler $ss_sched($i)
    [$m_node_($i) set mac_(0)] set-channel [expr $i*2]

    #add the interfaces to supernode
    $multiFaceNode($i) add-interface-node $m_node_($i)

    puts "InitNode: tcl=$m_node_($i); id=[$m_node_($i) id]; addr=[$m_node_($i) node-addr]
X=[expr 500.0 + $i*750.0] Y=1000.0"
}

#configure the MOBILE NODE
Mac/802_16 set debug_ 1
set wl_node  [$ns node 2.0.2]                                               ;# create the node
with given @.
$wl_node random-motion 0                                            ;# disable random motion
$wl_node base-station [AddrParams addr2id [$bstation(0) node-addr]]   ;# attach mn to
basestation
$wl_node set X_ 500.0
$wl_node set Y_ 1310
$wl_node set Z_ 0.0
set clas_wl [new SDUClassifier/Dest]
[$wl_node set mac_(0)] add-classifier $clas_wl
#set the scheduler for the node. Must be changed to -shed [new $opt(sched)]
set ss_sched_wl [new WimaxScheduler/SS]
[$wl_node set mac_(0)] set-scheduler $ss_sched_wl
[$wl_node set mac_(0)] set-channel 0
$multiFaceNode_wl  add-interface-node $wl_node
puts "Mobile Node: tcl=$wl_node; id=[$wl_node id]; addr=[$wl_node node-addr] X=500.0 Y=1000.0"

# add link to backbone
for {set i 0} {$i < $opt(nbBS) } {incr i} {
    # add link to backbone
    $ns duplex-link $bstation($i) $router1 100MBit 15ms DropTail 1000
}

# configure each Base station 802.16
for {set i 0} {$i < $opt(nbBS) } {incr i} {
    set nd_bs($i) [$bstation($i) install-nd]
    $nd_bs($i) set-router TRUE
    $nd_bs($i) router-lifetime 250
    $ns at 1 "$nd_bs($i) start-ra"
    set mih_bs($i) [$bstation($i) install-mih]
    set tmp2($i) [$bstation($i) set mac_(0)] ;#in 802.16 one interface is created
    $tmp2($i) mih $mih_bs($i)
    $mih_bs($i) add-mac $tmp2($i)
}

# configure each MN 802.16
for {set i 0} {$i < $opt(nbBS)} {incr i} {
    set nd_mn($i) [$m_node_($i) install-nd]
    set handover($i) [new Agent/MIHUser/IFMNGMT/MIPV6/Handover/Handover1]
    $nd_mn($i) set-ifmanager $handover($i)
    $multiFaceNode($i) install-ifmanager $handover($i)

    # install MIH in multi-interface node
```

```
    set mih($i) [$multiFaceNode($i) install-mih]
    set nd_mn($i) [$m_node_($i) install-nd]
    $handover($i) connect-mih $mih($i)
    set tmp3($i) [$m_node_($i) set mac_(0)]
    $handover($i) nd_mac $nd_mn($i) $tmp3($i)
    $tmp3($i) mih $mih($i)
    $mih($i) add-mac $tmp3($i)
}

set nd_mn_wl [$wl_node install-nd]
set handover_wl [new Agent/MIHUser/IFMNGMT/MIPV6/Handover/Handover1]
$multiFaceNode_wl install-ifmanager $handover_wl
$nd_mn_wl set-ifmanager $handover_wl

#**********************
proc record {} {
    global sink0 sink1 sink2 f0 f1 f2
    set ns [Simulator instance]
    #Set the time after which the procedure should be called again
    set time 0.001
    #How many bytes have been received by the traffic sinks?
    set bw0 [$sink0 set npkts_]
    set bw1 [$sink1 set bytes_]
#    set bw2 [$sink2 set bytes_]
    #Get the current time
    set now [$ns now]
    #Calculate the bandwidth (in Mbit/s) and write it to the files
    puts $f0 "$now [expr $bw0]"   ;# /$time*8/1000000]"
    puts $f1 "$now [expr $bw1]"   ;# /$time*8/1000000]"
#    puts $f2 "$now [expr $bw2]"   ;# /$time*8/1000000]"
    #Reset the bytes_ values on the traffic sinks
#    $sink0 set bytes_ 0
    $sink1 set bytes_ 0
#    $sink2 set bytes_ 0
    #Re-schedule the procedure
    $ns at [expr $now+$time] "record"
}
#**************************

# install MIH in multi-interface node
set mih_wl [$multiFaceNode_wl install-mih]
$handover_wl connect-mih $mih_wl
set tmp_wl [$wl_node set mac_(0)]
$handover_wl nd_mac $nd_mn_wl $tmp_wl
$tmp_wl mih $mih_wl
$mih_wl add-mac $tmp_wl

#Create a UDP agent and attach it to node n0
set udp_ [new Agent/UDP]
$udp_ set packetSize_ 1500

set quiet 0

if {$quiet == 0} {
    puts "udp on node : $udp_"
}

# Create a CBR traffic source and attach it to udp0
set cbr_ [new Application/Traffic/CBR]
$cbr_ set packetSize_ 1500
$cbr_ set interval_ 0.01
$cbr_ attach-agent $udp_

# Create the Null agent to sink traffic
```

```
set null_ [new Agent/Null]

if { $opt(mnSender) == 1} {
    #CN is receiver
    $ns attach-agent $CN $null_

    #Multiface node is transmitter
    $multiFaceNode_wl attach-agent $udp_ $wl_node
    $handover_wl add-flow $udp_ $null_ $wl_node 1 2000.
} else {
    #Multiface node is receiver
    $multiFaceNode_wl attach-agent $null_ $wl_node
    $handover_wl add-flow $null_ $udp_ $wl_node 1 2000.

    #CN is transmitter
    $ns attach-agent $CN $udp_
}

#******************
#Create three traffic sinks and attach them to the nodes
set sink0 [new Agent/LossMonitor]
set sink1 [new Agent/LossMonitor]
#set sink2 [new Agent/LossMonitor]

$ns attach-agent $bstation(0) $sink0
$ns attach-agent $bstation(1) $sink1
#$ns attach-agent $bstation(2) $sink2

$ns connect $sink0 $udp_
$ns connect $sink1 $udp_
#$ns connect $sink2 $udp_
#******************

#Start the recording of the received bandwidth
$ns at 0.0 "record"

#Start the application 1sec before the MN is entering the WiMAX cell
$ns at [expr $moveStart - 1] "$cbr_ start"

#Stop the application according to another poisson distribution (note that we don't leave the
802.16 cell)
$ns at [expr $moveStop  + 1] "$cbr_ stop"

#calculate the speed of the node
$ns at $moveStart "$wl_node setdest $X_dst $Y_dst $moveSpeed"
$ns at $moveStop "finish"
puts "Running simulation for $opt(nbMN) mobile nodes..."
$ns run
puts "Simulation done."
```