WiMAX™ System Evaluation Methodology

Version 2.1

July 7, 2008

Acknowledgments

The material presented in this document represents the combined efforts of many people from several WiMAX Forum member organizations. WiMAX Forum member organizations that have made a substantial contribution to the material presented in this document are:

- Alcatel-Lucent
- Alvarion
- Arraycom
- AT&T
- BAE Systems
- Beceem
- Clearwire
- Comsys
- Intel
- Kozo Keikaku Engineering, Inc.
- Motorola
- POSDATA
- Rensselaer Polytechnic Institute (RPI)
- Samsung
- Siemens
- Sprint
- Telsima
- Venturi Wireless
- Washington University in Saint Louis

The following individuals from these organizations have provided significant contributions to this document:

- Ali Koc/Intel
- Arun Ghosh/AT&T
- Arvind Raghavan/Arraycom
- Bong Ho Kim/POSDATA
- Honghai Zhang/NEC
- Hyunjeong hannah Lee/Intel
- Jaeyoung Kim/Samsung
- Jeff Andrews/University of Texas
- Jie Hui/Intel (Editor for the MAC Section)
- John Kim/Sprint
- Jungnam Yun/POSDATA
- Krishna Kamal Sayana/Motorola
- Krishna Ramadas/Venturi Wireless
- Maruti Gupta/Intel
- Louay Jalloul/Beceem
- Maruti Gupta/Intel
- Mineo Takai/Kozo Keikaku Engineering, Inc.

- Muthaiah Venkatachalam/Intel
- Nat Natarajan/Motorola
- Pete Gelbman/Clearwire
- Raj Iyengar/RPI
- Rok Preseren/Telsima
- Roshni Srinivasan/Intel
- Sherry Chen/Intel
- Shyam Parekh/Alcatel-Lucent
- Teck Hu/Siemens
- Tom Tofigh/AT&T
- Vafa Ghazi/CoWare
- Xiangying Yang/Intel
- Yaron Alpert/Comsys
- Zee'v Roth/Alvarion

The comments and feedback should be sent to Krishna Ramadas, Kramadas@venturiwireless.com and Raj Jain, Jain@wustl.edu

Table of Contents

List of Figures

List of Tables

List of Acronyms

3GPP	3G Partnership Project
3GPP2	3G Partnership Project 2
AAS	Adaptive Antenna System also Advanced Antenna System
ACK	Acknowledge
AES	Advanced Encryption Standard
AG	Absolute Grant
AMC	Adaptive Modulation and Coding
A-MIMO	Adaptive Multiple Input Multiple Output (Antenna)
ASM	Adaptive MIMO Switching
ARQ	Automatic Repeat reQuest
ASN	Access Service Network
ASP	Application Service Provider
BE	Best Effort
CC	Chase Combining (also Convolutional Code)
CCI	Co-Channel Interference
CCM	Counter with Cipher-block chaining Message authentication code
CDF	Cumulative Distribution Function
CINR	Carrier to Interference + Noise Ratio
CMAC	block Cipher-based Message Authentication Code
CP	Cyclic Prefix
CQI	Channel Quality Indicator
CSN	Connectivity Service Network
CSTD	Cyclic Shift Transmit Diversity
CTC	Convolutional Turbo Code
DL	Downlink
DOCSIS	Data Over Cable Service Interface Specification
DSL	Digital Subscriber Line

DVB	Digital Video Broadcast
EAP	Extensible Authentication Protocol
EESM	Exponential Effective SIR Mapping
EIRP	Effective Isotropic Radiated Power
ErtVR	Extended Real-Time Variable Rate
FBSS	Fast Base Station Switch
FCH	Frame Control Header
FDD	Frequency Division Duplex
FFT	Fast Fourier Transform
FTP	File Transfer Protocol
FUSC	Fully Used Sub-Channel
HARQ	Hybrid Automatic Repeat reQuest
HHO	Hard Hand Over
HMAC	keyed Hash Message Authentication Code
HO	Hand Over
HTTP	Hyper Text Transfer Protocol
IE	Information Element
IEFT	Internet Engineering Task Force
IFFT	Inverse Fast Fourier Transform
IR	Incremental Redundancy
ISI	Inter-Symbol Interference
LDPC	Low-Density-Parity-Check
LOS	Line of Sight
MAC	Media Access Control
MAI	Multiple Access Interference
MAN	Metropolitan Area Network
MAP	Media Access Protocol
MBS	Multicast and Broadcast Service
MDHO	Macro Diversity Hand Over

MIMO	Multiple Input Multiple Output (Antenna)
MMS	Multimedia Message Service
MPLS	Multi-Protocol Label Switching
MS	Mobile Station (same as SS)
MSO	Multi-Services Operator
NACK	Not Acknowledge
NAP	Network Access Provider
NLOS	Non Line-of-Sight
NRM	Network Reference Model
nrtPS	Non-Real-Time Polling Service
NSP	Network Service Provider
OFDM	Orthogonal Frequency Division Multiplex
OFDMA	Orthogonal Frequency Division Multiple Access
PER	Packet Error Rate
PF	Proportional Fair (Scheduler)
PKM	Public Key Management
PUSC	Partially Used Sub-Channel
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RG	Relative Grant
RR	Round Robin (Scheduler)
RRI	Reverse Rate Indicator
RTG	Receive/transmit Transition Gap
rtPS	Real-Time Polling Service
RUIM	Removable User Identify Module
SCM	Spatial Channel Model [3GPP]
SDMA	Space (or Spatial) Division (or Diversity) Multiple Access
SF	Spreading Factor
SFN	Single Frequency Network

SGSN	Serving GPRS Support Node
SHO	Soft Hand Over
SIM	Subscriber Identify Module
SINR	Signal to Interference + Noise Ratio
SISO	Single Input Single Output (Antenna)
SLA	Service Level Agreement
SM	Spatial Multiplexing
SMS	Short Message Service
SNIR	Signal to Noise + Interference Ratio
SNR	Signal to Noise Ratio
S-OFDMA	Scalable Orthogonal Frequency Division Multiple Access
SS	Subscriber Station (same as MS)
STC	Space Time Coding
TDD	Time Division Duplex
TEK	Traffic Encryption Key
TTG	Transmit/receive Transition Gap
TTI	Transmission Time Interval
TU	Typical Urban (as in channel model)
UE	User Equipment
UGS	Unsolicited Grant Service
UL	Uplink
UMTS	Universal Mobile Telephone System
USIM	Universal Subscriber Identify Module
VoIP	Voice over Internet Protocol
VPN	Virtual Private Network
VSF	Variable Spreading Factor
WiFi	Wireless Fidelity
WAP	Wireless Application Protocol
WiBro	Wireless Broadband (Service)

Copyright Notice, Use Restrictions, Disclaimer, and Limitation of Liability.

1. Introduction

This document captures important aspects of system simulation methodology for a WiMAX™ network. Many simplifications are discussed in modelling the various aspects of the end-to-end system. Simplified models are necessary to reduce the computational complexity and achieve shorter run times. However, simplifications sacrifice some precision. The WiMAX vendor community has extensive experience with implementing both system and link level simulations and have contributed their ideas about effective models. In many cases alternate models have been received as contributions from different vendors. The document captures all contributions. One of the contributions is chosen as the primary method both for the purpose of discussion and also as a default so that the results from different models can be compared. Alternate methods are all documented as a reference so that the v so that vendors can chose an alternate method when required.

The document is structured similar to system evaluation methodology documents from any other standard's body such as 3GPP or 3GPP2. Structural similarity should help experts on other technology to easily apply their knowledge to develop/modify WiMAX system simulation modules. The introduction section describes the basics of a WiMAX system. The introduction section also describes the system profiles elements that have been defined by the WiMAX Forum. Knowledge of system profiles is important to understand the need for related models and also to understand profile elements that are not included in the evaluation methodology. Version 1 of this document does not discuss, for example, any network and control plane aspects of the WiMAX system. Models for the network layers will be incorporated in a future revision of this document. The simplification is driven from a need to publish a version of the document that addresses the system methodology aspects that are key to comparable wireless technologies such as 3GPP and 3GPP2.

The simulation methodology described in this document is general so that it can be used with any modelling platform such as NS2, OPNET, OMNET, Qualnet, etc. A complete list of known WiMAX modelling platforms is provided in Annex J. WiMAX Forum is developing an NS-2 Model of WiMAX using this methodology for general use by all WiMAX members. The selection of NS-2 as the modelling platform is based not on its technical superiority but on the fact that it is a public domain platform available under GNU Public License (GPL) and can be freely distributed. We encourage members to use this methodology with other platforms as well.

Section 2 highlights the "system level" approach. The basic difference between link layer simulation and system simulation are listed in this section. The standard approach to modelling the end user application behaviour includes the use of traffic models, protocol models and a more detailed MAC layer. Physical layer is abstracted as much as possible to retain the most important consequences of physical layer impairments on the probability of getting MAC PDUs across a WiMAX link. The recommended methodology is based on "center cell" approach, where impact to user traffic within the center cell is closely observed in relation to the impact of traffic on surrounding cells. This methodology document addresses these topics in a manner similar to how other technologies such as 3GPP/3GPP2 address them.

Section 2 also provides a list of recommended values for important configuration variables. There are three sources for these recommended values. First, the values specified in the profile documents were used to set ranges. Second, the values used in recent requests for proposals by leading service providers were discussed to be used as default. Third, the members decided that this document should be such that it can be easily be adopted for upcoming waves and the next generation of WiMAX.

IEEE has already started 802.16m study group for the next generation. This third requirement lead us to not limit the options to strictly follow the current profiles but be a superset of the current WiMAX profile. The default values are shown in **bold** and are compliant with current WiMAX profiles. Section 2 also establishes the output metrics from a standard system simulation environment.

Section 3 captures the application traffic models. Most of these applications and traffic models are new and have been developed by WiMAX Forum® members. The 3GPP and 3GPP2 documents describe models for Web, FTP, and near real time video applications. Since the user behaviour has changed over the last few years, these models may no longer be applicable. However, these models are useful in comparing WiMAX technology to existing technologies and so we have included them in this document along with newer versions of these applications.

Section 4 covers the MAC layer models and contains many key areas that differentiate WiMAX implementations from other wireless implementations.

Section 5 addresses PHY modem abstraction for system simulation. It also covers channel models and interference models for system simulation.

Finally, a set of Appendices addresses the following topics essential to system simulation: channel models, PHY abstraction, and modelling PUSC.

Annex H covers NS2 protocol layer and common framework modules.

Annex I presents a list of known WiMAX simulation models.

1.1 OFDMA Basics

Orthogonal Frequency Division Multiplexing (OFDM) is a multiplexing technique that subdivides the bandwidth into multiple frequency sub-carriers as shown in Figure 1.1.1. In an OFDM system, the input data stream is divided into several parallel sub-streams of reduced data rate (thus increased symbol duration) and each sub-stream is modulated and transmitted on a separate orthogonal sub-carrier. The increased symbol duration improves the robustness of OFDM to delay spread. Furthermore, the introduction of the cyclic prefix (CP) can completely eliminate Inter-Symbol Interference (ISI) as long as the CP duration is longer than the channel delay spread. The CP is typically a repetition of the last samples of data portion of the block that is appended to the beginning of the data payload. The CP prevents inter-block interference and makes the channel appear circular and permits low-complexity frequency domain equalization. A perceived drawback of CP is that it introduces overhead, which effectively reduces bandwidth efficiency. While the CP does reduce bandwidth efficiency somewhat, the impact of the CP is similar to the “roll-off factor” in raised-cosine filtered single-carrier systems. Since OFDM has a very sharp, almost “brick-wall” spectrum, a large fraction of the allocated channel bandwidth can be utilized for data transmission, which helps to moderate the loss in efficiency due to the cyclic prefix.

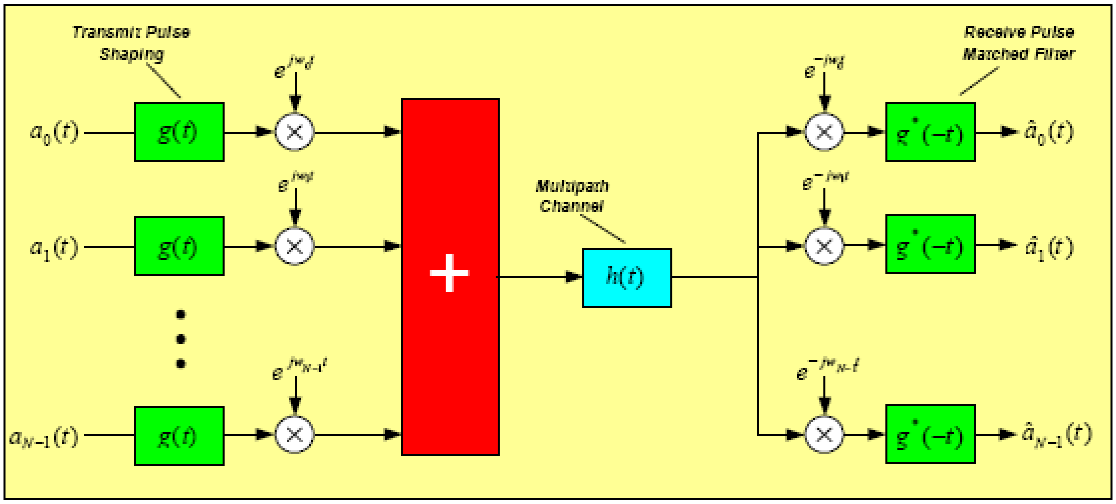


Figure 1.1.1: Basic Architecture of OFDMA system

OFDM exploits the frequency diversity of the multipath channel by coding and interleaving the information across the sub-carriers prior to transmissions. OFDM modulation can be realized with efficient Inverse Fast Fourier Transform (IFFT), which enables a large number of sub-carriers (up to 2048) with low complexity. In an OFDM system, resources are available in the time domain by means of OFDM symbols and in the frequency domain by means of sub-carriers. The time and frequency resources can be organized into sub-channels for allocation to individual users. Orthogonal Frequency Division Multiple Access (OFDMA) is a multiple-access/multiplexing scheme that provides multiplexing operation of data streams from multiple users onto the downlink sub-channels and uplink multiple access by means of uplink sub-channels.

References:
http://www.wimaxforum.org/news/downloads/Mobile_WiMAX_Part1_Overview_and_Performance.pdf

1.2 Scalable OFDMA

A scalable physical layer enables standard-based solutions to deliver optimum performance in channel bandwidths ranging from 1.25 MHz to 20 MHz with fixed sub-carrier spacing for both fixed and portable/mobile usage models, while keeping the product cost low. The architecture is based on a scalable subchannelization structure with variable Fast Fourier Transform (FFT) sizes according to the channel bandwidth. In addition to variable FFT sizes, the specification supports other features such as Adaptive Modulation and Coding (AMC) subchannels, Hybrid Automatic Repeat Request (HARQ), high-efficiency uplink subchannel structures, Multiple-Input-Multiple-Output (MIMO) diversity, and coverage enhancing safety channels, as well as other OFDMA default features such as different subcarrier allocations and diversity schemes. Coherence time, Doppler shift, and coherence bandwidth of the channel forms the basis for the consideration of a scalable structure where the FFT sizes scale with bandwidth to keep the subcarrier spacing fixed. The following table shows the scalability range proposed in the corresponding 802.16 standard.

The standard recommends the following sampling rates: For channel bandwidths that are a multiple of 1.75 MHz, n=8/7; for channel bandwidths that are a multiple of 1.25, 1.5, 2, or 2.75 MHz, n=28/25; for all other channel bandwidths, n=8/7.

Table 1.2.1: OFDMA scalability parameters

Parameters	Values						
System bandwidth (MHz)[1]	1.25	5	10	20	3.5	7	8.75
Sampling factor	28/25				8/7		
Sampling frequency (F_s, MHz)	1.4	5.6	11.2	22.4	4	8	10
Sample time ($1/F_s$, nsec)	714.3	178.6	89.3	44.6	250	125	100
FFT size (N_{FFT})	128	512	1024	2048	512	1024	1024
Subcarrier frequency spacing (Δf, kHz)	10.9375				7.8125		9.765625
Useful symbol time ($T_b=1/\Delta f$, μs)	91.4				128		102.4
Guard time ($T_g = T_b/8$, μs)[2]	11.4				16		12.8
OFDMA symbol time ($T_s=T_b+T_g$, μs)	102.8				144		115.2

Notes:

1) The Mobile WiMAX™ system profile does not include 1.25MHz and 20MHz.

2) Other possible cyclic prefix ratios are 1/4, 1/16, 1/32, however, 1/8 is the only mandatory value in the Mobile WiMAX system profile.

Note that the channel bandwidth and subcarrier spacing are related as follows:

Subcarrier spacing*FFT size = Channel Bandwidth * Sampling rate

For 10 MHz channel, with 28/25 sampling rate, and 1024 FFT:

Subcarrier spacing = 10*(28/25)/1024 = 10.9375 kHz

Table 1.2.2 shows the resulting frame sizes and frame durations for the scalable FFT sizes. Although the table lists multiple frame sizes, Mobile WiMAX system profile allows only 5ms frame size. The division of these symbols between DL and UL are shown in Table 1.2.3 for 10MHz (Default) and two other channel bandwidths.

Table 1.2.2: Scalable OFDMA frame sizes for 10 MHz

Frame Sizes (msec)	Frame Sizes (OFDM symbols)
2	19
2.5	24
4	39
5	**47**
8	79
10	99
12.5	124
20	198
WiMAX MTG Profiles support a frame size of 5ms only. 47 symbols allow for 1.6 symbol times for TTG+RTG. The frame size includes 1 symbol for preamble.	

Table 1.2.3: Number of OFDM symbols in DL and UL (5ms Frame)

Item	**Channel Bandwidth**	**(DL,UL) Symbols**
1	5 and 10 MHz	(47-n, n), $12 \leq n \leq 21$
2	8.75 MHz	(42-n, n), $12 \leq n \leq 18$
3	7 and 3.5 MHz	(33-n, n), $9 \leq n \leq 15$

1.3 OFDMA sub-carriers and sub-channels

There are three types of OFDMA subcarriers:

1. Data subcarriers for data transmission.

2. Pilot subcarriers for various estimation and synchronization purposes.

3. Null subcarriers for no transmission at all, used for guard bands and DC carriers.

Active subcarriers are divided into subsets of subcarriers called subchannels. The subcarriers forming one subchannel may be, but need not be, adjacent. Subchannels are the smallest granularity for resource allocation as indicated via MAP.

The pilot allocation is performed differently in different subcarrier allocation modes. For DL Fully Used Subchannelization (FUSC), the pilot tones are allocated first and then the remaining subcarriers are divided into data subchannels. For DL Partially Used Subchannelization (PUSC) and all UL modes, the set of used subcarriers, that is, data and pilots, is first partitioned into subchannels, and then the pilot subcarriers are allocated from within each subchannel. In FUSC, there is one set of common pilot subcarriers, but in PUSC, each subchannel contains its own set of pilot subcarriers.

In a DL, subchannels may be intended for different (groups of) receivers while in UL, Subscriber Stations (SS) may be assigned one or more subchannels and several transmitters may transmit simultaneously. There are two main types of subcarrier permutations: distributed and adjacent. In general, distributed subcarrier permutations perform very well in mobile applications while adjacent subcarrier permutations can be properly used for fixed, portable, or low mobility environments. These options enable the system designers to trade mobility for throughput. The subcarriers forming one subchannel may, but need not be, adjacent.

Figure 1.3.1 illustrates one sample OFDM frame structure for a Time Division Duplex (TDD) implementation (Section 4.7 provides several other possible frame structures with other variations). Each frame is divided into DL and UL sub-frames separated by Transmit/Receive and Receive/Transmit Transition Gaps (TTG and RTG, respectively) to prevent DL and UL transmission collisions. Each DL subframe starts with a preamble followed by the Frame Control Header (FCH), the DL-MAP, and a UL-MAP, respectively. In a frame, the following control information is used to ensure optimal system operation:

- **Preamble:** The preamble, used for synchronization, is the first OFDM symbol of the frame.
- **Frame Control Header (FCH):** The FCH follows the preamble. It provides the frame configuration information such as MAP message length and coding scheme and usable sub-channels.
- **DL-MAP and UL-MAP:** The DL-MAP and UL-MAP provide sub-channel allocation and other control information for the DL and UL sub-frames respectively.
- **UL Ranging:** The UL ranging sub-channel is allocated for mobile stations (MS) to perform closed-loop time, frequency, and power adjustment as well as bandwidth request
- **UL CQICH:** The UL CQICH channel is allocated for the SSto feedback channel state information.
- **UL ACK:** The UL ACK is allocated for the SSto feedback DL HARQ acknowledge.

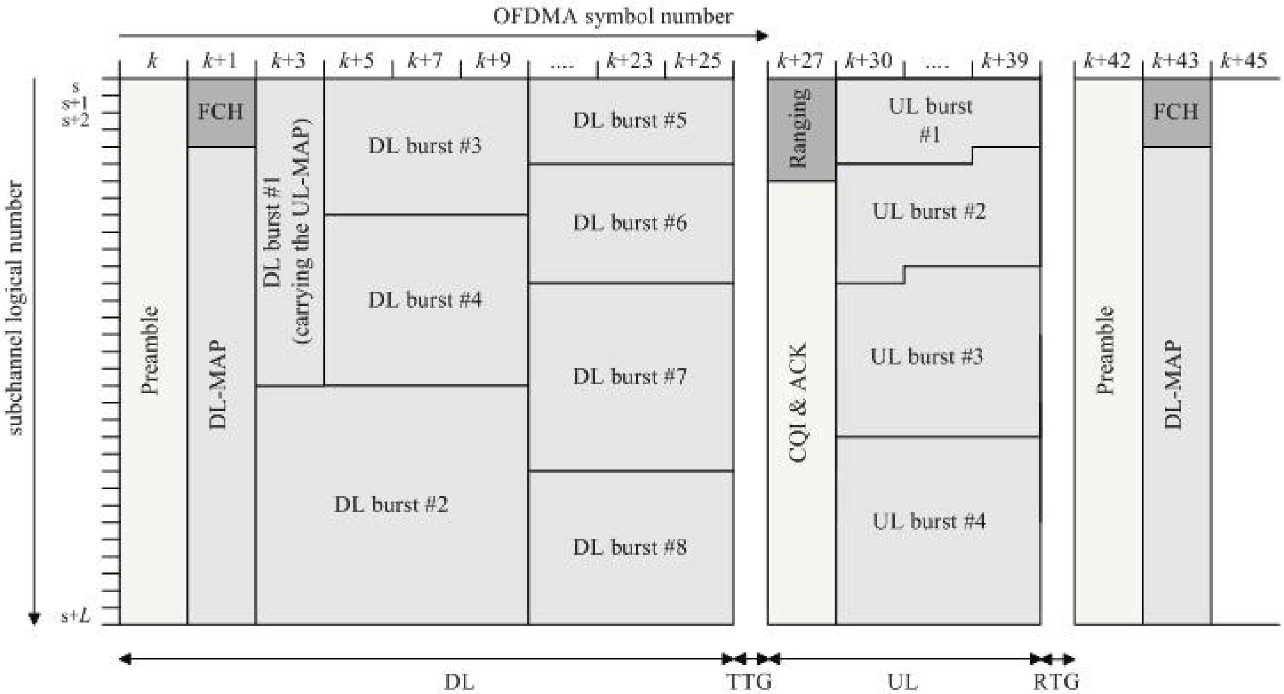


Figure 1.3.1: A Sample OFDMA Frame Structure

1.4 WiMAX Forum Profiles

The WiMAX Forum has specified a set of mobility system profiles for IEEE 802.16e-based systems as shown in Table 1.4.1. The profiles provide a direction for what to emphasize in the simulation methodology work in a phased manner, starting with the most basic features and progressively adding optional and/or advanced features. Thus the scope of system simulation work is bounded by the profile configuration set as recommended by the WiMAX Forum. Throughout the document a range of features and parameter values are specified. These ranges are generally a superset of what is specified in current WiMAX Forum profiles. For each parameter a default value is also specified that conforms to the current WiMAX Forum profile if applicable. The default values are specified in bold.

Table 1.4.1: WiMAX Forum Profiles

Item	Capability
1.	Prof1.A_2.3 – 8.75 MHz channel PHY (2.3-2.4 GHz)
2.	Prof1.B_2.3 – 5 AND 10 MHz channel PHY (2.3-2.4 GHz)
3.	Prof2.A_2.305 – 3.5 MHz channel PHY (2.305-2.320, 2.345-2.360 GHz)
4.	Prof2.B_2.305 – 5 MHz channel PHY (2.305-2.320, 2.345-2.360 GHz)
5.	Prof2.C_2.305 – 10 MHz channel PHY (2.305-2.320, 2.345-2.360 GHz)
6.	Prof3.A_2.496 – 5 AND 10 MHz channel PHY (2.496-2.69GHz)

7.	Prof4.A_3.3 – 5 MHz channel PHY (3.3-3.4 GHz)
8.	Prof4.B_3.3 – 7 MHz channel PHY (3.3-3.4 GHz)
9.	Prof4.C_3.3 – 10 MHz channel PHY (3.3-3.4 GHz)
10.	Prof5.A_3.4 – 5 MHz channel PHY (3.4-3.8 GHz)
	Prof5L.A_3.4 – 5 MHz channel PHY (3.4-3.6 GHz)
	Prof5H.A_3.4 – 5 MHz channel PHY (3.6-3.8 GHz)
11.	Prof5.B_3.4 – 7 MHz channel PHY (3.4-3.8 GHz)
	Prof5L.B_3.4 –7MHz channel PHY (3.4-3.6 GHz)
	Prof5H.B_3.4 – 7 MHz channel PHY (3.6-3.8 GHz)
12.	Prof5.C_3.4 – 10 MHz channel PHY (3.4-3.8 GHz)
	Prof5L.C_3.4 –10 MHz channel PHY (3.4-3.6 GHz)
	Prof5H.C_3.4 – 10 MHz channel PHY (3.6-3.8 GHz)

2. System Simulation Modelling

This section describes the procedure for simulating a multi-cell Mobile WiMAX network. It is intended to be used with an end-to-end network simulation with explicit models for application traffic, transport, network, and MAC protocols. As such, the downlink (DL) modelling method must be similar in philosophy and complexity to the uplink (UL) modelling method so that the performance of two-way traffic is analysed in a consistent manner.

2.1 System simulation procedure for center cell approach

This section describes the methodology for system level evaluation of IEEE 802.16e systems. It is based partly on the methodology developed by 3GPP study groups. It is applicable to scenarios with low/medium subscriber velocities such that a subscriber location does not change over several frames. Mobility/handoff is not supported. However, Doppler effects are simulated to model CQI aging, PHY impairments like channel estimation etc. It is assumed that perfect time and frequency synchronization is available, and channel estimation is ideal. Performance statistics are collected only for users associated with the BS in the center cell of a network, while users in the outer cells are explicitly simulated to produce two tiers of interference. We define this procedure as the center cell approach.

The simulation models the evolution of signal received by the subscriber and interference in time, and employs a PHY abstraction to predict link layer performance. The step-by-step procedure outlined below describes the procedure for system-level performance evaluation.

2.1.1 Network topology and deployment scenario specification

As a first step, the key parameters for simulations are specified. These include network configuration parameters, BS & SS equipment model parameters, OFDMA air interface parameters, propagation model parameters, methodology parameters, and dynamic system simulation features as described below:

Table 2.1.1: Network Configuration Parameters

Parameter	Description	Value Range
N_c	Number of cells.	**19**
S	Number of sectors/cell.	1, **3**, 4, 6
$N_s = SN_c$	Total number of sectors.	19, **57**, 76, 114
R	BS-to-BS distance	0.5 to 30 km **(1 km)**
ϕ_{BS}	Orientation (boresight angle) of each sector as defined by 3GPP-3GPP2 [10]	$S = 3 : \phi_{BS} = 30,150,270$ $S = 6 : \phi_{BS} = 0,60,120,...300$

K	Number of frequency allocations in the network.	**1**, 2, 3, 4, 6
F_{BS}	Frequency allocation (integer index) used in each BS sector.	**1**, 2, 3, 4, 5, 6
	Operating Frequency	2.0–3.5 GHz (**2.5 GHz**)
	Duplexing Scheme	**TDD**

Table 2.1.2: Base Station Equipment Model Parameters

Parameter	**Description**	**Value Range**
$P1_{BS}$	BS power amplifier 1dB compression point	39-60 dBm
PAR_{BS}	Peak-to-average backoff at BS	9-11 dB
P_{BS}	Rms transmit power per sector/carrier	30-51 dBm (**43 dBm**)
H_{BS}	Base station height	10-50m (**32 m**)
G_{BS}	Gain (boresight)	**16 dBi**
θ_{BS}	3-dB beamwidth as defined by 3GPP-3GPP2 [10]	$S = 3 : \theta_{BS} = 70^0$ $S = 6 : \theta_{BS} = 35^0$
G_{FB}	Front-to-back power ratio	**25 dB**
M_{TX}	Number of transmit antennas	1,**2**,3,4
M_{RX}	Number of receive antennas	1,**2**,3,4
d_{BS}	BS antenna spacing (ref: ULA)	$\lambda / 2$, 4λ, 10λ
NF_{BS}	Noise figure (transmit & receive)	4-6 dB (**5 dB**)
HW_{BS}	Hardware loss (cable, implementation, etc.)	**2 dB**

Table 2.1.3: Subscriber Station Equipment Model Parameters

Parameter	Description	Value Range
$P1_{SS}$	SS power amplifier 1dB compression point	29-54 dBm
PAR_{SS}	Peak-to-average backoff at SS	9-11 dB
P_{SS}	RMS transmit power/per SS	20-45 dBm (**23 dBm**)
H_{SS}	Subscriber station height	1.5-7 m (**1.5 m**)
G_{SS}	Gain (boresight)	**0 dBi**
$\{\theta_{SS}\}$, $G(\{\theta_{SS}\})$	Table of Gains as a function of Angle-of-arrival	Omni
N_{TX}	Number of transmit antennas	**1**,2
N_{RX}	Number of receive antennas	1,**2**,3,4
d_{SS}	SS antenna correlation	0-0.7 (**0.5**)
	SS antenna gain mismatch	0-5 dB (**3 dB**)
NF_{SS}	Noise figure (transmit & receive)	6-7 dB (**7 dB**)
HW_{SS}	Hardware loss (cable, implementation, etc.)	**2 dB**

Table 2.1.4: OFDMA Air Interface Parameters

Parameter	Description	Value Range
OFDMA symbol parameters		
BW	Total bandwidth	5, **10**, 20 MHz (See Table 1.2.1 for other values)
N_{FFT}	Number of points in full FFT	512, **1024**, 2048 (See Table 1.2.1)
Δ_f	Subcarrier spacing	10.9375 kHz
$T_S = 1 / \Delta_f$	OFDMA symbol duration	91.43 usec
CP	Cyclic prefix length (fraction of T_S)	1/2, 1/4, **1/8**, 1/16

T_O	OFDMA symbol duration w/ CP	102.86 usec for $CP = 1/8$
Frame parameters		
T_F	Frame length	2, **5**, 10, 20 msec
N_F	Number of OFDMA symbols in frame	18,**47**,95,193 (See Table 1.2.2)
R_{DL-UL}	Ratio of DL to UL	1:1, **2:1** (See also Table 1.2.3)
T_{duplex}	Duplex time between UL and DL	0.67 to 20ms
T_{class}	Classification of traffic	Control or Data
Permutation parameters		
DL_{Perm}	DL permutation type	**PUSC**, Band AMC, FUSC
UL_{Perm}	UL permutation type	**PUSC**, Band AMC,
BS_{Nused}	DL: number of subcarriers for BS TX	Depends on the bandwidth. For 10MHz, PUSC: $SS_{Nused} = 841$ FUSC: $SS_{Nused} = 851$ AMC: $SS_{Nused} = 865$
SS_{Nused}	UL: number of subcarriers for SS TX	Depends on the bandwidth. For 10 MHz, PUSC/FUSC: $SS_{Nused} = 24$ Band AMC: $SS_{Nused} = 18$
F_{sim}	Data subcarriers explicitly simulated	
$SubCh_{MAX,DL}$	Maximum number of subchannels in DL permutation	Depends on the bandwidth. For 10 MHz, Band AMC (48), **PUSC (30)**, FUSC (16)
$SubCh_{MAX,UL}$	Maximum number of subchannels in UL permutation	Depends on the bandwidth. For 10 MHz, Band AMC (48), **PUSC (35)**

Table 2.1.5: Propagation Model Parameters

Parameter	Description	Value Range
f	Carrier frequency	2.3-3.8 GHz **(2.5 GHz)** See also Table 1.4.1
PL	Path loss model	**COST-HATA-231,** Erceg
σ_{SF}	Log normal shadowing standard deviation	8-12dB (**8.9dB**)
ξ_{SF}	Shadowing correlation	$\xi_{SF} = 0.5$
L_P	Penetration and other losses	10-20dB **(10 dB)**
	Thermal noise density	**-174 dBm/Hz**

Table 2.1.6: Methodology Parameters

Parameter	Description	Value Range
	Simulation fading sample time granularity	**Symbol**, slot, frame
N_{Drop}	Number of times subscribers are dropped in the network.	50-500 (200)
T_{Drop}	Time duration of a drop	1 to 600 sec **(180s)**
$N_{\mathrm{Frames}} = \frac{T_{Drop}}{T_{Frame}}$	Number of frames in a drop	200 to 120000
N_{Sub}	Number of active subscribers/sector	TBD

Table 2.1.7: Dynamic System Simulation Features

Description	Value Range
Link adaptation	**Dynamic with delay feedback**
Channel estimation	**Realistic/ideal**

MIMO Support DL	**Alamouti STC, VSM**
MIMO Support UL	**Collaborative SM**
MIMO Switch	**Adaptive STC/ VSM Switch**
Coding	**CTC**
Control/MAP overhead modeling	**Dynamic**
Control/MAP channel reliability	**Realistic**
Hybrid ARQ	**Chase Combining with up to 3 retransmissions (total 4 transmissions)**
MAC ARQ	**Enabled for BE traffic**
Power Control	**UL**
Traffic models	**Realistic traffic mix,** Full buffer data only
Scheduling algorithm	**QoS Scheduling,** Proportional fair (data only)
ACK/CQI/BW-REQ Feedback channel error	**1%**
CQI Feedback delay	**2 frames**

2.1.2 Cell configuration and user placement

As shown in Figure 2.1.1, the system consists of a network of 19 hexagonal cells with six cells surrounding the center cell in the first tier and 12 cells surrounding the center cell in the second tier. For the default case, each cell has 3 sectors. SN_{Sub} mobile subscribers are randomly dropped over the 57 sectors such that each sector serves N_{Sub} subscribers. Each mobile subscriber corresponds to an active user session that runs for the duration of a drop. A *drop* is defined as a simulation run for a given set of subscribers over a specified number of time frames, N_{Frames}. At the beginning of each drop, the subscribers are associated with a specific BS and sector, henceforth referred to as the serving BS. The association is based on both the path loss and shadow fading, which are fixed for the duration of the drop. These network entry mechanisms are described in subsequent sections.

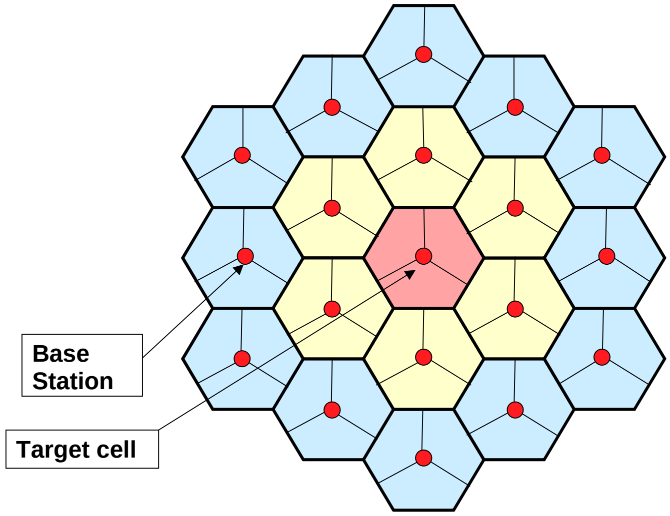


Figure 2.1.1: Configuration of adjacent tiers of neighbouring cells, sectors, and Base Stations

2.1.3 Interference Modelling

One of the key benefits of the OFDMA air interface is its possibility in enabling frequency reuse of 1. This will allow the same frequency channel to be shared within a cell, and ease deployment since no frequency planning is required. This attribute is already available in CDMA networks, and is highly desirable in Mobile WiMAX based on IEEE 802.16e. High frequency reuse patterns however cause the system to become interference limited. The interference seen by SS in downlink and BS in uplink is typically frequency selective and time selective. Therefore, it is not accurate to model interference as a white noise process with flat spectrum. The network simulator models interference using a realistic channel model, which includes both slow fading and fast time-frequency selective fading components. It is recognized that interference may have a different channel model than desired signal due to different propagation conditions. For simplicity however, we limit interference model to have the same characteristics.

Also, as an optional simplification, the time-frequency channel of only the strongest B interferers may be modelled to reduce simulation complexity. The remaining interferers are then modelled as (spatially) white and non-frequency selective AWGN processes whose variances are changing in time based on a Raleigh fading process. B can be that subset of interferers that contributes y% (e.g., 95%) of the total interference power as computed from the bulk path loss component.

Another consideration in studying interference is how network loading is modelled and the effect on system performance. OFDMA resource allocation is on a time-frequency grid, where the minimum allocation unit is a slot comprising of 1 or more frequency subchannels times 1 or more symbol duration. A slot duration consists of one, two, or three OFDMA symbols, depending on the permutation.

In order to allocate slots that interfere with a slot of interest at the desired receiver, two approaches are possible. One approach uses the notion of fractional loading to calculate the number of interfering subchannels in DL sub-frames from co-channel sectors. In this approach, the interfering allocations are not explicitly generated using schedulers, but rather by randomly selecting a set of ‘loaded’ subchannels based on the factional loading factor. However, this approach does not translate well to modeling uplink interference because the position and size of allocation for each interfering SSin co-channel sectors cannot be modeled in an appropriate manner. Therefore, we propose a methodology that requires schedulers to be modeled in all co-channel sectors to explicitly generate dynamic interference.

2.1.3.1 Modeling Frequency Reuse

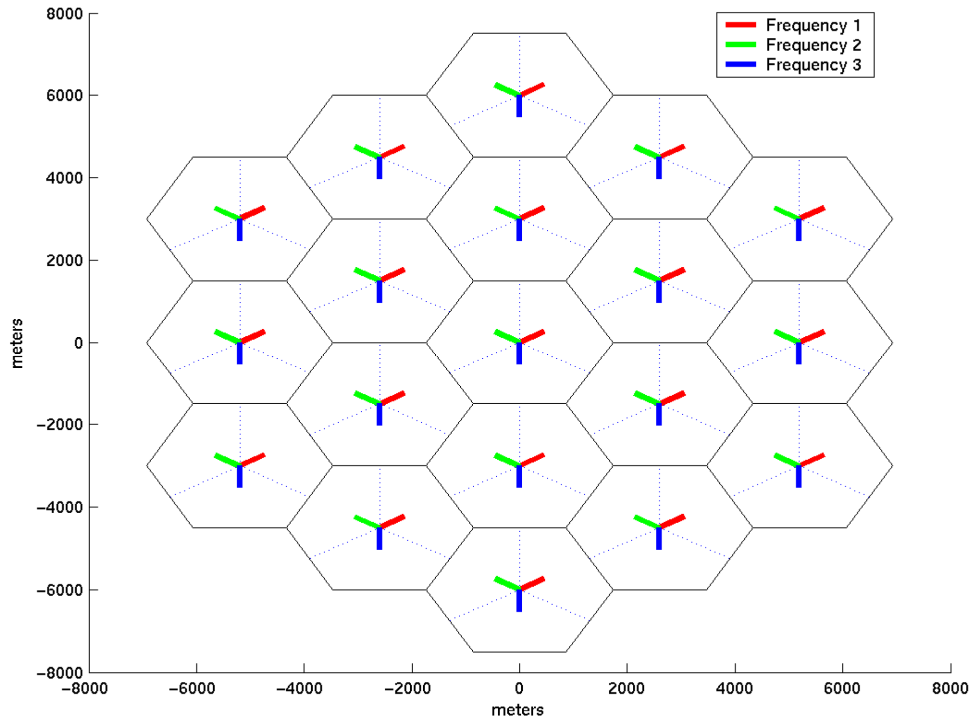


Figure 2.1.2 Network topology for tri-sector (19 cell) configuration with 1×3×3 reuse

The network is divided in to clusters of N cells (each cell in the cluster has a different frequency allocations), S sectors per cell, and K different frequency allocations per cell. This yields a frequency reuse pattern of N×S×K. Figure 2.1.2 shows a network topology with reuse pattern 1×3×3. The colored markings in the center of each cell indicate sectors and point in the boresight direction. The red markings correspond to sectors deployment in the same frequency allocation. Similarly, the blue and green markings indicate the other two frequency allocations for a reuse three network. Networks with universal frequency reuse 1×3×1 have the same network topology except the same frequency allocation is deployed in all sectors throughout the network. Thus, an operator using , say, 10 MHz channelization would require a total of 10 MHz of spectrum to support UL & DL for a TDD system with 1×3×1 reuse. To reduce interference, a frequency reuse pattern of 1×3×3 can be implemented by either sharing the available subchannels (say 1/3) in a 10 MHz channel or using 30 MHz of spectrum with 10 MHz in each sector. Additional reuse patterns are shown in the following figure.

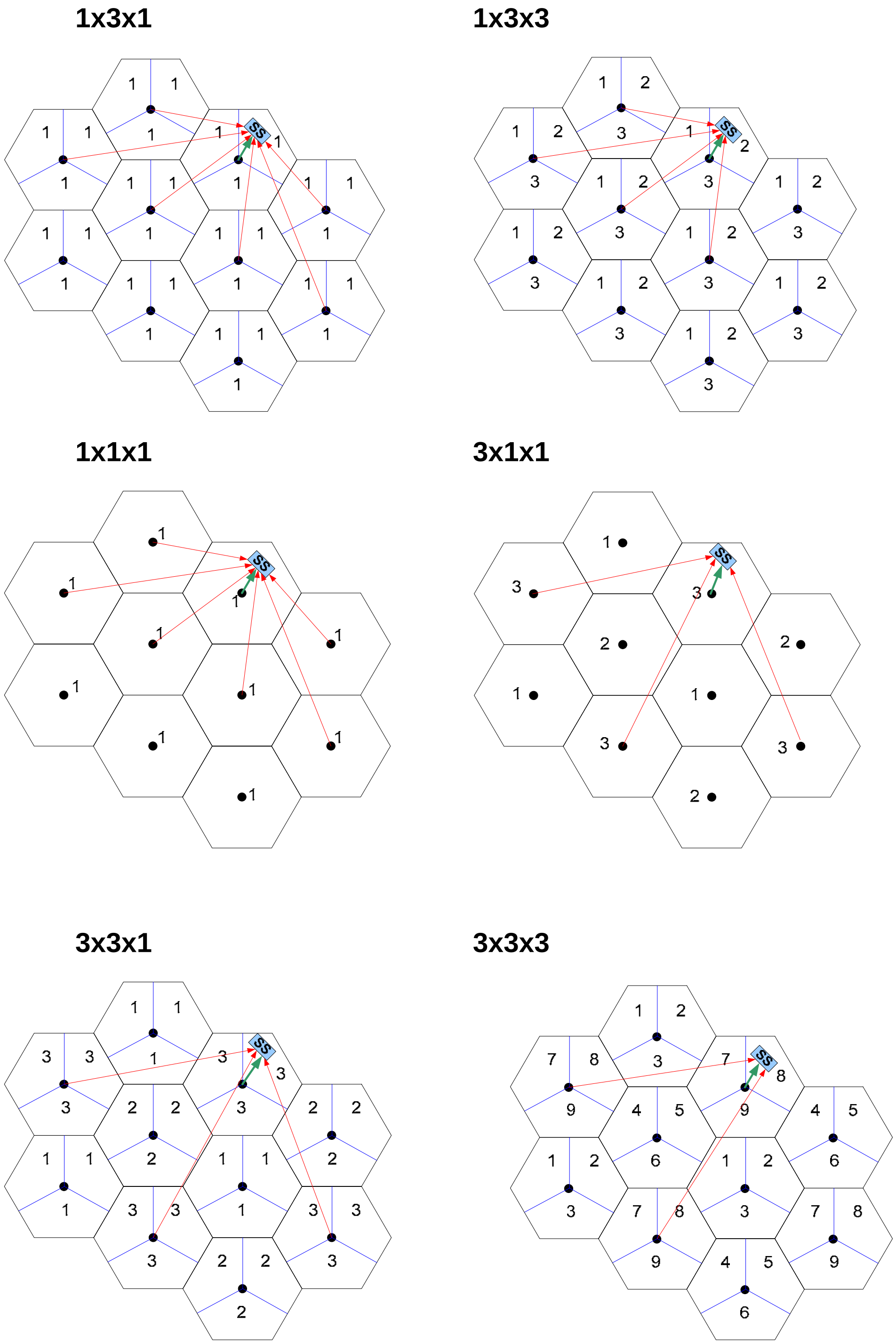


Figure 2.1.3: Additional Reuse Patterns

2.1.3.1 Fading Models for Useful and Interfering Signals

At each drop, a fast fading channel model [See Section 2.4] based on the required channel mix probability is assigned to each active and base station antenna, and is evolved in time. The bulk parameters of the model such as SS location, statistical characteristics of shadowing, delay profile and

Doppler rate remain fixed. Due to TDD channel reciprocity, the UL and DL channel are time-correlated. The time variation of channel from UL to DL depends on T_{duplex}, the duplexing time between UL to DL transmission.

In DL, a fast fading channel is instantiated for each pair of antennas between a base station and a co-channel subscriber that is allocated resources on the DL sub-frame. In the UL, a fast fading channel is instantiated for each pair of antennas between a subscriber with an UL allocation and a co-channel BS. Optionally, to reduce simulation complexity, only B strongest interferers may be modeled for downlink. For uplink, it may be necessary to model more interferers (B'>B), so that the number of active interferers at any given time is greater than B (TBD). (A simpler approach could be to keep track of fixed effective path loss (PL+shadowing etc) of B' interferers, and only simulate B interfering channels to be applied to active interferers at any instant.)

2.1.3.2 SINR computation for SISO

For each BS and SS link, the simulator computes the channel and interference power on the loaded data subcarriers. The received signal power level at the k th subcarrier for m th target user is calculated as:

$$P_{RX}^{m}(k) = \frac{P_{TX}^{m} \cdot P_{loss}^{m} \cdot |h_m(k)|^2}{N_{load}(j)},$$

Where

- P_{TX}^{m} is the total transmit power from BS (per sector) or m th SS
- P_{loss}^{m} is the path loss including shadowing and antenna gains
- $N_{load}(j)$ is the total number of co-channel subcarriers for the jth OFDMA symbol,

The co-channel interference power level (from the remaining BS in downlink and SS in uplink) at the k th subcarrier of m th target user is calculated as:

$$P_{CCI}^{m}(k) = \sum_{l=2}^{N_{CCI}} P_{CCI}^{l,m}(k) = \sum_{l=2}^{N_{CCI}} \frac{|h_{l,m}(k)|^2 \cdot P_{loss}^{l,m} \cdot P_{TX}^{l}}{N_{load}(j)}$$

where $P_{CCI}^{l,m}(k)$ is CCI power level from the l th interferer to m th target user, and N_{CCI} is a number of co-channel interferers. Finally, the SINR of m th target user at the k th subcarrier is:

$$SINR^{m}(k) = \frac{P_{RX}^{m}(k)}{\sigma_n^2 + P_{CCI}^{m}(k)} = \frac{P_{RX}^{m}(k)}{\sigma_n^2 + \sum_{l=2}^{N_{CCI}} P_{CCI}^{l,m}(k)},$$

Where σ_n^2 is the per subcarrier AWGN power level.

The set of SINRs per subcarrier, $\{SINR(k)\}_{k=1}^{Nused}$ where $Nused = 48$ subcarriers in a slot, capture frequency selectivity of channel and interference. Link performance, that is decoded BER/PER for the slot, depends on this SINR distribution and modulation & coding scheme (MCS). Since there is no known closed-form expression for capturing the decoder performance, we use a mapping function $PER = f(\{SINR(k)\}, MCS)$ to predict the slot error rate. This mapping function is referred to as the *PHY abstraction*, and is described in a subsequent section.

Note that for non-SISO cases, the SINR computation will depend on the multi-antenna processing algorithms employed at the transmitter and the receiver. The detailed description of this model is TBD.

2.1.3.3 SINR Computation for MIMO: See Section 5.4

2.1.4 Modelling Control and Signalling Information

The system level simulation must dynamically model DL and UL MAP overhead. This means that the MAP overhead calculation must be based on the MAP permutation, modulation/coding and instantaneous system load, on per frame basis.

The system level simulation must accurately model FCH/MAP channel reliability based on the permutation, modulation/coding and instantaneous channel condition. The impact of this FCH/MAP channel reliability on the system performance must be reflected in the simulation results. Further details of MAP overhead modelling are described in MAC chapter.

2.1.5 Number of simultaneous users serviced

In each cell, the scheduler in each sector will serve N_{Sub} subscribers. As discussed in Sec 8, the OFDMA physical resource available for scheduling is a frame. Hence, the scheduler will transmit to $N_a \leq N_{Sub}$ active user(s) at each scheduling decision boundary. Scheduler simulation is described in Section 4.6.

2.1.6 Traffic Modelling

Packets arrive as per the traffic models and traffic mixes specified in section (See Chapter 3) and packets are scheduled with a packet scheduler.

2.1.7 ARQ and HARQ

The ARQ process is modeled explicitly by rescheduling a packet after the ARQ feedback delay period. When HARQ is enabled, retransmissions are modeled based on the HARQ option chosen. HARQ can be configured as synchronous/asynchronous with adaptive/non-adaptive modulation and coding schemes for Chase combining or incremental redundancy operation. Chapter 4 on MAC includes further discussion on ARQ and HARQ modeling.

2.1.8 Modelling Feedback Delay

The scheduler at the BS uses the SINR reported by an SS via the CQICH to make a decision on which user to transmit based on some metric that takes into account a fairness criterion. See Section 4.6 for further discussion on scheduling metric and algorithm. In general for the UL and DL, the SINR reported by the each SS can be characterized by two parameters: the feedback period, and the feedback processing delay. The feedback period accounts for the periodicity of the CQICH allocation to an SS and is measured in frames. The feedback processing delay accounts for the latency associated with the measurement of SINR at the receiver, the decoding of the feedback channel, and the lead-time between the scheduling decision and actual transmission. Note that for UL scheduling, the SINR is measured directly on previous transmissions from the same MS, and therefore, the CQICH period, and feedback channel decoding delay does not apply.

2.1.9 Performance statistics calculation

Multiple time frames provide temporal averaging over channel and interference variations, and multiple drops provide averaging over placement of users within the network, shadowing parameters and fast fading impairments. Hence, output statistics such as throughput and spectral efficiency are collected over N_{Drop} drops. These statistics are discussed in Section 2.4 on performance metrics. Note that the traffic sessions must be configured in such a way that a sufficient number of traffic sessions can complete in the duration of a drop. Table 2.1.6 summarizes some key methodology parameters.

2.2 Sector Assignment

Sector assignment is based on the received power at a SS from all potential serving sectors. The sector with best path to SS, taking into account slow fading characteristics (path loss, shadowing, and antenna gains) is chosen as the serving sector. The following procedure, under the control of the parameters in Tables 2.1.1 and 2.1.6, is used to drop subscribers in the network.

1) Create a regular, three tier hexagonal tessellation of N_c cells.
2) For each of N_{drop} iterations, place $N_s N_{\mathrm{sub}}$ subscribers in the network:
    a) Drop subscribers randomly in the network with uniform (area) probability.
    b) Compute the total path loss (PL) from the subscriber to all sectors. Total PL includes the effects of shadowing and the directional antenna gain to all sectors in the network.
3) Assign the subscriber to the sector to which it has the smallest total path loss. If the best sector is already serving N_{sub} subscribers, then the subscriber is excluded from the simulation.

Continue to drop users until every sector is serving exactly N_{sub} subscribers.

2.3 Mobility Model

Mobility of subscribers can be considered in the system simulation in order to examine the effects of handoff on the overall system performance.

N_{sub} subscribers are decomposed into three groups; R_{fixed} N_{sub} subscribers are stationary, R_{ped} N_{sub} subscribers move at a low speed V_{low} to represent pedestrians, and R_{veh} N_{sub} subscribers move at a high speed V_{high} to represent vehicles where $0 \le R_{fix}, R_{ped}, R_{veh} \le 1$ and $R_{fix} + R_{ped} + R_{veh} = 1$ (to be in a table with recommended values).

Mobile subscribers move based on a mobility model that is a combination of "random walk" and "random waypoint" mobility models; a subscriber randomly chooses an adjacent cell, and then randomly chooses a waypoint to move to within the chosen adjacent cell.

For the ease of description, all the cells are categorized into the following 4 types:

C_0: the center cell

C_1: 6 cells surrounding the center cell in the first tier

C_2: 6 cells in the second tier with 3 adjacent cells

C_3: 6 cells in the second tier with 4 adjacent cells

The probability of choosing an adjacent cell (C_1) is uniform if a subscriber is currently in C_0, but non-uniform if it is in the other 18 cells in order to avoid an edge effect that makes the subscriber density non-uniform (higher density towards the center cell) over simulation time. Let P_{xy} denote the probability of a subscriber in C_x to choose a waypoint in C_y. The following 4 equations on incoming subscribers to each cell type need to be met in order to keep a uniform density over the 19 cell area:

$$6 P_{10} = 1$$

$$P_{01} + 2P_{11} + 2P_{31} + P_{21} = 1$$

$$P_{12} + 2P_{32} = 1$$

$$2P_{13} + 2P_{23} = 1$$

Assuming $P_{xy} = P_{yx}$ and $P_{01} = P_{11}$, these probabilities become as follows:

$$P_{01} = P_{10} = P_{11} = 1/6$$

$$P_{12} = P_{21} = 1/4$$

$$P_{13} = P_{31} = 1/8$$

$$P_{23} = P_{32} = 3/8$$

Further, when a subscriber in C2 (C3) chooses a cell in C3 (C2), it is possible that the subscriber goes out of service area. In such cases, the next waypoint needs to be reselected within the same destination cell to avoid the problem.

2.4 Channel Models for System Simulation

2.4.1 Fading and Mobility Channel Models

This annex provides recommendations for fading, mobility, and path models to be used for evaluation of different solution and options for WiMAX mobility profiles.

2.4.1.1 Path Loss Models

The channel models described in this Section could be used in different environments for system level simulation. The environments are Suburban macro-cell, Urban macro-cell, Urban micro-cell, and Indoor pico-cell. We recommend using Macro-Cell model for all system level simulations. For each environment, different Log-normal shadow fading standard deviation, and path loss models are used. The following summarize these values:

- **Macrocell**: 10 dB Log-Normal Shadow Fading, and modified COST231 Hata urban propagation path loss model is assumed for the macrocell environment. This model is described in the following formula:

$$PL[dB] = (44.9 - 6.55 Log_{10}(h_b)) Log_{10}\left(d/_{1000}\right) + 45.5 + (35.46 - 1.1 h_m) Log_{10}(f_c) - 13.82 Log_{10}(h_b) + 0.7 h_m + C$$

  where h_b and h_m are the mobile and BS antenna heights, f_c is the carrier frequency in MHz, d is the distance between the SS and BS, and C is a parameter which is 0 dB for suburban and 3 dB for urban environments.

- **Microcell**: The microcell NLOS path loss is based on the COST 231 Walfish-Ikegami NLOS model with 10 dB Log-Normal shadow fading. Assuming BS antenna height to be 12.5 m, building height 12 m, building to building distance 50 m, street width 25 m, SS antenna height 1.5 m, orientation 30 deg for all paths, and selection of metropolitan center, the equation simplifies to:

$$PL[dB] = -55.9 + 38 * Log_{10}(d) + (24.5 + 1.5 * fc/925) * Log_{10}(f_c)$$

  The microcell LOS path loss is based on the COST 231 Walfish-Ikegami street canyon model with 4 dB shadow fading. Assuming the same parameters as for the NLOS case, the path loss is

$$PL[dB] = -35.4 + 26 * Log_{10}(d) + 20 * Log_{10}(f_c),$$

  where h_b and h_m are the mobile and BS antenna heights, f_c is the carrier frequency in MHz, d is the distance between the SS and BS, and C is a parameter that is 0 dB for suburban and 3 dB for urban environments.

- **Indoor Picocell**: The indoor path loss is based on the COST 231 model:

$$PL[dB] = 37 + 30 * Log_{10}(d) + 18.3 * n^{\left(\frac{n+2}{n+1} - 0.46\right)},$$

  where n is the number of penetrated floor ($n = 4$ is an average for indoor office environments).

2.4.1.2 Channel Models

This section describes baseline channel models to be used as part of unified framework for evaluating the performances of different proposed features and profiles. Appropriate delay spread, Doppler spread, and spatial characteristics that are typical for the licensed bands below 6GHz are captured.

The channel delay profiles are complying with the standardized ITU SISO models. Proposed mobility profiles should be evaluated and compared based on the channel model methodologies described in this document or a similar agreed up on framework adopted by 3GPP/3GPP2. This facilitates a consistent reference for comparison among different proposals and with documented performance results of existing systems using the ITU models.

Table 2.4.1: Fading and Mobility Channel Model

Power Delay Profile		Pedestrian-A		Vehicular-A		Pedestrian-B	
Number of Paths		4		6		6	
Power of the each path (dB)	Path Delay (ns)	0	0	0	0	0	0
		-9.7	110	-1	300	-0.9	200
		-19.2	190	-9	700	-4.9	800
		-22.8	410	-10	110	-8	1200
				-15	1700	-7.8	2300
				-20	2500	-23.9	3700
Speed (km/h)		3		60		3	

In Table 2.4.1**Error! Reference source not found.**, "Path Delays" refer to the rms values of the path delays. Note that the delay profile of a channel is continuous, but for the simulation purposes, the rms values are chosen. Also, "Power Profiles" refer to the power (variance) of each path relative to the first path.

A stationary channel model similar to Pedestrian A, and B have been considered, where the mobile is stationary or very low speed (portable), and the models could be used with Rician (LOS) components. The simulation description is in the context of a downlink transmission; however the methodology described here can be applied to the uplink as well.

The performance metrics such as data throughputs should be obtained form a system level simulation models consisting of multiple cells/sectors, BSs, and MSs (at least two tiers of BSs), as opposed to link level simulations wherein only a single BS transmitting to a single SS is considered. During one simulation run, the channel undergoes fast fading due to frequency selectivity, and Doppler Fading according to the speed of MSs.

When using multiple antennas, the performance metrics are expected to be adjusted to the situation where there is 50% spatial correlation among the channels constituting the link between each pair of Tx/Rx antennas.

Note that the model outlined in this annex is a proposed framework for the channel models to be used as a testbed for evaluating the performances of different features and profiles. The models are similar to those used in 3GPP/3GPP2 with different parameters. As long as it is consistent with the models described in this section, the 3GPP/3GPP2 channel model may be used.

2.5 Power Control

Basic *open-loop* power control is applied in the uplink at the beginning of each drop. The transmit power for the i^{th} user is adjusted such that the noise-limited SNR at its *serving* sector is no greater than a SNR target. The target SNR, SNR_{target}, should be based on maximum SNR requirement for the UL MCS. The TX power is computed as

$$Power_{tx}^{i} = \min\{P_{SS}, PL_i + SNR_{\text{target}}\}$$

Where PL_i is the total path loss (including shadowing and directional antenna gain) for the i^{th} user to its serving sector, and P_{SS} is the maximum subscriber station TX power. This basic UL power control algorithm is *open loop* in the sense that it is based only on total path loss and does not track variations in the channel (fading).

2.6 PHY Abstraction

In a system simulation, it is important to device a model that accurately predicts the link-layer performance in a computationally simple way. The requirement for an such a model stems from the fact that simulating the physical layer links between multiples BSs and SSs in a network/system simulator can be computationally prohibitive. This model is referred to as the *PHY abstraction*. The role of the abstraction is to predict the decoded PDU error rate (PER) given a set of SINR values for each sub-carrier k in the PDU, and the MCS, that is $PER = f(\{SINR(k)\}, MCS)$.

There are two concepts that come into play when modelling PDU errors. First, a method is required to assign an equivalent SINR to an FEC block. Secondly, a method is needed to model PDU error based on the errors in the FEC blocks contained within a PDU. These methods are described in the following subsections.

2.6.1 Computation of Equivalent SINR for an FEC Block

Methods such as EESM or MIC can be used to assign the equivalent SINR to an FEC block. These methods are described in the Annex B through E.

2.6.2 Computation of PDU Errors

Modelling PDU errors is based on AWGN curves of block error rate versus SNR for all block sizes and MCS values allowed in the WiMAX profile. The method to model PDU errors is described below:

1) Compute number of FEC blocks and block sizes for received PDU. The slot-concatenation computation rule along with a sample calculation is given in the appendix. If repetition is used, combine the repeated blocks to get the composite SINR on each non-repeated FEC block

2) Based on the block sizes, look up $P_{blk,\,i}$ in the appropriate curve for the each block “i” in the PDU

3) Compute the probability of packet success:

$$P_{pkt} = 1 - \prod_{i=1}^{N}\left(1 - P_{blk,i}\right)$$

Use a "Bernoulli toss" with probability P_{pkt} to decide whether the packet is successful or not

2.7 Performance Metrics

This section describes the performance statistics that are generated as an output from the system-level simulations. System-level simulations can be used to study the performance using infinite buffer models or real-traffic as described in Section 3. In this section, we therefore, present metrics for both types of studies. In each case, a performance curve given as a function of the number of users per sector is generated.

2.7.1 Output Metrics for Infinite Buffer Models

The following statistics related to data traffic should be generated and included in the evaluation report for each scheme.

1. **Average Sector throughput [kbps/cell]** is used to study the network throughput performance, and is measured as

$$R = \frac{b}{k \cdot T}$$

   Where b is the total number of correctly received data bits by all SSs in the simulated system over the whole simulated time, k is the number of cells in the simulation and T is the simulated time. In the case of only evaluating the center cell site, k is the number of sectors.

2. **Average Connection throughput [kbps]** for user i is defined as

$$R_{CID}(i) = \frac{\sum_{k} \text{good bits in CID } k \text{ of user } i}{\sum_{k}\left(t_{end\_k} - t_{arrival\_k}\right)}$$

   Where the connection id (CID) k denotes the k^{th} connection from a group of K connections where the K connections can be for a given user i, $t_{arrival\_k}$ = first block of connection k arrives in queue, and t_{end_k} = last block of connection k is received by the subscriber station. Note for uncompleted connections, t_{end_k} is set to simulation end time. The mean, standard deviation, and distribution of this statistic are to be provided.

3. **Block Error Rate** is calculated for all the application sessions. It is defined as the ratio

$$BLER_{session} = \frac{n_{erroneous\_blocks}}{n_{blocks}},$$

   Where $n_{erroneous\_blocks}$ is the total number of erroneous FEC blocks in the application session and n_{blocks} is the total number of blocks in the session. These individual session BLERs from all applica sessions form the distribution for this statistic. The mean, standard deviation, and the distribution of this statistic are to be provided.

A Definition of an Application Session: An application session contains one or several connections depending on the application. Application session starts when the transmission of the first block of the first connection of a given service begins and ends when the last block of the last connection of that service has been transmitted. Note, that BLER statistics are only collected from those frames during which SS is receiving data.

4. **The residual BLER** is calculated for each user for each application session. An application session residual BLER is defined by the ratio

$$BLER_{residual} = \frac{n_{dropped\_blocks}}{n_{blocks}},$$

Where $n_{dropped\_blocks}$ is the total number of dropped blocks in the application session and n_{blocks} is the total number of FEC blocks in the session. A dropped block is one in which the maximum ARQ or HARQ re-transmissions have been exhausted without the block being successfully decoded. The mean, standard deviation, and distribution of this statistic over all the application sessions in the simulation are to be provided.

5. **The averaged block delay per sector** is defined as the ratio of the accumulated delay for all blocks for all SSs received by the sector and the total number of blocks. The delay for an individual block is defined as the time between when the block enters the queue at transmitter and the time when the block is received successively by the SS. If a packet is not successfully delivered by the end of a run, its ending time is the end of the run.

6. **Throughput Outage** is defined as percentage of users with data rate less than some minimum data rate. This outage is computed by plotting CDF of user throughputs across all drops, and determining percentage of users that are unable to meet minimum requirements.

7. **User outage** is defined as the event where a connection of ith subscriber has short-term packet error rate (PER) higher than target value PER_{out} more often than in t_{out} percent of all frames. The number N_{out} is defined as the number of users in outage.

8. **System outage percentage** is estimated as the fraction of users in outage. The *system outage* occurs when the system outage percentage is higher than target system outage value T_{out}. Table 2.5.1 contains typical parameters for system outage calculation in system level simulations.

Table 2.5.1: Parameters for system outage calculation

Outage parameters	Values
Target PER for MCS selection (PER_{MCS})	PER_{MCS} = 1%
Target PER for user outage (PER_{out})	PER_{out} = 15%

Target value t_{out}. of user outage	$t_{out} = 1\%$
Target value T_{out}. of system outage	$T_{out} = 3\%$

2.7.2 Output Metrics for Real-Traffic Models

In the case of real-traffic, the throughput and latency can be measured at the application transaction level, application PDU level, transport PDU level (TCP or UDP segment), IP datagrams, link level SDU (Ethernet frames), MAC PDUs, or FEC blocks. A transaction is defined as an application level activity that has some user significance. For example, transfer of a file using FTP, displaying of a web page using HTTP. Each transaction results in a number of application PDUs that are sent via TCP or UDP segments, which in turn result in a number of IP datagrams, etc. It is important provide performance statistics both at the application layer and at the link layer. The higher layer statistics provide a measure of user experience while the lower layer statistics provide the impact of various link layer features.

1. **Transaction Completion Time** is measured as the time between the end of user request to the end of the system response. For example, for HTTP application it is the time to fetch and display a web page specified by the user.

2. **Transactions Per Second** is a measure of the application level throughput. It is measured by

$$TransactionsPerSecond = \frac{TransactionsCompleted}{TotalTime}$$

   Note that the transactions that were incomplete or in process when the simulation ended are not included in the computation above.

3. **Fairness** Among Similar Users is measured by variation of throughput among users with similar demands. Thus, if there are n users and the average transaction completion time for the ith user is Ci, then the fairness index for the transaction completion time is [Jain91]:

$$FairnessIndex = \frac{\left[\sum_{i}^{n} C_i\right]^2}{n\sum_{i}^{n} C_i^2}$$

   This fairness index has the important property that if only k of n users are given service and get equal service then the fairness index is k/n. The index is 1 (or 100%) if all n users get equal service.**Probability of Transactions in Error** is measured by the fraction of transactions that are not completed satisfactorily, for example, files damaged during FTP, web pages with missing images, etc. Statistics on the time between errors is also of interest.

3 Application Traffic Models

The application traffic models are categorized in to two types of traffic modelled: foreground and background traffic. The foreground traffic model represents a specific user behaviour or interaction, or in other word, application traffic is generated mainly through a user interaction with a device. One of the goals of the simulator to implement the foreground traffic model is to evaluating a user perspective application performance in detail as specified in [2]. The background traffic, however, is not directly related to a user interaction. For certain application or service generates network traffic as soon as a session starts regardless of a user interaction with a device and the amount of network traffic is significantly larger than foreground traffic for the application or service. IM application and VPN service are the examples.

Each model in the following subsections is represented for both user level and IP packet level. The usages of these two levels of model are the following:

- **User level traffic model:**
  - User behaviour interactions in an application is modelled, and this model can be used with a simulation which includes detailed application layer, transport layer and IP layer model on top of the Layer 1(PHY) and Layer 2 (MAC) models.
  - Application performance metrics specific to the application can be evaluated (ex, web page download time, email download time, FTP download time, etc), and a scheduling mechanism for the application QoS in the MAC layer can be evaluated or optimised.
  - Since the IP level traffic will be generated according to the upper layer protocols, the generated UL and DL traffic will be correlated.
- **IP packet level traffic model:**
  - This model generally obtained from a network traffic measurement and represented as statistical packet distributions, such as packet size distribution and packet inter-arrival time distribution at the IP layer, and this model can be used with a simulation which does not includes detailed protocol layers above Layer 2 (MAC) models.
  - The IP packet level traffic model can be directly applied to the Layer 2 and Layer 1 model, and evaluating both application performance and scheduling mechanism in the MAC layer are not easy task with this type of model. In general, this model is used in System Level Simulation (SLS) for the air link resource allocation.
  - UL and DL traffic may not be correlated since DL and LU traffic may independently modelled.

For an application, which is difficult to define or implement at the user level, only IP level model is discussed.

The WiMAX Forum has identified several applications for 802.16e-based systems and is developing traffic and usage models for them. These applications can be broken down into five major classes. These application classes are summarized in the following table together with guidelines for latency and jitter to assure a quality user experience.

Table 3.1: WiMAX Application Classes

Class	Application	Bandwidth Guideline		Latency Guideline		Jitter Guideline	
1	Multiplayer Interactive Gaming	Low	50-85 kbps	Low	< 150 msec	Low	<100 msec
2	VoIP & Video Conference	Low	4-384 kbps	Low	< 150 msec	Low	<50 msec
3	Streaming Media	Low to High	5 kbps to 2 Mbps	N/A		Low	<100 msec
4	Web Browsing & Instant Messaging	Moderate	10 kbps to 2 Mbps	N/A		N/A	
5	Media Content Downloads	High	> 2 Mbps	N/A		N/A	

3.1 Internet Game Traffic Model (Class 1)

Internet gaming is a fast growing business and most of the recent computer game releases support multiplayer options over Internet connections. However, there exist no publicly available standardized protocols for exchanging network gaming data. This means that different game software manufacturers utilize either their own or licensed protocols for network gaming. There exist several different types of network games with varying requirements for network support. Especially real time gaming traffic will most probably be highly sensitive to the Quality of Service of the network layer. One of the major types of games is action games. The amount of traffic generated by different games can vary heavily. Action games usually contain virtual persons moving (and probably shooting at each other) in real-time virtual environments. We have shown Quake II, Xbox Halo2, and Toon Town in this document corresponding to [3], *WiMAX Application Usage Profile*, and analysed the traffic behaviour in terms of packet size distribution, packet inter-arrival time distribution and bandwidth usage.

3.1.1 Internet Game: User level model

Since there exist no publicly available standardized protocols for exchanging network gaming data, it is not trivial to model the application at the user level.

3.1.2 Internet Game: IP level model

3.1.2.1 Quake II Traffic Model

Table 3.1.1 describes detailed packet distribution information and session duration distribution. The session duration is assuming that each user would play about one hour on average but no more than 2 hours considering wireless environment. Figure 3.1.1 and Figure 3.1.2 depict packet size and packet inter-arrival probability density function generated with the information in Table 3.1.1. In this table and many subsequent tables, the packet inter-arrival times are modelled with Extreme value distribution. Its cumulative distribution function is given by:

$$F(x) = 1 - e^{-e^{(x-a)/b}}$$

Here a is the location and b is the scale parameter. Random variates for this distribution can be easily generated by using inverse transformation as follows [Jain91]:

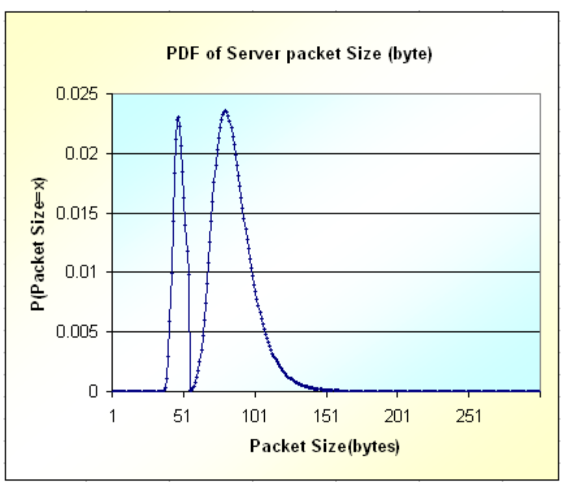
$$x = a + b \ln \ln u$$

Where u is a uniform U(0,1) random number. Session durations follow a truncated extreme distribution, that is, if the generated extreme random variate is outside the specified range, the value is discarded and another value is generated.

Table 3.1.1: Quake II Traffic Model

Internet Game: Quake II Traffic Model			
Session Duration (hour)	Extreme (a=1, b=0.1), Truncated (0, 2)		
Client/Server	Component	Distribution	Parameter

Client to Server	Packet Inter-arrival time (msec)	Lower 4.5%, x<18:Extreme	a=6.57, b=0.517
		Upper 95.5%, x>= 18: Extreme	a=37.9, b=7.22
	Packet Sizes (byte)	Seven Distinct values	10.6%:36, 26.4%: 42, 6.26%: 44, 13.9%: 45, 4.95%: 46, 16.3%: 48, 21.5%: 51
Server to Client	Packet Inter-arrival time (sec)	Lower 4.8%, x<60:Extreme	a=58.2, b=7.47
		Upper 95.2%, x>= 60: Normal	a=100, b=17.7
	Packet Sizes (byte)	Lower 27.6%, x<55:Extreme	a=46.7, b=4.39
		Upper 72.4%, x>= 55: Extreme	a=79.7, b=11.3



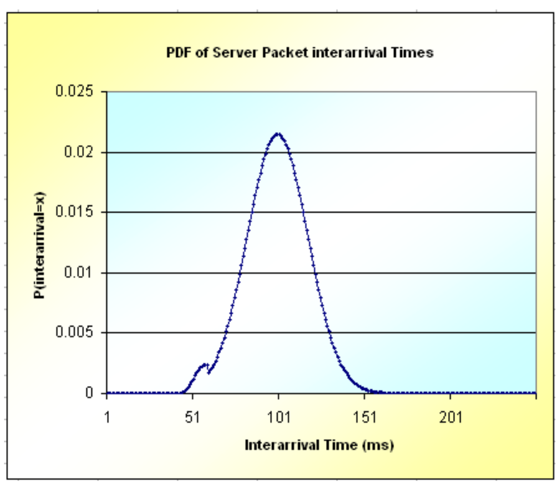


Figure 3.1.1 Server packet size and packet inter-arrival distribution

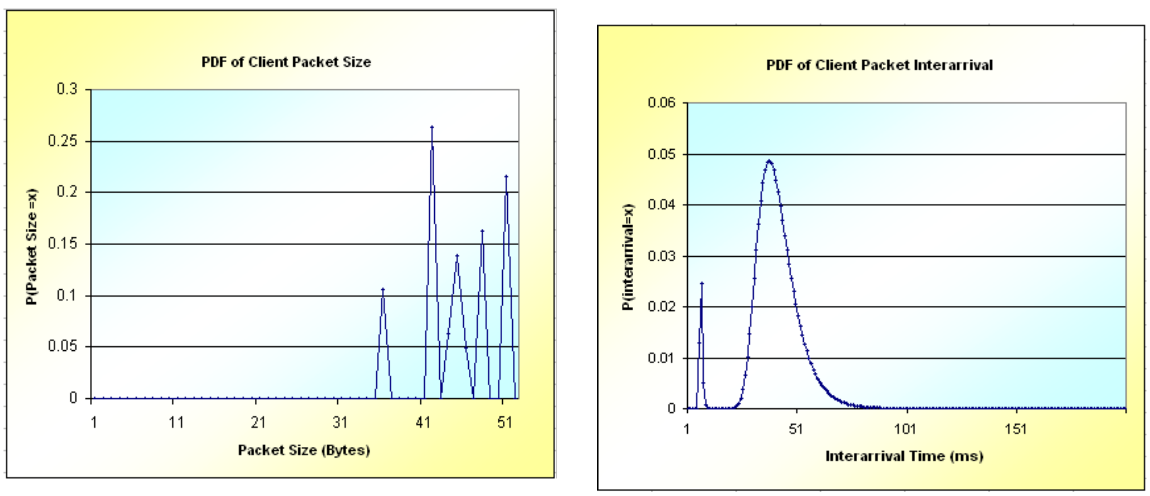


Figure 3.1.2 Client packet size and packet inter-arrival distribution

3.1.2.2 Xbox game (Halo2) Traffic Model

Based on the observed traffic characteristics, traffic models for the IAT and packet size distributions of Halo 2 clients and servers has been developed, and the session duration is assuming that each user would play about one hour on average but no more than 2 hours considering wireless environment. Table 3.1.2 summarizes the traffic model parameters. The packet size from the model need to be round to nearest 8 bytes packet size since the characteristic of packet size are always multiple of 4 bytes and spaced 8 bytes apart. See section 3.1.2.1 for discussion on generation of extreme and truncated extreme distributed random variates.

Table 3.1.2: Halo 2 Traffic Model

Internet Game: Halo 2 Traffic Model	
Session Duration (hour)	Extreme (a=1, b=0.1), Truncated (0, 2)

Client/Server	Component	# Players	Distribution	Parameter
Client to Server	Packet Inter-arrival time (msec)	> 0 player	Normal	a=40, b=1
	Packet Sizes	1 Player	Extreme	a=71.2, b=5.7
		2 Player	Extreme	a=86.9, b=5.1
		3 Player	Extreme	a=111.5, b=7.7
		4 Player	Extreme	a=127.7, b=8.2
Server to Client	Packet Inter-arrival time (msec)	> 0 player	Extreme	a=39.7, b=1.9
	Packet Sizes	3 Player	Extreme	a=126.9, b=20.4
		4 Player	Extreme	a=146.7, b=22.3
		5 Player	Extreme	a=167.6, b=25.1
		6 Player	Extreme	a=180.5, b=31.2
		7 Player	Extreme	a=241.3, b=31.6
		8 Player	Extreme	a=236.4, b=35.8
		9 Player	Extreme	a=226.8, b=26.2
		10 Player	Extreme	a=249.8, b=28.9
		11 Player	Extreme	a=271.0, b=33.0
NOTE	Packet Size = round to nearest 8 byte packet size			

3.1.2.3 Toon Town Traffic Model

Table 3.1.3 describes detailed packet distribution information and session duration distribution. The session duration is assuming that each user would play about one hour on average but no more than 2 hours considering wireless environment.

Table 3.1.3: Toon Town Traffic Model

Internet Game: Toon Town Traffic Model		
Session Duration (hour)	Extreme (a=1, b=0.1), Truncated (0, 2)	
Client to Server	Packet Inter-arrival time (msec)	Extreme (a= 0.13, b= 0.30), Truncated(0, 16.6)

	Packet Sizes (byte)	Extreme (a= 100, b= 20.6), Truncated(94, 1554)
Server to Client	Packet Inter-arrival time (msec)	Extreme (a= 0.13, b= 0.30), Truncated(0,16.62)
	Packet Sizes (byte)	Exponential(104.8), Truncated(100, 1554)

3.2 VoIP Traffic Model (Class 2)

There are a variety of encoding schemes for voice (i.e., G.711, G.722, G.722.1, G.723.1, G.728, G.729, and AMR) that result in different bandwidth requirements. Including the protocol overhead, it is very common for a VoIP call to require between 5 Kbps and 64 Kbps of bi-directional bandwidth.

The Adaptive multi Rate (AMR) codec is the most important vocoder in wireless applications being the newest vocoder for the existing GSM networks and it has been adopted as a mandatory speech codec speech processing function in UMTS [4].

The VoIP traffic model in Table 3.2.1 assumes AMR codec and RTP/UDP/IP header is included in the packet size calculation. Most likely, header compression will be used and hence the actual VoIP packet size will be 33 bytes plus 3 bytes compressed RTP/UDP/IP header. For the simplicity, signalling traffic is not modelled.

3.2.1 VoIP traffic model: User level model and IP level model

Voice call activities generate a pattern of talk spurt and silence (or ON and OFF) intervals by means of a speech activity detector so that it can be modelled as a two state Markov chain. The experimental measurements with ten conversations of 15 minutes in length showed the activity rate was 0.4717 with AMR, and the mean duration of the ON periods was 1026 ms, while the OFF periods mean duration was 1171 ms [6].

In the voice traffic model 1026 ms for a talk spurt and 1171 ms silence period are used followed by [6]. The number of voice packet during a talk spurt (1026 sec) and silence period (1171 sec) can be calculated by Eq. (1) [7].

$$PktCount_{Talk} = \left(\frac{Duration_{Talk}(1026ms)}{TTI(20ms)} + Hangover(7)\right) = 58$$

$$PktCount_{ComfortNoise} = \frac{\dfrac{Duration_{Silence}(1171ms)}{TTI(20ms)} - Hangover(7)}{8} = 6.44 \tag{1}$$

3.2.1.1 Adaptive Multi Rate (AMR) Codec Model

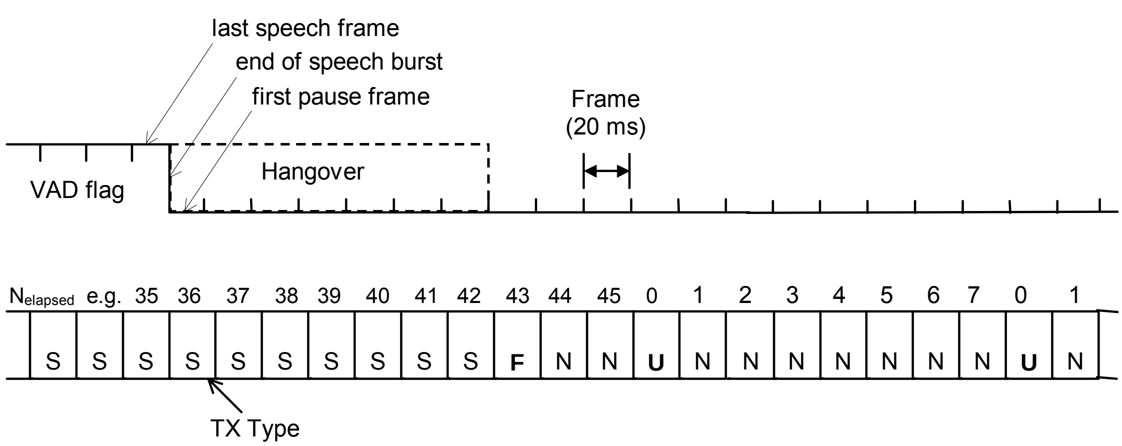
The Adaptive multi Rate (AMR) codec is the most important vocoder in wireless applications being the newest vocoder for the existing GSM networks and it has been adopted as a mandatory speech codec speech processing function in UMTS. Source controlled rate (SCR), called "Discontinuous transmission" in GSM, functionality is also part of the standard.

The two alternative generic frame formats for both the speech and comfort noise frames of the AMR speech codec are described in [5]. The Voice traffic model in this paper considers only one of the specified rates during the ON state (talk spurt) of the AMR codec (12.2 kbps) and the comfort noise (AMR_SID) during the OFF state of the AMR codec.

AMR Source Controlled Rate (SCR) Operation in AMR Codec

The AMR codec supports silence detection and it consists of "ON" and "OFF" states due to source controlled rate (SRC) operation. During the silence period (OFF state), the transmitter periodically generate "comfort noise" frame and these frames are called silence descriptor (AMR_SID) frame, and it can improve the speech quality by reproducing the speaker's background noise. At the end of a talk spurt (ON state) the transmitter generate additional seven frames at the same rate as speech to ensure the correct estimation of comfort noise parameters at the receiver side even if it is a silence period (called hangover period)

After the hangover period, a SID_FIRST frame is sent and it does not contain data. The transmission is cut after the transmission of a SID_FIRST frame. During the silence period the transmission is resumed at regular intervals for transmission of one SID_UPDATE frame, in order to update the generated comfort noise on the receiver side.



("S" = SPEECH; "F" = SID_FIRST; "U" = "SID_UPDATE; "N" = NO DATA,
Nelapsed: No. Of elapsed frames since last SID_UPDATE)

Figure 3.2.1: SRC procedure

Once the SID_FIRST frame has transmitted, the transmitter shall at regular intervals compute and transmit updated SID_UPDATE (Comfort Noise) frames during the silence period. SID_UPDATE frames shall be generated every 8th frame but the first SID_UPDATE shall be sent as the third frame after the SID_FIRST frame. In addition to the ARM, SRC feature including hangover, SID_FIRST, NO_DATA, and SID_UPDATE frames are modelled.

Table 3.2.1: VoIP Traffic Model

VoIP Traffic Model	
Average Call Holding Time (sec)	Exponential: μ =210 sec
Voice CODEC	AMR (12.2kbps)
Frame Length	20 msec
Talk spurt length	Exponential: μ =1026 ms
Silence length	Exponential: μ =1171 ms
Silence suppression	Yes
Protocols	RTP/UDP/IP
Header compression	RTP/UDP/IP header compression
Speech Activity	47.17%
UL:DL Ratio	1:1
Total MAC PDU size during a talk spurt	MAC header (6bytes) + compressed RTP/UDP/IP header (3 bytes) + voice packet (33 bytes) = 42 bytes
Total MAC PDU size during a silence	MAC header (6bytes) + compressed RTP/UDP/IP header (3 bytes) + voice packet (7 bytes) = 16 bytes
Average BW usage at the MAC layer	((58*42Bytes + 6.44*16Bytes)*8)/2.198sec = 9.25kbps (w/o HARQ CRC 2Bytes), 9.71kbps(w/ HARQ CRC 2Bytes)
Note: VoIP characteristics may vary for a particular codec	

3.3 Video Conference Traffic Model (Class 2)

Video conferencing has differing bandwidth requirements for the audio and the video components. For example, the audio component of a video-conference requires between 16 and 64 Kbps and the video component of a video-conference requires between 320 Kbps and 1 Mbps. The typical sustainable send/receive throughput requirements range is from a low of 32 Kbps to a high of 1 Mbps. A typical business-quality videoconference runs at 384 Kbps and can deliver TV-quality video at 25 to 30 frames per second.

H.264 is the next-generation video compression technology in the MPEG-4 standard, and it can match the best possible MPEG-2 quality at up to half the data rate. H.264 also delivers excellent video quality across the entire bandwidth spectrum — from 3G to HD and everything in between (from 40

Kbps to upwards of 10 Mbps). HD MPEG-2 content at 1920x1080 traditionally runs at 12-20 Mbps, while H.264 can deliver 1920x1080 content at 7-8 Mbps at the same or better quality. H.264 provides DVD quality at about half the data rate of MPEG-2. Because of this efficiency, H.264, an ISO standard, stands to be the likely successor to MPEG-2 in the professional media industry. For a mobile device, the MPEG encoder may not generate *B* frames and the distance between *I* frames could be longer than 10 sec [need reference]

The H.264 surpasses H.261 and H.263 in terms of video quality, effective compression and resilience to transmission losses, giving it the potential to halve the required bandwidth for digital video services over the Internet or 3G Wireless networks. H.264 is likely to be used in applications such as Video Conferencing, Video Streaming, Mobile devices, Tele-Medicine etc. Current 3G mobiles use a derivate of MPEG-4.

MPEG compressed videos are composed of pictures (frames) that are separated into three different types: I, B, and P. I frames are intraframes that encode the current picture, while B and P frames interpolate from previous and future frames. When transmitted over an IP network, these frames are segmented into one or more IP packets.

Krunz and Tripathi describe MPEG4 video traffic modelling, and they separate the video trace into I, B, and P frames, model each set of frames, and combine them together to form a final model that looks very similar to the original in [8]. Matrawy, Lambdaris, and Huang describe MPEG4 video traffic modeling with separate set of I, B, and P frames in [9].

An important feature of common MPEG encoders is the manner in which frame types are generated. Although not required by the standards, typical encoders use a fixed Group-of-Pictures (GOP) pattern when compressing a video sequence (the GOP pattern specifies the number and temporal order of P and B frames between two successive I frames). Quite often the fixed GOP pattern is “regular” in the sense that the number of B frames between two reference frames (I or P) is fixed. Such a GOP pattern can be characterized by two parameters: the I-to-I frame distance (N), and the I-to-P frame distance (M) (if no P frames are used, then M = N).

3.3.1 Video Conference traffic model: User/Application level model

Based on the compression efficiency and market acceptance as described above, MPEG-4 has been selected for the video codec. The estimated values, for the parameters to model a video stream, vary from one trace to another. The scene lengths have a lognormal distribution, which means that the log of the scene length has a normal distribution. The parameters μ and σ are the mean and the standard deviations of log lengths. Lognormal random variates with parameters μ and σ can be generated by generating Normal N(0,1) random variates x and then returning $e^{\mu+\sigma x}$.

For parameters associated with the lognormal distributions, the estimates depend strongly on the dimensions of the captured frames. For the video conferencing traffic model, high quality-ARD Talk (German Sunday morning talk show) trace is described in Table 3.3.1. In the model we have considered two different resolutions for the display; 176x144 for a small device and 320x240 for a large device. The required bandwidth for the uncompressed video stream with 176x144 pixels and 8-bit color depth is about 7.6 Mbps and with 320x240 pixels and 8-bit color depth is about 23 Mbps.

Table 3.3.1: Video Conference Traffic Model

Video Conference Traffic Model		
Session Duration (sec)	3600 sec	
Video Codec	MPEG4	
Protocols	RTP/UDP/IP	
Header compression	RTP/UDP/IP header compression	
Scene Length (sec)	Lognormal(μ =5.1 sec , σ =9.05 sec)	
Direction	Bi-direction (UL and DL)	
GOP	N=12, M=2	
Display size	176x144	320x240
Color depth (bit)	8	8
Subsampling method	4:1:1	4:1:1
Mean BW for Uncompressed stream	7.6 Mbps	23 Mbps
Compression ratio	13.95	13.95
Mean BW for compressed stream	0.54 Mbps	1.65 Mbps
I frame size	Lognormal(μ =6210 ,σ =1798)	Lognormal(μ =18793 ,σ =5441)
P frame size	Lognormal(μ =2826 ,σ =1131)	Lognormal(μ =8552 , σ =3422)
B frame size	Lognormal(μ =1998 ,σ =716)	Lognormal(μ = 6048 , σ =2168)
AR coefficient	a_1=0.39, a_2=0.15, $\sigma_\varepsilon = 4.36$	a_1=0.39, a_2=0.15, $\sigma_\varepsilon = 4.36$

3.4 PTT Traffic Model (Class 2)

Push to talk (PTT) is basically a wireless service that allows cell phones to act like long-range walkie-talkies. Instead of calling someone, you simply push a button and talk to them. It happens in less than a second, as with walkie-talkies. Because there is no time spent dialling or making a connection to a network, calls are shorter and less expensive than usual. Currently, there are many PTT solutions in the market from various vendors and most of the detailed implementations or protocols of those PTT solutions are proprietary except the Push to talk over Cellular (PoC) specification, which is based on the IP Multimedia Subsystem (IMS) as defined by 3GPP and jointly developed by Siemens, Ericsson, Motorola and Nokia. Because of this reason, the PTT protocol in this document is mostly based on the PoC specification as a starting point [10][11].

3.4.1 PTT traffic model: User/application level model

The detailed PTT application traffic model is summarized in the following Table 3.4.1, and for the simplicity of the traffic model, we assumed that a talk burst is consist of only talk spurt in the model.

There are some important terminologies to understand before reading the following table since some terminology is used in different places with different definitions.

- **Talk-burst:** The voice of the granted participant from the time the user is granted until the user releases the permission to talk.
- **Talk spurt:** A part of the speech signal that starts with a speech onset and ends when the speech coder goes down in DTX-mode. Hence a talk burst can consist of several talk spurts.
- **Talk session:** This is an established connection between PTT users where the users can communicate one at a time in a half duplex manner.

Table 3.4.1: Push-to-Talk Traffic Model

Push to Talk (PTT) Traffic Model		
Call Type Mix	Personal Talk (One-to-one)	Group Talk
	90%	10%
Average number of people/group	5	
Voice CODEC	AMR (12.2 kbps)	
Speech Activity	40%	
Protocols	SIP/RTP/UDP/IP/	
Number of Voice frames per RTP packet	10 (this is equivalent to 100msec of voice)	
Header compression	RTP/UDP/IP header compression	
SIP Message Compression	Shall support SigComp. Compression ratio = 1:8	
Call Type	Personal Talk (One-to-one)	Group Talk
Network Configuration	Early session procedure and auto answer	
Inactivity Timer Expire (sec)	15 sec	
Number of Transactions (talk and listen)	5 (ex, 2.5 times talk + 2.5 times listen on average)	5 (ex, 1 talk + 4 listens)
Talk burst duration (sec)	Exponential: mean=6 sec	Exponential: mean=6 sec
Each Volley Latency (sec)	2.75 sec	2.75 sec

3.5 Music/Speech Traffic Model (Class 3)

Listening tests have attempted to find the best-quality lossy audio codecs at certain bit rates. At high bit rates (>128 kbit/s), most people do not hear significant differences among most of the audio codecs. What is considered 'CD quality' is quite subjective; for some 128 kbit/s MP3 is sufficient, while for others 200 kbit/s or higher MP3 is necessary.

Though proponents of newer codecs such as WMA and RealAudio have asserted that their respective algorithms can achieve CD quality at 64 kbit/s, listening tests have shown otherwise; however, the quality of these codecs at 64 kbit/s is definitely superior to MP3 at the same bit rate.

MP3, which was designed and tuned for use alongside MPEG-1/2 Video, generally performs poorly on monaural data at less than 48 kbit/s or in stereo at less than 80 kbit/s

Because MP3 is a lossy format, it is able to provide a number of different options for its "bit rate" — that is, the number of bits of encoded data that are used to represent each second of audio. Typically, rates chosen are between 128 and 320 kilobit per second. By contrast, uncompressed audio as stored on a compact disc has a bit rate of 1411.2 kbit/s (16 bits/sample × 44100 samples/second × 2 channels). For average signals with good encoders, some listeners accept the MP3 bit rate of 128 kbit/s and the CD sampling rate of 44.1 kHz as near enough to compact disc quality for them, providing a compression ratio of approximately 11:1

3.5.1 Music/Speech traffic model: Application level model

Table 3.5.1: Music/Speech Traffic Model

Music/Speech Traffic Model	
Session Duration (sec)	1800 sec
Bit rate (kbps)	128 kbps
Protocols	TCP
Direction	Uni-direction (DL only)
Frames/sec	10 frames/sec

3.6 Video Clip Traffic Model (Class 3)

Based on the compression efficiency and market acceptance as described above, MPEG-4 has been selected for the video codec. The estimated values, for the parameters to model a video clip, vary from one trace to another. For parameters associated with the lognormal distributions, the estimates depend strongly on the dimensions of the captured frames. For the video clip traffic model, high quality-ARD Talk (German Sunday morning talk show) trace is described in [1]. In the model we have considered two different resolutions for the display; 176x144 for a small device and 320x240 for a large device. The required bandwidth for the uncompressed video stream with 176x144 pixels and

8 bit color depth is about 7.6 Mbps and with 320x240 pixels and 8 bit color depth is about 23 Mbps. The stored video clip will be transmitted like file transmission.

3.6.1 Video Clip traffic model: User/Application level model

Table 3.6.1: Video Clip Traffic Model

Video Clip Traffic Model		
Video clip length (sec)	Truncated exponential(15), Max = 60 sec	
Video Codec	MPEG4	
Protocols	TCP	
Direction	Uni-direction (UL or DL)	
Display size	176x144	320x240
Color depth (bit)	8	8
Subsampling method	4:1:1	4:1:1
Mean Uncompressed frame size (Kbyte)	38.016 Kbytes	115 Kbytes
Compression ratio	13.95	13.95
Mean compressed frame size (Kbyte)	2.725 Kbytes	65.950 Kbytes
Ex) file size for the 15 sec of video clip : – for (176x144) resolution = 2.725kB*25frames/sec*15sec = 1.022 MB – for (320x240) resolution = 8.243*25frames/sec*15sec = 3.91 MB		

3.7 Movie Streaming Traffic Model (Class 3)

Streaming media application can be classified into two types in terms of bandwidth requirement such as low (20 kbps~ 384 kbps) and moderate/high bandwidth (> 2 Mbps).

Based on the compression efficiency and market acceptance as described above, MPEG-4 has been selected for the video codec. The estimated values, for the parameters to model a video stream, vary from one trace to another. For parameters associated with the lognormal distributions, the estimates depend strongly on the dimensions of the captured frames. For the Movie Streaming traffic model, high quality movie trace, *Silence of the Lambs,* was used. In the model we have considered two different resolutions for the display; 176x144 for a small device and 320x240 for a large device. The required bandwidth for the uncompressed video stream with 176x144 pixels and 8 bit color depth is about 7.6 Mbps and with 320x240 pixels and 8 bit color depth is about 23 Mbps.

3.7.1 Movie Streaming traffic model: User/Application level model

Table 3.7.1: Movie Streaming Traffic Model

Movie Streaming Traffic Model		
Session Duration (sec)	3600	
Video Codec	MPEG4	
Protocols	TCP	
Scene Length (sec)	Lognormal(μ =5.1 sec , σ =9.05 sec)	
Direction	Uni-direction (DL only)	
Display size	176x144	320x240
Color depth (bit)	8	8
Subsampling method	4:1:1	4:1:1
Mean BW for uncompressed stream	7.6 Mbps	23 Mbps
Compression ratio	13.22	13.22
Mean BW for compressed stream	0.58 Mbps	1.74 Mbps
I frame size (byte)	Lognormal(μ =5640 , σ =2632)	Lognormal(μ =17068 ,σ =7965)
P frame size (byte)	Lognormal(μ =3037 , σ =2315)	Lognormal(μ =9190 ,σ =7005)
B frame size (byte)	Lognormal(μ =2260 ,σ =1759)	Lognormal(μ =6839 ,σ =5323)
AR coefficient	a_1=0.39, a_2=0.15, σ_ε = 4.36	a_1=0.39, a_2=0.15, σ_ε = 4.36

3.8 MBS Traffic Model (Class 3)

Based on the compression efficiency and market acceptance as described above, MPEG-4 has been selected for the video codec. The estimated values, for the parameters to model a video stream, vary from one trace to another. For parameters associated with the lognormal distributions, the estimates depend strongly on the dimensions of the captured frames. For the MBS traffic model, high quality Sports event trace, *Soccer game* (European championship 1996), was used. In the model we have considered two different resolutions for the display; 176x144 for a small device and 320x240 for a large device. The required bandwidth for the uncompressed video stream with 176x144 pixels and 8 bit color depth is about 7.6 Mbps and with 320x240 pixels and 8 bit color depth is about 23 Mbps.

3.8.1 MBS traffic model: User/Application level model

Table 3.8.1: MBS Traffic Model

MBS Traffic Model		
Session Duration (sec)	1800	
Number of channels	2	
Video Codec	MPEG4	
Protocols	RTP/UDP/IP	
Header compression	RTP/UDP/IP header compression	
Scene Length	Lognormal(μ =5.1 sec , σ =9.05 sec)	
Direction	Uni-direction (DL only)	
Display size	176x144	320x240
Color depth (bit)	8	8
Subsampling method	4:1:1	4:1:1
Mean BW for uncompressed stream	7.6 Mbps	23 Mbps
Compression ratio	6.87	6.87
Mean BW for compressed stream	1.1 Mbps	3.35 Mbps
I frame size (byte)	Lognormal(μ =19504 , σ =2213)	Lognormal(μ =59025 ,σ =6697)
P frame size (byte)	Lognormal(μ =9891, σ =2310)	Lognormal(μ =29933,σ =6990)
B frame size (byte)	Lognormal(μ =6496 , σ =1896)	Lognormal(μ =19658,σ =5737)
AR coefficient	a_1=0.39, a_2=0.15, σ_ε = 4.36	a_1=0.39, a_2=0.15, σ_ε = 4.36

3.9 IM Traffic Model (Class 4)

To understand IM traffic behavior we have conducted a lab test. The results are based on measurements performed on accessing the MSN IM server. Through the experiments, the key findings are:

1) There are enormous number of packets exchange between the user equipment and the IM server, even when the user is not sending or receiving any messages. The frequency of the packet exchange on the background is very high.
2) The periodic traffic pattern is very simple except advertisement related packets

Background Traffic

The following is the observed traffic behavior, which caused the periodic pattern (let's call it as heartbeat)

- The client sends a MSN packet (PNG) to a server every 40 to50 sec and the IP packet size is 45 bytes and the Server send an Ack packet back to the client
- After the Ack packet, the server sends QNG packet to the client about 3 sec later. The client also replies to this QNG packet (48 bytes) with Ack (40 bytes) packet to the server.
- If PNG message does not get Ack then the client retransmit the PNG packet after 5 sec, 10sec, and 20 sec.

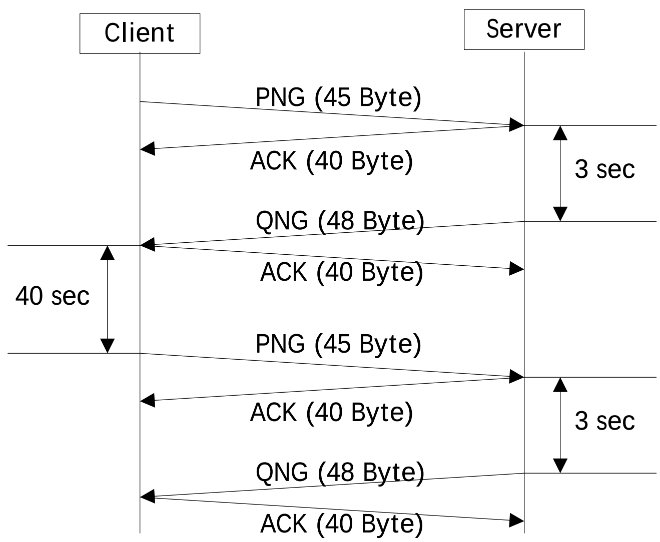


Figure 3.9.1: IM Heartbeat Flow Every 40 sec

Foreground User Traffic

The user traffic is composed of two main events, which triggers the packet exchange.

- User Message send
- User status change (ex, online. away, busy, on the phone, etc)

User Message:

While a user is typing a message, a typing indication packet is sent to the other party every 5 sec (In Exchange 2000 Server these notifications are sent every four seconds) and the typing indication packet transmission will be stopped when the user stops typing or the message has sent. 187 bytes of IP packet was observed when a user sent a message of 10 characters and the sender received an ACK packet from the server (40 Bytes of IP packet). This implies that the message packet size is 177 bytes of message header plus actual user message size (10 bytes in this case). The 177 bytes of message header size may vary depend on the length of receiver's ID.

User Status Change:

The presence update is not actively updated by the server but a MSN client probe the server every 40 sec to update the other member's status. The number of packets exchanged related to any type of status change is 4 packets and those 4 packets are composed of:

1. Client sends status change packet to server 317 bytes of IP packet
2. Server sends ACK 40 bytes of IP packet
3. Server sends the exactly same packet which client has sent back to the client (317 Bytes)
4. Client sends an ACK packet (40 Bytes)

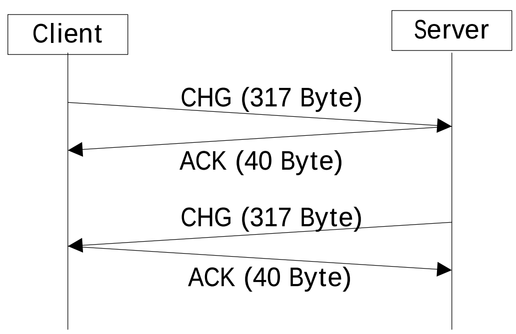


Figure 3.9.2: User Status Change Packet Flow Whenever the Event Occurs

3.9.1 IM traffic model: IP level model

This model, Table 3.9.1, is a summary of the IM model above and this model does not include the user traffic model because the frequency of the events is rare compare to the IM background traffic. As an example there are about 20 user traffic related events (instant messages and status change messages) per day and there are more than 2000 heartbeats per day.

Table 3.9.1: Instant Messaging Traffic Model

Instant Messaging Traffic Model	
Avg. BHDSA	N/A (Automatic login when a user device is ON)

Number of Instant Messages/session	Ignore (negligible)
Presence Update or Status Check	Exchange four packets between client and server every 40 seconds (see Figure 3.9.2)

3.10 Web Browsing (HTTP) Traffic Model

The web model is one of the most complex models of all. Measurements of HTTP traffic show that the large majority of page responses consist of relatively small objects. The distribution of page sizes however is heavy-tailed, i.e., infrequent but very large page objects constitute a significant proportion of overall transmitted bytes. Each web page consists of a number of web objects such as a main page, embedded images, style sheets and executable java applets or plug-ins. The time between accessing two pages is denoted the think time (reading time) and includes time for the user to read all or parts of the page.

Since the web browsing model in [4] might be outdated, a recent study has been incorporated in [1].

3.10.1 Web Browsing: User level model

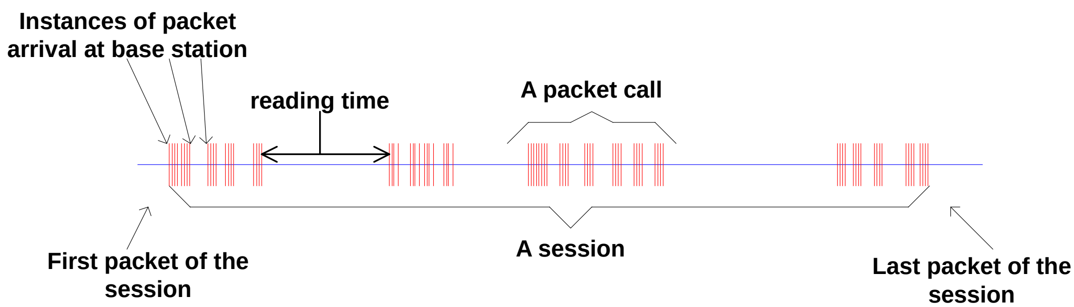


Figure 3.10.1: Packet Trace of a Typical Web Browsing Session

Figure 3.10.1 shows the packet trace of a typical web browsing session. The session is divided into ON/OFF periods representing web-page downloads and the intermediate reading times, where the web-page downloads are referred to as packet calls. These ON and OFF periods are a result of human interaction where the packet call represents a user's request for information and the reading time identifies the time required to digest the web-page.

As is well known, web-browsing traffic is self-similar. In other words, the traffic exhibits similar statistics on different timescales. Therefore, a packet call, like a packet session, is divided into ON/OFF periods as in Figure 3.10.1. Unlike a packet session, the ON/OFF periods within a packet call are attributed to machine interaction rather than human interaction. A web-browser will begin serving a user's request by fetching the initial HTML page using an HTTP GET request. The retrieval of the initial page and each of the constituent *objects* is represented by ON period within the packet call while the parsing time and protocol overhead are represented by the OFF periods within a packet call. For simplicity, the term "page" will be used in this paper to refer to each packet call ON period.

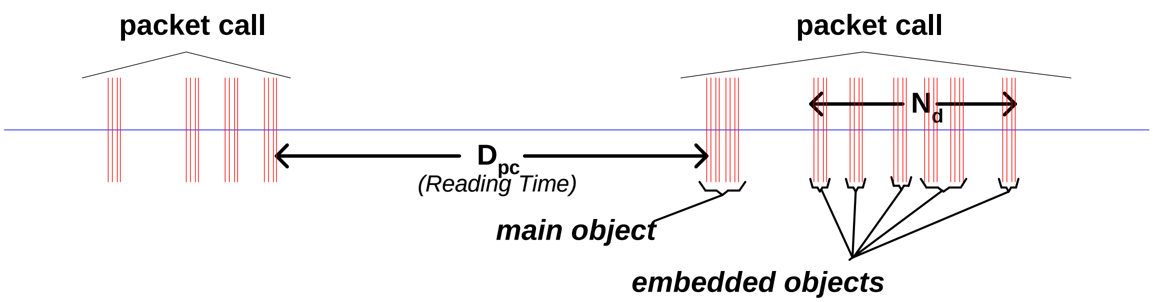


Figure 3.10.2: Contents in a Packet Call

The parameters for the web browsing traffic are as follows:

- - S_M: Size of the main object in a page
- - S_E: Size of an embedded object in a page
- - N_d: Number of embedded objects in a page
- - D_{pc}: Reading time
- - T_p: Parsing time for the main page

HTTP/1.1 persistent mode transfer is used to download the objects, which are located at the same server and the objects are transferred serially over a single TCP connection. The distributions of the parameters for the web browsing traffic model are described in Table 3.10.1 which has been obtained through a set of measurements using a recent online-traffic analysis[1].

Table 3.10.1: Web Browsing Traffic Model

Web Browsing Traffic Model		
Component	**Distribution**	**Parameters**
Number of Pages per session	Lognormal	Mean = 17 pages SD = 22 pages
Page request size	constant	350 B
Main object size (S_M)	**Truncated Lognormal**	**Mean = 52390bytes** **SD= 49591bytes** **Min = 1290bytes** **Max = 0.25Mbytes**
Embedded object size (S_E)	Truncated Lognormal	**Mean = 8551bytes** **SD = 59232bytes** **Min = 5bytes** **Max = 6Mbytes**
Number of embedded objects per page (N_d)	**Truncated Pareto**	**Mean = 51.1** **Max = 165**
Reading time (D_{pc})	Exponential	Mean = 30 sec
Parsing time (T_p)	Exponential	Mean = 0.13 sec

3.10.2 Web Browsing: IP packet level model

HTTP/1.1 persistent mode transfer is used to download the objects, which are located at the same server and the objects are transferred serially over a single TCP connection as modelled in [4]. Based on observed packet size distributions, 76% of the HTTP packet calls should use an MTU of 1500 bytes, with the remaining 24% of the HTTP packet calls using an MTU of 576 bytes. These two potential packet sizes also include a 40 byte IP packet header (thereby resulting in useful data payloads of 1460 and 536 bytes, respectively), and this header overhead for the appropriate number of packets must be added to the object data sizes calculated from the probabilistic distributions in Table 3.10.1.

3.10.3 HTTP Traffic Model [3GPP]

The HTTP traffic parameters used by 3GPP are listed in Table 3.10.2. Results for 3GPP and 3GPP2 are already available for this model and thus this model has been included here to allow comparing WiMAX performance with existing results.

Based on observed packet size distributions, 76% of the HTTP packet calls should use an MTU of 1500 bytes, with the remaining 24% of the HTTP packet calls using an MTU of 576 bytes. These two potential packet sizes also include a 40 byte IP packet header (thereby resulting in useful data payloads of 1460 and 536 bytes, respectively), and this header overhead for the appropriate number of packets must be added to the object data sizes calculated from the probabilistic distributions in Table 3.10.2.

Table 3.10.2: HTTP Traffic Model Parameters [3GPP]

Component	Distribution	Parameters	PDF
Main object size (S_M)	Truncated Lognormal	Mean = 10710 bytes Std. dev. = 25032 bytes Minimum = 100 bytes Maximum = 2 Mbytes	$f_x = \frac{1}{\sqrt{2\pi}\sigma x}\exp\left[\frac{-(\ln x-\mu)^2}{2\sigma^2}\right], x \ge 0$ $\sigma = 1.37, \mu = 8.35$
Embedded object size (S_E)	Truncated Lognormal	Mean = 7758 bytes Std. dev. = 126168 bytes Minimum = 50 bytes Maximum = 2 Mbytes	$f_x = \frac{1}{\sqrt{2\pi}\sigma x}\exp\left[\frac{-(\ln x-\mu)^2}{2\sigma^2}\right], x \ge 0$ $\sigma = 2.36, \mu = 6.17$
Number of embedded objects per page (N_d)	Truncated Pareto	Mean = 5.64 Max. = 53	$f_x = \frac{\alpha k^\alpha}{x^{\alpha+1}}, k \le x < m$ $f_x = \left(\frac{k}{m}\right)^\alpha, x = m$ (NOTE) $\alpha = 1.1, k = 2, m = 55$
Reading time (D_{pc})	Exponential	Mean = 30 sec	$f_x = \lambda e^{-\lambda x}, x \ge 0$ $\lambda = 0.033$
Parsing time (T_p)	Exponential	Mean = 0.13 sec	$f_x = \lambda e^{-\lambda x}, x \ge 0$ $\lambda = 7.69$
Note: Subtract k from the generated random value to obtain N_d			

3.11 Email Traffic Model (Class 4)

The mostly used email protocols are POP3 and MAPI which is supported by Outlook and Exchange server. We describes a behavior of Outlook and Exchange Server for the e-mail application in this section since this is one of the major email applications in the market and the behavior is not widely known compare to POP3 protocol. At the session layer of the protocol stack, Messaging Application Programming Interface (MAPI) protocol is central to the operation of Microsoft Outlook and Exchange Server.

Outlook and Exchange Server use a Microsoft proprietary protocol (Messaging Application Programming Interface) in the application layer and the description is not available. Thus, we need to understand the details of the traffic behavior specific to the Outlook application. Through the experiments, our key findings are 1) Both Outlook and Exchange server compress e-mail body. 2) An email transaction is consists of multiple sub-transactions.

First of all, Outlook depends upon Remote Procedure Call (RPC) for its transport and uses a somewhat chatty protocol. During the Outlook invoke phase, a lot of the chattiness is observed. There are eleven active TCP connections during the invoking phase. Each Outlook email transaction is consisting of multiple MAPI segment transactions in series and each MAPI fragment is segmented into smaller segments again. During the test, most of the time, the maximum MAPI fragment was 16896 bytes and this information is indicated in the first packet of a MAPI fragment. The maximum size of the transmitted and received MAPI sub-fragment is negotiated during the invoking phase (5840 bytes was observed). Outlook finishes all the ACK packet transmission for the current MAPI segment and the Exchange server waits for the MAPI fragment completion indication packet before sending the next one. The last packet in the MAPI fragment sets the "PUSH" bit in the TCP packet to transmit all of the packets in the TCP buffer to the application layer at the receiver side.

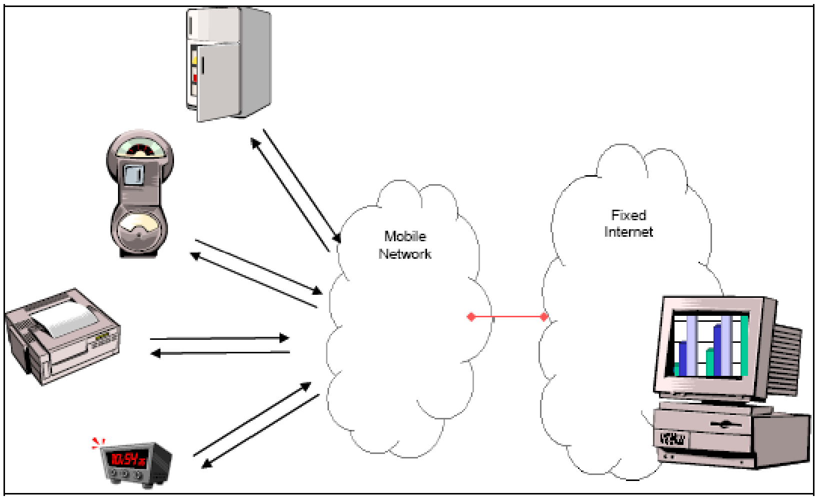
3.11.1 Email: User/Application level model

Table 3.11.1: E-mail traffic model

Email Traffic Model		
Component	**Distribution**	**Parameters**
E-mail Protocol	N/A	POP3, MAPI
Number of email receive	Lognormal	Mean = 30, SD.=17
Number of email send	Lognormal	Mean = 14, SD.=12
Reading Time (sec)	Pareto	Mean = 60,
Writing Time (sec)	Pareto	Mean = 120
Average E-mail header size	Constant	1 Kbytes
E-mail size (This includes e-mails w/ and w/o file attachment)	Cauchy distribution	Mean = 22.7 Kbytes, SD=200.3 Kbytes 90%-tile = 80 KBytes

3.12 Telemetry Traffic Model (Class 4)

Telemetry is defined as low bandwidth Machine-to-Machine initiated communications, monitoring or tracking of stationary objects. Telemetry does not include people and machines, and telemetry service is one way. The assumed symmetry of Telemetry is illustrated in Figure 3.12.1, and the traffic model is described in Table 3.12.1.



Source: Telecompetition, Inc. May 2003.

Figure 3.12.1 Telemetry symmetry

3.12.1 Telemetry: User/Application level model

Table 3.12.1: Telemetry Traffic Model

Telemetry Traffic Model	
Frequency of use	24 messages per day (every hour 24x7)
Message Size (byte)	10

3.13 FTP Traffic Model (Class 5)

3.13.1 FTP traffic model: User level model

In FTP applications, a session consists of a sequence of file transfers, separated by *reading times*, and The underlying transport protocol for FTP is TCP. The two main parameters of an FTP session are:

- 1. S : the size of a file to be transferred

- 2. D_{pc}: reading time, i.e., the time interval between end of download of the previous file and the user request for the next file.

The parameters for the FTP application sessions are described in Table 3.13.1.

Table 3.13.1: FTP Traffic Model Parameters [3GPP2]

Component	Distribution	Parameters	PDF
File size (S)	Truncated Lognormal	Mean = 2Mbytes Std. Dev. = 0.722 Mbytes Maximum = 5 Mbytes	$f_x = \frac{1}{\sqrt{2\pi}\sigma x}\exp\left[\frac{-(\ln x-\mu)^2}{2\sigma^2}\right], x \geq 0$ $\sigma = 0.35, \mu = 14.45$
Reading time (D_{pc})	Exponential	Mean = 180 sec.	$f_x = \lambda e^{-\lambda x}, x \geq 0$ $\lambda = 0.006$

3.13.2 FTP traffic model: IP level model

The underlying transport protocol for FTP is TCP. The model of TCP connection described in [4] will be used to model the FTP traffic. The packet trace of an FTP session is shown in Figure 3.13.1.

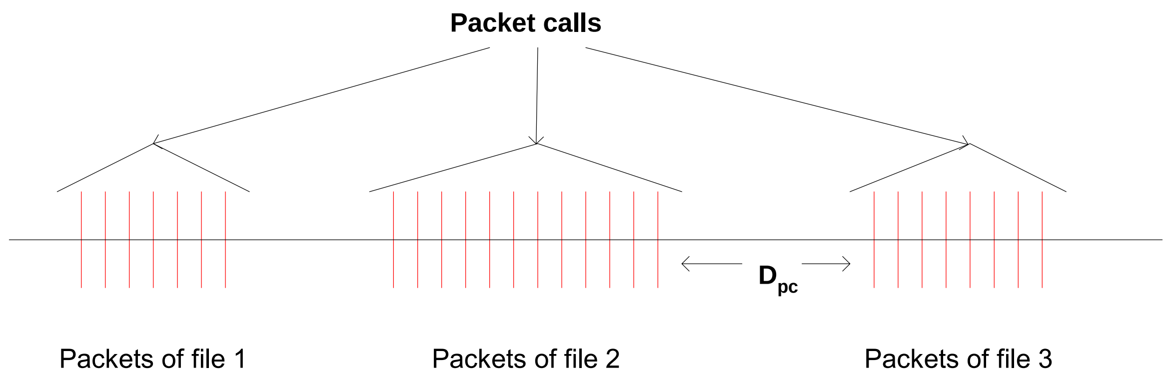


Figure 3.13.1: Packet Trace in a Typical FTP Session

Based on the results on packet size distribution, 76% of the files are transferred using and MTU of 1500 bytes and 24% of the files are transferred using an MTU of 576 bytes. Note that these two packet sizes also include a 40 byte IP packet header (thereby resulting in useful data payloads of 1460 and 536 bytes, respectively) and this header overhead for the appropriate number of packets must be added to the file sizes calculated from the probabilistic distributions in Table 3.13.1. For each file transfer a new TCP connection is used whose initial congestion window size is 1 segment (i.e., MTU).

3.14 P2P Traffic Model (Class 5)

A peer-to-peer (or P2P) computer network is a network that relies primarily on the computing power and bandwidth of the participants in the network rather than concentrating it in a relatively low number of servers. P2P networks are typically used for connecting nodes via largely ad hoc connections. Such networks are useful for many purposes. Sharing content files containing audio, video, data or anything in digital format is very common, and real-time data, such as telephony traffic, is also passed using P2P technology.

File sharing is the practice of making files available for other users to download over the Internet and smaller networks and usually it follows the peer-to-peer (P2P) model, where the files are stored on and served by personal computers of the users. Most people who engage in file sharing are also downloading files that other users share. Sometimes these two activities are linked together. P2P File sharing is distinct from file trading in that downloading files from a P2P network does not require uploading, although some networks either provide incentives for uploading such as credits or force the sharing of files being currently downloaded. We assume that a P2P user disables uploading for wireless device, and assuming that the average DL rate is about 500 kbps.

3.14.1 P2P traffic model: User/Application level model

Table 3.14.1: P2P traffic model

P2P Traffic Model	
Session Duration (sec)	1800 sec
Average DL bit rate	500 kbps
Direction	DL only
Protocol	TCP
Note: use FTP protocol for the application layer.	

3.15 VPN Service

VPN service is common for enterprise, and Network service providers and enterprise network managers face the same sticky problem: how to overcome bottlenecks to increase throughput of information over virtual private networks (VPNs). Bandwidth bottlenecks typically occur at points of traffic concentration or a link which has a limited bandwidth such as wireless link.

A significant amount of VPN bandwidth originates either from telecommuters and mobile workers. Effectively, the load placed on a networking device or a VPN is the total amount of IPSec bandwidth needing to be processed. This bandwidth is the number of traffic packets per second that need to be encrypted and authenticated and the number of individual user sessions that need to be maintained.

Data compression is a vital element in VPN services for making data packets as small as possible. Compression is not a cryptographic function like encryption and authentication. However, it is highly desirable for VPNs that use IPSec encryption and/or authentication.

To the service provider, data compression provides four primary benefits, which in turn, are passed on to the VPN user.

- First, compressed packets consume less bandwidth.
- Next, compression reduces the latency of packets as they traverse the network, since packet length is shorter.
- Thirdly, performance is significantly enhanced.
- And lastly, applying compression to the data before it is encrypted improves its resistance to cryptanalysis. Cryptanalysis is the process of attempting to find a shortcut method, not envisioned by the designer, for decrypting an encrypted message when the key used to encrypt the message is not known.

LZS is one of the IP payload compression methods, and the strength of LZS is that it produces an optimum combination of compression and performance. LZS achieves "lossless compression" which simply means no data is lost during compression and decompression. Lossless compression reduces data typically by about 1/2 - but nothing is ever lost.

3.15.1 Compression efficiency versus datagram size

The following table offers some guidance on the compression efficiency that can be achieved as a function of datagram size.

Table 3.14.2: Compression Efficiency versus Datagram Size

Datagram Size(Byte)	64	128	256	512	1024	2048	4096	8192	16384
Compression Ratio	1.18	1.28	1.43	1.58	1.74	1.91	2.04	2.11	2.14

Source: RFC 2395

3.15.2 VPN Background Traffic Model

Understanding VPN background traffic behaviour is very importnat when deciding WiMAX network parameters such as sleep mode and idle mode related timeout values which are critical for the power saving. The key findings are:

1) There are several key network elements communicating with user equipment, even when a user is not utilizing any applications. The amount of background traffic is significant and should be accounted for in estimating WiMAX network/system capacity.

2) The traffic has complex and mixed periodicities, but still it is not random.

3) The observed background traffic from different types of access networks differ both in terms of periodicity and volume. Thus, the VPN background traffic study should be done through the specific type of access network of interest.

4) The results indicate that the VPN background traffic is strongly dependent on how an Enterprise network is configured. Hence, the traffic measurement should be done over the real target network.

3.16 NRTV (Near Real Time Video) Traffic Model [3GPP]

This traffic model used in 3GPP system level simulations is included here so that it can be used in comparing WiMAX systems with 3GPP systems. This section describes a model for streaming video traffic on the forward link. Figure 3.16.1 describes the steady state of video streaming traffic from the network, as seen by the base station. Latency at call start-up is not considered in this steady-state model.

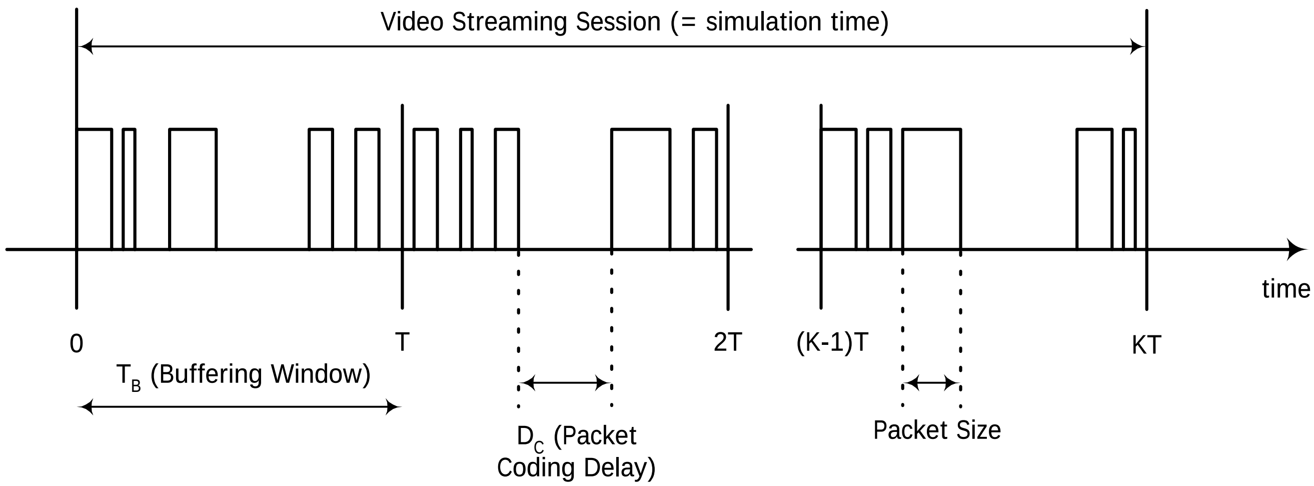


Figure 3.16.1: Video Streaming Traffic Model

A video streaming session is defined as the entire video streaming call time, which is equal to the simulation time for this model. Each frame of video data arrives at a regular interval T determined by the number of frames per second (fps). Each frame is decomposed into a fixed number of slices, each transmitted as a single packet. The size of these packets/slices is distributed as a truncated Pareto distribution. Encoding delay, Dc, at the video encoder introduces delay intervals between the packets of a frame. These intervals are modelled by a truncated Pareto distribution.

The parameter T_B is the length (in seconds) of de-jitter buffer window in the mobile station, and is used to guarantee a continuous display of video streaming data. This parameter is not relevant for generating the traffic distribution, but it is useful for identifying periods when the real-time constraint of this service is not met. At the beginning of the simulation, it is assumed that the mobile station de-jitter buffer is full with (T_B x source video data rate) bits of data. Over the simulation time, data is "leaked" out of this buffer at the source video data rate and "filled" as forward link traffic reaches the mobile station. As a performance criterion, the mobile station can record the length of time, if any, during which the de-jitter buffer runs dry. The de-jitter buffer window for the video streaming service is 5 seconds.

Using a source video rate of 64 kbps, the video traffic model parameters are defined in Table 3.16.1.

Table 3.16.1: Video Streaming Traffic Model Parameters.

Information types	Inter-arrival time between the beginning of each frame	Number of packets (slices) in a frame	Packet (slice) size	Inter-arrival time between packets (slices) in a frame
Distribution	Deterministic (Based on 10fps)	Deterministic	Truncated Pareto (Mean= 50bytes, Max= 250bytes)	Truncated Pareto (Mean= 6ms, Max= 12.5ms)
Distribution Parameters	100ms	8	K = 40 bytes $\alpha = 1.2$	K = 2.5ms $\alpha = 1.2$

References

[1] Bong Ho Kim, Application Traffic Model, WiMAX Forum AWG Contribution, May 2007

[2] Bong Ho Kim et al., Performance metrics and guidelines for the five application classes, WiMAX Forum AWG Contribution, Aug. 2005

[3] Bong Ho Kim, Michael Librizzi, WiMAX Application Usage Profile, WiMAX Forum Application Business Working Group, April 2006

[4] 3GPP2-TSGC5, HTTP and FTP Traffic Model for 1xEV-DV Simulations, 3GPP2-C50-EVAL-2001022-0xx

[5] 3GPP TS 26.101, AMR Speech Codec Frame Structure (Release 1999)

[6] A. Estepa, R. Estepa, and J. Vozmediano, :A new Approach for VoIP Traffic Characterization," IEEE Communications Letters, Vol. 8, pp 644-646, 2004.

[7] Bong Ho Kim, David Houck, Insup Lee et al."Performance Analysis of AAL2 switching/multiplexing in the Iub interface in UMTS Network", International conference on Computer, Communication and Control Technologies, pp. 133-138, 2003

[8] M. Krunz and S. Tripathi, "On the Characterization of VBR MPEG Stream," ACM, pp. 192-202, 1997

[9] A. Matrawy, I. Lambadaris, and C. Huang, "MPEG4 Traffic Modeling Using The Transform Expand Sample Methodology," Proceedings of the IEEE 4th International Workshop on Networked Appliances, pp. 249-256, 2001.

[10] Ericsson, Motorola, Nokia, Siemens, Push-to-Talk over Cellular (PoC) Signalling Flow; Release 1.1.3, August 2003

[11] Push-to-Talk over Cellular (PoC) User Plane; PoC release 1.0

4. MAC Layer Modelling

The main focus of MAC layer is to manage the resources of the air link in an efficient manner. The MAC layer schedules the usage of the air link resources and provides QoS differentiation. It performs link adaptation and ARQ functions to maintain target BER while maximizing the data throughput. The MAC layer also handles network entry for SSs that enter and leave the network, and it performs MAC PDU creation tasks. Finally, the MAC layer provides a convergence sub layer that supports ATM and packet-based network layers [White Paper Part I].

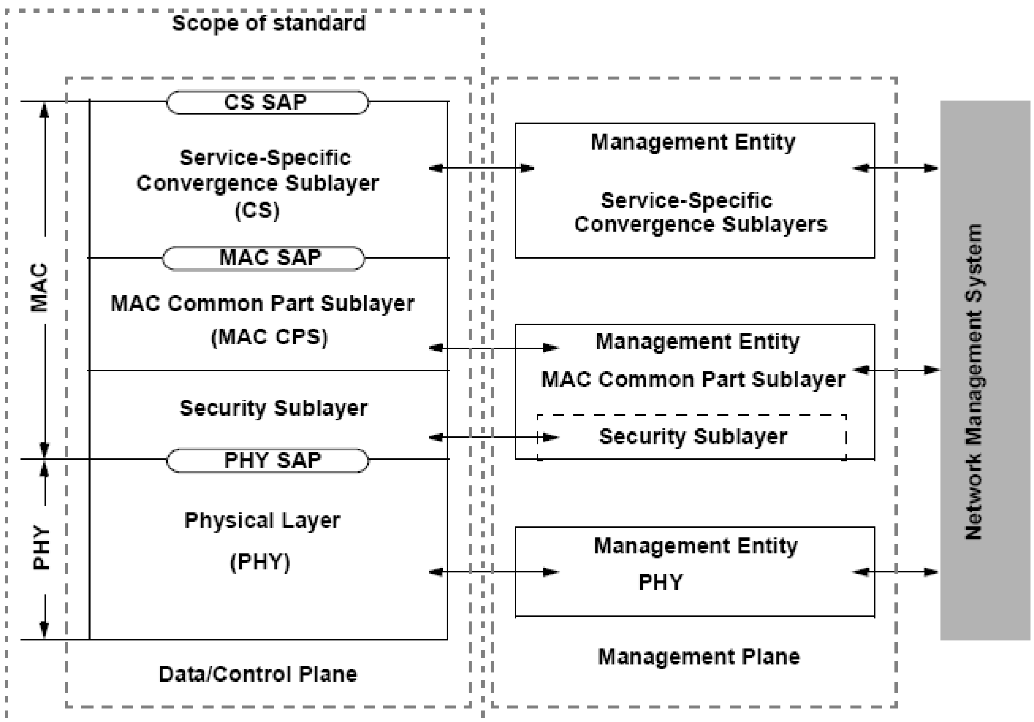


Figure 4.1: MAC function diagram [802.16-2004]

The 802.16 standard was developed from the outset for the delivery of broadband services including voice, data, and video. The MAC layer can support bursty data traffic with high peak rate demand while simultaneously supporting streaming video and latency-sensitive voice traffic over the same channel. The resource allocated to one terminal by the MAC scheduler can vary from a single time slot to the entire frame, thus providing a very large dynamic range of throughput to a specific user terminal at any given time. Furthermore, since the resource allocation information is conveyed in the MAP messages at the beginning of each frame, the scheduler can effectively change the resource allocation on a frame-by-frame basis to adapt to the bursty nature of the traffic [White Paper Part I].

With fast air link, asymmetric downlink/uplink capability, fine resource granularity and a flexible resource allocation mechanism, Mobile WiMAX networks can meet QoS requirements for a wide range of data services and applications. In the Mobile WiMAX MAC layer, QoS is provided via service flows as illustrated in Figure 4.1. This is a unidirectional flow of packets that is provided with a particular set of QoS parameters. Before providing a certain type of data service, the base station and user-terminal first establish a unidirectional logical link between the peer MACs called a connection. The outbound MAC then associates packets traversing the MAC interface into a service flow to be delivered over the connection. The QoS parameters associated with the service flow define the transmission ordering and scheduling on the air interface. The connection-oriented QoS therefore, can provide accurate control over the air interface. Since the air interface is usually the bottleneck, the connection-oriented QoS can effectively enable the end-to-end QoS control. The service flow parameters can be dynamically managed through MAC messages to accommodate the dynamic service demand. The service flow based QoS mechanism applies to both DL and UL to provide improved QoS in both directions. Mobile WiMAX systems support a wide range of data services and applications with varied QoS requirements.

4.1 Convergence Sublayer

The 802.16 MAC layer provides a convergence sublayer (CS) for the transport of ATM cells and IP packets, namely ATM CS and packet CS.

The packet CS resides on top of the IEEE Std 802.16 MAC CPS. The CS performs mainly the following functions, utilizing the services of the MAC:

a) Classification: classifying of the higher-layer protocol PDU into the appropriate transport connection, and associate them to the proper MAC service flow identifier (SFID) and connection identifier (CID).
b) Payload header Suppression (PHS) and de-suppression (optional)
c) Delivery and receipt of the resulting CS PDU (MAC SDU) to and from the peer MAC

Once classified and associated with a specific MAC connection, higher-layer PDUs shall be encapsulated in the MAC SDU format as illustrated in Figure below. The 8-bit Payload Header Suppression Index (PHSI) field is only present when a PHS rule has been defined for the associated connection.

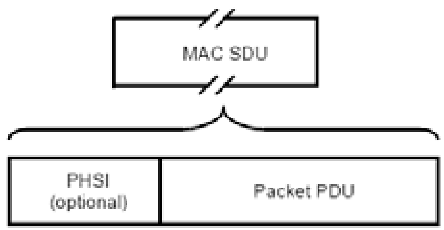


Figure 4.1.1: MAC SDU format [802.16]

4.1.1 Classification

Classification is the process by which a MAC SDU is mapped onto a particular transport connection and service flow for transmission between MAC peers. The mapping process associates a MAC SDU with a connection (CID) and a service flow (SFID) with the service flow characteristics of that connection. This process facilitates the delivery of MAC SDUs with the appropriate QoS constraints.

Classification is implemented by downlink classifiers at the BS (Figure 4.1.2) to packets it is transmitting and uplink classifiers (figure 4.4) at the SS. A classifier is a set of matching criteria. It consists of some protocol-specific packet matching criteria, a classifier priority, and a reference to a CID, SFID and its service flow characteristics.

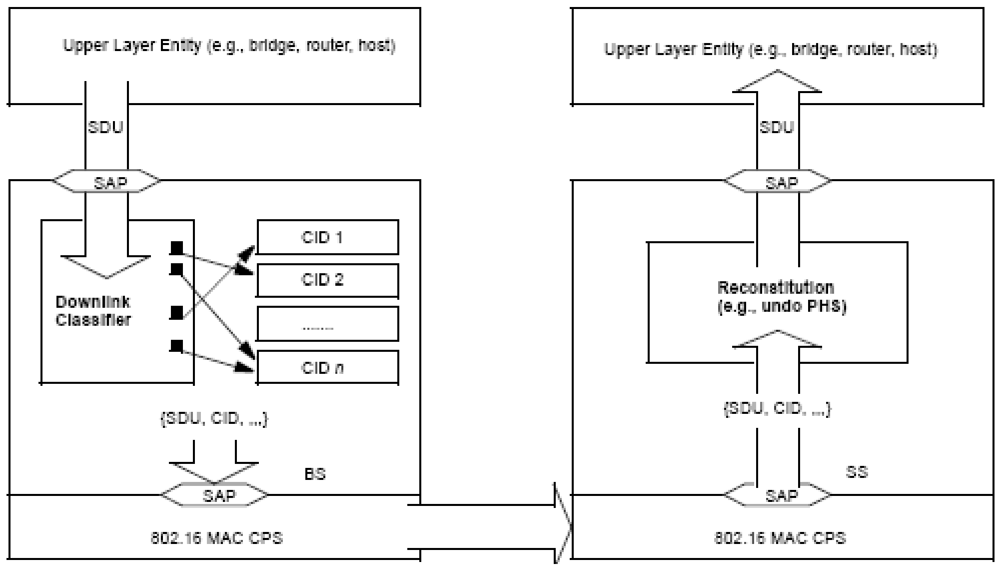


Figure 4.1.2 Classification and CID Mapping (BS to SS) [1]

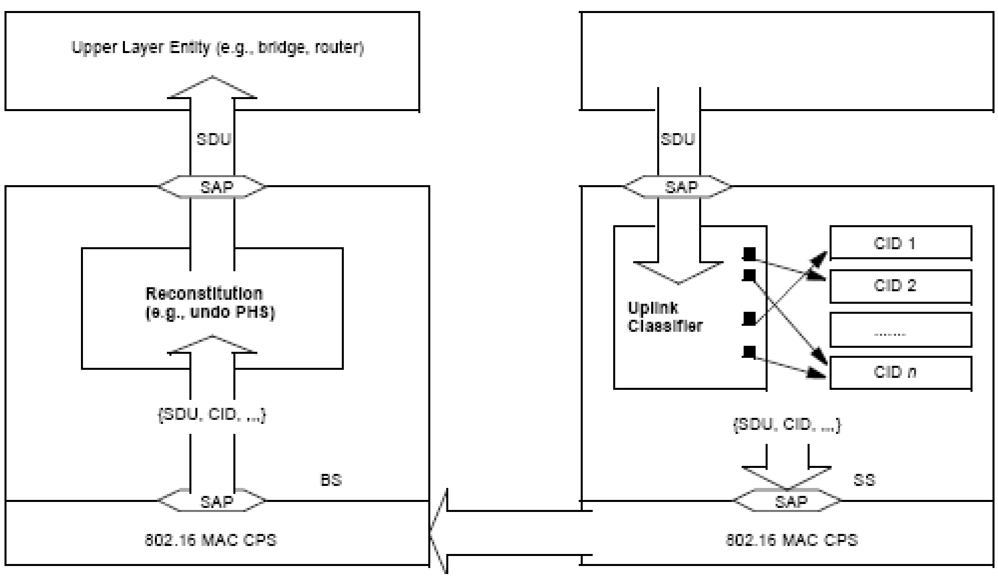


Figure 4.1.3: Classification and CID Mapping (SS to MS) [1]

Packet CS is used for transport for all packet-based protocols as specified in spec [802.16e-2005] as followings.

- Packet, IPv4 (Support Only)
- Packet, Ipv6
- Packet, IEEE 802.3/Ethernet
- Packet, IEEE 802.1Q VLAN

- Packet, Ipv4 over IEEE 802.3/Ethernet
- Packet, Ipv6 over IEEE 802.3/Ethernet
- Packet, Ipv4 over IEEE 802.1Q VLAN
- Packet, Ipv6 over IEEE 802.1Q VLAN
- Packet, IEEE 802.3/Ethernet with ROHC header compression
- Packet, IEEE 802.3/Ethernet with ECRTP header compression
- Packet, IP2 with ROHC header compression
- Packet, IP2 with ECRTP header compression

Only packet Ipv4 classifier is supported in the current simulator. IP classifiers operate on the fields of the IP header and the transport protocol. The parameters may be used in IP classifiers are defined in [802.16-2004] (11.13.19.3.4.2–11.13.19.3.4.7).

- IP ToS/DSCP range and mask
- IP masked source address
- IP masked destination address
- Protocol source port range
- Protocol destination port range

4.1.2 Packet Header Suppression

In PHS, a repetitive portion of the payload headers of the higher layer is suppressed in the MAC SDU by the sending entity and restored by the receiving entity. Implementation of PHS capability is optional. On the uplink, the sending entity is the SS and the receiving entity is the BS. On the downlink, the sending entity is the BS and the receiving entity is the SS. If PHS is enabled at MAC connection, each MAC SDU is prefixed with a PHSI, which references the Payload Header Suppression Field (PHSF).

The sending entity uses classifiers to map packets into a service flow. The classifier uniquely maps packets to its associated PHS Rule. The receiving entity uses the CID and the PHSI to restore the PHSF. Once a PHSF has been assigned to a PHSI, it shall not be changed. To change the value of a PHSF on a service flow, a new PHS rule shall be defined, the old rule is removed from the service flow, and the new rule is added. When a classifier is deleted, any associated PHS rule shall also be deleted.

PHS has a Payload Header Suppression Valid (PHSV) option to verify or not verify the payload header before suppressing it. PHS has also a Payload Header Suppression Mask (PHSM) option to allow select bytes not to be suppressed. The PHSM facilitates suppression of header fields that remain static within a higher-layer session (e.g. IP addresses), while enabling transmission of fields that change from packet to packet (e.g. IP Total Length).

The BS shall assign all PHSI values just as it assigns all CID values. Either the sending or the receiving entity shall specify the PHSF and the Payload Header Suppression Size (PHSS). This provision allows for preconfigured headers or for higher level signaling protocols outside the scope of this standard to establish cache entries.

It is the responsibility of the higher-layer service entity to generate a PHS Rule that uniquely identifies the suppressed header within the service flow. It is also the responsibility of the higher-layer service entity to guarantee that the byte strings that are being suppressed are constant from packet to packet for the duration of the active service flow.

PHS operation

A packet is submitted to the packet CS. The SS applies its list of Classifier rules. A match of the rule shall result in an Uplink Service Flow, CID, and a PHS Rule. The PHS Rule provides PHSF, PHSI, PHSM, PHSS, and PHSV. If PHSV is set or not present, the SS shall compare the bytes in the packet header with the bytes in the PHSF that are to be suppressed as indicated by the PHSM. If they match, the SS shall suppress all the bytes in the Uplink PHSF except the bytes masked by PHSM. The SS shall then prefix the PDU with the PHSI and present the entire MAC SDU to the MAC SAP for transport on the uplink.

When the MAC PDU is received by the BS from the air interface, the BS MAC layer shall determine the associated CID by examination of the generic MAC header. The BS MAC layer sends the PDU to the MAC SAP associated with that CID. The receiving packet CS uses the CID and the PHSI to look up PHSF, PHSM, and PHSS. The BS reassembles the packet and then proceeds with normal packet processing. The reassembled packet contains bytes from the PHSF. If verification was enabled, then the PHSF bytes equal the original header bytes. If verification was not enabled, then there is no guarantee that the PHSF bytes match the original header bytes.

A similar operation occurs on the downlink. The BS applies its list of Classifiers. A match of the Classifier shall result in a Downlink Service Flow and a PHS Rule. The PHS Rule provides PHSF, PHSI, PHSM, PHSS, and PHSV. If PHSV is set or not present, the BS shall verify the Downlink Suppression Field in the packet with the PHSF. If they match, the BS shall suppress all the bytes in the Downlink Suppression Field except the bytes masked by PHSM. The BS shall then prefix the PDU with the PHSI and present the entire MAC SDU to the MAC SAP for transport on the downlink.

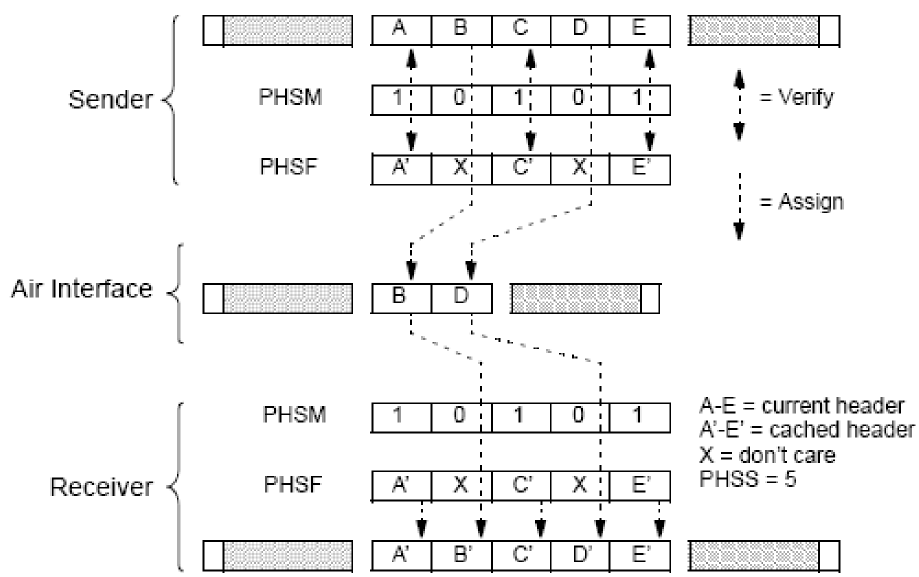


Figure 4.1.4: PHS operation [802.6-2004]

The format of the IP CS PDU shall be as shown in Figure 4.6 (when header suppression is enabled at the connection, but not applied to the CS PDU) or Figure 17 [802.16-2004] (with header suppression). In the case where PHS is not enabled, PHSI field shall be omitted.

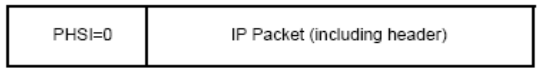


Figure 16—IP CS PDU format without header suppression

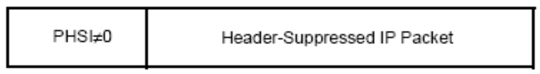


Figure 17—IP CS PDU format with header suppression

Figure 4.1.5: IP CS PDU [802.16-2004]

4.2 MAC PDU formats

MAC PDUs shall be of the form illustrated in Figure 4.7. Each PDU shall begin with a fixed-length generic MAC header. The header may be followed by the Payload of the MAC PDU. If present, the Payload shall consist of zero or more subheaders and zero or more MAC SDUs and/or fragments thereof. The payload information may vary in length, so that a MAC PDU may represent a variable number of bytes. This allows the MAC to tunnel various higher-layer traffic types without knowledge of the formats or bit patterns of those messages.

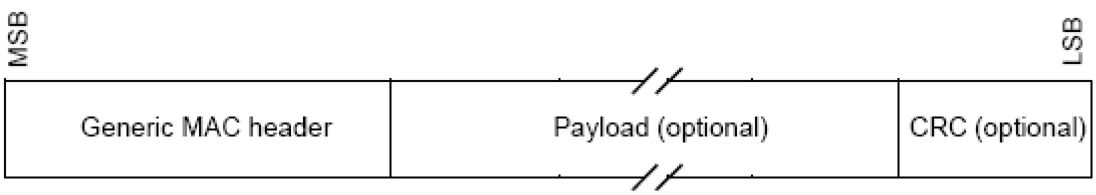


Figure 4.2.1: MAC PDU Format

4.2.1 MAC Headers and Sub-headers

Headers

For DL, generic MAC header (6 bytes) begins each MAC PDU containing either MAC management messages or CS data.

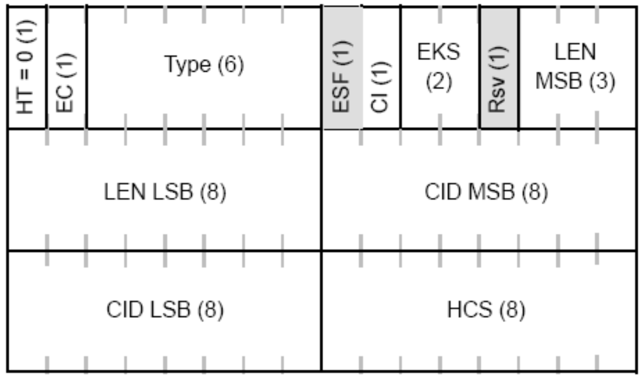


Figure 4.2.2: Generic MAC header format (Fig 19 [802.16-2005])

For UL, There are two defined UL MAC header formats. The first is the generic MAC header that begins each MAC PDU containing either MAC management messages or CS data, where HT is set to 0. The second is the MAC header format without payload where HT is set to 1. For the latter format, the MAC header is not followed by any MAC PDU payload and CRC. It can be MAC signaling header type I or type II, both of 6 bytes long.

Bandwidth Request Header is one of the most commonly used UL MAC signaling header type I.

- The CID shall indicate the connection for which uplink bandwidth is requested.
- The Bandwidth Request (BR) field shall indicate the number of bytes requested.
- The allowed types for bandwidth requests are “000” for incremental and “001” for aggregate.

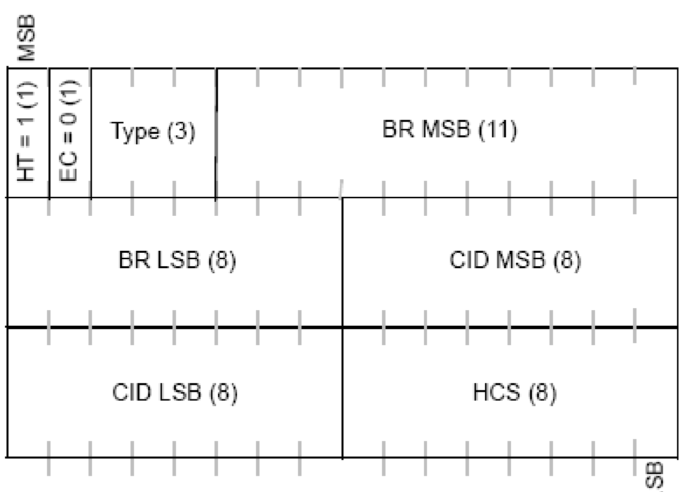


Figure 4.2.3: Bandwidth Request header format ([802.16-2004])

Subheaders

Six types of subheaders may be present in a MAC PDU with generic MAC header. The Type field (6 bits) of generic MAC header (GMH) will state if there is sub-header immediately following GMH in this MAC PDU. Each bit of the Type field indicates different organizations of subheaders as in Table 4.2.1.

Table 4.2.1: Type Encodings

Type bit	Value
#5 most significant bit (MSB)	Mesh subheader (not in WiMAX profile) 1 = present, 0 = absent
#4	ARQ Feedback Payload 1 = present, 0 = absent
#3	Extended Type Indicates whether the present Packing or Fragmentation subheaders, is Extended, for non ARQ-enabled connections 1 = Extended 0 = not Extended For ARQ-enabled connections, this bit shall be set to 1.
#2	Fragmentation subheader 1 = present, 0 = absent
#1	Packing subheader 1 = present, 0 = abset
#0 least significant bit (LSB)	Downlink: FAST-FEEDBACK Allocation subheader Uplink: Grant Management subheader 1 = present, 0 = absent

The per-PDU subheaders (i.e., extended subheader field, Mesh, Fragmentation, Fast-feedback allocation and Grant Management) may be inserted in MAC PDUs immediately following the Generic MAC header. If both the Fragmentation subheader and Grant Management subheader are indicated, the Grant Management subheader shall come first. If the Mesh subheader is indicated, it shall precede all other subheaders except for the extended subheader. In the downlink, the Fast-feedback allocation subheader shall always appear as the last per-PDU subheader. The ESF bit in the Generic MAC header indicates that the extended subheader is present. Using this field, a number of additional subheaders can be used within a PDU. The extended subheader shall always appear immediately after the Generic MAC header, and before all other subheaders. All extended subheaders are not encrypted.

The only per-SDU subheader is the Packing subheader. It may be inserted before each MAC SDU if so indicated by the Type field. The Packing and Fragmentation subheaders are mutually exclusive and shall not both be present within the same MAC PDU. When present, per-PDU subheaders shall always precede the first per-SDU subheader.

4.2.2 MAC Management Messages

A set of MAC management messages are defined. These messages shall be carried in the Payload of the MAC PDU. All MAC management messages begin with a Management Message Type field and may contain additional fields. MAC management messages on the Basic, Broadcast, and Initial

Ranging connections shall neither be fragmented nor packed. MAC management messages on the Primary Management Connection may be packed and/or fragmented. MAC management messages on the Fragmentable Broadcast connection may be fragmented. For the SCa, OFDM, and OFDMA PHY layers, management messages carried on the Initial Ranging, Broadcast, Fragmentable Broadcast, Basic, and Primary Management connections shall have CRC usage enabled. The format of the management message is given in Figure 4.2.4.

Management Message Type	Management Message Payload

Figure 4.2.4: MAC Management message format (Fig 21 of [802.16-2004]

The encoding of the Management Message Type field is given in Table 14 of [802.16-2005]. MAC management messages shall not be carried on Transport Connections. MAC management messages that have a Type value specified in Table 14 as "*Reserved,*" or those not containing all required parameters or containing erroneously encoded parameters, shall be silently discarded.

4.2.3 Fragmentation

Fragmentation is the process by which a MAC SDU (or MAC Management message) is divided into one or more MAC PDUs. The authority to fragment traffic on a connection is defined when the connection is created by the MAC SAP. Fragmentation may be initiated by a BS for downlink connections and by an SS for uplink connections.

The Fragmentation subheader shown in Table 8 of [802.16-2004] will show in front of each fragment if fragmentation happens.

Table 4.2.2: Fragmentation subheader format

Syntax	Size	Notes
Fragmentation Subheader() {		
FC	2 bits	Indicates the fragmentation state of the payload: 00 = no fragmentation 01 = last fragment 10 = first fragment 11 = continuing (middle) fragment
if (ARQ-enabled Connection)		
BSN	11 bits	Sequence number of first block in the current SDU fragment.
else {		
if (Type bit Extended Type)		See Table 6
FSN	11 bits	Sequence number of the current SDU fragment. This field increments by one (modulo 2048) for each fragment, including unfragmented SDUs.
else		
FSN	3 bits	Sequence number of the current SDU fragment. This field increments by one (modulo 8) for each fragment, including unfragmented SDUs.
}		
reserved	3 bits	Shall be set to zero
}		

The size of the FSN field in fragmentation subheaders is fixed per connection. The fragmentable broadcast connection shall use 11-bit FSN. BS and SS shall support 11-bit FSN. BS and SS may support 3-bit FSN. All management connections shall use 11-bit FSN. The size of the FSN used on non-ARQ fragmentable transport connections is determined during connection set-up (see 11.13.22 of [802.16-2004]).

Non-ARQ Connections

For non-ARQ connections, fragments are transmitted once and in sequence. The sequence number assigned to each fragment allows the receiver to recreate the original payload and to detect the loss of any intermediate packets. A connection may be in only one fragmentation state at any given time. Upon loss, the receiver shall discard all MAC PDUs on the connection until a new first fragment is detected or a non-fragmented MAC PDU is detected.

ARQ-Enabled Connections

For ARQ-enabled connections, fragments are formed for each transmission by concatenating sets of ARQ blocks with adjacent sequence numbers (see 4.3.4.2). The BSN value carried in the fragmentation subheader is the BSN for the first ARQ block appearing in the segment.

4.2.4 Packing

If packing is turned on for a connection, the MAC may pack multiple MAC SDUs into a single MAC PDU. The transmitting side has full discretion whether or not to pack a group of MAC SDUs in a single MAC PDU. The capability of unpacking is mandatory.

Connection attribution "Fixed-length versus variable-length SDU indicator" IE in DSA-REQ, DSA-RSP or DSA-ACK message (11.13.15of [802.16-2004]) will indicate if the connection carries fixed-length or variable-length packets.

0 = variable-length SDUs
1 = fixed-length SDUs
default = 0

Packing for non-ARQ connections with fixed-length MAC SDUs

For packing with fixed-length blocks, the Request/Transmission Policy (11.13.12) shall be set to allow packing and prohibit fragmentation, and the SDU size (11.13.16) shall be included in DSA-REQ message when establishing the connection. The length field of the MAC header implicitly indicates the number of MAC SDUs packed into a single MAC PDU. If the MAC SDU size is n bytes, the receiving side can unpack simply by knowing that the length field in the MAC header will be $n \times k+j$, where k is the number of MAC SDUs packed into the MAC PDU and j is the size of the MAC header and any prepended MAC subheaders.

A MAC PDU containing a packed sequence of fixed-length MAC SDUs would be constructed as in Figure 26 [802.16-2004]. Note that there is no added overhead due to packing in the fixed-length MAC SDU case, and a single MAC SDU is simply a packed sequence of length *1*.

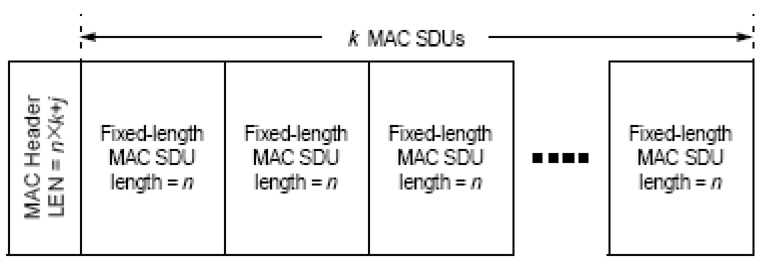


Figure 4.2.5: Packing fixed-length MAC SDUs into a single MAC PDU

Packing for non-ARQ connections with variable-length MAC SDUs

When packing variable-length SDU connections, such as 802.3/Ethernet, the $n \times k+j$ relationship between the MAC header's length field and the higher-layer MAC SDUs no longer holds. This necessitates indication of where one MAC SDU ends and another begins. In the variable-length MAC SDU case, the MAC attaches a Packing subheader to each MAC SDU.

Table 4.2.3: Packing Subheader Format

Syntax	Size	Notes
Packing Subheader() {		
FC	2 bits	Indicates the fragmentation state of the payload: 00 = no fragmentation 01 = last fragment 10 = first fragment 11 = continuing (middle) fragment
if (ARQ-enabled Connection)		
BSN	11 bits	Sequence number of first block in the current SDU fragment.
else {		
if (**Type** bit **Extended Type**)		See Table 6.
FSN	11 bits	Sequence number of the current SDU fragment. This field increments by one (modulo 2048) for each fragment, including unfragmented SDUs.
else		
FSN	3 bits	Sequence number of the current SDU fragment. This field increments by one (modulo 8) for each fragment, including unfragmented SDUs.
}		
Length	11 bits	
}		

Note: Length of the SDU fragment in bytes including the packing subheader.

A MAC PDU containing a packed sequence of variable-length MAC SDUs is constructed as shown in Figure 4.2.6. If more than one MAC SDU is packed into the MAC PDU, the type field in the MAC header indicates the presence of Packing subheaders (PSHs). Note that unfragmented MAC SDUs and MAC SDU fragments may both be present in the same MAC PDU (see Figure 4.2.6).

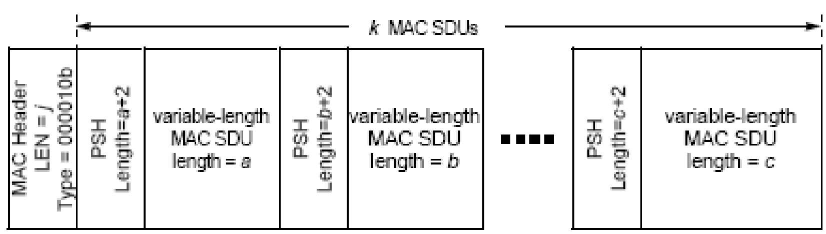


Figure 4.2.6: Packing variable-length MAC SDUs into a single MAC PDU

Simultaneous fragmentation and packing allows efficient use of the airlink, but requires guidelines to be followed so it is clear which MAC SDU is currently in a state of fragmentation. To accomplish this, when a Packing subheader is present, the fragmentation information for individual MAC SDUs or MAC SDU fragments is contained in the corresponding Packing subheader. If no PSH is present, the fragmentation information for individual MAC SDU fragments is contained in the corresponding Fragmentation subheader (FSH). This is shown in Figure 4.2.7 (Figure 28 of [802.16-2004]).

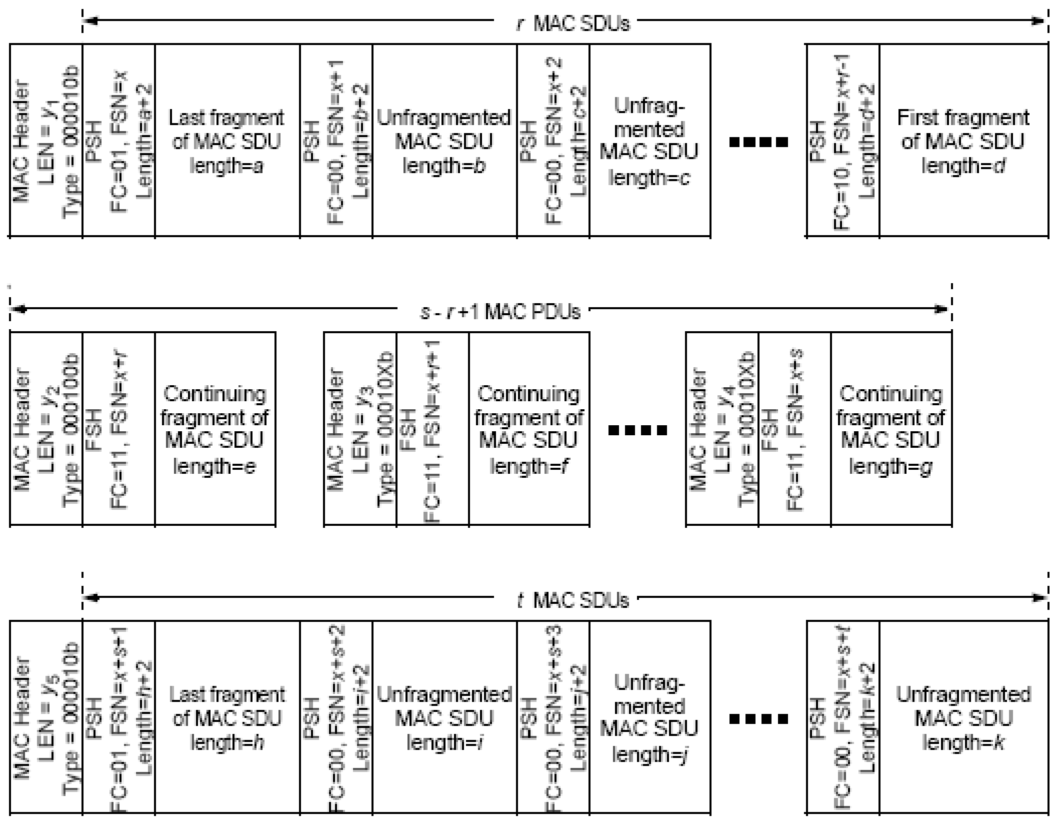


Figure 4.2.7: Packing with fragmentation

Packing for ARQ-enabled connections

The use of Packing subheaders for ARQ-enabled connections is similar to that for non-ARQ connections as described, except that ARQ-enabled connections shall set the Extended Type bit in the generic MAC header to 1.

The packing of variable-length MAC SDUs for the ARQ-enabled connections is similar to that of non-ARQ connections, when fragmentation is enabled. The BSN of the Packing subheader shall be used by the ARQ protocol to identify and retransmit ARQ blocks.

For ARQ-enabled connections, when the type field indicates Packing subheaders are in use, fragmentation information for each individual MAC SDU or MAC SDU fragment is contained in the associated Packing subheader. When the type field indicates that packing is not in use, fragmentation information for the MAC PDU's single payload (MAC SDU or MAC SDU fragment) is contained in the fragmentation header appearing in the message. Figure 4.2.8 (Figure 29 of [802.16-2004]) illustrates the use of Fragmentation subheader without packing.

Generic MAC Header	Other subheaders	Fragmentation subheader	Payload (One SDU or fragment of an SDU)	CRC-32

Figure 4.2.8: Example MAC PDU with Extended Fragmentation Subheaders

Figure 4.2.9 (Figure 30 of [802.16-2004]) illustrates the structure of a MAC PDU with ARQ Packing subheaders. Each of the packed MAC SDU or MAC SDU fragments or ARQ feedback payload requires its own Packing subheader and some of them may be transmissions while others are retransmissions.

Generic MAC header	Grant Management subheader (UL only)	Packing subheader	Payload (One SDU or SDU fragment or a set of ARQ Feedback IEs)	Packing subheader	Payload (One SDU or SDU fragment)	CRC-32

Figure 4.2.9: Example MAC PDU with ARQ Packing Subheader

A MAC SDU may be partitioned into multiple fragments that are then packed into the same MAC PDU for the first transmission. MAC PDUs may have fragments from the same or different SDUs, including a mix of first transmissions and retransmissions. The 11-bit BSN and 2-bit FC fields uniquely identify each fragment or non-fragmented SDU.

Packing ARQ Feedback

An ARQ Feedback Payload (see Table 4.2.3) consists of one or more ARQ Feedback IEs (see 4.3.4.2). The ARQ Feedback Payload may be sent on an ARQ or non-ARQ connection; however, policies based on implementation and/or QoS constraints may restrict the use of certain connections for transporting ARQ Feedback Payload. The ARQ Feedback Payload is treated like any other payload (SDU or fragments) from the packing perspective, except that only one ARQ Feedback Payload shall be present within a single MAC PDU.

Table 4.2.4: ARQ Feedback Payload Format

Syntax	Size	Notes
ARQ_Feedback_Payload_Format() {		
do		
ARQ_Feedback_IE(last)	*variable*	Insert as many as desired, until last==TRUE. See 6.3.4.2.
until (last)		
}		

The presence of an ARQ Feedback Payload in a MAC PDU is indicated by the value of the ARQ Feedback Payload bit in the Type field in the generic MAC header. When present, the first packed payload shall be the ARQ Feedback Payload. The Packing subheader preceding the ARQ Feedback Payload indicates the total length of the payload including the Packing subheader and all ARQ Feedback IEs within the payload. The FSN/BSN field of the Packing subheader shall be ignored for the ARQ Feedback Payload and the FC bits shall be set to 00.

4.3 ARQ Mechanisms

Automatic Repeat reQuest (ARQ) is an optional MAC feature in [802.16-2004] standard. However, the implementation of ARQ is mandatory for Mobile WiMAX systems.

When implemented, ARQ may be enabled on a per-connection basis. The per-connection ARQ shall be specified and negotiated during connection creation. A connection cannot have a mixture of ARQ and non-ARQ traffic. The scope of a specific instance of ARQ is limited to one unidirectional connection.

4.3.1 ARQ Operations

Basic ARQ functions.

1. Fragmentation. When ARQ is enabled, SDU is treated as fragmented into logical ARQ blocks with fixed block size defined by ARQ_BLOCK_SIZE as in Figure 4.3.1. Fragmentation shall occur only on ARQ block boundaries. When the length of the SDU is not an integer multiple of ARQ_BLOCK_SIZE, the final block of the SDU is formed using the SDU bytes remaining after the final full block has been determined. Fragmentation sub-header is attached to each fragmentation boundary block and contains a BSN of the first ARQ block of each SDU. When a PDU is packed, packing sub-header is attached instead of fragmentation sub-header.

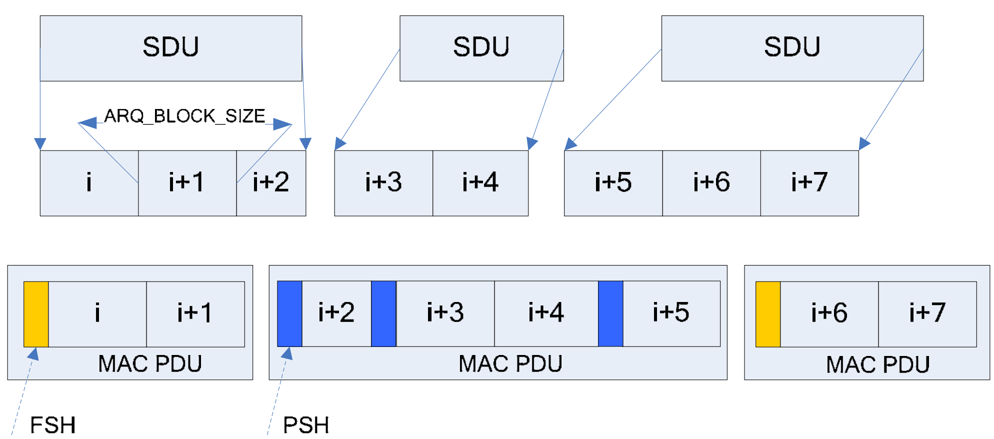


Figure 4.3.1: ARQ mapping

2. Buffer management. ARQ transmitter and receiver each manages blocks in a sliding window. Due to the sliding window mechanism, the minimum buffer size shall be ARQ_WINDOW_SIZE * ARQ_BLOCK_SIZE.

3. ARQ feedback. ARQ feedback is per-block ACK/NAK response sent from ARQ receiver to ARQ transmitter. ARQ feedback should be treated as a regular MAC payload, i.e., resource allocation will go through scheduling and uplink bandwidth request if necessary. ARQ feedback information can be sent as a standalone MAC management message on basic management connection, or piggybacked on an existing transport connection. ARQ feedback cannot be fragmented. ARQ feedback method combines both cumulative and selective acknowledgement mechanisms. There are four different ARQ feedback types.

4. ARQ retransmissions. Retransmission will be triggered by an ARQ NACK or a transmitter-side retransmission timer defined by ARQ_RETRY_TIMEOUT.

5. CRC-32. CRC-32 shall be used for error detection of PDUs for all ARQ-enabled connections.

Advanced ARQ functions.

1. ARQ discard. In rare cases, ARQ transmitter discards blocks when their delay reaches the threshold defined by ARQ_BLOCK_LIFETIME before receiving acknowledgements. ARQ Discard management message shall be sent by ARQ transmitter.

2. ARQ purge. ARQ receiver shall wait until ARQ_RX_PURGE_TIMEOUT after successful reception of a block that does not result in advancement of ARQ Rx window.

3. ARQ in-order delivery. When ARQ_DELIVER_IN_ORDER is enabled, a MAC SDU is handed to the upper layers as soon as all the ARQ blocks of the MAC SDU have been correctly received within the defined time-out values and all blocks with sequence numbers smaller than those of the completed message have either been discarded or delivered to the upper layers. If ARQ_DELIVER_IN_ORDER is not enabled, MAC SDUs are handed to the upper layers as soon as all blocks of the MAC SDU have been successfully received. ARQ_DELIVER_IN_ORDER shall be enabled especially for TCP connections.

4. ARQ synchronization. ARQ transmitter and receiver maintain independent timer, which is reset whenever sliding window start pointer advances. If this timer expires at either side, as defined by ARQ_SYNC_LOSS_TIMEOUT, ARQ reset shall be initiated either by ARQ transmitter or by ARQ receiver.

5. ARQ reset. All signalling between ARQ transmitter and receiver are associated with a retry timer defined by T22 up to some maximum retry number. ARQ Reset management messages shall be exchanged. The action sequence is explained in Figure 4.3.2.

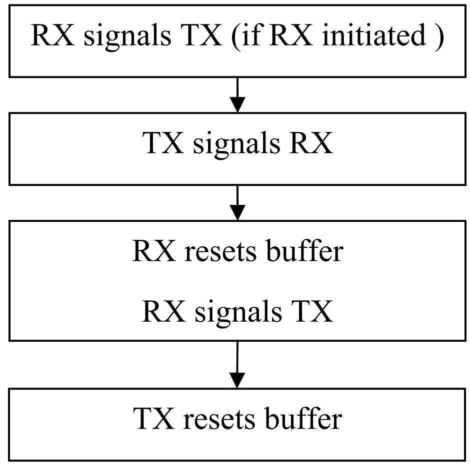
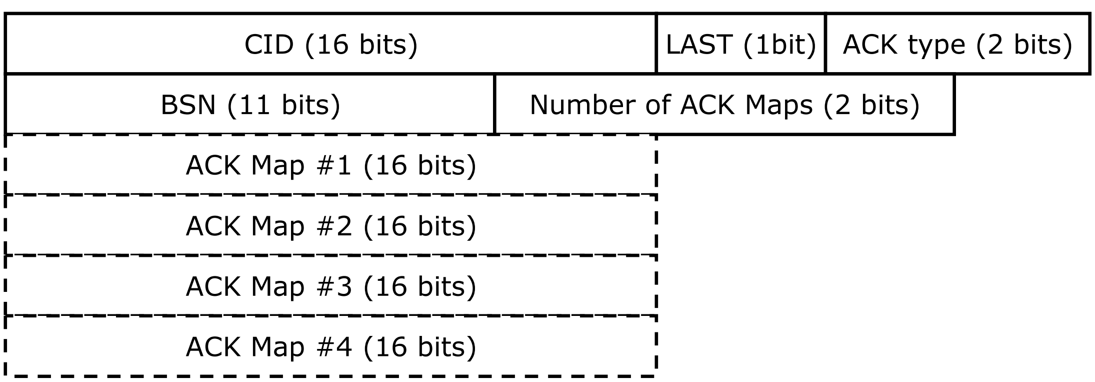


Figure 4.3.2: ARQ Reset operation

4.3.2 ARQ Feedback Methods

ARQ feedback information is formatted as ARQ Feedback IE and a number of ARQ Feedback IEs can be concatenated to create an ARQ feedback payload that are transferred either by a separate basic management connection or by an existing transport connection in a piggybacked way.

The ARQ Feedback IE used by ARQ receiver to signal positive or negative acknowledgements for the specified transport connection is defined as in Figure 4.3.3.



- CID: TCID for this feedback
- LAST: 0 = more ARQ feedback IE follows, 1 = this is the last
- ACK Type: four different feedback types
- Number of ACK Maps: 0~3 means 1~4 ACK maps follows
- ACK Maps
  - No ACK map for cumulative feedback
  - 1~4 ACK maps otherwise

Figure 4.3.3: ARQ Feedback IE

One of four different ACK types can be specified for each ARQ Feedback IE. The detailed descriptions for each ACK types are summarized in Table 4.3.1.

Table 4.3.1: ARQ ACK Types

ACK Type	Name	Operation
0x0	Selective ACK	BSN value corresponds to the most significant bit of the first 16-bit ARQ ACK map and follows an MSB first approach with the BSN incremented by 1 for each bit in the ARQ ACK map. Each bit of ARQ ACK map represents ACK (1) or NAK (0) of corresponding ARQ block. This ACK type is of no use due to excessive number of ACK/NAK bits.
0x1	Cumulative ACK	BSN value indicates that its corresponding block and all blocks with lesser values within the transmission window have been successfully received. No ACK map is used.
0x2	Cumulative ACK with Selective ACK Map	Combines the functionality of types 0x0 and 0x1.
0x3	Cumulative ACK with Block Sequence ACK Map	Combines the functionality of type 0x1 with the ability to acknowledge reception of ARQ blocks in terms of block sequences. A block sequence is defined as a set of ARQ blocks with consecutive BSN values. ACK map is interpreted differently as two block sequences (NAKs/ACKs) or three block sequences (NAKs/ACKs/NAKs).

ARQ feedback payload consists of one of more number of ARQ Feedback IEs. The ARQ feedback payload can be sent as a standalone MAC management message on basic management connection, or piggybacked on an existing transport connection as described in Figure 4.3.4.

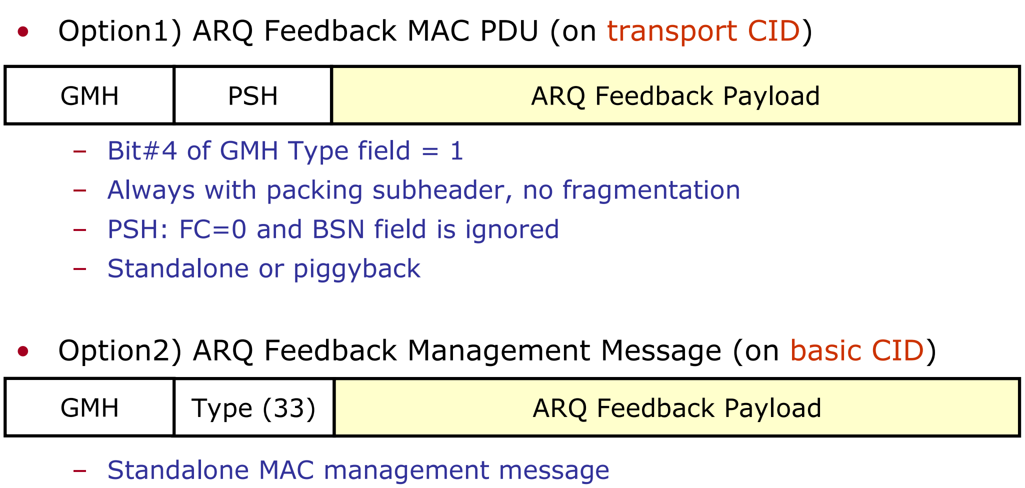


Figure 4.3.4: Transmitting ARQ Feedback Payload

The decision for which transmission method is used is open to simulation users. For example, the basic management message can be always used for simple modeling or it can be dynamically chosen whenever sending ARQ feedback payloads.

4.3.3 ARQ Parameters

Table 4.3.2 lists the ARQ configuration parameters exchanged when an ARQ connection is established.

Table 4.3.2: ARQ Configuration Parameters

Description	Value	Comments
ARQ_BSN_MODULUS	2048	fixed to 2^{11}
ARQ_WINDOW_SIZE	> 0 and ≤ (ARQ_BSN_MODULUS / 2)	
ARQ_BLOCK_SIZE	0-15: reserved 16, 32, 64, 128, 256, 512, or 1024 1025 ~ 65535: reserved	A set of desired/agreed size in bytes is specified.
ARQ_RETRY_TIMEOUT	0 ~ 6555.35 sec (100 µs granularity[1])	
ARQ_BLOCK_LIFETIME	0 = infinite 0.1 ~ 6553.5 msec (100 µs granularity)	
ARQ_SYNC_LOSS_TIMEOUT	0 = infinite 0.1 ~ 6553.5 msec (100 µs granularity)	
ARQ_RX_PURGE_TIMEOUT	0 = infinite 0.1 ~ 6553.5 msec (100 µs granularity)	
ARQ_DELIVER_IN_ORDER	0 = disabled 1 = enabled	

Table 4.3.3 lists the key MAC parameters particularly those for ARQ simulation. [Mobile WiMAX system profile v1.2.2]

Table 4.3.3: MAC Parameters

Item	Description	Status	Min	Default	Max	Comments
1.	Tproc	BS	Frame Length			Time provided between arrival of the last bit of a UL-MAP at an SS and effectiveness of that MAP
2.	Maximum SDU size		1522			Recommended value to derive Maximum

[1] The granularity has been changed from 10 to 100 in 802.16 Cor2.

			bytes			Transmission Unit (MTU) from
3	T22	MS/BS			0.5s	Wait for ARQ-Reset
4	ARQ_RESET_MAX_RETRIES			2		
5	ARQ_TRANSMITTER_DELAY	BS/MA	0 ms		6553.5 ms	ARQ_RETRY_TIMEOUT is the sum of ARQ_TRANSMITTER_DELAY and ARQ_RECEIVER_DELAY
6.	ARQ_RECEIVER_DELAY	BS/MS	0 ms		6553.5 ms	
7	ARQ_SYNC_LOSS_TIMEOUT for non HARQ connections	BS/MS	100ms			Used in DSA-REQ and DSA-RSP to indicate timeout for ARQ. 5msec granularity.
8	ARQ_RX_PURGE_TIMEOUT for non HARQ connections	BS/MS	100ms			Used in DSA-REQ and DSA-RSP to indicate timeout for ARQ. 5msec granularity.

4.4 MAC Support of PHY layer

In this section, the mapping of the MAC PDUs (MPDU) to the PHY SDUs (PSDU) is described. In addition, the processing of the MAC SDUs into MAC PDUs is also explained. The MAC interfaces with the PHYs through the PHY SAP. As introduced in the previous section, the MAC CPS provides the following services:

- Segmentation and Packing of the MAC SDUs into the MAC PDUs
- Concatenation of the MAC PDUs into a PHY SDU;

Figure 4.4.1.shows an overview of the processing chains from a MAC SDU into a PHY SDU. The MAC SDUs are packed and fragmented into MAC PDUs, as necessary. It should be noted that the decisions on how extensive on the packing and segmentation to be performed would be made by the MAC in conjunction with the specific PHY type and parameters and at the same time maintaining minimal overhead. While both packing and segmentation is supported for the processing from MAC SDUs into MAC PDUs, only concatenation is in the process of the PHY SDUs construction.

The PHY SDU would then proceed to be converted into FEC Blocks, which would then be mapped to the rectangular resource allocated to the SS or burst, as indicated in the DL-Map at the start of the frame. If there are insufficient MAC PDUs to fill the entire resources then padding MAC PDU may be included. Therefore, the following properties would apply:

- A MAC PDU can be sent in separate FEC Blocks although it cannot be sent in separate burst;
- A MAC SDU can be sent in separate burst through the MAC SDU segmentation process although this event would not be likely due to the increased overhead;
- In the event when HARQ is used, the FEC blocks would be sent in separate burst dependent on whether Chase or IR combining is used.

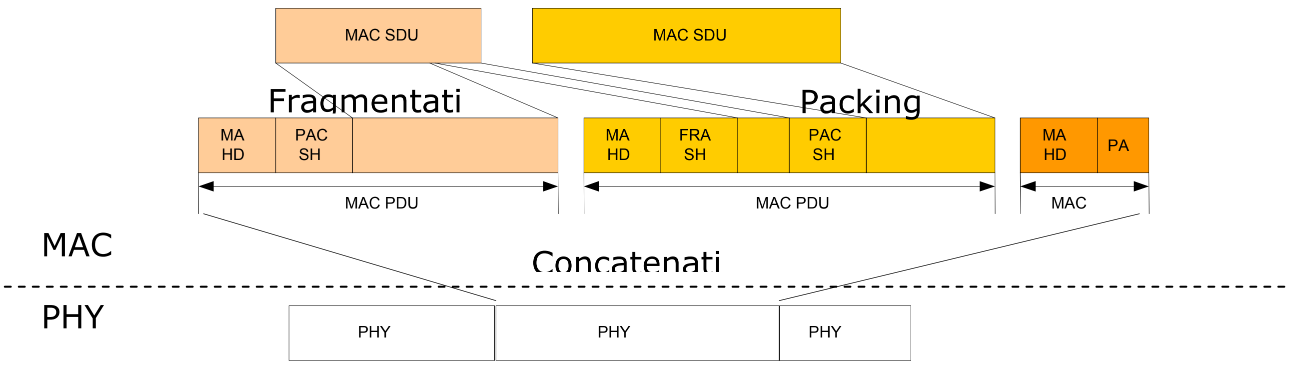


Figure 4.4.1: MAC PDU processing

The following figure shows the construction of MAC PDU and its transmission. Mapping a MAC-PDU to FEC Blocks depends on:

- PHY allocation size (burst size)
- Modulation and code rate (MCS) of burst
- MAC PDU concatenation

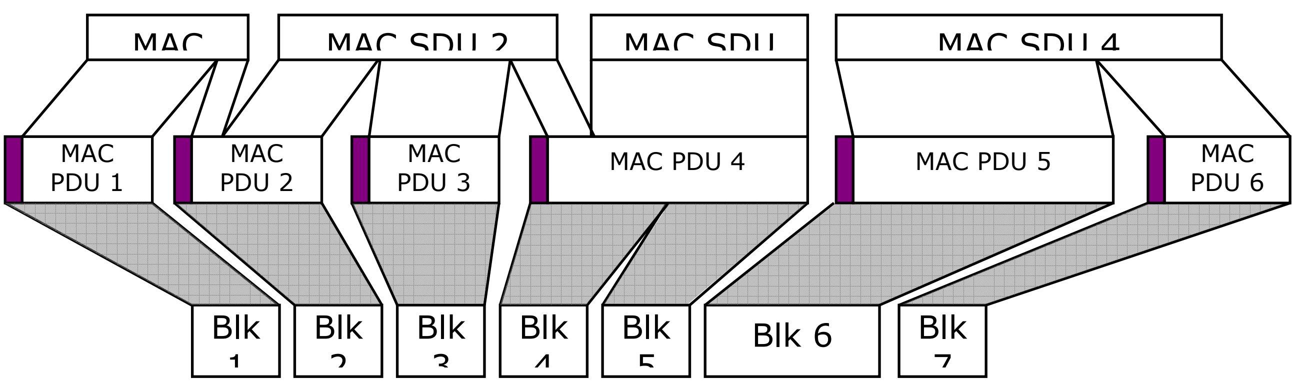


Figure 4.4.2: MAC PDU mapping to FEC blocks

The receiver applies the following logic to simulate PDU errors to account for the corresponding channel conditions.

1. Compute number of FEC blocks and block sizes for received PDU (see Appendix for computation rule and example)
2. If repetition is used, combine the repeated blocks to get the composite SINR on each non-repeated FEC block
3. Based on the block sizes, look up $P_{blk,i}$ in the appropriate curve for the each block "i" in the PDU
4. Compute the Probability of packet success:
5. Use a "Bernoulli toss" with probability P_{pkt} to decide whether the packet is successful or not.

4.5 Service Flow Operation

A series of MAC management messages DSA, DSC and DSD create, change or delete the service flows. The DSA messages create a new service flow. The DSC messages change an existing service flow. The DSD messages delete an existing service flow. They can be both BS initiated and SS initiated. Only BS initiated DSA and DSD are simulated to establish and tear down the service flow (connection).

A service flow should maintain a state machine (refer to Figure 99 in [802.16-2004]). BS's local decision of adding or deleting a service flow, and the following DSA or DSD events will trigger the transition among the states.

a) DSA events: DSA Succeeded, DSA Failed, DSA-ACK Lost, DSA Erred, DSA Ended
b) DSD events: DSD Succeeded, DSD Erred, DSD Ended.

4.5.1 BS initiated DSA

A BS establishes either a UL or a DL service flow with an SS performs the following operations:

- BS should first perform *admission control algorithm* to see if the QoS requirements of the requested class of service can be supported based on the current resource information.
- If the service can be supported, the BS generates a new SFID and a new CID with the required service class and informs the SS using a DSA-REQ message. BS set timer T7.
- SS should respond with a DSA-REP message, and enable the corresponding service flows.
- BS receiver DSA-REP. Timer T7 stops. BS enables the corresponding service flow.
- BS sends the DSA-ACK back.

Detailed process can be seen in Table 126 of [802.16-2004] and state transition diagrams in Figure 106-114 of [802.16-2004]. The service flow parameters including QoS parameters, ARQ parameters are defined in 11.13. of 802.16d 2004 and 802.16e 2005.

4.5.2 BS initiated DSD

A BS deletes either a service flow with an SS performs the following operations:

- BS deletes service flow
- BS send DSD-REQ to SS

- SS deletes the service flow
- sends DSD-REP to BS
- BS receive DSD-REP

Detailed process can be seen in Table 130 of [802.16-2004] and state transition diagrams in Figure 124-128 of [802.16-2004].

Table 4.5.1: MAC Parameters for Service Flow Operation

Item	Description	Status	Min	Default	Max	Comments
1.	T7	MS/BS			1s	Wait for DSA/DSC/DSD Response timeout
2.	T8	MS/BS			100ms	Wait for DSA/DSC Acknowledge timeout

4.6 MAC Scheduler

The MAC scheduler must efficiently allocate the available radio resources in response to bursty data traffic, time-varying channel conditions and specified scheduler criteria, if any. The MAC provides for contention-free access to the shared medium and for the most part it prevents simultaneous transmission from subscriber stations. The 802.16e air-interface provides support for a variety of QoS classes associated with service flows.

4.6.1 Scheduling Mechanisms

The Mobile WiMAX MAC scheduling service is designed to efficiently deliver broadband data services including voice, data, and video over time varying broadband wireless channel. The MAC scheduling service has the following properties that enable the broadband data service:

Fast Data Scheduler: The MAC scheduler must efficiently allocate available resources in response to bursty data traffic and time-varying channel conditions. The scheduler is located at each base station to enable rapid response to traffic requirements and channel conditions. The data packets are associated to service flows with well defined QoS parameters in the MAC layer so that the scheduler can correctly determine the packet transmission ordering over the air interface. The CQICH channel provides fast channel information feedback to enable the scheduler to choose the appropriate coding and modulation for each allocation. The adaptive modulation/coding combined with HARQ provide robust transmission over the time varying channel.

Scheduling for both DL and UL: The scheduling service is provided for both DL and UL traffic. In order for the MAC scheduler to make an efficient resource allocation and provide the desired QoS in the UL, the UL must feedback accurate and timely information as to the traffic conditions and QoS requirements. Multiple uplink bandwidth request mechanisms, such as bandwidth request through ranging channel, piggyback request and polling are designed to support UL bandwidth requests. The UL service flow defines the feedback mechanism for each uplink connection to ensure predictable UL scheduler behavior. Furthermore, with orthogonal UL sub-channels, there is no intra-cell interference. UL scheduling can allocate resource more efficiently and better enforce QoS.

Dynamic Resource Allocation: The MAC supports frequency-time resource allocation in both DL and UL on a per-frame basis. The resource allocation is delivered in MAP messages at the beginning of each frame. Therefore, the resource allocation can be changed on frame-by-frame in response to traffic and channel conditions. Additionally, the amount of resource in each allocation can range from one slot to the entire frame. The fast and fine granular resource allocation allows superior QoS for data traffic.

QoS Oriented: The MAC scheduler handles data transport on a connection-by connection basis. Each connection is associated with a single data service with a set of QoS parameters that quantify the aspects of its behavior. With the ability to dynamically allocate resources in both DL and UL, the scheduler can provide superior QoS for both DL and UL traffic. Particularly with uplink scheduling - the uplink resource is more efficiently allocated, performance is more predictable, and QoS is better enforced.

Frequency Selective Scheduling: The scheduler can operate on different types of sub-channels. For frequency-diverse sub-channels such as PUSC permutation, where sub-carriers in the sub-channels are pseudo-randomly distributed across the bandwidth, sub-channels are of similar quality. Frequency-diversity scheduling can support a QoS with fine granularity and flexible time-frequency resource scheduling. With contiguous permutation such as Band AMC permutation, the sub-channels may experience different attenuation. The frequency-selective scheduling can allocate mobile users to their corresponding strongest sub-channels. The frequency-selective scheduling can enhance system capacity with a moderate increase in CQI overhead in the UL.

Scheduling

- Determine packet transmission ordering over the air interface
- The MS must feedback accurate and timely information as to the traffic conditions and QoS requirements
- The MAC supports frequency-time resource allocation in both DL and UL on a per-frame basis.
- The resource allocation is delivered in MAP messages at the beginning of each frame. Therefore, the resource allocation can be changed on frame-by-frame in response to traffic and channel conditions.
- Q**oS Oriented:** The MAC scheduler handles data transport on a connection-by-connection basis. Each connection is associated with a single data service with a set of QoS parameters that quantify the aspects of its behavior.

Resource Allocation

- Choose the appropriate coding and modulation for each allocation
- Mapping of individual service flows to specific regions of the 2-D TDD frame (“slots”)
  - Generalized 2-dimensional bin packing problem (NP-hard)
  - Heuristic schemes are used for practical system, typically vendor-specific and proprietary implementation

Creation of DL and UL Maps

- The adaptive modulation/coding combined with HARQ provides robust transmission over the time-varying channel.

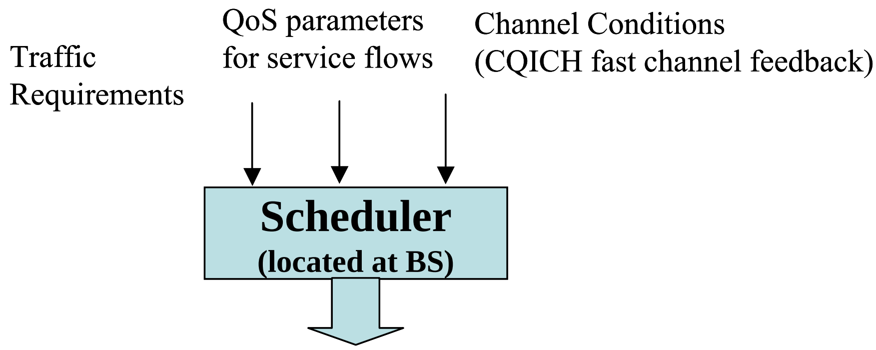


Figure 4.6.1: Scheduler Inputs

The following are basic assumptions for operation of a generic scheduler that is resident at the Base Station.

- The MAC supports frequency-time resource allocation in both UL and DL on a frame-by-frame basis.
- The MS must feedback accurate and timely information as to the traffic conditions and QoS requirements of service flows at subscriber stations
- The resource allocation is delivered in MAP messages at the beginning of each frame and can be changed on a per frame basis in response to traffic and channel conditions
- The MAC scheduler handles data transport on a per connection basis. Each connection is associated with a single flow with QoS parameters that quantify its requirements

The scheduler can operate on different types of sub-channels. Wave-1 simulator development needs to consider:

- Frequency diverse scheduling
  - For frequency diverse mode, such as the Partial Usage of Sub-Channels (PUSC), the sub-carriers assigned to each logical sub-channel are pseudo-randomly distributed across the available sub-carrier set which may only represent a fraction of the total sub-carrier resources according to a frequency re-use pattern.
  - The random allocation of sub-carriers in frequency diverse mode provides frequency diversity in a frequency selective fading channel.

A later phase of simulator needs to consider:

- Frequency selective scheduling:
  - For this mode of scheduling with a contiguous permutation of sub-carriers (aka as the Band Adaptive Modulation and Coding (Band AMC) mode in WiMAX systems), the sub-channels may experience different attenuation. The scheduler allocates adjacent sub-carriers for transmission based on the channel feedback information (e.g., allocate mobile flows to their corresponding strongest sub-channels).

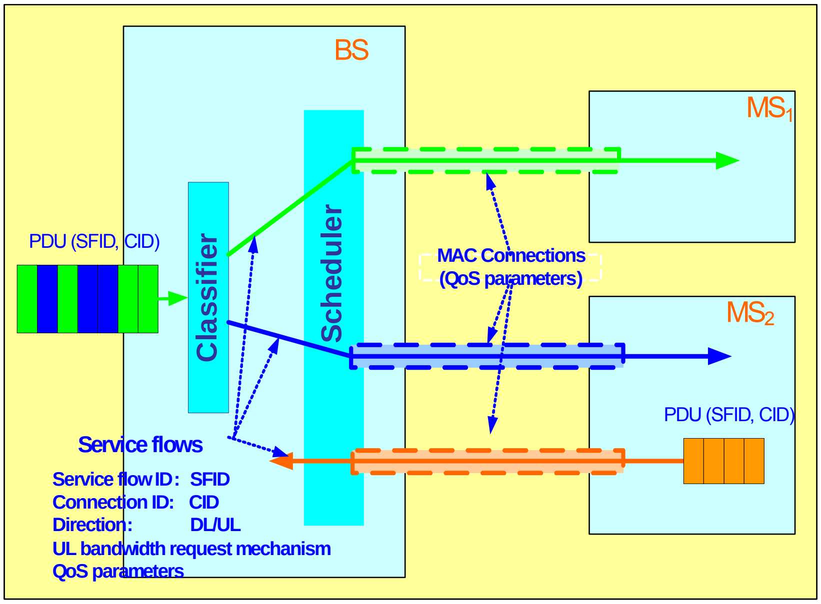


Figure 4.6.2: Scheduler Components

Definitions (Ref: 802.16e standard)

- **Connection**: A unidirectional mapping between base station (BS) and subscriber station (SS) medium access control (MAC) peers for the purpose of transporting a service flow's traffic. Connections are identified by a connection identifier (CID). All traffic is carried on a connection, even for service flows that implement connectionless protocols such as Internet Protocol (IP). The MAC defines two kinds of connections: management connections and transport connections.

- **Connection identifier (CID)**: A 16-bit value that identifies a transport connection or a uplink (UL)/downlink (DL) pair of associated management connections (i.e., belonging to the same subscriber station (SS)) to equivalent peers in the MAC of the base station (BS) and subscriber station (SS). The connection identifier (CID) address space is common (i.e., shared) between UL and DL. It maps to a service flow identifier (SFID), which defines the Quality of Service (QoS) parameters of the service flow associated with that connection.

- **Quality of Service (QoS) parameter set:** Is associated with a service flow identifier (SFID). The contained traffic parameters define scheduling behavior of uplink or downlink flows associated with transport connections.

DL Packet Scheduler

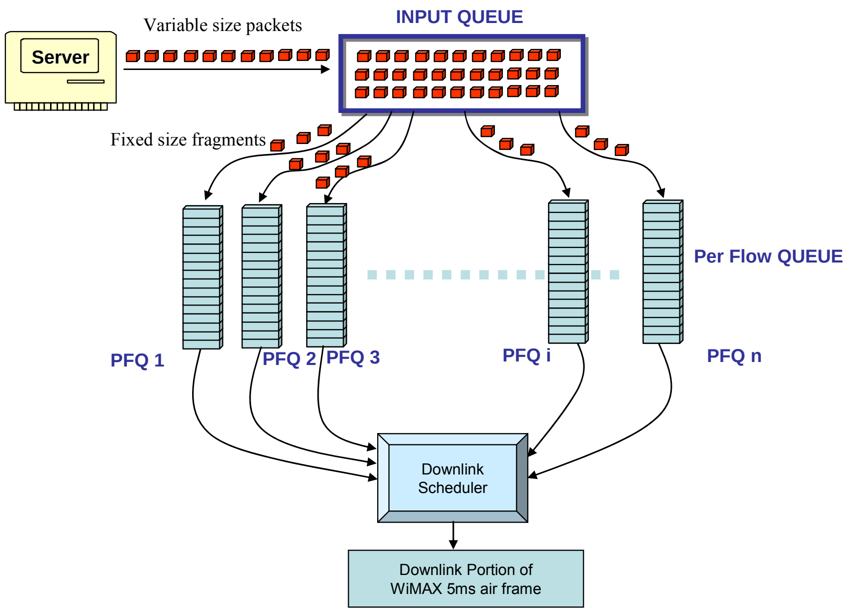


Figure 4.6.3: Downlink Packet Scheduler

UL Packet Scheduler

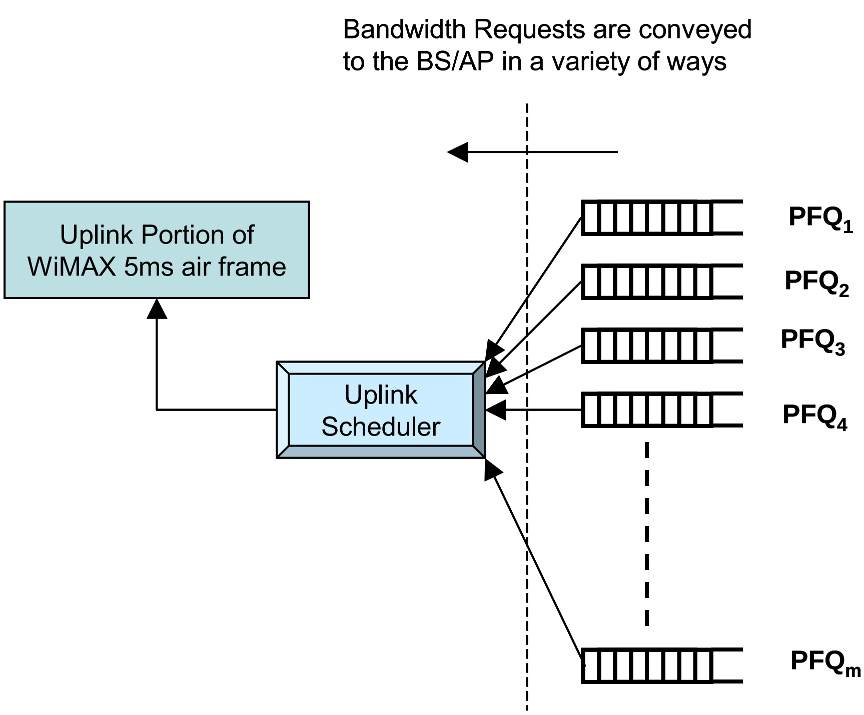


Figure 4.6.4: Uplink Packet Scheduler

UL scheduler is very similar to Downlink Scheduler. UL scheduler maintains the request-grant status of various uplink service flows. Bandwidth requests arriving from various uplink service flows at the WiMAX BS will be granted in a similar fashion as the downlink traffic. One additional consideration is that of available subscriber station power. This power is divided among allocated subchannels and therefore, affects the number of subchannels that the subscriber station can use.
For UL scheduling, multiple uplink bandwidth request mechanisms exist:

- Bandwidth request through ranging channel
- Piggyback request
- Polling

In its most generic form, a MAC scheduler may calculate a metric M_i per Flow I that is a function *f* of many attributes specific to the flow, and serve the flows in descending order of the metric values, where:

$\boldsymbol{M_i} = f\,(\,QoS_i,\ CINR_i,\ Delay_i,\ Throughput_i,\ other\_parameter_i\,)$

There are several options for constructing a basic scheduler, and each may serve a useful purpose depending on the design objectives. Some of the basic schemes are listed below along with some characteristic attributes:

1) Max C/I scheduler – order the flows in decreasing order based on their C/I ratio values. This scheme will have the tendency to maximize throughput but may not be fair to all the flows.

2) Round robin scheduler – Serve the flows in a round robin manner until all flows are served. This is simple to implement but is often inadequate to meet objectives such as QoS and throughput maximization

3) FIFO with priorities – Classify the flows according to priority classes (based on QoS requirements) and serve them in First In First Out order

4) Round robin with priorities – Classify the flows according to priority classes (based on QoS requirements) and serve them in a round robin manner

5) Proportional Fair Scheduler **(PFS)**– This offers a good compromise between many conflicting objectives, including, throughput maximization, fairness to all flows, ease of implementation, etc. This scheme provides an approach to maximize throughput in a greedy manner.

   At time t, let Flow i have an average realized throughput of $H_i(T)$ over the time interval < **t-T, t** >. Let the real throughput that can be achieved by Flow i at time t be $L_i(t)$. The real throughput is estimated based on the current channel conditions associated with Flow i.

   The PFS scheme will seek to achieve proportional fairness by selecting that flow which maximizes the ratio: $L_i(t)/H_i(T)$.

   Once the flows with non-zero requirements have been ordered for transmission based on the criterion above, the remaining steps are carried out in sequence (using heuristic considerations that are specific to individual scheduler implementation).

   - Determine how many fragments to transmit for a chosen flow
   - Resource allocation in the DL or UL sub-frames

- Creation of DL-MAP and UL-MAP (fixed as well as variable portions)
- Post-processing and packing the TDD frame to capacity with MAP, control and bearer traffic portions

A generic scheme based on Proportional Fair Scheduler principles offers a reasonable starting point, especially if there is a single class of traffic (Best Effort class) and throughput maximization is the primary objective. However, the MAC scheduler, in the general deployment case, should be capable of handling different QoS classes that are enabled by the 802.16e air interface and supported by the WiMAX Forum MTG system profile.

In particular, the following different classes need to be supported:

- Unsolicited Grant Service (UGS)
- Real Time Variable Rate Service (RT-VR)
- Non Real Time Variable Rate Service (NRT-VR)
- Best Effort Service (BE)
- Enhanced Real-Time Variable Rate Service (ERT-VR)

It is assumed that an Admission Control Policy exists for determining how many flows of each class are admitted at a given time (a function outside the scope of MAC scheduler). A generic MAC scheduler, may classify flows based on service classes and schedule users within classes, subject to:

- Delay constraints of real-time flows
- Maximizing system throughput
- Jitter constraints
- Power status of the subscriber station (The station may be power limited).

Scheduler Mappings – logical and physical views

It should be noted that the 802.16e standard offers a multitude of features – some mandatory and many optional in scope. A number of DL/UL maps are presented below. The MAC scheduler should determine the size and physical position (within a TDD frame) of mandatory logical components and optional logical components, if any for the appropriate map.

4.7 UL/DL Maps

The following diagram shows the basic mandatory logical components. A list of mandatory and optional information elements is presented later in this section.

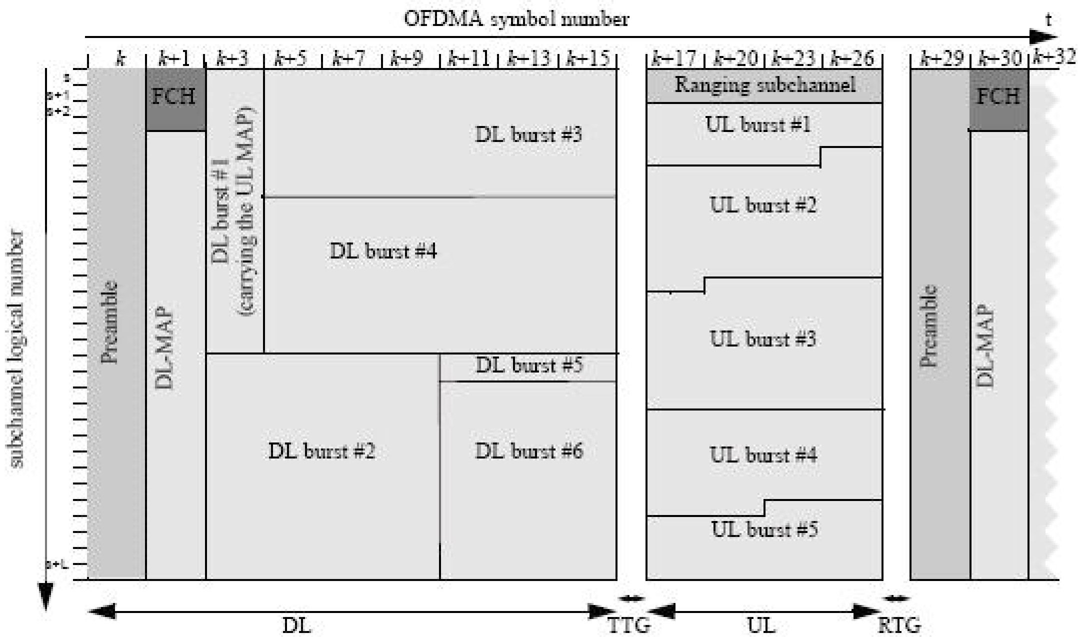


Figure 4.7.1: Example of an OFDMA frame (with only mandatory zone) in TDD mode

The DL control information is transmitted at the beginning of each frame that includes a preamble, Frame Control Header (FCH) and MAP messages. The first symbol Preamble is used for synchronization, DL channel estimation etc. FCH provides the information required to decode the subsequent MAP message, e.g. sub-channels used by the sector in the frame, coding and length of MAP etc. The MAP indicates the DL and UL transmission format and resource allocation for the current frame. MAP message has a fixed part plus a variable part that depends on number of DL and UL flows being scheduled in the frame. The size of this message can be several symbol durations. Sub-MAP messages may be present with different data rates for transmitting control messages to flows under different channel conditions. The UL control channels include ACK/NACK, channel feedback and the ranging messages.

OFDMA PHY also supports multiple subcarrier allocation zones within the same frame to enable the possibility of support for and coexistence of different types of SS's. The following figure illustrates zone switching within the DL and UL subframes. The figure includes several optional information elements which are not part of WiMAX Forum profiles. The switching is performed using an information element included in DL-MAP and UL-MAP. DL and UL subframes both start in PUSC mode where groups of subchannels are assigned to different segments by the use of dedicated FCH messages. The PUSC subcarrier allocation zone can be switched to a different type of subcarrier allocation zone through a directive from the PUSC DL-MAP.

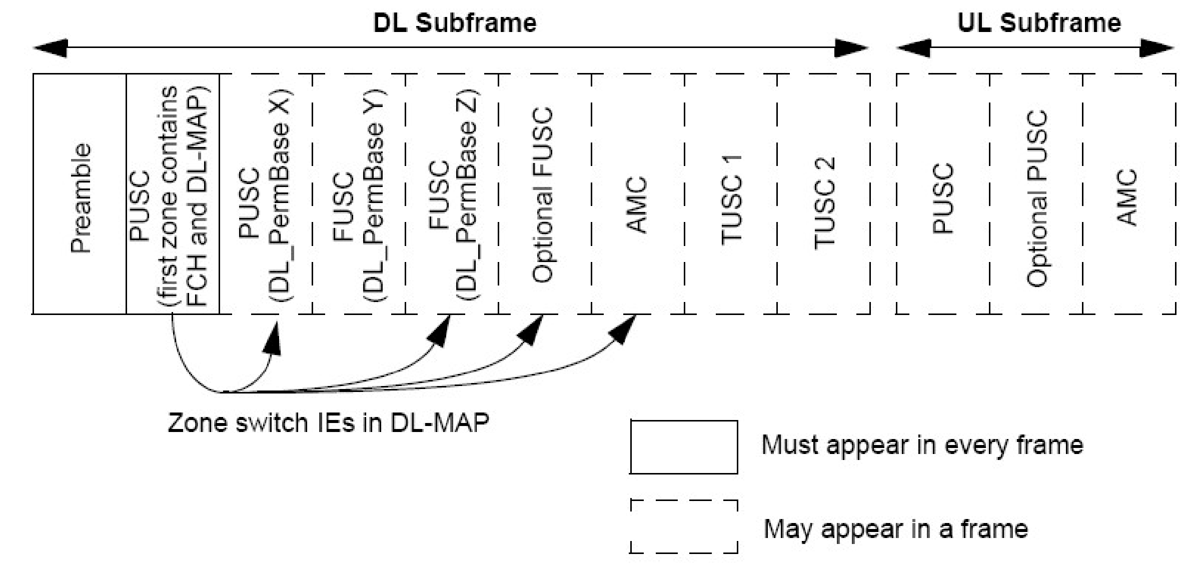


Figure 4.7.2: Example of OFDMA frame with multiple zones

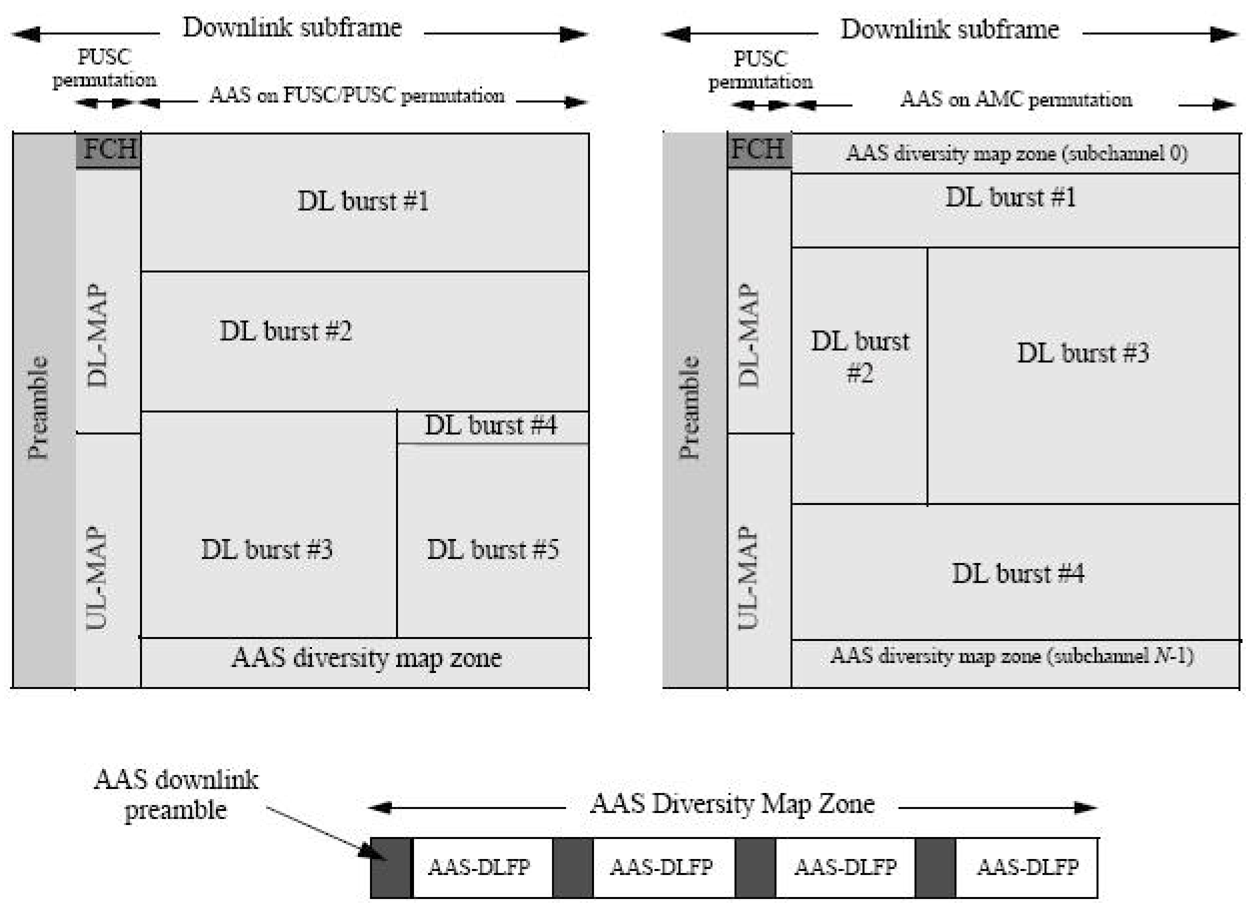


Figure 4.7.3: AAS Diversity Map Frame Structure

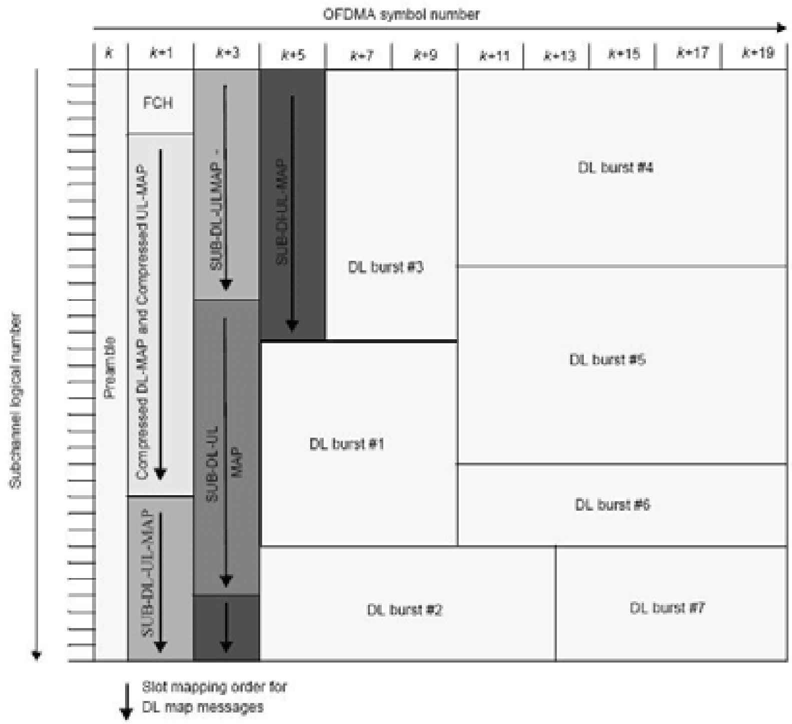


Figure 4.7.4: Sub-MAP Burst

4.7.1 DL/UL MAP Information Elements

The information elements affect the overhead. The recommended map elements and their sizes are as follows.

Table 4.7.1: DL/UL MAP Information Elements

MAP	IE size (bits)
Fixed Compressed MAP (DL + UL + CRC)	N_1 (152)
Ranging region allocation IE (3IEs)	N_2 (168)
Fast feedback allocation IE	N_3 (52)
HARQ ACK region allocation IE	N_4 (36)
Fixed overhead in HARQ DL MAP IE	N_5 (68)
Fixed overhead in HARQ UL MAP IE	N_6 (44)
Interference and Noise IE	N_7 (24)
UL HARQ/user	N_8 (33)
DL HARQ/user	N_9 (44)
Additional Optional Fields	??

Note: This is offered as an example of MAP elements and their sizes. Additional IEs **may/will** be present in certain frames (depending on options implemented). WMF members are welcome to propose & agree on a baseline representation prior to simulator development. Above list offers a viable starting point.

Given below is a more complete list of Information Elements (including mandatory and optional logical elements) as per the 802.16e standard. Depending on the system being simulated, and the corresponding optional features implemented in it, the MAP elements and the associated overhead can be accounted for in the simulator. However, it is recommended a generic simulator include at the least all the mandatory IEs that were identified above.

DL MAP Information Elements (Ref: 802.16e standard):

- DIUC Allocation (Downlink Interval Usage Code)
- DL-MAP extended IE
- AAS downlink IE
- Space-Time Coding (STC) / DL Zone switch IE
- Data location in another BS IE
- CID Switch IE
- MIMO DL basic IE
- MIMO DL enhanced IE
- HARQ and Sub-MAP Pointer IE
- DL-MAP Physical modifier IE
- Multicast and Broadcast Service MAP IE
- DL PUSC Burst Allocation in Other Segment IEHO anchor active DL MAP IE
- HO active anchor DL MAP IEHO CID Translation MAP IE
- MIMO in another BS IE
- Macro-MIMO DL Basic IE
- UL interference and noise level extended IE
- Dedicated DL control IE
- Reduced CID IE
- Skip IE
- HARQ DL MAP IE
- DL HARQ ACK IE
- Enhanced DL MAP IE
- Closed-loop MIMO DL enhanced IE
- Broadcast Control Pointer IE
- AAS_SDMA_DL_IE
- PUSC ASCA Allocation
- . . . . . .

UL-MAP IE format (Ref: 802.16e standard)

- UIUC Allocation – Uplink Interval Usage Code
- PAPR Reduction/Safety zone/Sounding zone allocation IE
- CDMA Allocation UL-MAP IE
- UL-MAP Extended IE
- Power Control IE
- AAS Uplink IE
- UL Zone Switch IE
- Mini-sub-channel allocation IE
- FAST-FEEDBACK allocation IE
- MIMO UL Basic IE

- CQICH Allocation IE
- UL-MAP Physical Modifier IE
- UL allocation start IE
- CQICH Enhanced Allocation IE
- UL PUSC Burst Allocation in Other Segment IE
- HO Anchor active UL MAP IE
- HO Active anchor UL MAP IE
- MIMO UL Enhanced IE
- OFDMA Fast_Ranging_IE
- UL_MAP_Fast_Tracking IE
- Anchor BS Switch IE
- HARQ UL MAP IE
- HARQ ACK Region Allocation IE
- UL-Sounding_Command_IE
- AAS_SDMA_UL_IE
- Feedback polling IE

The following IEs can be added in later releases:

a) DL HARQ and Sub-MAP Pointer IE

b) CID Switch IE

For wave-2 features:

a) UL Sounding Command IE

b) UL Sounding Zone Allocation IE

c) UL/DL MIMO basic IE

d) UL/DL MIMO Chase HARQ Sub-burst IE

4.8 HARQ

4.8.1 HARQ operation

- HARQ sub-burst are generated by concatenating multiple MAC PDUs and attaching a CRC.
- When HARQ enabled, the frame structure will include HARQ zone, with up to 16 sub-bursts vertically packed inside one zone. Each sub-burst is referred by HARQ_CONTROL_IE.
- Only chase combining is used in the current profile. Chase combining requires all retransmissions to send the exact same information and stick to the original modulation coding scheme.
- HARQ retransmission is asynchronous, in the sense that all HARQ burst will go through opportunistic scheduling.
- The maximum number of retransmissions is determined by target residual PER. Typically it is set to 4 for a PER of 1E-4.

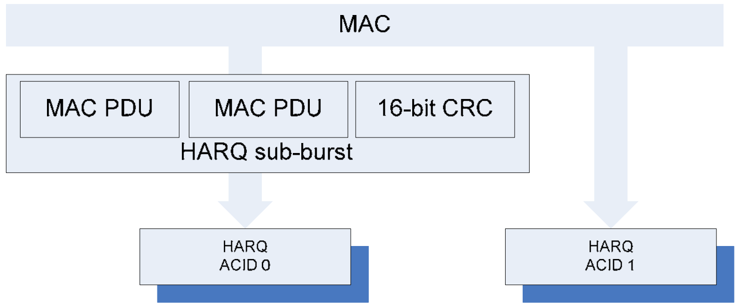


Figure 4.8.1: HARQ ACID

HARQ feedback.

- HARQ feedback is synchronous. Feedback channel is pre-allocated. Typically if burst is sent at frame i, feedback will be available at i+2 frame.
- UL ACK channel is specified by UL_HARQ_ACK_IE. One slot is shared by 2 ACKs.
- DL ACK channel is implicitly in the HARQ_Control_IE by toggling AI_SN bit. AI_SN is toggled whenever a successful transmission happens. Implementing DL HARQ ACK bitmap is unnecessary and can lead to ambiguity.

Key parameters for HARQ are listed in Table 4.8.1 and various categories of HARQ are listed in Table 4.8.2.

Table 4.8.1: HARQ Parameters

Timing	Synchronous/AsynchscyncSyhronous
Mode	Adaptive/NonAdaptive
Type	Chase Combining / Incremental Redundancy
Target PER	10%
Maximum number of ACIDs per user	16
Maximum number of ACIDs per frame	4
Retransmission delay for synchronous HARQ	3 frames
Maximum retransmission delay for asynchronous HARQ	5 frames
Maximum number of retransmissions	3 (Total 4 transmissions)

Table 4.8.2: HARQ Categories

HARQ Categories	1	2	3	4
DL buffer size per channel	16384 (K=20)	8192 (K=16)	16384 (K=20)	23170 (K=22)
UL buffer size per	16384 (K=20)	16384 (K=20)	16384 (K=20)	16384 (K=20)

channel				
DL aggregate flag	ON or OFF	ON	ON	ON
UL aggregate flag	OFF	ON	ON	ON
Number of DL HARQ channels supported by MS	4	16	16	16
Number of UL HARQ channels supported by MS	4	8	8	8
Maximum number of bursts in DL subframe with HARQ	2	5	5	5
Maximum number of bursts in UL subframe with HARQ	2	2	2	2

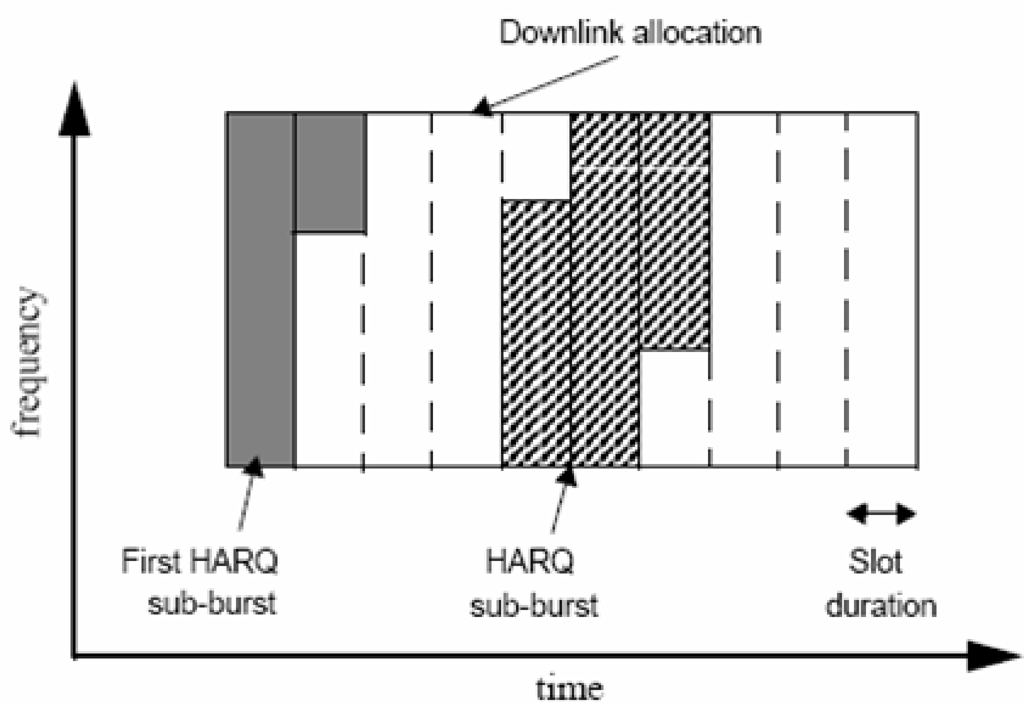


Figure 4.8.2: HARQ Downlink Allocation (PUSC Example)

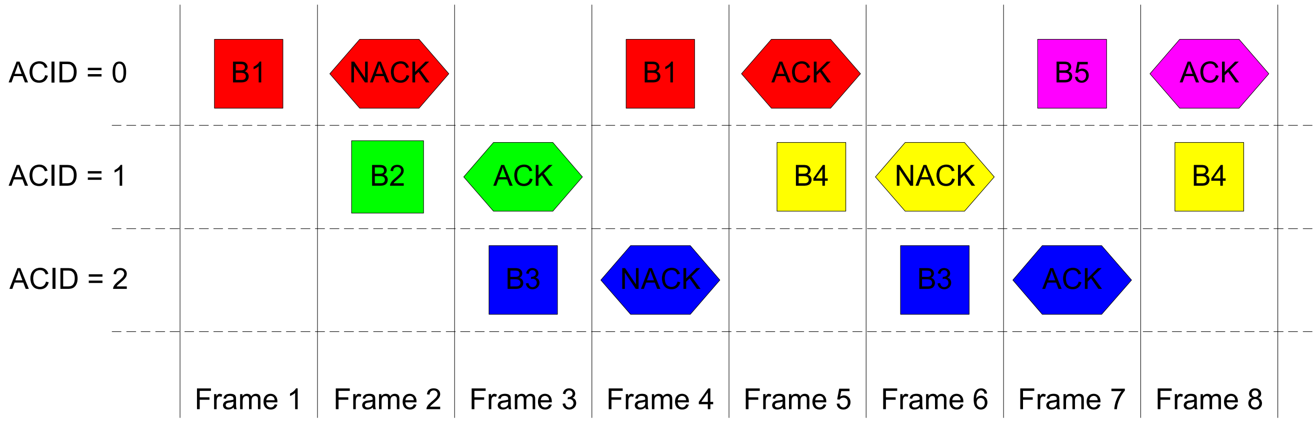


Figure 4.8.3: HARQ Multichannel

4.9 Mobility Management

4.9.1 Network entry and initialization

All network entry procedures described in this section are applied only to the PMP (point-to-multipoint) operation. The following features will be modelled in the simulation for SS network entry and initialization process as specified in the following sequence.

- Channel scanning and synchronization
- Obtain DL and UL transmit parameters
- Initial ranging
- Negotiate basic capabilities
- Authorization and key exchange
- Registration
- Establish IP connectivity
- Establish time of day
- Transfer operational parameters
- Setup connection

The procedure for initialization of an SS is shown in the figure below.

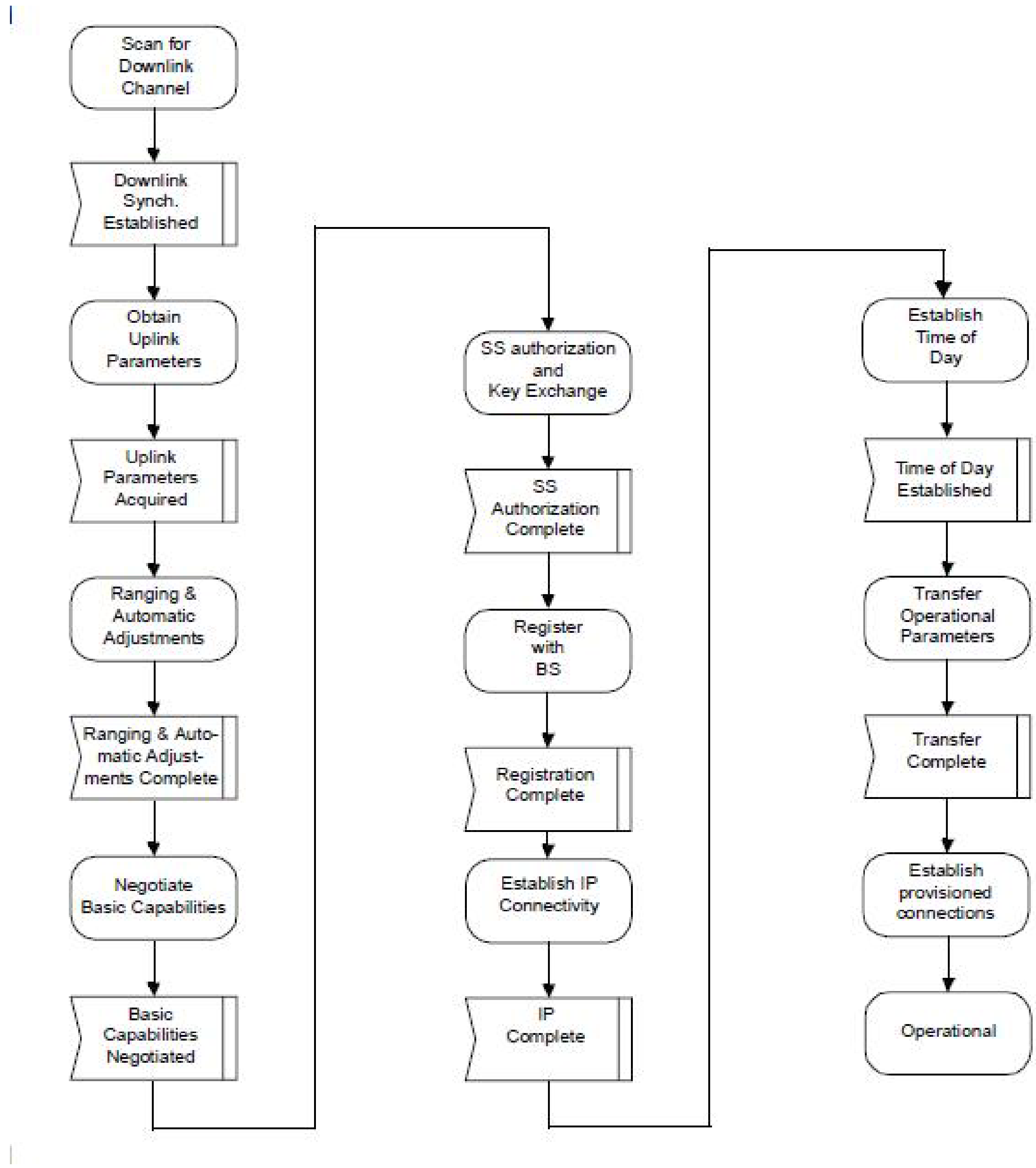


Figure 4.9.1 SS Initialization Overview [802.16-2004]

4.9.1.1 Channel scanning and synchronization

SS should acquire the downlink channel by scan the possible channels of the downlink frequency. The SS should memorize the last operation parameters and try to reacquire this downlink channel. If it fails, SS should continuously scan the possible downlink frequency band of operation until it finds a valid downlink frequency.

Once the PHY synchronization has achieved, the MAC should acquire downlink channel control parameters and then the uplink.

4.9.1.2 Obtain DL and UL transmit parameters

The BS generates DL-MAP and UL-MAP at a fixed interval. Also DCD (Downlink Channel Descriptor) and UCD (Uplink Channel Descriptor) are transmitted periodically by the BS. DCD includes BS transmit power, PHY type, and FDD/TDD frame duration, as described in [[802.16-2004]. The SS MAC shall search for the DL-MAP on the detected channel. SS decodes the DL-MAP on the detected channel. The most significant 24 bits of the "Base Station ID" shall be used as Operator ID. Once SS receive one valid DL-MAP message, it achieves MAC synchronization and remains synchronized as long as the DL-MAP and DCD messages are continuous successfully received.

After synchronization, the SS retrieve a set of transmission parameters from UCD MAC management message for a possible uplink channel. Then, the MS shall start initial ranging.

The SS need restart from scanning to find another DL channel, if any one of situation occurred:

- If the channel is not suitable.
- If the initial ranging failed.
- If SS doesn't receive a valid DL-MAP at expiry of the Lost DL-MAP interval (600ms, as specified in [[802.16-2004] Table 342)
- Without a valid DCD message at the T1 (5 * DCD interval maximum value, 50sec) interval.
- If SS doesn't receive a valid UL-MAP at expiry of the of the Lost UL-MAP interval (600ms, as specified in [802.16-2004] Table 342)
- Without a valid UCD message at the T12 (5 * UCD interval maximum value, 50sec) interval.

4.9.1.3 Ranging process

After the SS synchronize to the DL and obtain the UL channel characteristics through the UCD message, the SS shall scan the UL-MAP message to find an initial ranging interval. The ranging request opportunity size is specified in the UCD TLV. When the Initial Ranging transmission opportunity occurs, the SS start the ranging process by transmitting a randomly selected CDMA initial ranging code. Upon receiving the CDMA initial ranging code, the BS replies with a ranging response message that contains the required transmission adjustment (power, time, frequency) and a status indication. If the status is continue, the SS will apply the transmission adjustment and perform CDMA initial ranging again. If the status is success, the BS will make a bandwidth allocation using CDMA_Allocation_IE and the SS will transmit a RNG-REQ (Ranging Request) using the allocated bandwidth. Once the BS has successfully received the RNG-REQ message, the BS will then reply with a RNG-RSP (Ranging Response) message which contains the allocation of basic and primary management CID. The SS will timeout if a RNG-RSP is not received after a CDMA ranging code or a RNG-REQ. The timeout value is defined by T3 (200 ms).

If the SS doesn't receive RNG-RSP, the SS resends the RNG-REQ at the next appropriate Initial Ranging transmission opportunity at one step higher power level. If the SS receive a response containing the frame number of the transmitted RNG-REQ message, the transmission is unsuccessful. The SS need implement the corrections specified in the RNG-RSP and issue another RNG-REQ after the backoff delay. The backoff window size and the possible choice of initial ranging code are also obtained from the UCD message.

Format of RNG-REQ and RNG-RSP messages can be found in Table 19 and Table 20 of [1 [1]

4.9.1.4 Negotiate basic capabilities

Upon successful completion of ranging, the SS sends the SBC-REQ (SS Basic Capability Request) message. The BS responds with SBC-RSP (SS Basic Capability Response) message. The SBC-RSP should be uniform across all BS in the same NAP. If the SS doesn't receive SBC-RSP at expiry of T18 (50 ms), it need retransmit the SBC-REQ message. Those basic capabilities include the maximum available power, the current transmitted power (reported in dBm), SS supported demodulation types, SS supported FEC (Forward Error Correction) types, and more. Please see Table 51 and Table 52 in section 11.8 of reference [802.16-2004] for details about SBC-REQ/RSP management message encodings.

4.9.1.5 SS Authorization and key exchange

If PKM (Privacy Key Management) is enabled, the BS and SS shall perform authorization and key exchange. The detail messages flow is described in the following figure. To simplify the simulation model, all the steps behind BS can be simulated inside the BS. Another option is having the authenticator be a module inside the BS.

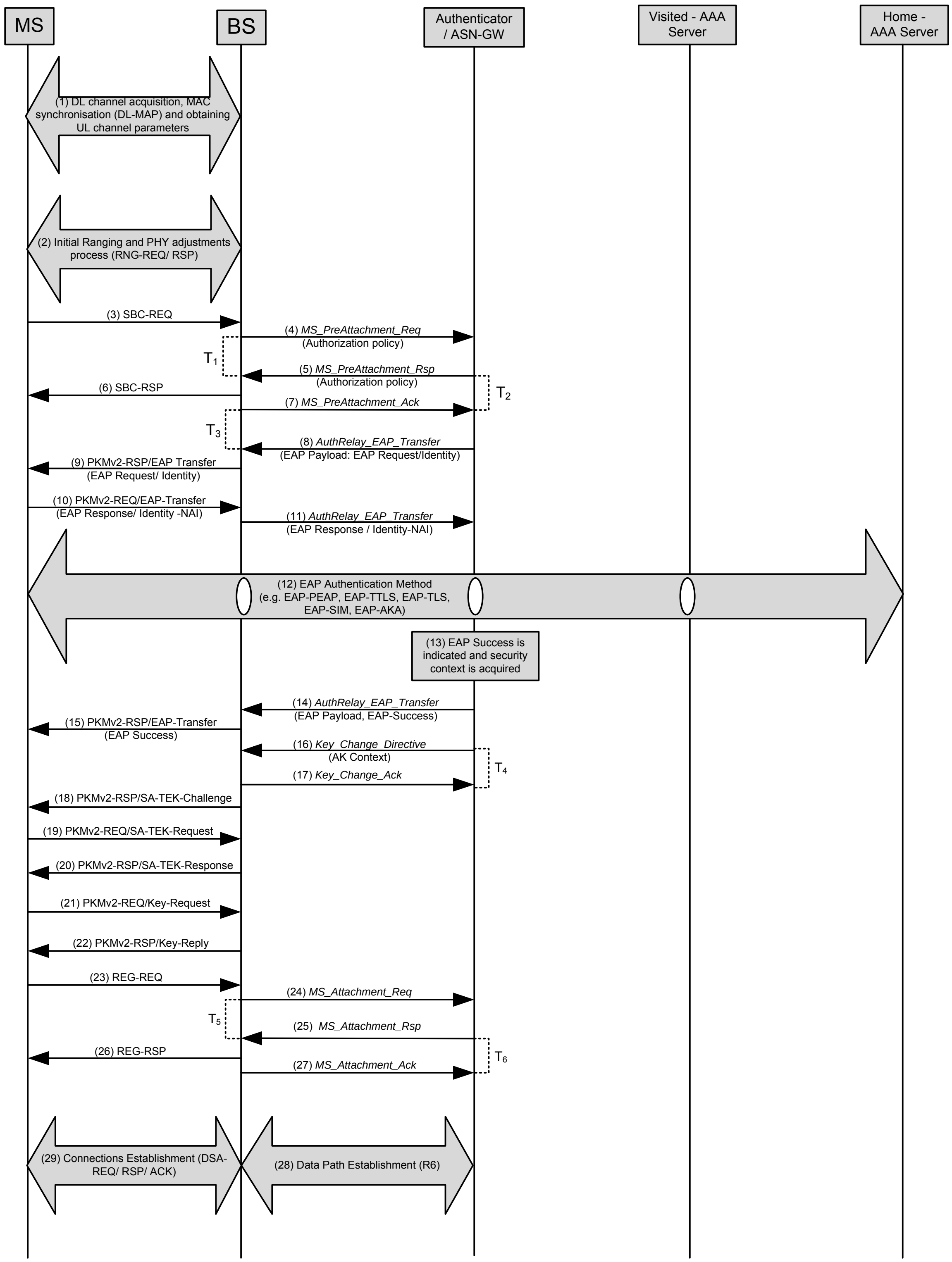


Figure 4.9.2 SS Initial Network Entry [WiMAX NWG Stage 3 Rel 1.0.0.]

The detail description of sequence flow specific to SS authentication and key exchange is as following. The step numbers are matching the same sequence number in the figure of SS Initial Network Entry.

STEP 4 – 5: After BS receiving SBC-REQ, BS informs the Authenticator the new SS entering the network.

STEP 6: BS sends SBC-RSP message to MS enforcing the authentication framework policy (PKMv.2, single EAP, CMAC mode).

In the case MS does not receive SBC-RSP, it will retransmit SBC-REQ.

STEP 7 – 8: BS confirms with the Authenticator the receiving SBC-RSP from the SS, and initiates EAP authentication procedure with the SS.

STEP 9: The BS sends the EAP Request/ Identity payload in the PKMv2-RSP/EAP-Transfer message to the SS.

STEP 10: The SS responds with EAP Response/ Identity message providing NAI. This message is transferred to BS over PKMv2 EAP-Transfer/ PKM-REQ message.

STEP 11: The BS relays EAP payload received in PKMv2 EAP-Transfer to the Authenticator over Authentication Relay protocol (*AR_EAP_Transfer* message).

STEP 12: The SS is using RADIUS Access-Request message. The EAP authentication process (tunneling EAP authentication method) is performed. The detail RADIUS messages are detailed in NWG R1.0.0 Stage 3, section 5.4.1.

STEP 13, 14, 15: The Authenticator receives indication about the successful completion of EAP-based authentication, the SS authorization profile and the required security context (i.e. MSK (Master Session Key) key and its lifetime). The BS sends EAP payload to the SS in PKMv2 EAP-Transfer/ PKM-RSP message.

STEP 16, 17: The BS receives the completion of authentication process from the Authenticator.

In the case authentication failure signal is received from the Authenticator (RADIUS Access-Reject with EAP-Failure), the Authenticator may decide to restart EAP authentication process (by sending the new EAP Request Identity) or bring down the user. In the latter case, the Authenticator proceeds with MS Network Exit procedure.

STEP 18, 19, 20: The BS and the SS starts PKMv2 3-way handshake (SA-TEK-Challenge/Request/Response exchange) to verify the AK to be used and to establish the Security Association(s) pre-provisioned for the SS.

STEP 21, 22: The SS acquires the valid TEK keys by using PKMv2 Key-Request/ Reply messages.

4.9.1.6 Registration

When PKMv2 3-way handshake is completed, the SS proceeds with Registration procedure. The purpose of the registration is to set up the provisioned data connections between MS and BS. It starts with a REG-REQ (Table 21 of [802.16-2004]) from the mobile and ends with a REG-RSP (Table 22 of [802.16-2004]) from the BS.

4.9.1.7 Establish IP connectivity

Establish IP connectivity should be performed on the SS's secondary Management Connection. The SS SHALL support the DHCP client function as defined in RFC 2131. In real deployment, it requires a DHCP server in the network. A DHCP module is needed for this purpose. In order to acquire an IPv4 address, the SS SHALL send a DHCPDISCOVER message to the network (the DHCP module) over the initial service flow. Upon receiving the DHCPOFFER message from the network (the DHCP module), the SS SHALL follow the procedures defined in RFC 2131 to select and configure an IPv4 address included in the DHCPOFFER message.

4.9.1.8 Establish time of day

Establish time of day should be performed on the SS's secondary Management Connection. IETF RFC 868 – time protocol is used to establish time of day. The request and response are transferred using UDP. The time is the number of seconds since 00:00 (midnight) 1 January 1900 GMT, and only need to accurate to the nearest second. The time retrieved from the universal coordinated time (UTC) server needs to be combined with the time offset received from the DHCP response to create the current local time.

4.9.1.9 Transfer operational parameters

When DHCP is successful, the SS need download the SS Configuration File from the BS by using TFTP on the SS's secondary Management Connection. The SS Configuration File is specified in [802.16-2004] section 9.2. SS needs process all standard configuration settings. The SS sends TFTP-CPLT message to the BS. When the configuration file completed download successfully, the SS need notify the BS by transmitting a TFTP-RSP (Config File TFTP Complete Response) message on the primary management connection. The transmission need continue periodically until the BS receives TFTP-RSP with "OK" response from the SS; or terminated due to retry exhausted.

4.9.1.10 Setup connection

After transfer operational parameters (for managed SS), or after registration (for unmanaged SS), the BS sends DSA-REQ (Table 38 of [802.16-2004]) message to the SS to setup the connection. The SS responds with DSA-RSP (Table 39 of [802.16-2004]) message. A DSA-REQ (Dynamic Service Addition Request) message contains a service flow reference and QoS parameter set (marked either for admission-only or for admission and activation). If the transaction is failed, the DSP-RSP needs include the Service Flow Error Set.

4.9.2 Handover

The system simulation defined elsewhere in the document deals with throughput, spectral efficiency, and latency. User experience in a mobile broadband wireless system is also influenced by the performance of handover. This section focuses on the methods to study the performance of handover which affects the end-users experience. Only intra-radio access technology handover is considered; inter-radio access technology handover is not considered.

The handover procedure consists of cell reselection via scanning, handover decision and initiation, and network entry including synchronization and ranging with a target BS. During scanning interval, MSs scan available neighbour BSs with or without association (i.e. association levels 0, 1, and 2) as specified on IEEE 802.16e-2006. Because association is an optional feature and only the scanning without association is mandated by WiMAX Forum profile, scanning without association is considered as a default.

Latency is a key metric to evaluate and compare various handover schemes as it has direct impact on application performance perceived by a user. Total handover latency is decomposed into several latency elements. Further, data loss rate and unsuccessful handover rate are important metrics. For simplicity, one mobile MS and multiple fixed MSs can be modeled as a baseline for the mobility simulations. The mobility related performance metrics shall be computed only for this mobile terminal. The mobile speed is fixed in every simulation drop and aligned with mobility mix associated with the specified channel models.

For parameters such as cell size, DL&UL transmit powers, number of users in a cell, traffic models, and channel models; the simulation follows the simulation methodology defined elsewhere in the document. For a network topology, as shown in Figure 4.9.5, a 10 cell topology with each cell of three sectors is used. In the topology, both serving and target cells in the center should have one tier of neighboring cells as interferers. In contrast to the 19 cell topology (shown in Section 2 based on one center cell with two tier of neighboring cells as interferers), the 10 cell topology approach serves a specific handover performance evaluation with a reduced simulation complexity.

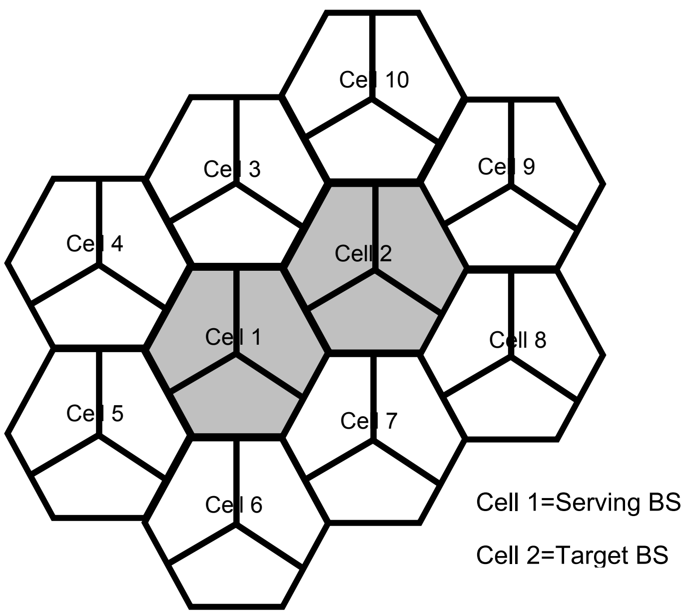


Figure 4.9.3 Cell Topology

Given the topology shown in Figure 4.9.3, the movement of the single mobile terminal is constrained to one of the following Trajectories 1 and 2.

Trajectory 1: In this trajectory, the MS moves from Cell 1 to Cell 2 along the arrow shown in Figure 4.9.4. The trajectory starts from the center of Cell 1 to the center of Cell 2 while passing through the midpoint of the sector boundaries. The purpose of this trajectory is to evaluate handover performance in a scenario where the signal strength from the serving sector continuously decreases whereas the signal strength from the target sector continuously increases.

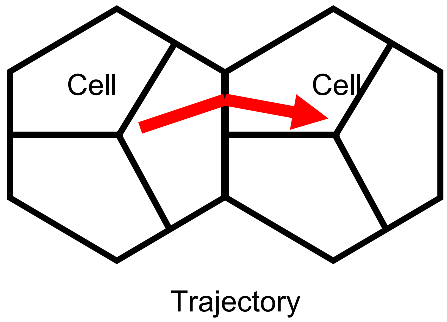


Figure 4.9.4 Trajectory 1

Trajectory 2: In this trajectory, the MS moves from Cell 1 to Cell 2 along the arrow. The MS moves along the sector boundary between Cell 1 and Cell 2 until the midpoint of the cell boundary between Cell 1 and Cell 2. The purpose of this trajectory is to evaluate handover performance when the MS moves along the boundary of two adjacent sectors.

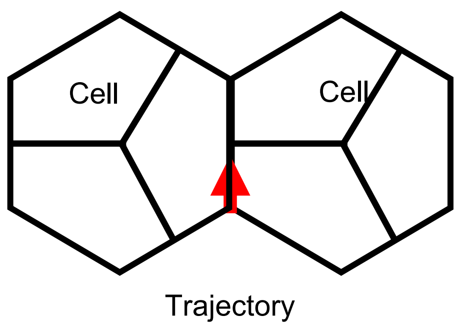


Figure 4.9.5 Trajectory 2

Handover Evaluation Procedure

1. The system may be modeled using the 10 cell topology as illustrated in Figure 4.9.3 for the evaluation of handover performance. Each cell has three sectors and frequency reuse is modeled by planning frequency allocations in different sectors in the network.
2. *N* MSs are dropped uniformly in each cell. Different load levels in the network are simulated by changing the number of MSs and the traffic generated.
3. Path loss, shadow fading and fast fading models for each MS should be consistent with the models defined elsewhere in the document. Fading signal and fading interference are computed from each mobile station into each sector and from each sector to each mobile for each simulation interval.
4. The trajectoriy 1 and trajectory 2 defined above should be used to model the movement of a single MS associated with the center cell. The locations of all other MSs are assumed to be fixed and the serving sector for the fixed MSs does not change for the duration of the drop.

5. Path loss, shadow fading and fast fading are updated based on location and velocity of a moving MS. As the MS moves along the specified trajectory, the target sector is chosen according to the metric used to perform handover.
6. Traffic from the fixed MSs constitutes background load. Any traffic type among the traffic model specified in Section 3 can be chosen according to the simulation objective. For the moving MS, one of the traffic type (e.g., VoIP, HTTP or FTP) should be generated to measure the application performance metrics affected by handover. Start times for each traffic type for each user should be randomized as specified in the traffic model being simulated.
7. Statistics related to handover metrics are collected for the moving MS only.
8. Packets are not blocked when they arrive into the system (i.e. queue depths are infinite).
9. Sequences of simulation are run, each with a different random seed. For a given drop the simulation is run for this duration, and then the process is repeated again, with the MSs dropped at new random locations. A sufficient number of drops are simulated to ensure convergence in the system performance metrics.

Handover Performance Metrics

The following parameters should be collected in order to evaluate the performance of different handover schemes. These statistics defined in this section should be collected in relation to the occurrence of handovers. A CDF of each metric may be generated to evaluate a probability that the corresponding metric exceeds a certain value. .

For a simulation run, we assume:

- Total number of successful handovers occurred during the simulation time = $N_{HO\_success}$
- Total number of failed handover during the simulation time = N_{HO_fail}
- Total number of handover attempts during the simulation time = $N_{attempt}$, where $N_{attempt} = N_{HO\_success} + N_{HO\_fail}$

1- **Radio Layer Latency:** This value measures the delay between the time instance $T_{1,i}$ that an MS transmits a serving BS its commitment to HO (for a HHO, this is the time that the MS disconnects from the serving BS) and the time instance $T_{2,i}$ that the MS achieves the success of the PHY layer synchronization (i.e., frequency and DL timing synchronization) due to handover occurrence i. The exact thresholds for successful PHY synchronization are for further study. For this metric, the following will be measured.

- Average Radio Layer Latency = $\dfrac{\sum_{i=1}^{N_{HO\_success}} (T_{2,i} - T_{1,i})}{N_{HO\_success}}$
- Maximum Radio Layer Latency = $\underset{1 \le i \le N_{HO\_success}}{Max} [Radio\ Layer\ Latency\ of\ handover\ occurrance\ i\ ]$]

2- **Network Entry Time:** This value represents the delay between an MS's radio layer synchronization at $T_{2,i}$, and its completion of a Layer 2 network entry procedure at $T_{3,i}$ due to handover occurrence i.

This consists of ranging, UL resource request processes (contention or non-contention based), negotiation of capabilities, and registration. All HO MAC messages success/failure rates must be consistent with the packet error rate used for data.

- Average Network Entry Time = $\dfrac{\sum_{i=1}^{N_{HO\_success}}(T_{3,i} - T_{2,i})}{N_{HO\_success}}$
- Maximum Network Entry Time = $\underset{1 \le i \le N_{HO\_success}}{Max} [Network\ Entry\ Time\ of\ handover\ occurance\ i\ ]$

3- **Connection Setup Time:** This value represents the delay between the completion of Layer 2 network entry procedure at $T_{3,i}$ and the transmission of first data packet from new BS (target BS) at $T_{4,i}$ due to handover occurrence i.. This consists of DL-UL packet coordination and a path switching time. A path switching time, as a simulation input parameter, may vary depending on network architecture.

- Average Connection Setup Time = $\dfrac{\sum_{i=1}^{N_{HO\_success}}(T_{4,i} - T_{3,i})}{N_{HO\_success}}$
- Maximum Connection Setup Time = $\underset{1 \le i \le N_{HO\_success}}{Max} [Connection\ Setup\ Time\ of\ handover\ occurance\ i\ ]$

4- **Service Disruption Time:** This value represents time duration that a user can not receive any service from any BS. It is defined as the sum of Radio Layer Latency, Network Entry Time and Connection Setup Time due to handover occurrence i. This metric can be expressed by the CDF in order to obtain various statistics such as mean, variance, etc.

5- **Data Loss**: This value represents the number of lost bits during the handover processes. $D_{RX,i}$ and $D_{TX,i}$ denotes the number of received bits by the MS and the number of total bits transmitted by the serving and the target BSs during the MS performs handover occurrence *i,* respectively. Traffic profiles used for the simulation experiments to compare different handover schemes need to be identical.

Data Loss = $\dfrac{\sum_{i=1}^{N_{HO\_success}} D_{TX,i} - D_{RX,i}}{N_{HO\_success}}$

6- **Handover Failure Rate:** This value represents the ratio of failed handover to total handover attempts. A failed handover happens if a handover is executed while the reception conditions are inadequate.

Handover Failure Rate = $\dfrac{N_{HO\_fail}}{N_{attempt}}$

Figure 4.9.6 illustrates an example of signalling diagram for performance metrics 1, 2, and 3 (i.e., radio layer latency, network entry time, and connection setup time)

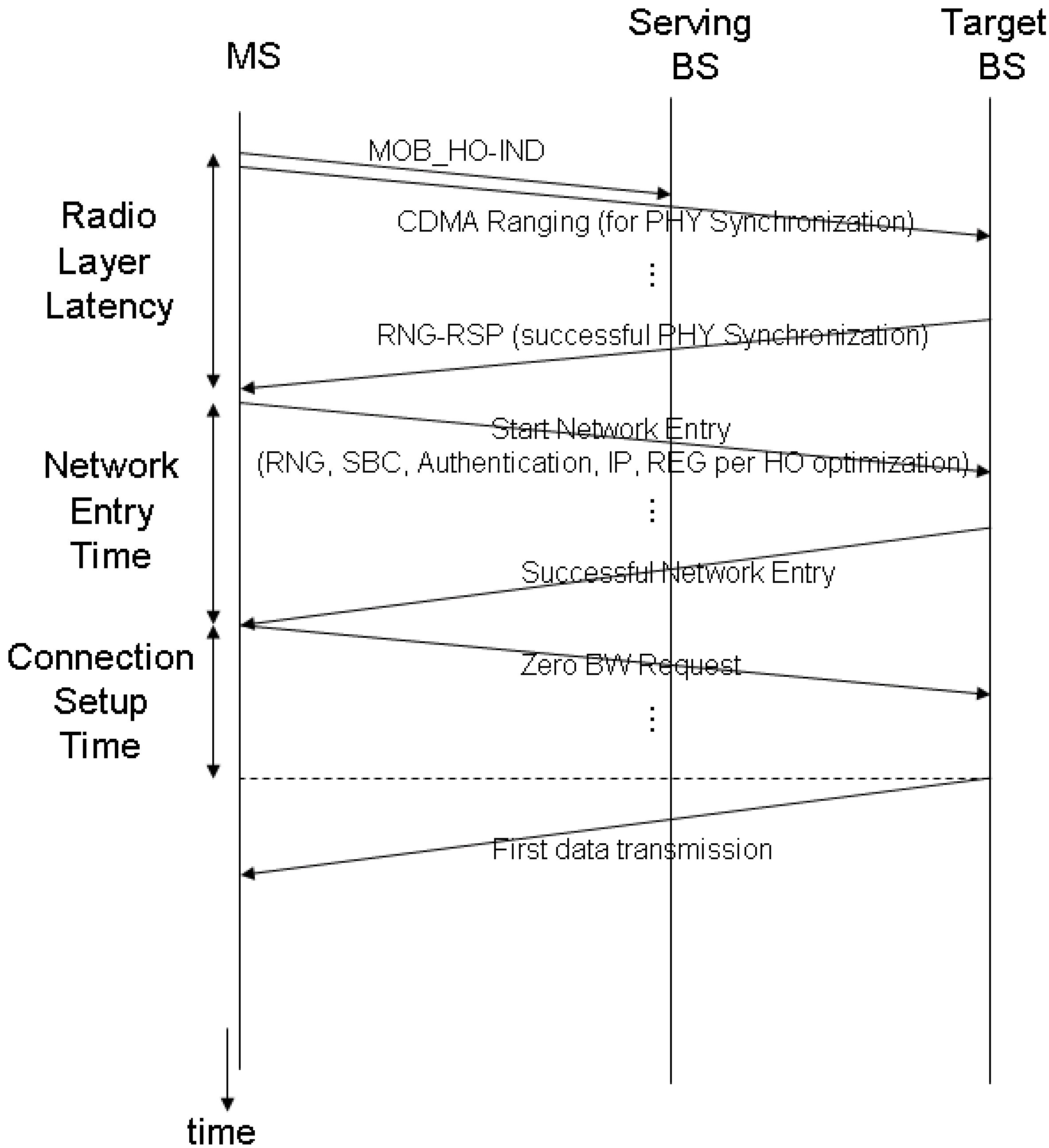


Figure 4.9.6 : An example of signalling flow for performance metrics

4.10 Power management - Sleep-Idle Mode

The 802.16e standard specifies two states for power management at the MS, namely sleep mode and idle mode to enable power efficiency and to decrease air link resource usage.

4.10.1 Sleep Mode

Sleep Mode is a state in which the MS can remain absent for certain pre-negotiated intervals from the BS air interface, while still actively connected to the BS. The periods of MS absence are characterized by MS unavailability to the serving BS, to DL or UL traffic. In Sleep mode, the BS retains all the information related to the connections that currently belong to the MS as it does during Normal Operations.

For each involved MS, the BS needs to keep context related to a certain Power Saving Class (PSC). A PSC can contain one or more connections. A PSC may be repeatedly activated or deactivated. To activate a PSC means to start sleep/listening windows associated with the class.

The 802.16e-2005 standard specifies three different PSCs, each of them specifying different parameter sets, procedures of activation/deactivation and policies for MS availability for data transmission. We will only describe PSC 1 here since that's the only class which has been certified for implementation at the BS by WiMAX Forum [WiMAX Forum Mobile System Profile Rel 1.0]. Thus, all connections in the simulation will map to the PSC of type 1. The time MS spends in sleep mode as defined by a given PSC is characterized by sleep and listening windows.

Figure 4.10.1 shows an illustration of sleep mode. A *sleep window* is the period of absence negotiated by the MS for a given PSC. A *listen window* is the period during which the MS can receive traffic associated with the connections mapped to that class. *Unavailability interval* is a time interval that does not overlap with any listening window of any active PSC. *Availability interval* is a time interval that does not overlap with any Unavailability Interval. The figure below shows the difference between sleep window and unavailability interval and is valid if all the connections in the MS map to either PSC_ID 1 or PSC_ID 2. The MS will maintain a listen/sleep window with each PSC and then calculate its unavailability and availability intervals.

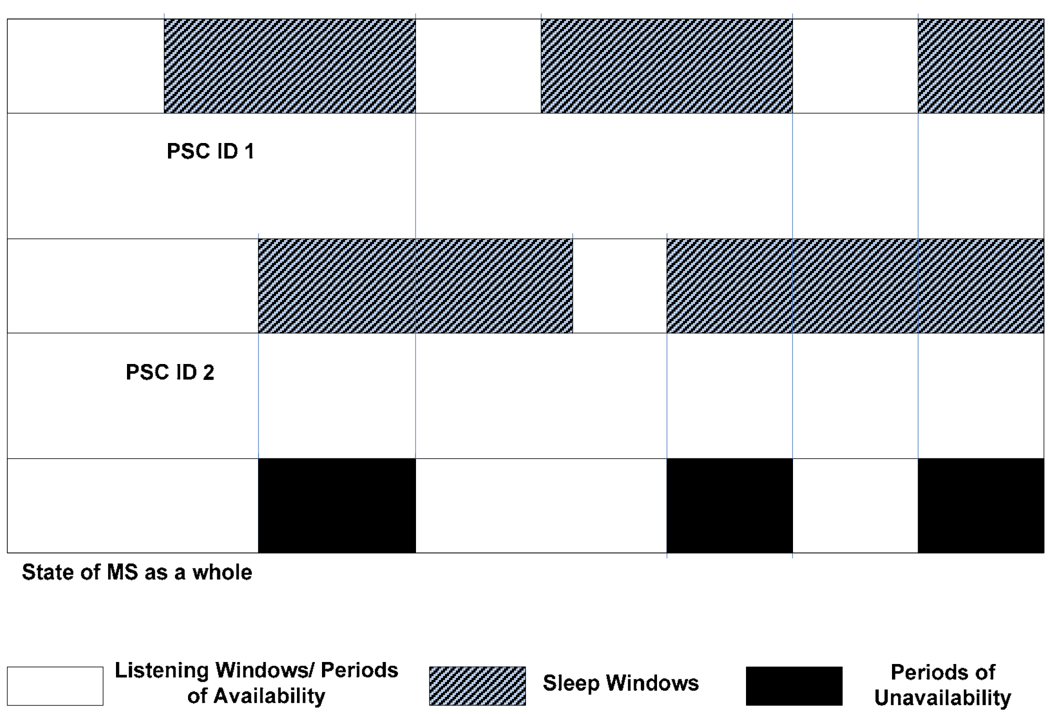


Figure 4.10.1: Sleep Mode in 802.16e

During Unavailability interval the BS shall not transmit to the MS, so the MS may power down one or more physical components or perform other operations that do not involve communication with the BS —scanning neighbor BSs, associating with neighbor BSs, etc. *If there is a connection at the MS, which is not associated with any active PSC, the MS will be considered available by the BS on permanent basis.*

During Availability interval the MS is expected to receive all DL transmissions same way as in the state of normal operations (no sleep). In addition, the MS needs to be synchronized to the BS. This it will do by checking the DCD and UCD change counts and the frame number of the DL-MAP PHY Synchronization field to verify synchronization with the BS. Upon detecting a changed DCD and/or UCD count in the DL MAP, unless using the Broadcast Control Pointer IE for tracking and updating DCD and/or UCD changes, the MS shall continue reception until receiving the corresponding updated message.

If the BS transmits the Broadcast Control Pointer IE, the MS shall read and react to this message according to the following:
1) If the DCD_UCD Configuration Change Counter has changed since MS last decoding of this IE, even if scheduled to be in a sleep interval the MS shall awaken at DCD_UCD Transmission Frame in time to synchronize to the DL and decode the DCD and UCD message in the frame, if present. If the MS fails to decode one or both of DCD and UCD, or no DCD or UCD was transmitted by the BS, the MS shall continue decoding all subsequent frames until it has acquired both updated DCD and UCD. Upon successful completion of DCD and UCD decoding, the MS shall immediately return to regular sleep mode operation.
2) If Skip Broadcast_System_Update is set to '0', even if scheduled to be in a sleep interval, the MS shall awaken at Broadcast_System_Update_Transmission_Frame in time to synchronize to the DL and decode and read the DL-MAP and any message, if present. Upon completion, the MS shall immediately return to regular sleep mode operation.

During Unavailability intervals for MS, the BS may buffer or drop MAC SDUs addressed to unicast connections bound to an MS, this depends upon the service flow parameters for the given connection and the resources available at the BS and is left to implementation. The BS may choose to delay transmission of SDUs addressed to multicast connections until the following Availability Interval, common for all MSs participating in the multicast connection, but the mechanism for this is not specified in the standard. Also, the algorithm to determine the length of sleep/listen window is up to the implementation.

4.10.1.1 Power Saving Class of type 1

PSC type I is recommended for BE and NRT_VR types of connections, but can be used for others as well.

For definition and/or activation of one or several PSCs of type I by the MS, the following 2 ways are supported:

a) MS shall send MOB_SLP-REQ (for definition and activation) and the BS shall respond with MOB_SLP_RSP message. The MS may retransmit MOB_SLP-REQ message if it does not receive the MOB_SLP-RSP message within the T43 timer (table 342 in 802.16e).
b) MS shall send Bandwidth request and UL sleep control header (for activation only) and the BS will respond with DL Sleep control extended subheader.

It is also possible for the BS to define and activate a PSC by sending an unsolicited MOB_SLP_RSP message to the MS.

Note: PSC definition is not the same as activation. To define a Power Saving Class means to specify the various parameters related to the Power Saving Class such as length of sleep window etc, for example, when the MS sends the MOB_SLP_REQ with the parameters in it, it is defining the PSC. To activate a PSC means that the MS will actually start its sleep cycle at a given frame.

The MOB_SLP_REQ, MOB_SLP_RSP and MOB_TRF_IND message formats are defined in tables 109c, 109d and 109e in the specification [802.16-2004 Cor2/D4] respectively. The DL sleep control extended subheader and the UL sleep control header message formats are defined in tables 13e and 7e respectively in 802.16e standard.

The following are relevant parameters in the definition of sleep/listen windows associated with PSC type 1:

- Initial-sleep window
- Final-sleep window base
- Listening window
- Final-sleep window exponent
- Start frame number for first sleep window
- TRF-IND required
- Traffic triggered wakening flag

The length of the sleep window is calculated using the following formula:

Sleep window = min (2(Previous sleep window), Final-sleep window base * 2 ^ (Final-sleep window exponent)).*

In addition to the above parameters, the standard specifies following constants (table 342 in 802.16e):

1) Min_Sleep_Interval = 2 Frames
2) Max_Sleep_Interval = 1024 Frames
3) Max_Listening_Interval = 64 Frames

Default Parameters for Timers/Sleep windows

1) T43 Timer = 2 Frames. Resend MOB_SLP_REQ within T43 timer, if no response received.
2) Sleep window size = 2 Frames
3) Listen Window Size = 2 Frames.

Sleep windows are interleaved with listening windows of fixed duration.

Note: It is possible to keep all sleep window lengths of same size by specifying the final-sleep window exponent as 1 and final sleep window base the same as the initial-sleep window. Thus, it is possible to use the PSC I for connections that require a regular wake-up interval such as NRT-PS or ERT-PS types of connections and thus use PSC type I for other types of connections as well.

To deactivate a PSC, the MS may send a MOB_SLP_REQ message or UL Sleep control extended header (with the BW_REQ message) or the BS may send MOB_SLP_RSP or DL Sleep control extended subheader.

The PSC shall be deactivated if traffic triggering wakening flag is set to 1 and if any of the following conditions is met:
— MS receives a MAC PDU over any connection belonging to the Power Saving Class during the listening window.
— MS transmits a bandwidth request with BR set to a value other than 0 on any connection belonging to the Power Saving Class. This may occur in the slot provided by BS during listening window or the MS can also send a CDMA code during availability or unavailability interval.

If the TRF_IND_Required flag was set to 1 in MOB_SLP_RSP, this means that the MS must receive a MOB_TRF_IND message during each listening window. If no message was received, the PSC will be deactivated. The MOB_TRF_IND message is sent by the BS on broadcast CID or Sleep mode multicast CID during listening window. When an MS receives an UL allocation after receiving a positive MOB_TRF-IND message indication, the MS shall transmit at least BR message (if there is no data to transmit, BR field of the BR PDU shall be set to 0).

The implementation of MS and BS initiated Sleep Mode is shown in the flowcharts below in the case where TRF_IND_Required flag is set to 1 and Traffic Triggered Wakening Flag is set to 1. In case Traffic Triggered Wakening Flag is set to 0, then the PSC is not de-activated even if the TRF_IND is positive and the MS may send/receive data during this window. However, if it sends a BR with value set greater than 1, than the PSC will be deactivated.

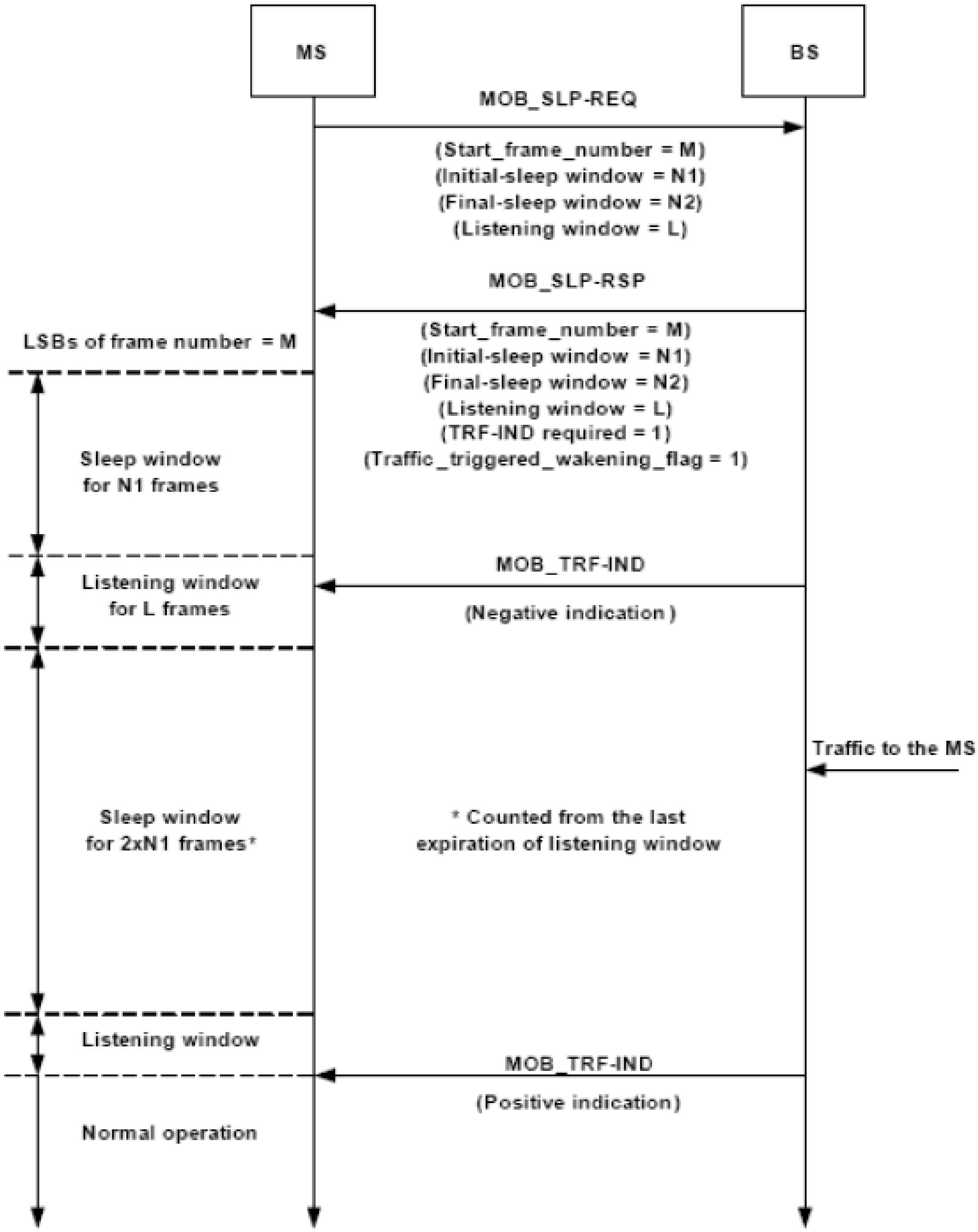


Figure D.7—Example sleep mode — MS initiated <u>for the case of TRF_IND required=1 and Traffic_triggering_wakening_flag=1</u>

Figure 4.10.2: MS-Initiated Sleep-Mode Messaging [figure D.7 in [802.16-2004 Cor2/D4]]

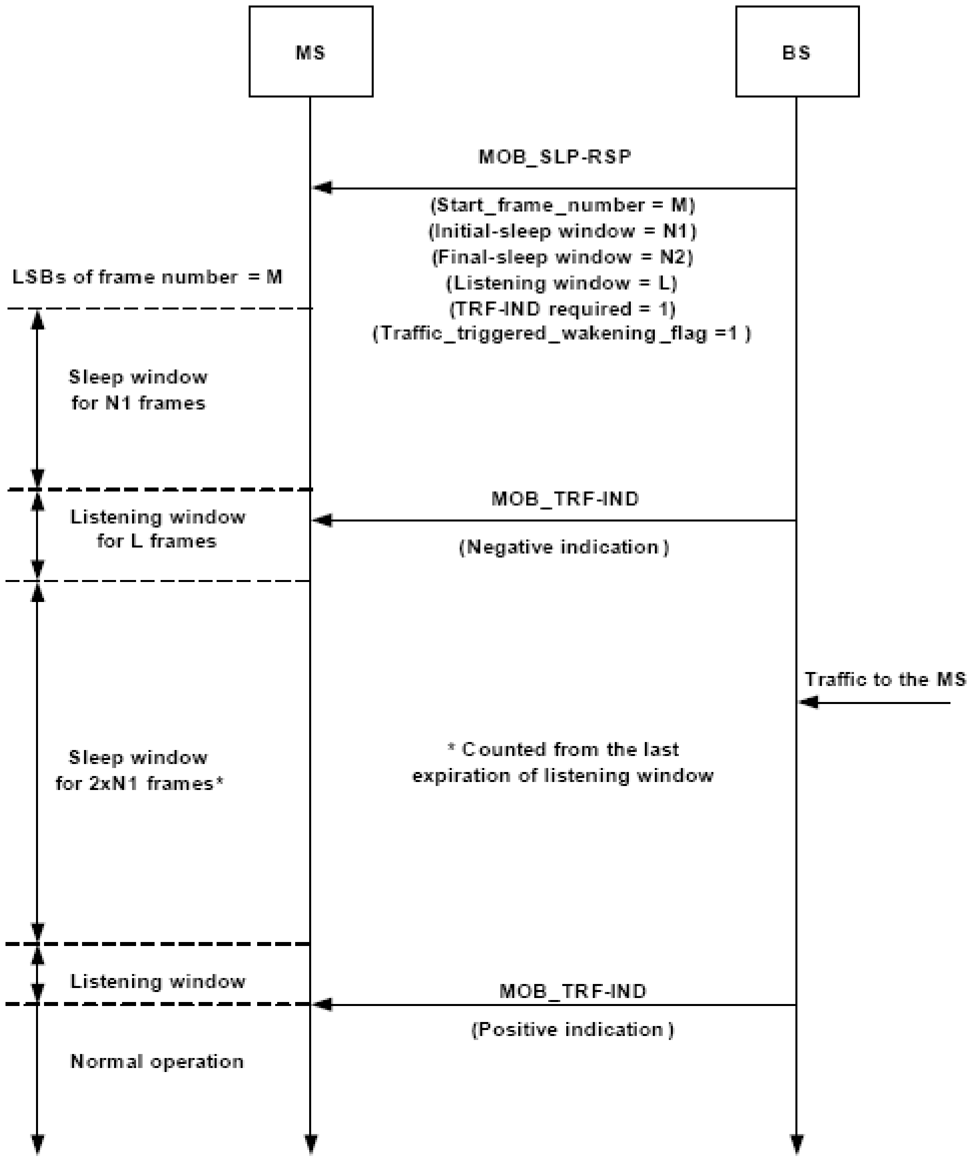


Figure 4.10.3: BS-initiated Sleep Mode in case of Traffic Triggered Wakening Flag = 1 and TRF_IND_REQUIRED = 1. [Figure D.8 in [802.16-2004 Cor2/D4]]

4.10.1.2.1 Periodic Ranging during Sleep Mode of OFDMA PHY

Upon expiration of MS Timer T4 shown in Figure 90 in 802.16e specification, MS may perform CDMA-based periodic ranging. CDMA code-based Periodic Ranging can be performed during any interval or MS may skip the periodic ranging depending on implementation. MS may anonymously extend its Availability interval in order to wait the reception of RNG-RSP message with ranging status = Success or find an appropriate ranging opportunity for sending/resending of CDMA code-based ranging request. Such temporary extension of Availability interval is not known by the BS and does not affect the previously negotiated status of the sleep mode between MS and BS.

4.10.1.3 Trigger-based actions during Sleep Mode

The MS in Sleep mode may perform certain actions based on events such as CINR, RSSI, RTD etc. triggers received in DCD message or in MOB_NBR_ADV message. For this purpose, the MS may include the Enabled-Action-Triggered TLV [802.16-2004, Cor/D4, section 11.1.8.1]. MS may include this TLV item in MOB_SLP-REQ message to request an activation of type of Power Saving Class by setting the bit corresponding to the action to 1. BS shall include this TLV in MOB_SLP-RSP message transmitted in response to the MOB_SLP-REQ message.

4.10.1.4 Updating SLPID

Each PSC is identified by a SLPID. To identify if there is any traffic in the MOB_TRF_IND message, the MS will go through the list of SLPIDs and try to find its own. If it finds a SLPID_Upate_TLV in the message containing its current SLPID, then it will add it to the new one contained in the TLV as shown in section 11.1.8.2 in [802.16-2004 Cor2/D4]. This TLV may include multiple Old_New_SLPID values for the MSs negatively indicated in MOB_TRF-IND message.

4.10.1.5 Handovers and Sleep Mode

The MS shall stop the sleep mode before any controlled handover to the target BS. After the MS completes handover to target BS, the MS shall discard all the sleep mode related information associated with previous serving BS. Also, new Serving BS shall regard any MS performing handover as operating in normal operation without entering sleep mode first. MS may include Power_Saving_Class_Parameters in RNG-REQ message. The scope of Power Saving Class parameters is confined to the HO process. MS may enter sleep mode after HO. If the MS enters sleep mode, it shall transmit MOB_SLP-REQ message or Bandwidth request and uplink sleep control header to activate the previously defined Power Saving Class. Also, BS may transmit MOB_SLP-RSP or DL Sleep control extended subheader in unsolicited manner in order to activate the previously defined Power Saving Class.

4.10.1.6 Evaluation methodology for Sleep mode

While sleep mode is primarily used to conserve power consumption of the MS, it can impact QoS parameters such as application latency, dropped packets, delay and the QoS of traffic, besides also conserving air-link resources by freeing the BS to schedule other clients during a client's sleep period.

Sleep Mode Evaluation Procedure

1. The system may be modeled using the default simulation topology as specified elsewhere in the document.
2. *N* MSs are dropped uniformly in each cell. Different load levels in the network are simulated by changing the number of MSs and the traffic generated.
3. Path loss, shadow fading and fast fading models for each MS should be consistent with the models defined elsewhere in the document. Fading signal and fading interference are computed from each mobile station into each sector and from each sector to each mobile for each simulation interval.
4. The MS performing sleep mode is stationary and the possibility of handover during sleep mode is thus avoided, the location of this MS in the cell may be determined randomly in the center

cell in different simulation runs. The locations of all other MSs are assumed to be fixed and the serving sector for the fixed MSs does not change for the duration of the drop.

5. The traffic models used should be consistent with those specified in chapter 3 of this document. Sleep mode should be evaluated for different applications (HTTP, VoIP, Email and traffic mixes, each as a separate case). Start times for each traffic type for each user should be randomized as specified in the traffic model being simulated.
6. Statistics related to sleep mode metrics are collected for the moving MS only.
7. Packets may be dropped when the QoS constraints on delay (depending upon application QoS parameters) are not met.
8. Sequences of simulation are run, each with a different random seed. For a given drop the simulation is run for this duration, and then the process is repeated again, with the MSs dropped at new random locations. A sufficient number of drops are simulated to ensure convergence in the system performance metrics.

The performance metrics to take into account should be as follows:

1) MS Unavailability ratio - MS_{UAI}: The total percentage of time in terms of the number of frames spent in sleep window.

$$MS_{UAI} = N_{UAI} / N_{TOT},$$

where, N_{UAI} = Total number of frames spent by MS in Unavailability interval

N_{TOT} = Total number of frames simulated

2. Application Downlink delay DL_{DELAY} : Average application air-link delay (in frames) caused due to sleep. This should be calculated from the time T_i when the BS would be ready to transmit a frame, but has to wait until the listen window interval to send it instead, T_{i+n} (where n is the # of frames spent waiting in buffer until Listen window).

$$DL_{DELAY} = T_{i+n} - T_i$$

3. Sleep mode message overhead: This is the message overhead (in bits) that is required for defining/activating/de-activating sleep mode for the MS.

4.10.2 Idle Mode

Idle Mode is intended as a mechanism to allow the MS to become periodically available for DL broadcast traffic messaging without registration at a specific BS as the MS traverses an air link environment populated by multiple BSs, typically over a large geographic area. Idle Mode benefits MS by removing the active requirement for HO, and all Normal Operation requirements. By restricting MS activity to scanning at discrete intervals, Idle Mode allows the MS to conserve power and operational resources and also allows BS to alert MS to pending DL traffic directed towards it.

For Idle mode, it is necessary to understand paging. The BSs are divided into logical groups called paging groups. The purpose of these groups is to offer a contiguous coverage region in which the MS does not need to transmit in the UL, yet can be paged in the DL if there is traffic targeted at it. The paging groups should be large enough so that most MSs will remain within the same paging group most of the time, and small enough such that the paging overhead is reasonable. Figure 4.10.2-1 (figure 130i from 802.16e std.) shows an example of four paging groups defined over multiple BS arranged in a hexagonal grid.

A BS may be a member of one or more Paging Groups. The paging-groups are defined in the management system.

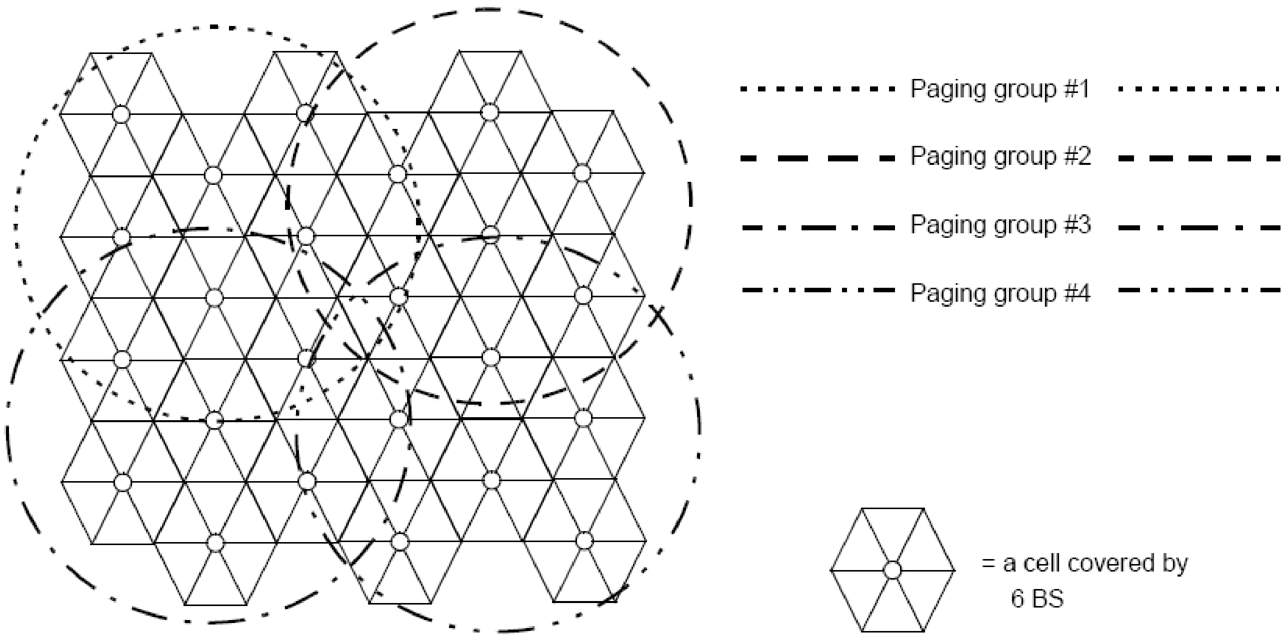


Figure 4.10.4 Paging Groups Example [Figure 130i in 802.16e]

In the simulation, the MS will maintain a state called Idle mode during which the MS will either be in the Unavailable Interval or in the Listening Interval (as shown in the figure 4.10.2-2). The Listening Interval shall overlap with the BS Paging Interval during which the BS will send out a Broadcast Paging Message.

Idle Mode is comprised of the following activities/stages:
— MS Idle Mode Initiation
— Cell Selection
— MS Broadcast Paging Message time synchronization
— MS Paging Unavailable Interval
— MS Paging Listening Interval
— BS Paging Interval
— BS Broadcast Paging message
— Paging Availability Mode Termination

4.10.2.1 MS Idle Mode Initiation Implementation

Idle mode is initiated using the one of the following processes:

1) MS initiated:

- MS sends DREG_REQ message De-Registration Request Code = 0x01 and other encoded TLVs namely Paging Information and Paging Controller Id and Idle Mode Retain Information TLVs and optionally the HMAC/CMAC Tuple and MAC_HASH_Skip_Threshold TLV. The DREG_REQ message format is shown in Table 87 from [802.16-2004 Cor2/D4].
- MS receives DREG_CMD message from the BS within T45 Timer expiry. The DREG_CMD message will contain Action code = 0x06 (action codes and actions can be referred to Table 55 in [802.16-2004 Cor2/D4]).
- If timer expires, DREG_REQ message is sent until DREG_Request_Retry_count is exhausted. T45 Timer default value is 1 sec and the default number of DREG_Request_Retry_count is 3, maximum is 16.
- When the BS responds with DREG_CMD message (section 11.14 in 802.16e), it starts the Management_Resource_Holding_Timer (default value = 500 ms, max = 1sec) until which it retains all MS state during Idle mode. Once this timer expires, all information will be deleted unless otherwise specified by the Idle Mode Retain Information TLV.

2) BS-initiated:

- BS sends unsolicited DREG_CMD message with Action Code = 0x05 to the MS which contains the Paging Cycle and Idle Mode Retain Information TLV. BS starts Management_Resource_Holding_Timer.
- MS responds with DREG_REQ message with De-Registration Request Code = 0x02 within the T46 Timer.
- If timer expires, the Management_Resource_Holding_Timer is reset. The BS keeps resending the DREG_CMD message until DREG_Command_Retry_count is exhausted. The values of the timers are similar to the T45 and the DREG_Request_Retry_count.

To maintain Idle Mode, the following steps should be implemented.

1) BS maintains a global constant called the Idle Mode System Timer and MS maintains a global constant called the Idle Mode Timer, both of which have the same values (default in the spec. is 4096s, min is 128s).
2) The timers are both started when the BS responds with a DREG_CMD message directing MS transition into Idle Mode. The timer values are set in the DREG_CMD message
3) The MS needs to provide a Location Update message before the Idle Mode Timer Expires during any of the BS Paging Intervals. On a successful Location Update, the Idle Mode System Timer is restarted.
4) If there Idle Mode System Timer expires without any Location Updates, then the Paging Controller will discard all MS related Idle Mode information.

4.10.2.2 Cell Selection

Once the MS is successfully initiated into Idle Mode, it finds a preferred BS using the same process as during Handover, refer to section 4.9 of this document.

4.10.2.3 MS Broadcast Paging Message Synchronization

When the MS has successfully identified the preferred BS, it gets relevant information about the preferred BS paging information, which is provided in the RNG_RSP message (if using a different BS than the serving BS or from the DREG_CMD message otherwise). The relevant information includes: Page Group ID, Page Controller ID (may be same as BS ID, if BS is the Paging Controller), Paging Cycle, Paging Offset and Paging Interval. Using this information, the MS sets its own Unavailability Interval and Listening Interval as well as the Idle Mode Timer. The MS Unavailability Interval will be calculated by first calculating the frame number at which the next regular BS Paging Interval starts. The MS Unavailability Interval will then be the time until the Next Paging Interval deducted by the amount of time the MS spends scanning, decoding and synchronizing DL channels. The BS Paging Interval lasts for N frames and begins with the frame whose frame number N_{frame}, meets the following condition:

N_{frame} modulo PAGING_CYCLE == PAGING_OFFSET

on each BS, where N is Paging Interval Length (a global constant defined in Table 342 in 802.16e).

4.10.2.4 MS Paging Unavailable Interval

MS goes into MS Paging Unavailable Interval until the BS Paging Interval starts.

4.10.2.5 MS Paging Listening Interval/ BS Broadcast Paging Interval

MS side:

During the Listening Interval, the MS will basically synchronize to its Preferred BS in time to listen to the MOB_PAG_ADV messages that the BS sends out on the Broadcast CID or Idle Mode Multicast CID (defined in table 345 of 802.16e-2005 standard). The BS Broadcast Paging message includes an Action Code directing each MS (multiple MS's may be notified using the same message) notified via the inclusion of its MS MAC Address Hash field to do either:

0b00: no action required
0b01: perform Ranging to establish location and acknowledge message
0b10: perform initial network entry
0b11: Reserved

The BS can send a Broadcast Control Pointer IE and the MS will react to this message in the same manner is it does when it is in Sleep Mode as specified in section 4.10.1.

If the BS sends a Paging Message with Action Code "Perform Ranging", or "Enter Network" and if the BS does not receive RNG_REQ message from the MS in the MS Paging Listening Interval, the BS will retransmit the Broadcast Paging Message until it receives a response of the Broadcast Paging Retry Count (global constant defined in Table 342) decreases to zero. If no response is received and the retries are exhausted, the BS marks the MS as unavailable and the sends a backbone message to the Paging Controller to update the list of MSs that belong to the paging group. The MOB_PAG_ADV message format is shown in the Table 109p of [802.16-2004 Cor2/D4].

4.10.2.6 Paging Availability Mode Termination

Idle Mode may only be terminated through:

- MS re-entry to the network
- Paging Controller detection of the MS unavailability when the Paging Retry Count goes to 0
- Expiration of the Idle Mode System Timer

The MS can initiate termination of the Idle Mode at any time by sending a RNG-REQ message to the BS and setting Bit 0 of the Ranging Purpose Indication field as 1. Or on the other hand, the BS may inform the MS in the MOB_PAG_ADV message using the MS MAC Address Hash and Action Code 0b10 to “Enter the Network”. In this event, the MS will perform a network re-entry. If the BS sends a message with Action Code 0b01 “Perform Ranging”, the MS will perform MS Idle Mode Location Update using either the Secure Location Update or the Unsecure Location Update process, described later in this section.

In both cases for the OFDMA PHY, if a PHY specific ranging code and transmission opportunity is assigned to the MS in the MOB_PAG-ADV message, the MS shall transmitting the code at on the dedicated ranging region assigned in the UL-MAP-IE (UIUC = 12 and dedicated ranging indicator bit set to 1).

The procedure for PHY specific ranging code operation is described as follows:

1) After receiving the MOB_PAG-ADV and within the Page-Response window, the MS shall transmit the assigned ranging code at the transmission opportunity in the frames where dedicated ranging regions are assigned in the UL-MAP_IE (UIUC = 12 and dedicated ranging indicator bit set to 1). The assigned ranging code transmission can be terminated early if the MS receives a RNG-RSP message with ‘success’ status before the end of the Page-Response window.
2) In the case where RNG-RSP message with ‘continue’ status is received and within the Page-Response window, the MS shall transmit the assigned ranging code at the transmission opportunity in the next frame where the dedicated ranging region is assigned.
3) In the case where RNG-RSP message with ‘success’ status is not received within the Page-Response window, the MS shall continue with the normal initial ranging procedure for Network Re-entry from Idle (4.10.2.8) or Idle Mode Location Update (4.10.2.7).
4) In the case where no RNG-RSP message is received or no dedicated ranging region is assigned within the Page-Response window to the MS, the MS shall continue with the normal initial ranging procedure for Network Re-entry from Idle (4.10.2.8) or Idle Mode Location Update (4.10.2.7)
5) In all other cases, the MS shall use normal network re-entry or Idle Mode Location Update procedure, as described in 4.10.2.8.

To prevent collisions from multiple MSs trying to wake from Idle mode at the same time, the MS shall use random backoff the Initial_ranging_backoff_start and Initial_ranging_backoff_end described in Table 349 in 802.16e standard.

4.10.2.7 Location Update

Location Update is comprised of condition evaluation and update processing. An MS in Idle mode shall perform a Location Update process operation if any Location Update condition is met. There are four location update evaluation conditions of which the following 3 are supported in Ref [802.16-2004

Cor2/D4]: Paging Group Update, Timer Update and Power Down Update. All location updates involve sending a RNG-REQ message with the appropriate encoding (see table 364 in 802.16e standard). For simulation purposes, we will only implement 2 of them, namely Paging Group Update and Timer Update and that the BS always keeps the corresponding HMAC/CMAC tuple of the MS, so it's always a secure location update process.

Paging Group update: The MS performs Location Update process when the MS detects a change in paging group by monitoring the paging group identifier, PG_ID, which is transmitted by the Preferred BS in the DCD message or MOB_PAG-ADV broadcast message during the MS Paging Listening Interval. If the PG_ID detected does not match the Paging Group to which the MS belongs, the MS shall determine that paging group has changed and perform the Location Update process.

Timer update: The MS maintains Idle mode by periodically performing Location Update process prior to the expiration of the Idle Mode Timer.

4.10.2.7 Network Re-Entry from Idle Mode

The MS initiates Network Re-entry by sending a RNG-REQ including Ranging Purpose Indication TLV with Bit #0 set to 1 and Paging Controller ID TLVs (section 11.1.9.2 from 802.16e). For the purpose of simulation, we can assume that the MS and BS share a valid security context to minimize simulation time and effort.
When the target BS receives RNG-REQ and if it doesn't already have MS information from the backbone network, then it may acquire it from the Paging Controller. Network re-entry then proceeds the same as in initial network entry except that it may be shortened by target BS's possession of previous MS information from the backbone network. The target BS can omit some messages during network re-entry process by including an HO Optimization Process TLV in the RNG-RSP message that will specify the messages to be omitted.

During network re-entry, the target BS may notify the MS, through the Bit #7 MS DL data pending element of the HO Process Optimization TLV item in RNG-RSP, of post- network reentry MS DL data pending. Upon MS successful re-entry at target BS, now new serving BS, and new serving BS completing reception of any network re-entry pending MS DL data retained and forwarded, MS may re-establish IP connectivity and the new serving BS may send a backbone message over the backbone network to request the old serving BS or other network entity to stop forwarding pre-HO pending MS DL data.

Network entry/re-entry process completes with establishment of Normal Operations.

The target BS shall notify the Paging Controller via the backbone network of MS successful network reentry and the Paging Controller may send a backbone message over the backbone network to inform the BS at which the MS entered Idle Mode that the MS has resumed Normal Operations at the new serving BS.

4.10.2.8 Evaluation methodology for Idle mode

Idle mode helps conserve power in a MS when there are no applications running. In case of Idle mode, there are two primary scenarios to be considered, namely device-initiated network re-entry from Idle mode and Network-assisted re-entry from Idle Mode. In both cases, the transition latency of the device from Idle mode to Normal Operation is the primary statistic of importance.

Idle Mode Evaluation Procedure

1. The system will be modeled using the default simulation topology as specified elsewhere in the document.
2. *N* MSs are dropped uniformly in each cell. Different load levels in the network are simulated by changing the number of MSs and the traffic generated.
3. Path loss, shadow fading and fast fading models for each MS should be consistent with the models defined elsewhere in the document. Fading signal and fading interference are computed from each mobile station into each sector and from each sector to each mobile for each simulation interval.
4. The MS initiating Idle mode is assumed to be stationary; this eliminates the procedure to perform location updates during listen intervals etc. The locations of all other MSs are assumed to be fixed and the serving sector for the fixed MSs does not change for the duration of the drop.
5. Traffic from the fixed MSs constitutes background load. Any traffic type among the traffic model specified in Section 3 can be chosen according to the simulation objective. For the moving MS, one of the traffic types (e.g., VoIP, HTTP or FTP) should be generated to measure the application performance metrics affected by transition latency from Idle to Normal operations. Start times for each traffic type for each user should be randomized as specified in the traffic model being simulated.
6. Statistics related to Idle mode metrics are collected for the MS performing Idle mode transition only.
7. Packets are not blocked when they arrive into the system (i.e. queue depths are infinite).
8. Sequences of simulation are run, each with a different random seed. For a given drop the simulation is run for this duration, and then the process is repeated again, with the MSs dropped at new random locations. A sufficient number of drops are simulated to ensure convergence in the system performance metrics.

Performance metric for Idle Mode:

Case 1: For device-initiated Idle to Normal operation transition, the IDLE to ACTIVE_STATE transition, the latency, τ_d is given by

$$\tau_d = T_r + T_e$$

where T_r and T_e are the times required to execute ranging and network re-entry, respectively.

Case 2: For Network-initiated Idle to Normal operation transition, the IDLE to ACTIVE_STATE transition, the latency, τ_d is given by

$$\tau_d = Tp + T_r + T_e$$

Where, T_r and T_e are same as defined above and T_p is the time required to successfully page the MS.

4.11 Security (Later release)

4.12 MBS (Later release)

4.13 Buffer Management

A simple tail-drop policy is used for buffer management. If another drop policy, e.g., Random early drop (RED) or head drop for VOIP, is used, it should be explicitly specified.

5. PHY Layer Modelling

5.1 PHY Modem Abstraction for System Simulation

WiMAX PHY layer has many options. The following table captures the PHY features as listed in the WiMAX Forum Mobile Certification Waves System Profile Features V1.0. PHY models must cover each of the features that has a significant impact on the system simulation methodology

Table 5.1.1: Certification Wave Recommended PHY Features

Features		Wave 1		Wave 2	
Main Item	Sub Item	BS	MS	BS	MS
DL subscriber allocation	PUSC	Y	Y	Y	Y
	PUSC w/ all subchannels	Y	Y	Y	Y
	FUSC	Y	Y	Y	Y
	Band AMC 2X3	N	N	Y	Y
UL subscriber allocation	PUSC	Y	Y	Y	Y
	Band AMC 2X3	N	N	Y	Y
Ranging & BW request	Initial	Y	Y	Y	Y
	HO	Y	Y	Y	Y
	Periodic	Y	Y	Y	Y
	BW request	Y	Y	Y	Y
Fast-feedback	6-bit	Y	Y	Y	Y
Channel coding	Repetition	Y	Y	Y	Y
	Randomization	Y	Y	Y	Y
	CC	Y	Y	Y	Y
	CTC	Y	Y	Y	Y
	Interleaving	Y	Y	Y	Y
HARQ*	Chase Combing	Y	Y	Y	Y
Synchronization	BS time synchronization	Y	N/A	Y	N/A
	BS freq synchronization	Y	N/A	Y	N/A
	BS-BS freq synchronization	Y	N/A	Y	N/A
	SS synchronization	N/A	Y	N/A	Y
Power control	Closed-loop power control	Y	Y	Y	Y
	Open-loop power control	Y	Y	Y	Y
CINR measurement	Physical CINR using Preamble	Y	Y	Y	Y
	Physical CINR using Pilots	Y	Y	Y	Y

	Effective CINR using pilots	N	N	Y	Y
	RSSI measurement	Y	Y	Y	Y
Modulation	DL QPSK	Y	Y	Y	Y
	DL 16 QAM	Y	Y	Y	Y
	DL 64 QAM	Y	Y	Y	Y
	UL QPSK	Y	Y	Y	Y
	UL 16 QAM	Y	Y	Y	Y
	Pilot modulation	Y	Y	Y	Y
	Preamble modulation	Y	Y	Y	Y
	Ranging modulation	Y	Y	Y	Y
MAP Support	Normal MAP	Y	Y	Y	Y
	Compressed MAP	Y	Y	Y	Y
	Sub-DL-UL MAP	Y	Y	Y	Y
MIMO	All IO-MIMO items	N	N	IO_MIMO	Y
Beamforming	All IO-BF items	N	N	IO_BF	Y

*Note: On HARQ, a waiver is recommended for total DL buffer size of 16384 softbits over all 4 HARQ channels supported (equivalent to 4096 softbits (K=12) per channel) for CAT 1 in Wave 1, with requirements for CAT 2, 3, 4 unchanged.

5.2 Modelling advanced PHY features

OFDMA PHY is capable of supporting Advanced Antenna Systems and a set of other transmit diversity options such as 2nd, 3rd, and 4th order transmit diversity. Systems supporting AAS use multiple antennas for diversity, beamforming, and null steering to increase capacity and coverage, while minimizing the probability of outage.

Options available to DL in transmit diversity include techniques based on 2nd or 4th order diversity. In the UL transmit diversity options are based on 2nd order diversity. These diversity techniques may be flexibly employed in capacity and coverage tradeoff. The techniques support Spatial Multiplexing (SM) for maximum spectral efficiency, and include both open and closed loop systems.

5.2.1 Advanced Antenna Systems

The Direct Signalling Method and the Diversity-Map Scan are two optional AAS modes supported in OFDMA PHY. The Adjacent (Band AMC) subcarrier permutation and diversity (FUSC and PUSC) options are both supported by Diversity-Map Scan. Adjacent subcarrier permutation is supported by the Direct Signalling Method with less overhead in control signalling.

5.2.2 Transmit Diversity

In DL stream, OFDMA supports 2nd , 3rd, and 4th order transmit diversity options, whereas in UL stream only 2nd order transmit diversity is supported. Diversity and adjacent subcarrier permutations both support all diversity options.

5.2.2.1 Second-Order STC

DL stream supports 2nd order STC with coding rates 1 and 2 as shown in the matrices below.

$$A = \begin{bmatrix} S_i & S_{i+1}^* \\ S_{i+1} & S_i^* \end{bmatrix}$$

$$B = \begin{bmatrix} S_i \\ S_{i+1} \end{bmatrix}$$

Here S_k*'s* are OFDMA symbols in the frequency domain right before IFFT operation.

5.3 Channel Models for System Simulation

It is very challenging to completely simulate actual propagation environment in the System level simulations. For simplicity the propagation model to be used need just to meet the normal accuracy of the System level simulation (e.g., accuracy, symbol, frame duration, slot), and the model may not be excessively complex. Otherwise the simulations will take too long.

The following are three basic steps required to describe the overall procedure for generating the channel matrices:

- Describe the environment to be simulated whether it is: a) urban micro, b) suburban macro, or c) urban macro.
- From the chosen environment above, get its parameters.
- Based on the chosen parameters the channel coefficients are generated.

The roadmap for generating the channel coefficients is given in the figure below:

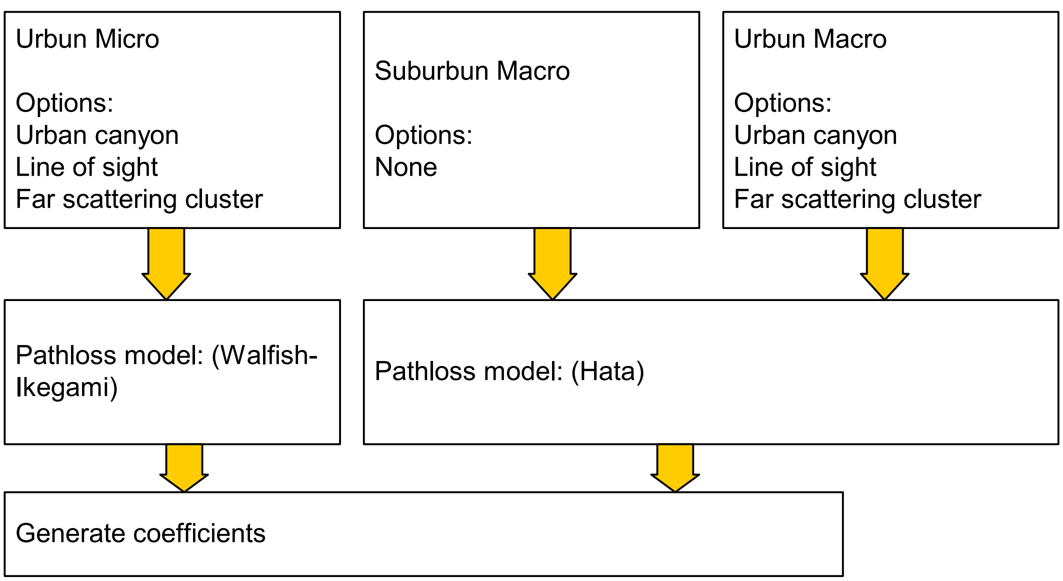


Figure 5.3.1: Roadmap for generating channel coefficients

Method for accounting for inter-cell interference is described in Chapter 2 on System Simulation. The SS receives N time-delayed multi-path copies of the transmitted signal. Each path consists of M subpaths, power and delay generated randomly according to the generation procedure.
Figure 5.3.1 shows the angular parameters used in the model. The following definitions are used:

Ω_{BS} BS antenna array orientation, defined as the difference between the broadside of the BS array and the absolute North (N) reference direction.

θ_{BS} LOS AoD direction between the BS and MS, with respect to the broadside of the BS array.

$\delta_{n,AoD}$ AoD for the *n*th (*n* = 1 … *N*) path with respect to the LOS AoD θ_0.

$\Delta_{n,m,AoD}$ Offset for the *m*th (*m* = 1 … *M*) subpath of the *n*th path with respect to $\delta_{n,AoD}$.

$\theta_{n,m,AoD}$ Absolute AoD for the *m*th (*m* = 1 … *M*) subpath of the *n*th path at the BS with respect to the BS broadside.

Ω_{MS} SS antenna array orientation, defined as the difference between the broadside of the SS array and the absolute North reference direction.

θ_{MS} Angle between the BS-SS LOS and the SS broadside.

$\delta_{n,AoA}$ AoA for the *n*th (*n* = 1 … *N*) path with respect to the LOS AoA $\theta_{0,MS}$.

$\Delta_{n,m,AoA}$ Offset for the *m*th (*m* = 1 … *M*) subpath of the *n*th path with respect to $\delta_{n,AoA}$.

$\theta_{n,m,AoA}$ Absolute AoA for the *m*th (*m* = 1 … *M*) subpath of the *n*th path at the BS with respect to the BS broadside.

v SS velocity vector.

θ_v Angle of the velocity vector with respect to the SS broadside: $\theta_v = \arg(\mathbf{v})$.

The angles shown in Figure 5.3.1 that are measured in a clockwise direction are assumed to be negative in value.

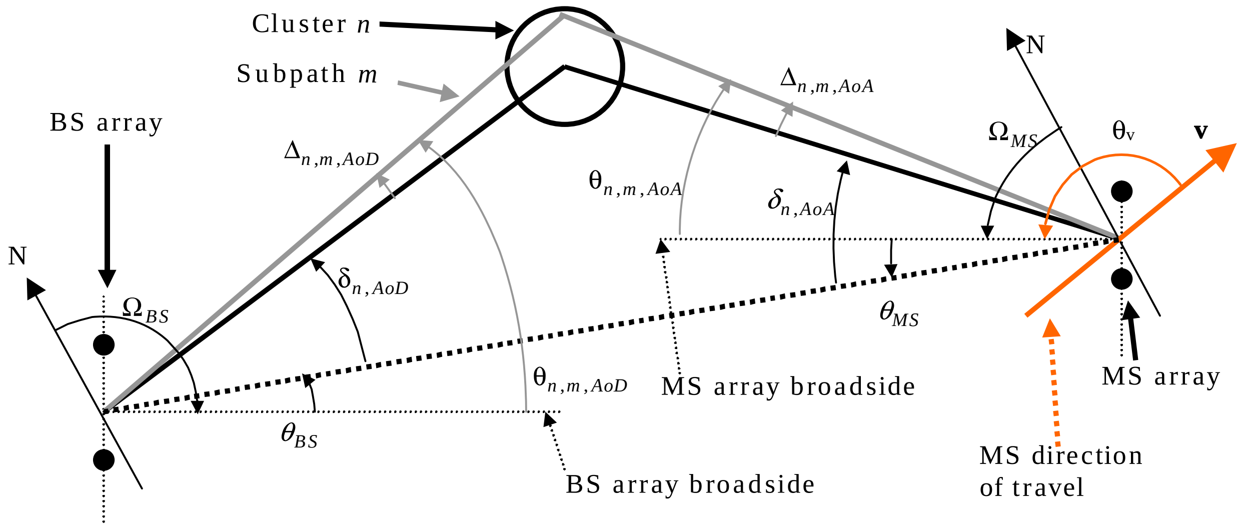


Figure 5.3.2: BS and SS angle parameters [3GPP]

Whereas the fast fading per-path will be evolved in time for system level simulation during the evaluation of a given MS, most other parameters such as: SS location, delay spread, angle spread, log normal shadowing will be fixed.

We need to consider the following three environments:

1. Suburban macro cell (BS to BS ≈ 3Km)
2. Urban macro cell (BS to BS ≈ 3Km)
3. Urban micro cell (BS to BS < 1Km)

It is assumed that BS antennas are installed higher than surrounding buildings in a macro cell, whereas for a micro cell BS antennas are installed at the rooftop level.

Shadowing effects and mean path loss for the user drop point will be indicated by the channel model. Which will be dependent the antenna height and the distance from the BS to the user.

Based on experience reported by vendors we recommend the use of two different path loss models in the system simulation, namely the SCM model and the Erceg model. The SCM model is included to enable a comparison of the performance of WiMAX technology against other technologies that have adopted SCM as the recommended path loss model. SCM path loss model has a penalty on the system performance and therefore the capacity prediction from this model is expected to be conservative. The current SCM model and the related path loss formula are being corrected for the initial frequency range that WiMAX systems are being evaluated in (namely 2.5 and 3.5 GHz). On the other hand simulation results applying the Erceg model have been observed to provide results closer to those from field deployments. Erceg model can be applied in system simulation to obtain results that are more aggressive and more representative of the WiMAX networks. The Erceg sub section covers three terrain cases all pertaining to suburban deployments

5.3.1 Erceg Model

The path loss equation with correction factors for higher frequencies (2.5 GHz) is given by:

$$PL = A + 10\gamma\log_{10}\left(\frac{d}{d_o}\right) + X_f + X_h + s \quad \text{for } d > d_o \qquad \textbf{(5.3.1)}$$

where $d_o = 100m$, d is the distance between the BS and SS antennas and s is the shadow fading component, which is lognormal distributed (typical variance = 8 to 10 dB). Other parameters are

$$A = 20\log_{10}\left(\frac{4\pi d_o}{\lambda}\right) \qquad \textbf{(5.3.2)}$$

The path loss exponent is

$$\gamma = a - bh_b + \frac{c}{h_b} \qquad \textbf{(5.3.3)}$$

where h_b is the base station antenna height. The parameters $a, b, c,$ are defined for three different terrain types as follows

Table 5.3.1: Parameters for Erceg Model for Different Terrain Types

Parameter	**Type A** (Hilly with Moderate-Heavy Tree Density)	**Type B** (Hilly with Light Tree Density OR Flat with Moderate-Heavy Tree Density)	**Type C** (Flat with Light Tree Density)
a	4.6	4	3.6
b	0.0075	0.0065	0.005
c	12.6	17.1	20

The correction factors for higher frequencies are given by

$$\begin{aligned} X_f &= 6\log 10(f/2000) \\ X_h &= -10.8\log(h/2) \quad \text{Type A and B} \\ &= -20\log(h/2) \qquad \text{Type C} \end{aligned} \qquad \textbf{(5.3.4)}$$

f – Carrier frequency (MHz),

h – MS antenna height (m) (between 2 and 10 m)

The Terrain Type B can be adopted as a typical path loss model among the three types.

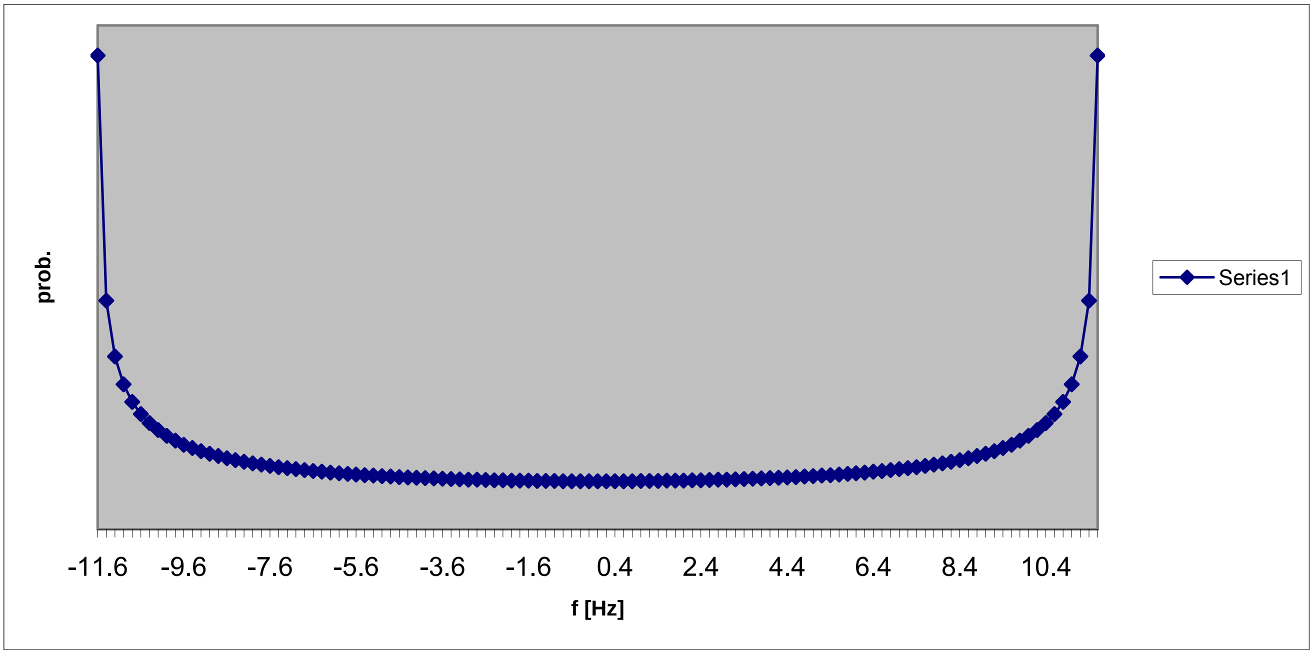


Figure 5.3.3: Probability density function of the random frequency due to Doppler associated with multipath

5.3.2 Other Channel Models

A channel model is specified by:

- A specific number of paths,
- A power profile giving the relative powers of these multiple paths (ITU multi-path models), and
- The fade rat specified by Doppler frequencies

WiMAX simulations should use the ITU Pedestrian B (3 km/hr) and Vehicular A (60 km/hr) mobility profiles specified by MWG. Table 5.3.1 lists the channel models used in 3GPP/3GPP2 for comparison purposes. The probability distribution listed in Table 5.3.2 randomly assigns the channel models (from 1 to 6). Once assigned to a specific user, the channel model remains the same for the duration of the simulation drop.

Table 5.3.1: Mixed User Channel Model for Performance Simulation

Channel Model	Numberof Paths	Speed (km/h)	Fading	Assignment Probability
ITU Veh. B. Ch-103	6	3	Jakes	0.60
ITU Veh. A. Ch-104	6	30	Jakes	0.30
	6	120	Jakes	0.10

Table 5.3.2: Channel Models and associated assignment probability distribution					
Channel Model	**Multi-path Model**	**# of Paths**	**Speed (km/h)**	**Fading**	**Assignment Probability**
Model 1	Ch-100	1	30	Jakes	0.1
Model 2	Ch-100	1	120	Jakes	0.1
Model 3	Ch-104	6	30	Jakes	0.1
Model 4	Ch-104	6	120	Jakes	0.1
Model 5	Ch-102	4	3	Jakes	0.3
Model 6	Ch-103	6	3	Jakes	0.3
Note: Fading model is Raleigh. The fading spectrum model is Jakes. Assignment probability is variable. The values in this table represent recommended defaults.					

Table 5.3.2 data (assignment probabilities, channel models, SS speeds (for fading rates)) has been created with the help of [1]. For each unique combination of SS velocity and channel model a separate link-level simulation must be performed. To simplify simulation while achieving a representative accurate result due to sufficient samples being taken, a combination of velocity and channel model must be chosen very carefully.

The relative contributions from each of the paths are shown in the table of the tap delay model in Table 5.3.3. The signal strength is calculated using the table and the multipath effect. The speed in the table is used to apply mobility effect in the drop.

The Received signal strength depends on the following:

1. propagation
2. multi-path
3. mobility for the drop point

For the users dropped in the central cell, the signal strength is calculated. The signal transmissions from drops in other cells are simulated as interference contributions to the users in the central cell. In an OFDMA system, SNR (over the duration of a FEC block) to individual sub-carrier or a sub-carrier band for the drops in the central cell is calculated from the signal and the interference. Using a predetermined mapping between the observed SNR and BLER (Block Error Rate), an error probability is determined for the transmitted block.

Table 5.3.3 shows the normalized power profiles for the different channel models such as flat fading. For each standard ITU channel model the absolute power values have been normalized so that they sum up to 0 dB (unit energy).

Table 5.3.3: Multipath Effects – Tap Delay Line Parameters							
Channel Model		Path 1	Path 2	Path 3	Path 4	Path 5	Path 6

Flat fading	Average power (dB)	0	-	-	-	-	-
	Relative delay (ns)	-	-	-	-	-	-
ITU Vehicular A	**Average power (dB)**	**-3.14**	**-4.14**	**-12.14**	**-13.14**	**-18.14**	**-23.14**
	Relative delay (ns)	**0**	**310**	**710**	**1090**	**1730**	**2510**
ITU Pedestrian A	Average power (dB)	-0.51	-10.21	-19.71	-23.31	-	-
	Relative delay (ns)	0	110	190	410	-	-
ITU Pedestrian B	**Average power (dB)**	**-3.92**	**-4.82**	**-8.82**	**-11.92**	**-11.72**	**-27.82**
	Relative delay (ns)	**0**	**200**	**800**	**1200**	**2300**	**3700**

5.4 MIMO Abstraction

This section describes the MIMO abstraction methodology used for System Level Simulations. The methodology involves reducing the transmit/receive configuration to an equivalent SISO configuration, which can then be translated to an equivalent AWGN SINR using the EESM approximation and then to the BLER by a table lookup through an AWGN curve.

5.4.1 General Per-Tone Model

The per tone signal model is given by

$$y = \sqrt{P_0} H_0 x_0 + \sum_{k=1}^{K} \sqrt{P_k} H_k x_k + n \qquad \textbf{(1)}$$

where y is the $P \times 1$ receive vector,

P is the number of receive antennas,

x_k is the $Q \times 1$ transmit vector,

Q is the number of transmit antennas,

n is the $P \times 1$ noise vector, modeled as zero-mean complex Gaussian with variance σ_n^2,

H_k is the $P \times Q$ channel matrix,

K is the number of interferers, and

P_k is the received power from the k^{th} BS[2] which is given as follows

$$P_k = P_{\mathrm{T}} G_k \frac{10^{X_k/10}}{L(d_k)} \qquad \textbf{(2)}$$

where P_{T} is the transmit power, X_i the log-normal shadowing, $L(d_i)$ is the path loss at distance d_i. G_i is the aggregate antenna gain. It is assumed that the $k = 0$ BS is the desired BS.

For a linear receiver, the post-receiver signal is obtained by pre-multiplying y with w^*

$$z = w^* y \qquad \textbf{(3)}$$

The choice of the weight vector, w, depends on the receiver scheme, i.e. maximum ratio combining (MRC) or minimum mean squared error (MMSE). Other receiver abstractions will also be described in later sections of the document.

The per tone post-receiver SINR is given as shown below

$$\gamma^{DL} = \frac{P_0 f(H_0)}{\sum_{k=1}^{K} P_k g(H_k) + \sigma_n^2} \qquad \textbf{(4)}$$

In terms of w, the SINR (for SISO case) is given as

$$\gamma^{DL} = \frac{P_0 \left| w^* H_0 \right|^2}{\sum_{k=1}^{K} P_k \left| w^* H_k \right|^2 + \| w \|^2 \sigma_n^2} \qquad \textbf{(5)}$$

The remainder of this section describes $f(.)$ and $g(.)$ for various MIMO configurations.

Note: The remaining part of the document uses H with three indices which refer to the BS, receive antenna and transmit antenna (i.e. the last two indices refers to the row and column of the H matrix).

---

2 The formulation shown here is for the downlink, but the same modeling applies to the uplink.

5.4.2 SISO/MISO

We first consider both the SISO and MISO (Alamouti) cases, i.e. with a single receive antenna. This can serve as a baseline for the other MIMO configurations.

5.4.1.1 SISO

For the SISO scheme, the received vector is multiplied by w = H, so that

$$z = w^* y = H_0^* y = |H_0|^2 \sqrt{P_0} x_0 + \sum_{k=1}^{K} \sqrt{P_k} H_0^* H_k x_k + H_0^* n \qquad \textbf{(6)}$$

The SINR for this scheme is given as

$$\gamma = \frac{P_0 \left(|H_0|^2\right)^2}{\sum_{k=1}^{K} P_k \left|H_0^* H_k\right|^2 + |H_0|^2 \sigma_n^2} = \frac{P_0 |H_0|^2}{\sum_{k=1}^{K} P_k |H_k|^2 + \sigma_n^2} \qquad \textbf{(7)}$$

The functions $f(.)$ and $g(.)$ are thus defined as follows

$$f(.) = |h_{0,0,0}|^2 \qquad \textbf{(8)}$$

and

$$g(.) = |h_{k,0,0}|^2 \qquad \textbf{(9)}$$

5.4.2.2 2x1 Alamouti

In this case, there are two antennas at the BS transmitter and a SS with a single antenna receiver. The signal received at the receiver can be represented in the following form

$$\tilde{y} = \sqrt{\frac{P_0}{2}} \underbrace{\begin{bmatrix} h_{0,0,0} & h_{0,0,1} \\ h_{0,0,1}^* & -h_{0,0,0}^* \end{bmatrix}}_{\bar{H}} x_{0,m} + \sum_{k=1}^{K} \sqrt{\frac{P_k}{2}} \begin{bmatrix} h_{k,0,0} & h_{k,0,1} \\ h_{k,0,1}^* & -h_{k,0,0}^* \end{bmatrix} x_k + n \qquad \textbf{(10)}$$

where y and x are vectors. First, we pre-multiply the received vector with the transpose conjugate of the equivalent channel matrix, $\bar{H}$, i.e. here $w = \bar{H}$, and then we have that

$$z = w^* \tilde{y} = \sqrt{\frac{P_0}{2}} \begin{bmatrix} |h_{0,0,0}|^2 + |h_{0,0,1}|^2 & 0 \\ 0 & |h_{0,0,0}|^2 + |h_{0,0,1}|^2 \end{bmatrix} x_0 + \sum_{k=1}^{K} \sqrt{\frac{P_k}{2}} \begin{bmatrix} h_{0,0,0} & h_{0,0,1} \\ h_{0,0,1}^* & -h_{0,0,0}^* \end{bmatrix}^H \begin{bmatrix} h_{k,0,0} & h_{k,0,1} \\ h_{k,0,1}^* & -h_{k,0,0}^* \end{bmatrix} x_k + \begin{bmatrix} h_{0,0,0} & h_{0,0,1} \\ h_{0,0,1}^* & -h_{0,0,0}^* \end{bmatrix}^* n \quad \textbf{(11)}$$

The SINR per stream is given as follows

$$\gamma = \frac{P_0 \left(|h_{0,0,0}|^2 + |h_{0,0,1}|^2 \right)^2 \Big/ 2}{\sum_{k=1}^{K} P_k \left(|h_{0,0,0}^* h_{k,0,0} + h_{0,0,1} h_{k,0,1}^*|^2 + |h_{0,0,0}^* h_{k,0,1} - h_{0,0,1} h_{k,0,0}^*|^2 \right) \Big/ 2 + \left(|h_{0,0,0}|^2 + |h_{0,0,1}|^2 \right) \sigma_n^2}$$

$$= \frac{P_0 \left(|h_{0,0,0}|^2 + |h_{0,0,1}|^2 \right) \Big/ 2}{\sum_{k=1}^{K} P_k \left(|h_{0,0,0}^* h_{k,0,0} + h_{0,0,1} h_{k,0,1}^*|^2 + |h_{0,0,0}^* h_{k,0,1} - h_{0,0,1} h_{k,0,0}^*|^2 \right) \Big/ 2 \left(|h_{0,0,0}|^2 + |h_{0,0,1}|^2 \right) + \sigma_n^2}$$

Note the transmit power is shared between two antennas and hence the scaling by 2. The signal and interference powers are given as

$$f(.) = \frac{1}{2} \left(|h_{0,0,0}|^2 + |h_{0,0,1}|^2 \right) \quad \textbf{(12)}$$

and

$$g(.) = \frac{1}{2} \left(\frac{|h_{0,0,0}^* h_{k,0,0} + h_{0,0,1} h_{k,0,1}^*|^2 + |h_{0,0,0}^* h_{k,0,1} - h_{0,0,1} h_{k,0,0}^*|^2}{|h_{0,0,0}|^2 + |h_{0,0,1}|^2} \right) \quad \textbf{(13)}$$

5.4.3 Linear Receivers

The MRC is used here as a base line receiver. Also, considered is the linear MMSE receiver. Both the SIMO and the STC (Matrix A) cases are considered in this section.

5.4.3.1 SIMO, 1x*P* MRC

In this case, the SS has P receive antennas and MRC is used to combine the received signal from the antennas (the received vector is multiplied by transpose conjugate of the channel matrix, i.e. $w = H_0$), thus

$$z = w^* y = H_0^* y = \sqrt{P_0}\,\|H_0\|^2\, x_0 + \sum_{k=1}^{K} \sqrt{P_k}\, H_0^* H_k x_k + H_0^* n \qquad \textbf{(14)}$$

The SINR is given as

$$\begin{aligned}\gamma &= \frac{P_0\left(\sum_{p=0}^{P-1}\left|h_{0,p,0}\right|^2\right)^2}{\sum_{k=1}^{K} P_k \left|\sum_{p=0}^{P-1} h_{0,p,0}^* h_{k,p,0}\right|^2 + \left(\sum_{p=0}^{P-1}\left|h_{0,p,0}\right|^2\right)\sigma_n^2} \\ &= \frac{P_0\left(\sum_{p=0}^{P-1}\left|h_{0,p,0}\right|^2\right)}{\sum_{k=1}^{K} P_k \left|\sum_{p=0}^{P-1} h_{0,p,0}^* h_{k,p,0}\right|^2 \Big/ \left(\sum_{p=0}^{P-1}\left|h_{0,p,0}\right|^2\right) + \sigma_n^2}\end{aligned} \qquad \textbf{(15)}$$

The signal and interference powers are given below

$$f(.) = \sum_{p=0}^{P-1}\left|h_{0,p,0}\right|^2 \qquad \textbf{(16)}$$

and

$$g(.) = \frac{\left|\sum_{p=0}^{P-1} h_{0,p,0}^* h_{k,p,0}\right|^2}{\sum_{p=0}^{P-1}\left|h_{0,p,0}\right|^2} \qquad \textbf{(17)}$$

5.4.3.1 SIMO, 1x*P* MMSE

For the MMSE receiver the vector w is given as follows

$$w = R_{yy}^{-1} H_0 \qquad \textbf{(18)}$$

where R_{yy} is the covariance matrix of the receive vector y, i.e. $R_{yy} = \mathrm{E}\{yy^*\}$.

The post receiver signal model is given as

$$z = w^* y = \sqrt{P_0} H_0^* R_{yy}^{-1} H_0 x_0 + \sum_{k=1}^{K} \sqrt{P_k} H_0^* R_{yy}^{-1} H_k x_k + H_0^* R_{yy}^{-1} n \quad \textbf{(19)}$$

The SINR is thus given as

$$\gamma = \frac{P_0 \left| \sum_{p=0}^{P-1} w_{0,p,0}^* h_{0,p,0} \right|^2}{\sum_{k=1}^{K} P_k \left| \sum_{p=0}^{P-1} w_{0,p,0}^* h_{k,p,0} \right|^2 + \left(\sum_{p=0}^{P-1} \left| w_{0,p,0} \right|^2 \right) \sigma_n^2} \quad \textbf{(20)}$$

5.4.3.2 2x*P* Alamouti (Matrix A) with MRC

In this case, there are 2 transmit and P receive antennas. An MRC receiver is used to combine the signals from the P antennas. This scheme is just an extension to the 2x1 Alamouti case. Similar to how 1xP MRC is extended from SISO, the per-stream SINR is given by

$$\gamma = \frac{P_0 \sum_{p=0}^{P-1} \left(\left| h_{0,p,0} \right|^2 + \left| h_{0,p,1} \right|^2 \right) \Big/ 2}{\sum_{k=1}^{K} P_k \left(\left| \sum_{p=0}^{P-1} \left(h_{0,p,0}^* h_{k,p,0} + h_{0,p,1} h_{k,p,1}^* \right) \right|^2 + \left| \sum_{p=0}^{P-1} \left(h_{0,p,0}^* h_{k,p,1} - h_{0,p,1} h_{k,p,0}^* \right) \right|^2 \right) \Big/ 2 \sum_{p=0}^{P-1} \left(\left| h_{0,p,0} \right|^2 + \left| h_{0,p,1} \right|^2 \right) + \sigma_n^2}$$

Hence $f(.)$ and $g(.)$ are given by

$$f(.) = \frac{1}{2} \sum_{p=0}^{P-1} \left(\left| h_{0,p,0} \right|^2 + \left| h_{0,p,1} \right|^2 \right) \quad \textbf{(21)}$$

and

$$g(.) = \frac{1}{2} \left(\frac{\left| \sum_{p=0}^{P-1} \left(h_{0,p,0}^* h_{k,p,0} + h_{0,p,1} h_{k,p,1}^* \right) \right|^2 + \left| \sum_{p=0}^{P-1} \left(h_{0,p,0}^* h_{k,p,1} - h_{0,p,1} h_{k,p,0}^* \right) \right|^2}{\sum_{p=0}^{P-1} \left(\left| h_{0,p,0} \right|^2 + \left| h_{0,p,1} \right|^2 \right)} \right) \quad \textbf{(22)}$$

In the single symbol detection receiver for Alamouti code, the effects of Doppler manifest as self interference.

Assume that:

1) The Alamouti code uses two transmit antennas and two time slots for transmission and the correlation between the channel on these slots is defined as follows

$$\rho = E\left(h_0 \tilde{h}_0^*\right) = E\left(h_1 \tilde{h}_1^*\right) = J_0\left(2\pi f_d \Delta t\right) \qquad \textbf{(23)}$$

where the h tilde (~) represents the channel on the second slot and the simple h refers to the channel on the first slot. The indices represent the transmit antenna.

2) Independence between the two transmit antennas is assumed

3) Perfect channel estimation is assumed on both time slots and both antennas.

The received signal for the Alamouti scheme when the channel is not static is represented as given below

$$\tilde{y} = \sqrt{\frac{P_0}{2}} \begin{bmatrix} h_0^{(0)} & h_1^{(0)} \\ \tilde{h}_1^{(0)*} & -\tilde{h}_0^{(0)*} \end{bmatrix} x_0 + \sum_{k=1}^{K} \sqrt{\frac{P_k}{2}} \begin{bmatrix} h_0^{(k)} & h_1^{(k)} \\ h_1^{(k)*} & -h_0^{(k)*} \end{bmatrix} x_k + n$$

Since we are concerned only about the self interference part, the desired signal ($k = 0$) component alone, is shown below after receiver post-processing

$$\begin{bmatrix} h_0 & h_1 \\ \tilde{h}_1^* & -\tilde{h}_0^* \end{bmatrix}^H \begin{bmatrix} h_0 & h_1 \\ \tilde{h}_1^* & -\tilde{h}_0^* \end{bmatrix} = \begin{bmatrix} \left|h_0\right|^2 + \left|\tilde{h}_1\right|^2 & h_0^* h_1 - \tilde{h}_0^* \tilde{h}_1 \\ h_0 h_1^* - \tilde{h}_0 \tilde{h}_1^* & \left|\tilde{h}_0\right|^2 + \left|h_1\right|^2 \end{bmatrix}$$

The diagonal terms in the above equation are the signal part and the off-diagonal elements contribute to self-interference. As Doppler becomes zero ($\rho = 1$), the off-diagonal elements vanishes and self-interference becomes zero.

The signal part and interference part for the first stream are given by $s = \left|h_0\right|^2 + \left|\tilde{h}_1\right|^2$ and $t = h_0^* h_1 - \tilde{h}_0^* \tilde{h}_1$ respectively.

The ratio of signal power to self-interference power is given as

$$\frac{E\left(|s|^2\right)}{E\left(|t|^2\right)}=\frac{E\left(\left||h_0|^2+|\tilde{h}_1|^2\right|^2\right)}{E\left(\left|h_0^*h_1-\tilde{h}_0^*\tilde{h}_1\right|^2\right)}=\frac{3}{\left(1-\rho^2\right)} \quad (24)$$

where from (23), $E\left(h_0\tilde{h}_0^*h_1\tilde{h}_1^*\right)=\rho^2$.

Hence the original SINR equation for Alamouti gets modified to incorporate this factor. The new SINR is computed as follows.

$$\begin{aligned}\tilde{\gamma}_s&=\frac{P_0\left(\bar{H}^{(0)H}\bar{H}^{(0)}\right)_{ss}\Big/2}{\left(1-\rho^2\right)P_0\left(\bar{H}^{(0)H}\bar{H}^{(0)}\right)_{ss}\Big/6+\sum_{k=1}^{K}P_k\left(\bar{H}^{(k)H}\bar{H}^{(0)}\bar{H}^{(0)H}\bar{H}^{(k)}\right)_{ss}\Big/2\left(\bar{H}^{(0)H}\bar{H}^{(0)}\right)_{ss}+\sigma_n^2}\\&=\frac{1}{\frac{1}{\gamma_s}+\frac{\left(1-\rho^2\right)}{3}}\end{aligned} \quad (25)$$

5.4.3.3 2x*P* Alamouti (Matrix A) with MMSE

The received signal shown in Equation (10), is pre-multiplied by matrix w which is given by $w=R_{\tilde{y}\tilde{y}}^{-1}\bar{H}$ where $R_{\tilde{y}\tilde{y}}^{-1}$ is the covariance matrix of the receive vector $\tilde{y}$, i.e. $R_{\tilde{y}\tilde{y}}=\mathrm{E}\left\{\tilde{y}\tilde{y}^*\right\}$. The post receiver signal is thus given by

$$\begin{aligned}z=w^*\tilde{y}&=\sqrt{\frac{P_0}{2}}w^*\bar{H}_0x_0+\sum_{k=1}^{K}\sqrt{\frac{P_k}{2}}w^*\bar{H}_kx_k+w^*n\\&=\sqrt{\frac{P_0}{2}}\left(D+O\right)x_0+\sum_{k=1}^{K}\sqrt{\frac{P_k}{2}}w^*\bar{H}_kx_k+w^*n\end{aligned} \quad (26)$$

where D and O are the diagonal (desired component) and off-diagonal (self-interfering component) components of $w^*\overline{\mathrm{H}}$. The post receiver SINR on the s-th symbol is given as

$$\gamma_s=\frac{P_0\left(DD^*\right)_{ss}\Big/2}{\frac{1}{2}P_0\left(OO^*\right)_{ss}+\frac{1}{2}\sum_{k=1}^{K}P_k\left(w^*\bar{H}_k^*\bar{H}_kw\right)_{ss}+\left(w^*w\right)_{ss}\sigma_n^2} \quad (27)$$

$(.)_{ss}$ refers to the s[th] row and s[th] column element of the matrix within braces.

5.4.4 2x2 Spatial Multiplexing (Vertical Encoding, Matrix B)

In spatial multiplexing with vertical encoding, both antennas receive symbols from a single FEC block. Each tone is occupied by two symbols. In this section we provide several receiver abstraction methods based on: (1) Eigen Decomposition (for MLD approximation), (2) linear MMSE and (3) non-linear MMSE-SIC.

5.4.4.1 Using Eigen Decomposition

This abstraction method is an approximation to the maximum likelihood detector (MLD). It is assumed that there are 2 spatial streams from the 2 transmit antennas and there are *P* receive antennas at the SS. The channel is diagonalized by multiplying with unitary matrices (obtained by singular value decomposition [$H = U\Lambda V^H$] of the channel matrix *H*) at the transmitter and receiver. The post receiver signal is given by

$$\tilde{y} = \sqrt{\frac{P_0}{2}}\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}\begin{bmatrix}x_0^1\\ x_0^2\end{bmatrix} + \sum_{k=1}^{K}\sqrt{\frac{P_k}{2}}U_k^* H_k \tilde{V}_k\begin{bmatrix}x_k^1\\ x_k^2\end{bmatrix} + U_k^*\begin{bmatrix}n^1\\ n^2\end{bmatrix} \qquad \textbf{(28)}$$

Since *U* and *V* are unitary matrices the statistics of the interference and noise variables don't change. The SINR in this case is given as

$$\gamma = \frac{P_0\lambda_s^2/2}{\frac{1}{2P}\sum_{k=1}^{K}P_k\left\|H_k\right\|^2 + \sigma_n^2} \qquad \textbf{(29)}$$

Hence the channel gains are given by λ_s^2, where λ_s is the diagonal elements of the diagonal matrix Λ.

$$f(.) = \lambda_s^2/2 \qquad \textbf{(30)}$$

The interference, for simplicity, is modeled as given below

$$g(.) = \frac{\left\|H_k\right\|^2}{2P} \qquad \textbf{(31)}$$

5.4.4.2 Using MMSE Receiver

To illustrate the post processing SINR calculation for a MIMO system based on a linear MMSE receiver, as before, we assume a downlink transmission system with P receive antennas and Q transmit antennas. We assume that Q, spatial streams are transmitted, where $P \geq Q$. We also assume that interferers and the desired signal use the same MIMO scheme for transmission.

The MMSE weights can be specified as

$$w = H_0^* \left[H_0 H_0^* + \left(\frac{\sigma_n^2}{P_0} \right) I \right]^{-1} \quad \textbf{(32)}$$

The post processing SINR can be computed by defining the following expressions: $E = wH_0$ and $D = diag(E)$ (which denotes the desired signal component). We also define $I_{self} = E - D$, which is the self interference between MIMO streams. The post processing SINR for the sth MIMO stream is thus given as

$$\gamma_s = \frac{diag\left[P_0 DD^* \right]_{ss}}{diag\left[\sigma_n^2 ww^* + P_0 I_{self} I_{self}^* + \sum_{k=1}^{K} P_k w H_k H_k^* w^* \right]_{ss}} \quad \textbf{(33)}$$

5.4.4.3 Using MMSE-SIC Receiver

In this section, we describe a receiver model for predicting the performance of spatial multiplexing (using 2x2 Matrix-B) in a system-level simulation. This receiver abstraction entails use of a two-stage MMSE detector with serial-interference cancellation (SIC) employed between the two stages. In the first step the post-processed SINR is calculated for both streams at the output an MMSE filter. For the weaker of the two streams, the SINR is improved by assuming that ideal serial interference cancellation is possible. To this end, in the second step, the SINR for the weak stream is computed by using an MMSE filter after eliminating the interference due to the stronger stream.

For the weak stream, the performance approaches that of ML detection. For the strong stream, the performance of the MMSE detector deviates significantly from that of an ML detector only when the spatial correlation between the streams is high.

We write the received signal in vector form as follows:

$$\mathbf{y} = \sqrt{\frac{P_0}{2}} [\mathbf{h_1} \quad \mathbf{h_2}] \begin{bmatrix} x_0^{\ 1} \\ x_0^{\ 2} \end{bmatrix} + \sum_{k=1}^{K} \sqrt{\frac{P_k}{2}} \mathbf{H_k} \mathbf{x_k} + n \quad \textbf{(34)}$$

where h_i is the 2 x 1 channel vector associated with stream *i*.

The MMSE filter for stream *s* is:

$$\begin{aligned} \mathbf{w_s} &= \sqrt{\frac{P_0}{2}} \mathbf{R_{yy}}^{-1} \mathbf{h}_s \qquad \text{where,} \\ \mathbf{R}_{yy} &= \frac{P_0}{2} \mathbf{H_0 H_0}^H + \sum_{k=1}^{K} \frac{P_k}{2} \mathbf{H_k H_k}^H + \sigma_n^2 \mathbf{I} \end{aligned} \qquad \textbf{(35)}$$

The post-processing SINR for stream *s* is:

$$\gamma_s = \frac{\frac{P_0}{2} \left\| \mathbf{w}_s^* \mathbf{h}_s \right\|^2}{\frac{P_0}{2} \left\| \mathbf{w}_s^* \mathbf{h}_{i \neq s} \right\|^2 + \sum_{k=1}^{K} P_k \left\| \mathbf{w}_k^* \mathbf{H}_k \right\|^2 + \sigma_n^2 \left\| \mathbf{w}_s \right\|^2} \qquad \textbf{(36)}$$

At this stage, for the stream *s* with the higher SINR, γ_s is the output SINR from the MIMO detector. The weaker stream is processed further, with the interference due to the higher stream cancelled by an ideal serial interference cancellation algorithm. For this stream *i*, the received signal can be re-written as:

$$\tilde{\mathbf{y}} = \sqrt{\frac{P_0}{2}} \mathbf{h}_i \, x_0^i + \sum_{k=1}^{K} \sqrt{\frac{P_k}{2}} \mathbf{H_k x_k} + n, \qquad \textbf{(37)}$$

The MMSE filter for stream *i* is given by:

$$\begin{aligned} \mathbf{w}_i^{SIC} &= \sqrt{\frac{P_0}{2}} \tilde{\mathbf{R}}^{-1} \mathbf{h}_i \qquad \text{where,} \\ \tilde{\mathbf{R}} &= \frac{P_0}{2} \mathbf{h_i h_i}^* + \sum_{k=1}^{K} \frac{P_k}{2} \mathbf{H_k H_k}^* + \sigma_n^2 \mathbf{I} \end{aligned} \qquad \textbf{(38)}$$

Finally, the post-processing SINR for stream *i* is given by:

$$\gamma_i = \frac{\frac{P_0}{2} \left\| \mathbf{w}_i^{SIC\mathbf{H}} \mathbf{h}_i \right\|^2}{\sum_{k=1}^{K} P_k \left\| \mathbf{w}_k^{\mathbf{H}} \mathbf{H}_k \right\|^2 + \sigma_n^2 \left\| \mathbf{w}_s \right\|^2} \qquad \textbf{(39)}$$

5.4.4.4 Switching between Alamouti and Spatial Multiplexing

After computing the per-tone SINR for both transmission schemes (Matrix A and Matrix B), we need to compute which scheme yields a higher throughput. The per-tone SINR are converted to an effective

SINR (γ_{eff}), which is then used in a lookup (η_{AWGN}) to find the throughput. The throughputs are compared and the higher one is picked.

$$R_{\text{Alm}} \underset{\text{Rate2}}{\overset{\text{Rate1}}{\gtrless}} R_{\text{SM}} \tag{40}$$

The throughputs are computed as follows

$$R_{\text{Alm}} = \eta_{\text{AWGN}}\left(\gamma_{eff}^{\text{Alm}}\right) \tag{41}$$

and

$$R_{\text{SM}} = 2\eta_{\text{AWGN}}\left(\gamma_{eff}^{\text{SM}}\right) \tag{42}$$

5.4.5 *Q*x1 Beamforming

In beamforming case the signal to be transmitted is multiplied by the beamforming vector w before transmission. Thus, the actual transmitted signal is $x = w\tilde{x}$. If the BS has P transmit antennas and uses rate 1 transmission, then w is of size Qx1. The received waveform is given as follows

$$y = \sqrt{P_0} H_0 w_0 \tilde{x}_{0,m} + \sum_{k=1}^{K} \sqrt{P_k} H_k w_k \tilde{x}_k + n \tag{43}$$

w ideally should be equal to the normalized conjugate of the channel matrix, H. But in presence of noise/delays, we would have imperfect knowledge of the channel, given by $w_0 = \tilde{H}_0^* \big/ \left\|\tilde{H}_0\right\|$. The signal power is defined as

$$f\left(H_0, \tilde{H}_0\right) = \left(\frac{\left|\sum_{q=0}^{Q-1} \tilde{h}_{0,0,q}^* \tilde{h}_{0,0,q}\right|^2}{\sum_{q=0}^{Q-1} \left|\tilde{h}_{0,0,q}\right|^2} \right) \tag{44}$$

For the interference, w will be independent of the interfering channel and we can assume $w_k = 1\big/\sqrt{Q}$. This will non-coherently combine the transmitted signal and hence the interfering power is given as

$$g(.) = \frac{1}{Q}\left|\sum_{q=0}^{Q-1} h_{k,0,q}\right|^2 \tag{45}$$

*5.4.6 Q*x1 CDD (Cyclic Delay Diversity)

In CDD, cyclically delayed version of the same data is transmitted from multiple (Q) antennas. In frequency domain this translates to a multiplication by a phase ramp to each of those Q signals, the slope of the ramp being proportional to the cyclic delay. The equivalent channel for a 2Tx CDD configuration is given by

$$\tilde{h}_{k,p,0} = \frac{1}{\sqrt{Q}} \sum_{q=0}^{Q-1} h_{k,p,q} \, \mathrm{e}^{j2\pi\delta mq/M} \qquad ,k = 0,1,...,K \tag{46}$$

where δ is the CDD time domain sample delay between adjacent antennas and M is the FFT size and m the tone index. Any of the other transmit/receive scheme can also be used in conjunction with CDD.

5.4.7 Impact of Receiver Impairments

In case of receiver imperfections like channel estimation error, w is typically (for the SISO/MRC case) given as

$$w = \hat{H}_0 = H_0 + \Delta H_0 \tag{47}$$

where ΔH_0, is the estimation error component which will have power equal to the mean squared error (MSE) of the estimation.

References

[Erceg01] V. Erceg, K.V.S. Hari , *et al*, “Channel Models for Fixed Wireless Applications”. IEEE 802.16 Broadband Wireless Working Group, July 2001.

[Erceg99] V. Erceg, L.J.Greenstein, *et.al,* “ An empirically based path loss model for wireless channels in suburban environments”, *IEEE Journal on Selected Areas in Communications*, vol. 17, July 1999.

[Abhaya05] V. S. Abhayawardhana, Il.J.Wassell, *et.al,* “Comparison of Empirical Propogation Path Loss Models for Fixed Wireless Systems”, *IEEE VTC-Spring*, 2005

Appendix A: A Tutorial on Channel Models

Many readers of this simulation methodology document may be experts in modeling, programming, or higher layers of networking but may not be familiar with many PHY layer concepts. This tutorial on Channel Models has been designed for such readers. This information has been gathered from various IEEE and ITU standards and contributions and published books.

A.1 Basic Concepts

A.1.1 Channel

The term channel refers to the medium between the transmitting antenna and the receiving antenna as shown in Figure A.1.1

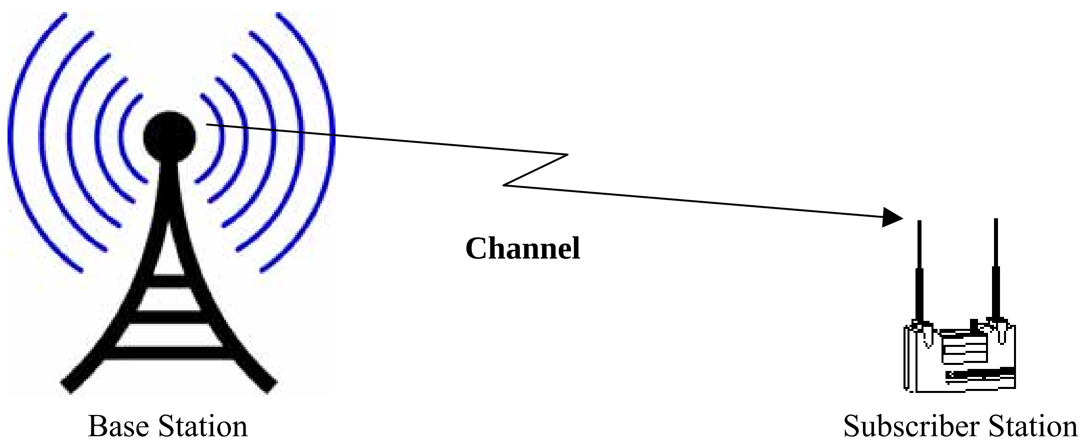


Figure A.1.1: Channel

The characteristics of wireless signal changes as it travels from the transmitter antenna to the receiver antenna. These characteristics depend upon the distance between the two antennas, the path(s) taken by the signal, and the environment (buildings and other objects) around the path. The profile of received signal can be obtained from that of the transmitted signal if we have a model of the medium between the two. This model of the medium is called **channel model**.

In general, the power profile of the received signal can be obtained by *convolving* the power profile of the transmitted signal with the impulse response of the channel. Convolution in time domain is equivalent to multiplication in the frequency domain. Therefore, the transmitted signal x, after propagation through the channel H becomes y:

$$y(f)=H(f)x(f)+n(f)$$

Here $H(f)$ is **channel response**, and $n(f)$ is the noise. Note that x, y, H, and n are all functions of the signal frequency f.

The three key components of the channel response are path loss, shadowing, and multipath as explained below.

A.1.2 Path Loss

The simplest channel is the free space line of sight channel with no objects between the receiver and the transmitter or around the path between them. In this simple case, the transmitted signal attenuates since the energy is spread spherically around the transmitting antenna. For this line of sight (LOS) channel, the received power is given by:

$$P_r = P_t \left[\frac{\sqrt{G_l}\lambda}{4\pi d} \right]^2$$

Here, P_t is the transmitted power, G_l is the product of the transmit and receive antenna field radiation patterns, λ is the wavelength, and d is the distance. Theoretically, the power falls off in proportion to the square of the distance. In practice, the power falls off more quickly, typically 3rd or 4th power of distance.

The presence of ground causes some of the waves to reflect and reach the transmitter. These reflected waves may sometime have a phase shift of $180°$ and so may reduce the net received power. A simple two-ray approximation for path loss can be shown to be:

$$P_r = P_t \frac{G_t G_r h_t^2 h_r^2}{d^4}$$

Here, h_t and h_r are the antenna heights of the transmitter and receiver, respectively. Note that there are three major differences from the previous formula. First, the antenna heights have effect. Second, the wavelength is absent and third the exponent on the distance is 4. In general, a common empirical formula for path loss is:

$$P_r = P_t P_0 \left(\frac{d_0}{d} \right)^{\alpha}$$

Where P_0 is the power at a distance d_0 and α is the path loss exponent. The path loss is given by:

$$PL(d)\, dB = \overline{PL}(d_0) + 10\alpha \log\left(\frac{d}{d_0} \right)$$

Here $\overline{PL}(d_0)$ is the mean path loss in dB at distance d_0. The thick dotted line in Figure A.1.2 shows the received power as a function of the distance from the transmitter.

A.1.3 Shadowing

If there are any objects (such buildings or trees) along the path of the signal, some part of the transmitted signal is lost through absorption, reflection, scattering, and diffraction. This effect is called shadowing. As shown in Figure A.1.3, if the base antenna were a light source, the middle building would cast a shadow on the subscriber antenna. Hence, the name shadowing.

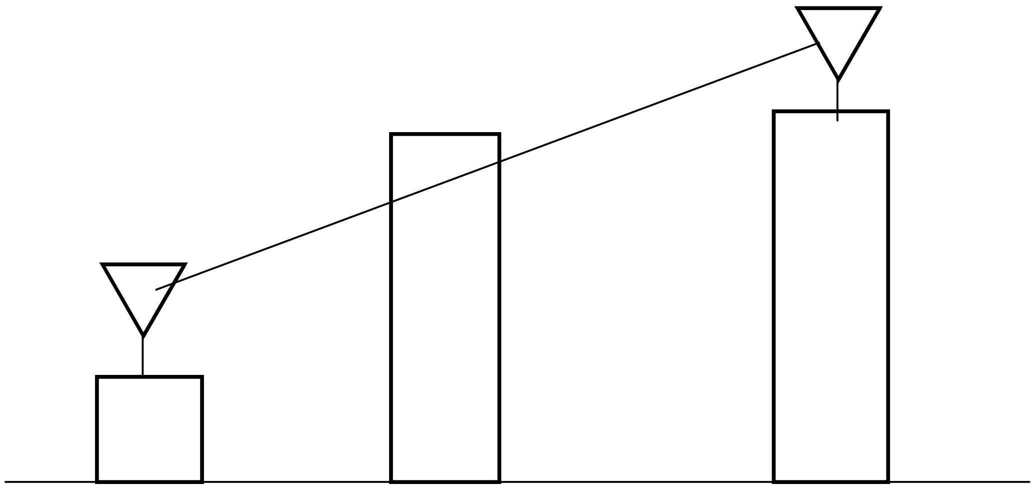


Figure A.1.2: Shadowing

The net path loss becomes:

$$PL(d)\,dB = \overline{PL}(d_0) + 10\alpha\log\left(\frac{d}{d_0}\right) + \chi$$

Here χ is a normally (Gaussian) distributed random variable (in dB) with standard deviation σ. χ represents the effect of shadowing. As a result of shadowing, power received at the points that are at the same distance d from the transmitter may be different and have a lognormal distribution. This phenomenon is referred to as **lognormal shadowing**.

A.1.4 Multipath

The objects located around the path of the wireless signal reflect the signal. Some of these reflected waves are also received at the receiver. Since each of these reflected signals takes a different path, it has a different amplitude and phase.

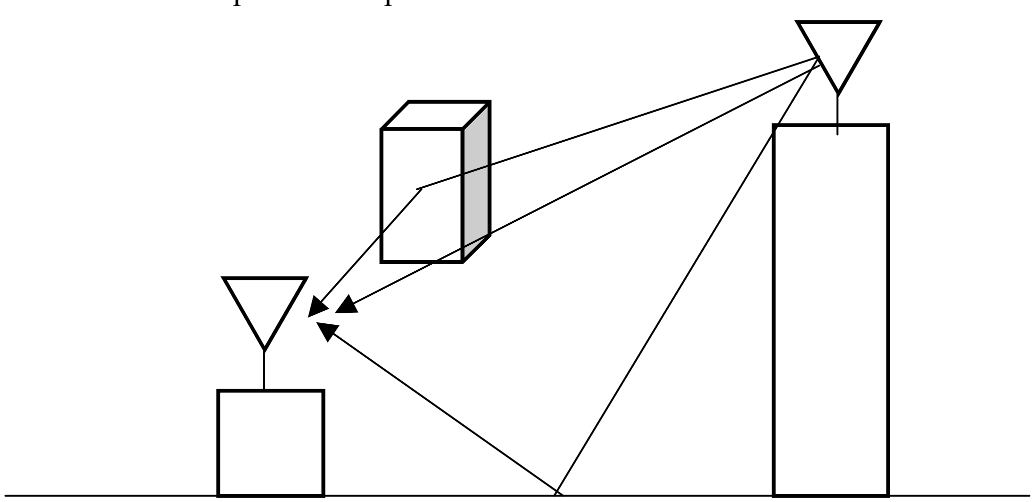


Figure A.1.3: Multipath

Depending upon the phase, these multiple signals may result in increased or decreased received power at the receiver. Even a slight change in position may result in a significant difference in phases of the signals and so in the total received power. The three components of the channel response are shown clearly in Figure A.1.4. The thick dashed line represents the path loss. The lognormal shadowing changes the total loss to that shown by the thin dashed line. The multipath finally results in variations shown by the solid thick line. Note that signal strength variations due to multipath change at distances in the range of the signal wavelength.

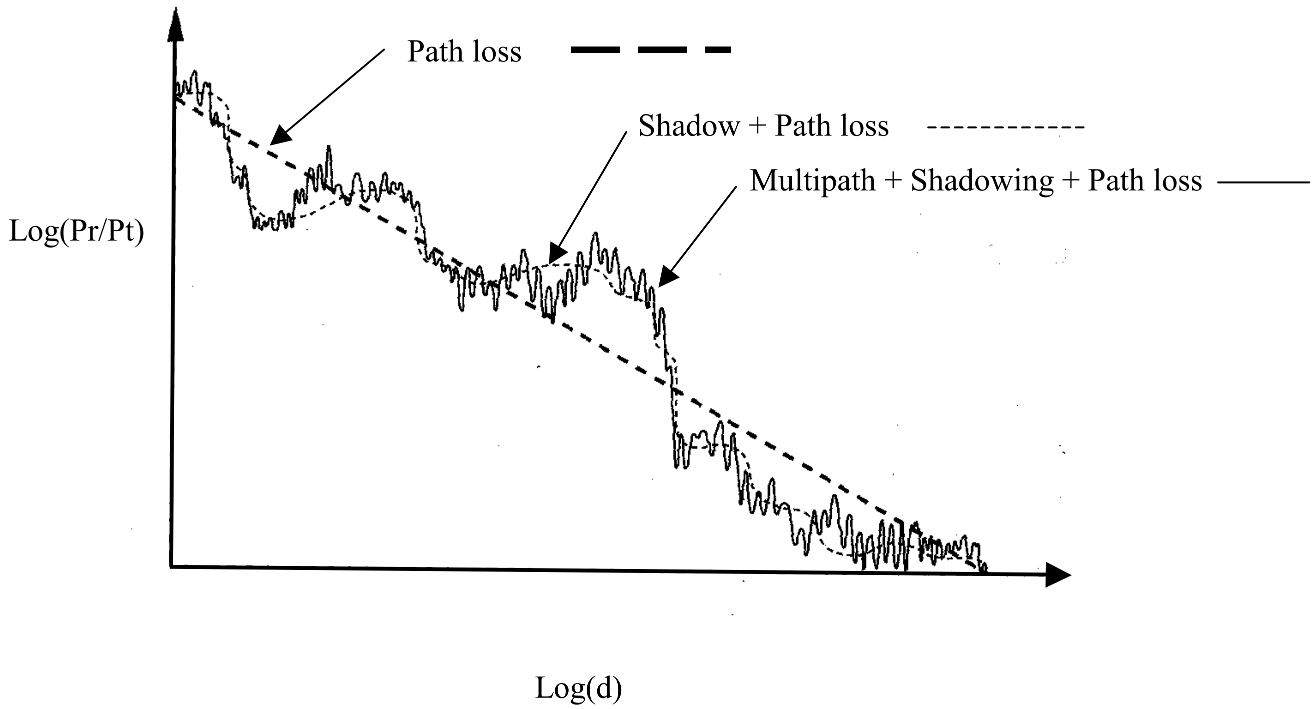


Figure A.1.4: Path loss, shadowing, and Multipath [Goldsmith2005]

Since different paths are of different lengths, a single impulse sent from the transmitter will result in multiple copies being received at different times as shown in Figure A.1.5

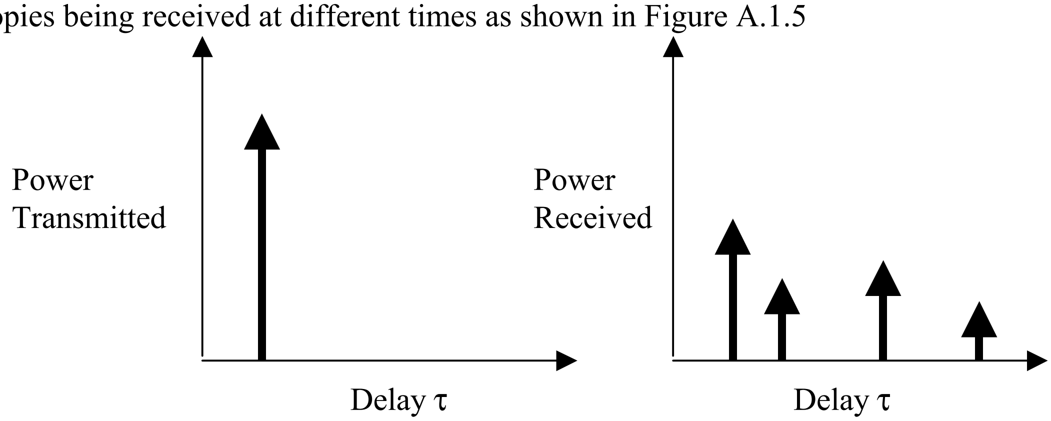


Figure A.1.5: Multipath Power Delay Profile

The maximum delay after which the received signal becomes negligible is called maximum **delay spread** τ_{max}. A large τ_{max} indicates a highly *dispersive* channel. Often root-mean-square (rms) value of the delay-spread τ_{rms} is used instead of the maximum.

A.1.5 Tapped Delay Line Model

One way to represent the impulse response of a multipath channel is by a discrete number of impulses as follows:

$$h(t,\tau) = \sum_{i=1}^{N} c_i(t)\delta(\tau - \tau_i)$$

Note that the impulse response h varies with time t. The coefficients $c_i(t)$ vary with time. There are N coefficients in the above model. The selection of the N and delay values τ_i depends upon what is considered a significant level. This model represents the channel by a delay line with N taps. For example, the channel shown in Figure A.1.5 can be represented by a 4-tap model as shown in Figure A.1.6.

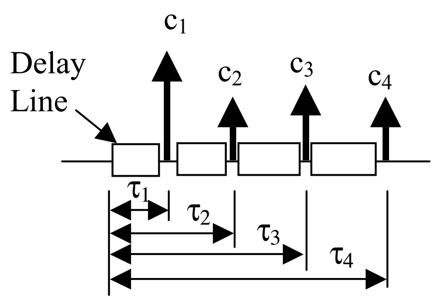


Figure A.1.6: Tapped Delay Line Model

If the transmitter, receiver, or even the other objects in the channel move, the channel characteristics change. The time for which the channel characteristics can be assumed to be constant is called **coherence time**. This is a simplistic definition in the sense that exact measurement of coherence time requires using the autocorrelation function.

For every phenomenon in the time domain, there is a corresponding phenomenon in the frequency domain. If we look at the Fourier transform of the power delay profile, we can obtain the frequency dependence of the channel characteristics. The frequency bandwidth for which the channel characteristics remain similar is called **coherence bandwidth**. Again, a more strict definition requires determining the autocorrelation of the channel characteristics. The coherence bandwidth is inversely related to the delay spread. The larger the delay spread, less is the coherence bandwidth and the channel is said to become more **frequency selective**.

A.1.6 Doppler Spread

The power delay profile gives the statistical power distribution of the channel over time for a signal transmitted for just an instant. Similarly, Doppler power spectrum gives the statistical power distribution of the channel for a signal transmitted at just one frequency f. While the power delay profile is caused by multipath, the Doppler spectrum is caused by motion of the intermediate objects in the channel. The Doppler power spectrum is nonzero for ($f-f_D$, $f+f_D$), where f_D is the maximum Doppler spread or **Doppler spread**.

The coherence time and Doppler spread are inversely related:

$$Coherence\,Time \approx \frac{1}{Doppler\,Spread}$$

Thus, if the transmitter, receiver, or the intermediate objects move very fast, the Doppler spread is large and the coherence time is small, i.e., the channel changes fast.

Table A.1.1 lists typical values for the Doppler spread and associated channel coherence time for two WiMAX frequency bands. Note that at high mobility, the channel changes 500 times per second, requiring good channel estimation algorithms

Table A.1.1: Typical Doppler Spreads and Coherence Times for WiMAX Transmissions [Andrews2007]

Carrier Freq	Speed	Max Doppler Spread	Coherence Time
2.5 GHz	2 km/hr	4.6 Hz	200 ms
2.5 GHz	45 km/hr	104.2 Hz	10 ms
2.5 GHz	100 km/hr	231.5 Hz	4 ms
5.8 GHz	2 km/hr	10.7 Hz	93 ms
5.8 GHz	45 km/hr	241.7 Hz	4 ms
5.8 GHz	100 km/hr	537 Hz	2 ms

A.2 Empirical Path Loss Models

Actual environments are too complex to model accurately. In practice, most simulation studies use empirical models that have been developed based on measurements taken in various real environments. In this section we describe a number of commonly used empirical models.

A.2.1 Hata Model

In 1968, Okumura conducted extensive measurements of base station to mobile signal attenuation throughout Tokyo and developed a set of curves giving median attenuation relative to free space path loss. To use this model one needs to use the empirical plots given in his paper. This is not very convenient to use. So in 1980, Hata developed closed-form expressions for Okumura's data. According to Hata model the path loss in an urban area at a distance d is:

$$P_{L,urban}(d)dB = 69.55 + 26.16\log_1 0(f_c) - 13.82\log_1 0(h_t) - a(h_r) + (44.9 - 6.55\log_1 0(h_t))\log_1 0(d)$$

Here, f_c is the carrier frequency, h_t is the height of the transmitting (base station) antenna, h_r is the height of the receiving (mobile) antenna, and $a(h_r)$ is a correction factor for the mobile antenna height based on the size of the coverage area.

Hata model well approximates the Okumura model for distances greater than 1km. This model is intended for large cells with BS being placed higher than the surrounding rooftops. Both models are designed for 150-1500 MHz and are applicable to the first generation cellular systems. They may not work well for WiMAX systems with smaller cell sizes and higher frequencies.

A.2.2 COST 231 Extension to Hata Model

The European Cooperative for Scientific and Technical (COST) research extended the Hata model to 2 GHz as follows:

$$P_{L,urban}(d)dB = 46.3 + 33.9\log_{10}(f_c) - 13.82\log_{10}(h_t) - a(h_r) + (44.9 - 6.55\log_{10}(h_t))\log_{10}(d) + C_M$$

Here, C_M is 0 dB for medium sized cities and suburbs and is 3 dB for metropolitan areas. The remaining parameters are same as before. This model is restricted to the following range of parameters:

Carrier Frequency	1.5 GHz to 2 GHz
Base Antenna Height	30 m to 300 m
Mobile Antenna Height	1m to 10 m
Distance d	1 km to 20 km

COST 321-Hata model is designed for large and small macro-cells, i.e., base station antenna heights above rooftop levels adjacent to base station.

A.2.3 COST 231-Walfish-Ikegami Model

In addition to the COST 231-Hata model, the COST 231 group also proposed another model for micro cells and small macro cells by combining models proposed by Walfisch and Ikegami [Walfisch1988]. This model considers additional characteristics of the urban environment, namely, heights of buildings h_{roof}, width of roads w, building separation b, and road orientation with respect to the direct radio path φ. These parameters are shown in Figure A.2.1 below.

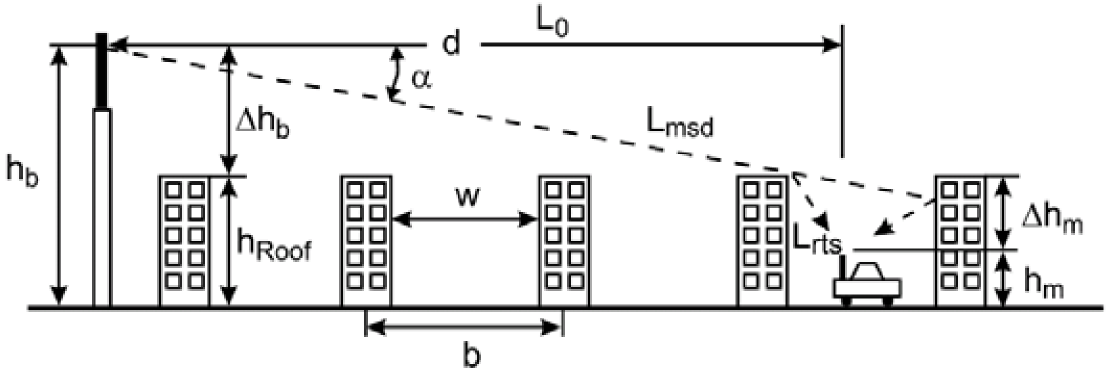


Figure A.2.1: Parameters of the COST-231 W-I Model [Molisch2005]

This model distinguishes between the line-of-sight (LOS) and non-line-of-sight (NLOS) cases. For LOS, the total path loss is:

$$P_L dB = 42.6 + 26\log(d) + 20\log(f_c)$$

for $d \geq 0.02$ km. Here d is in units of kilometers, and f_c is the carrier frequency in MHz.

For the non-LOS case, path loss consists of three terms: the free space path loss L_0, the multi-screen loss L_{msd} along the propagation path, and attenuation from the last roof edge to the MS, L_{rts} (rooftop-to-street diffraction and scatter loss):

$$P_L dB = \begin{cases} P_{L0} + L_{rts} + L_{msd} & for\ L_{rts} + L_{msd} > 0 \\ P_{L0} & for\ L_{rts} + L_{msd} \leq 0 \end{cases}$$

The free space path loss is:

$$P_{L0} dB = 32.4 + 20\log d + 20\log f_c$$

Ikegami derived the diffraction loss L_{rts} as:

$$L_{rts} = -16.9 - 10\log w + 10\log f_c + 20\log \Delta h_m + L_{ori}$$

Here w is the width of the street in meters, and Δh_m is the difference between the building height h_{Roof} and the height of the MS h_m

$$\Delta h_m = h_{Roof} - h_m$$

Orientation of the street is taken into account by an empirical correction factor L_{ori}:

$$L_{ori} = \begin{cases} -10 + 0.354\varphi & for \quad 0^\circ \le \varphi \le 35^\circ \\ 2.5 + 0.075(\varphi - 35) & for \quad 35^\circ \le \varphi \le 55^\circ \\ 4.0 - 0.114(\varphi - 55) & for \quad 55^\circ \le \varphi \le 90^\circ \end{cases}$$

Here φ is the angle between the street orientation and the direction of incidence in degrees as shown in Figure A.2.2. While L_{rts} expression is from Ikegami, the expressions for L_{ori} are different from those proposed by Ikegami.

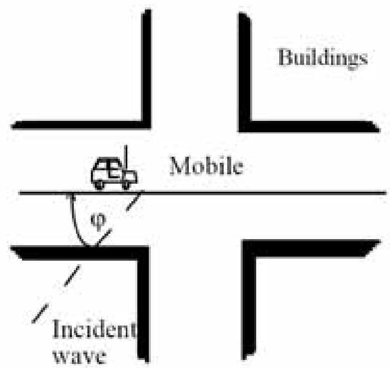


Figure A.2.2: Street Orientation Angle [Cichon]

The multi-screen loss L_{msd} is obtained by modeling building edges as screens. The multi-screen loss is [Walfish1988]:

$$L_{msd} = L_{bsh} + k_a + k_d \log d + k_f \log f_c - 9 \log b$$

Here b is the distance between two buildings (in meters), and:

$$L_{bsh} = \begin{cases} 18 \log(1 + \Delta h_b) & for \quad h_b > h_{Roof} \\ 0 & for \quad h_b \le h_{Roof} \end{cases}$$

$$k_a = \begin{cases} 54 & for \quad h_b > h_{Roof} \\ 54 - 0.8\Delta h_b & for \quad d \ge 0.5km \, and \, h_b \le h_{Roof} \\ 54 - 0.8\Delta h_b d/0.5 & for \quad d < 0.5km \, and \, h_b = h_{Roof} \end{cases}$$

Here

$$\Delta h_b = h_b - h_{Roof}$$

and h_b is the height of the BS. The path loss depends on frequency and distance as given via the parameters k_d and k_f:

$$k_d = \begin{cases} 18 & \text{for } h_b > h_{Roof} \\ 18 - 15 \Delta h_b / h_{Roof} & \text{for } h_b \le h_{Roof} \end{cases}$$

$$k_f = \begin{cases} 0.7\left(\frac{f_c}{925} - 1\right) & \text{for medium-size cities} \\ & \text{suburban areas with average vegetation density} \\ 1.5\left(\frac{f_c}{925} - 1\right) & \text{for metropolitan areas} \end{cases}$$

The validity range for this model is:

Carrier frequency f_c	800–2,000 MHz
Height of BS antenna h_b	4–50m
Height of MS antenna h_m	1–3m
Distance d	0.02–5km

The model assumes a Manhattan street grid (streets intersecting at right angles), constant building height, and flat terrain. The model does not include the effect of wave guiding through street canyons, which can lead to an underestimation of the received field strength.

The COST 231-WI model has been accepted by ITU-R and is included in Report 567-4. The estimation of the path loss agrees well with measurements for base station antenna heights above rooftop level. The prediction error becomes large for $h_{Base} = h_{Roof}$ compared to situations where $h_{Base} \gg h_{Roof}$. Also, the performance of the model is poor for $h_{Base} \ll h_{Roof}$. The prediction error for micro-cells may be quite large. The reliability of the estimation decreases also if terrain is not flat or the land cover is inhomogeneous.

A.2.4 Erceg Model

This model [Erceg1999] is based on extensive experimental data collected by AT&T Wireless Services across the United States in 95 existing macro cells at 1.9GHz. The terrains are classified in three categories. Category A is hilly terrain with moderate-to-heavy tree density and has a high path loss. Category C is mostly flat terrain with light tree density and has a low path loss. Category B is hilly terrain with light tree density or flat terrain with moderate-to-heavy tree density. Category B has an intermediate path loss. For all three categories, the median path loss at distance $d > d_0$ is given by:

$$P_L\, dB = 20\log_{10}(4\pi d_0/\lambda) + 10\gamma\log_{10}(d/d_0) + s \quad for\, d > d_0$$

Here, λ is the wavelength in meters, γ is the path-loss exponent with:

$$\gamma = a - bh_b + d/h_b$$

h_b is the height of the base station in meters (between 10 m and 80 m), d_0 = 100 m, and a, b, c are constants dependent on the terrain category. These parameters are listed in the table below.

Table A.2.1: Parameters of the Erceg Model

Model Parameter	Terrain Type A	Terrain Type B	Terrain Type C
a	4.6	4	3.6
b	0.0075	0.0065	0.005
c	12.6	17.1	20

s represents the shadowing effect and follows a lognormal distribution with a typical standard deviation of 8.2 to 10.6 dB.

The above model is valid for frequencies close to 2 GHz and for receive antenna heights close to 2 m. For other frequencies and antenna heights (between 2 m and 10 m), the following correction terms are recommended [Molisch2005]:

$$PL_{modified} = PL + \Delta PL_f + \Delta PL_h$$

Here, PL, is the path loss given earlier, ΔPL_f is the frequency term, and ΔPL_h is the receive antenna height correction terms given as follows:

$$\Delta PL_f = 6\log_{10}(f/2000)$$

$$\Delta PL_h = \begin{cases} -10.8\log_{10}(h/2) & for\, Categories\, A\, and\, B \\ -20\log_{10}(h/2) & for\, Category\, C \end{cases}$$

A.2.5 Stanford University Interim (SUI) Channel Models

This is a set of 6 channel models representing three terrain types and a variety of Doppler spreads, delay spread and line-of-sight/non-line-of-site conditions that are typical of the continental US as follows[Erceg2001]:

Table A.2.2: Terrain Type and Doppler Spread for SUI Channel Models

Channel	Terrain Type	Doppler Spread	Spread	LOS
SUI-1	C	Low	Low	High
SUI-2	C	Low	Low	High
SUI-3	B	Low	Low	Low
SUI-4	B	High	Moderate	Low
SUI-5	A	Low	High	Low
SUI-6	A	High	High	Low

The terrain type A, B,C are same as those defined earlier for Erceg model. The multipath fading is modeled as a tapped delay line with 3 taps with non-uniform delays. The gain associated with each tap is characterized by a Rician Distribution and the maximum Doppler frequency.

In a multipath environment, the received power r has a Rician distribution, whose pdf is given by:

$$pdf(r) = \frac{r}{\sigma^2} e^{\left[-fracr^2 + A^2 2\sigma^2\right]} I_0\left(\frac{rA}{\sigma^2}\right) \qquad 0 \le r \le \infty$$

Here, $I_0(x)$ is the modified Bessel function of the first kind, zero order. A is zero if there is no LOS component and the pdf of the received power becomes:

$$pdf(r) = \frac{r}{\sigma^2} e^{\left[-\frac{r^2}{2\sigma^2}\right]} \qquad 0 \le r \le \infty$$

This is the Raleigh distribution. The ratio $K = A^2/(2\sigma^2)$ in the Rician case represents the ratio of LOS component to NLOS component and is called the "K-Factor" or "Rician Factor." For NLOS case, K-factor is zero and the Rician distribution reduces to Raleigh Distribution.

The general structure for the SUI channel model is as shown below in Figure A.2.3. This structure is for Multiple Input Multiple Output (MIMO) channels and includes other configurations like Single Input Single Output (SISO) and Single Input Multiple Output (SIMO) as subsets. The SUI channel structure is the same for the primary and interfering signals.

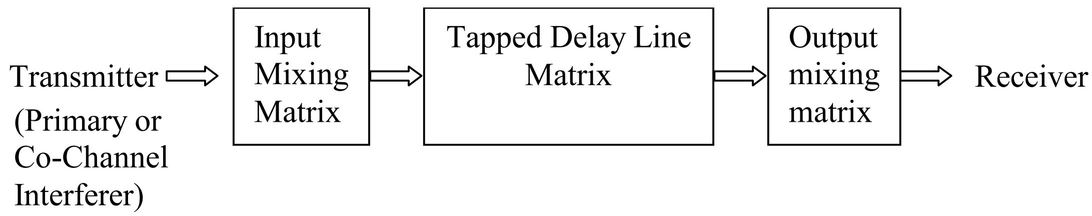


Figure A.2.3: Generic Structure of SUI Channel Models

A.2.5.1 Input Mixing Matrix

This part models correlation between input signals if multiple transmitting antennas are used.

A.2.5.2 Tapped Delay Line Matrix

This part models the multipath fading of the channel. The multipath fading is modeled as a tapped-delay line with 3 taps with non-uniform delays. The gain associated with each tap is characterized by a distribution (Rician with a K-factor > 0, or Raleigh with K-factor = 0) and the maximum Doppler frequency.

A.2.5.3 Output Mixing Matrix

This part models the correlation between output signals if multiple receiving antennas are used.

Using the above general structure of the SUI Channel and assuming the following scenario, six SUI channels are constructed which are representative of the real channels.

A.2.5.4 Scenario for modified SUI channels

Table A.2.3: Scenario for SUI Channel Models

Cell size	7 km
BTS Antenna Height	30 m
Receive Antenna Height	6 m
BTS Antenna beam Width	120°
Receive Antenna Beam Width	Omni directional (360°) and 30°.
Vertical Polarization Only	
90% cell coverage with 99.9% reliability at each location covered.	

In the following modles, the total channel gain is not normalized. Before using a SUI model, the specified normalization factors have to be added to each tap to arrive at 0dB total mean power. The specified Doppler is the maximum frequency parameter. The Gain Reduction Factor (GRF) is the total mean power reduction for a 30° antenna compared to an omni antenna. If 30° antennas are used the specified GRF should be added to the path loss. Note that this implies that all 3 taps are affected equally due to effects of local scattering. K-factors have linear values, not dB values. K-factors for the 90% and 75% cell coverage are shown in the tables, i.e., 90% and 75% of the cell locations have K-factors greater or equal to the K-factor value specified, respectively. For the SUI channels 5 and 6, 50% K-factor values are also shown.

Table A.2.3: SUI – 1 Channel Model

	Tap 1	Tap 2	Tap 3	Units
Delay	0	0.4	0.9	μs
Power (omni ant.) **90% K-factor (omni)** **75% K-factor (omni)**	0 4 20	-15 0 0	-20 0 0	dB
Power (30° ant.) **90% K-factor (30°)** **75% K-factor (30°)**	0 16 72	-21 0 0	-32 0 0	dB
Doppler	0.4	0.3	0.5	Hz
Antenna Correlation: $\rho_{ENV} = 0.7$ **Gain Reduction Factor:** GRF = 0 dB **Normalization Factor:** $F_{omni} = -0.1771$ dB, $F_{30°} = -0.0371$ dB		**Terrain Type:** C **Omni antenna:** $\tau_{RMS} = 0.111$ μs, overall K: K = 3.3 (90%); K = 10.4 (75%) **30° antenna:** $\tau_{RMS} = 0.042$ μs, overall K: K = 14.0 (90%); K = 44.2 (75%)		

Table A.2.4: SUI – 2 Channel Model

	Tap 1	Tap 2	Tap 3	Units
Delay	0	0.4	1.1	μs
Power (omni ant.) **90% K-factor (omni)** **75% K-factor (omni)**	0 2 11	-12 0 0	-15 0 0	dB
Power (30° ant.) **90% K-factor (30°)** **75% K-factor (30°)**	0 8 36	-18 0 0	-27 0 0	dB
Doppler	0.2	0.15	0.25	Hz
Antenna Correlation: $\rho_{ENV} = 0.5$ **Gain Reduction Factor:** GRF = 2 dB **Normalization Factor:** $F_{omni} = -0.3930$ dB, $F_{30°} = -0.0768$ dB		**Terrain Type:** C **Omni antenna:** $\tau_{RMS} = 0.202$ μs, overall K: K = 1.6 (90%); K = 5.1 (75%) **30° antenna:** $\tau_{RMS} = 0.069$ μs, overall K: K = 6.9 (90%); K = 21.8 (75%)		

Table A.2.5: SUI – 3 Channel Model

	Tap 1	Tap 2	Tap 3	Units

Delay	0	0.4	0.9	µs
Power (omni ant.)	0	-5	-10	dB
90% K-factor (omni)	1	0	0	
75% K-factor (omni)	7	0	0	
Power (30° ant.)	0 3 19	-11	-22	dB
90% K-factor (30°)		0	0	
75% K-factor (30°)		0	0	
Doppler	0.4	0.3	0.5	Hz

Antenna Correlation: $\rho_{ENV} = 0.4$ **Gain Reduction Factor:** GRF = 3 dB **Normalization Factor:** $F_{omni} = -1.5113$ dB, $F_{30°} = -0.3573$ dB	**Terrain Type:** B **Omni antenna:** $\tau_{RMS} = 0.264$ µs, overall K: K = 0.5 (90%); K = 1.6 (75%) **30° antenna:** $\tau_{RMS} = 0.123$ µs, overall K: K = 2.2 (90%); K = 7.0 (75%)

Table A.2.6: SUI – 4 Channel Model

	Tap 1	**Tap 2**	**Tap 3**	**Units**
Delay	0	1.5	4	µs
Power (omni ant.)	0	-4	-8	dB
90% K-factor (omni)	0	0	0	
75% K-factor (omni)	1	0	0	
Power (30° ant.)	0	-10	-20	dB
90% K-factor (30°)	1	0	0	
75% K-factor (30°)	5	0	0	
Doppler	0.2	0.15	0.25	Hz

Antenna Correlation: $\rho_{ENV} = 0.3$ **Gain Reduction Factor:** GRF = 4 dB **Normalization Factor:** $F_{omni} = -1.9218$ dB, $F_{30°} = -0.4532$ dB	**Terrain Type:** B **Omni antenna:** $\tau_{RMS} = 1.257$ µs overall K: K = 0.2 (90%); K = 0.6 (75%) **30° antenna:** $\tau_{RMS} = 0.563$ µs overall K: K = 1.0 (90%); K = 3.2 (75%)

Table A.2.7: SUI – 5 Channel Model

	Tap 1	**Tap 2**	**Tap 3**	**Units**
Delay	0	4	10	µs
Power (omni ant.)	0	-5	-10	dB
90% K-factor (omni)	0	0	0	
75% K-factor (omni)	0	0	0	
50% K-fact (omni)	2	0	0	
Power (30° ant.)	0	-11	-22	dB
90% K-factor (30°)	0	0	0	
75% K-factor (30°)	2	0	0	
50% K-factor (30°)	7	0	0	
Doppler	2	1.5	2.5	Hz

Antenna Correlation: $\rho_{ENV} = 0.3$ **Gain Reduction Factor:** GRF = 4 dB **Normalization Factor:** $F_{omni} = -1.5113$ dB, $F_{30°} = -0.3573$ dB	**Terrain Type:** A **Omni antenna:** $\tau_{RMS} = 2.842$ µs overall K: K = 0.1 (90%); K = 0.3 (75%); K = 1.0 (50%) **30° antenna:** $\tau_{RMS} = 1.276$ µs overall K: K = 0.4 (90%); K = 1.3 (75%); K = 4.2 (50%)

Table A.2.8: SUI – 6 Channel Model				
	Tap 1	**Tap 2**	**Tap 3**	**Units**
Delay	0	14	20	µs
Power (omni ant.) **90% K-factor (omni)** **75% K-factor (omni)** **50% K-factor (omni)**	0 0 0 1	-10 0 0 0	-14 0 0 0	dB
Power (30° ant.) **90% K-factor (30°)** **75% K-factor (30°)** **50% K-factor (30°)**	0 0 2 5	-16 0 0 0	-26 0 0 0	dB
Doppler	0.4	0.3	0.5	Hz
Antenna Correlation: ρ_{ENV} = 0.3 **Gain Reduction Factor:** GRF = 4 dB **Normalization Factor:** F_{omni} = -0.5683 dB, $F_{30°}$ = -0.1184 dB		**Terrain Type:** A **Omni antenna:** τ_{RMS} = 5.240 µs overall K: K = 0.1 (90%); K = 0.3 (75%); K = 1.0 (50%) **30° antenna:** τ_{RMS} = 2.370 µs overall K: K = 0.4 (90%); K = 1.3 (75%); K = 4.2 (50%)		

The MATLAB code for implementing these six SUI models is available in [Erceg2001].

A.2.6 ITU Path Loss Models

Another commonly used set of empirical channel models is that specified in ITU-R recommendation M.1225. The recommendation specifies three different test environments: Indoor office, outdoor-to-indoor pedestrian, and vehicular – high antenna. Since the delay spread can vary significantly, the recommendation specifies two different delay spreads for each test environment: low delay spread (A), and medium delay spread (B). In all there are 6 cases. For each of these cases, a multipath tap delay profile is specified. The number of multipath components in each model is different. The following three tables list the specified parameters [M.1225].

Table A.2.8: ITU Channel Model for Indoor Office

Tap	Channel A		Channel B		Doppler spectrum
	Relative delay (ns)	Average power (dB)	Relative delay (ns)	Average power (dB)	
1	0	0	0	0	Flat
2	50	–3.0	100	–3.6	Flat
3	110	–10.0	200	–7.2	Flat
4	170	–18.0	300	–10.8	Flat
5	290	–26.0	500	–18.0	Flat
6	310	–32.0	700	–25.2	Flat

Table A.2.9: ITU Channel Model for Outdoor to Indoor and Pedestrian Test Environment

Tap	Channel A		Channel B		Doppler spectrum
	Relative delay (ns)	Average power (dB)	Relative delay (ns)	Average power (dB)	
1	0	0	0	0	Classic
2	110	–9.7	200	–0.9	Classic
3	190	–19.2	800	–4.9	Classic
4	410	–22.8	1 200	–8.0	Classic
5	–	–	2 300	–7.8	Classic
6	–	–	3 700	–23.9	Classic

Table A.2.10: ITU Channel Model for Vehicular Test Environment

Tap	Channel A		Channel B		Doppler spectrum
	Relative delay (ns)	Average power (dB)	Relative delay (ns)	Average power (dB)	
1	0	0.0	0	–2.5	Classic
2	310	–1.0	300	0	Classic
3	710	–9.0	8.900	–12.8	Classic
4	1 090	–10.0	12 900	–10.0	Classic
5	1 730	–15.0	17 100	–25.2	Classic
6	2 510	–20.0	20 000	–16.0	Classic

Each of these 6 cases is expected to occur some percentage of time. The following table gives the expected percentage of occurance and associated average RMS delay spread as specified in the recommendation:

Table A.2.11: Percentage Occurrence and Associated RMS Delay Spread for ITU Channel Models

Test environment	Channel A		Channel B	
	r.m.s. (ns)	*P* (%)	r.m.s. (ns)	*P* (%)
Indoor office	35	50	100	45
Outdoor to indoor and pedestrian	45	40	750	55

Vehicular – high antenna	370	40	4 000	55

WiMAX Forum recommends using just two of these six models. These are Ped-A and Veh-B models.

References

[M.1225] ITU-R Recommendation M.1225, "Guidelines for evaluation of radio transmission technologies for IMT-2000," 1997

[Okumura1968] T. Okumura, E. Ohmore, and K. Fukuda, "Field strength and its variability in VHF and UHF land mobile service," Rev. Elec. Commun. Lab., pp. 825-73, September-October 1968.

[Hata1980] M. Hata, "Empirical formula for propagation kiss ub kabd nibuke radui services," IEEE Trans. Veh. Tech., pp. 317-25, August 1980.

[COST231] European Cooperative in the Field of Science and Technical Research EURO-COST 231, "Urban transmission loss models for mobile radio in the 900- and 1,800 MHz bands (Revision 2)," COST 231 TD(90)119 Rev. 2, The Hague, The Netherlands, September 1991, available at http://www.lx.it.pt/cost231/final_report.htm

[Ikegami1984] F. Ikegami, S. Yoshida, T. Takeuchi, and M. Umehira, "Propagation factors controlling mean field strength on urban streets," IEEE Trans. on Antennas Propagation, vol. 32, no. 8, pp. 822–829, Aug. 1984.

[Walfisch1988] J. Walfisch, H. L. Bertoni, "A theoretical model of UHF propagation in urban environments," IEEE Trans. on Antennas and Propagation, vol. 36, no. 12, pp. 1788-1796, Dec. 1988

[Erceg1999] V. Erceg et. al, "An empirically based path loss model for wireless channels in suburban environments," IEEE JSAC, vol. 17, no. 7, July 1999, pp. 1205-1211.

[Hari2000] K.V. S. Hari, K.P. Sheikh, and C. Bushue, "Interim channel models for G2 MMDS fixed wireless applications," IEEE 802.16.3c-00/49r2, November 15, 2000, available as www.ieee802.org/16/tg3/contrib/802163c-00_49r2.pdf

[Erceg2001] V. Erceg, et al, "Channel Models for Fixed Wireless Applications," IEEE 802.16.3c-01/29r4, July 2001, available as www.ieee802.org/16/tg3/contrib/802163c-01_29r4.pdf

[Molisch2005] Andreas F. Molisch, "Wireless Communications," Wiley 2005, 622 pp. and associated web page http://www.wiley.com/legacy/wileychi/molisch/supp/appendices/Chapter_7_Appendices.pdf

[Goldsmith2005] Andrea Goldsmith, "Wireless Communications," Cambridge University Press, 2005, 644 pp.

[Andrews2007] Jeffrey G. Andrews, Arunabha Ghosh, and Rias Muhamed, "Fundamentals of

WiMAX: Understanding Broadband Wireless Networking,", Prentice-Hall, February 2007, 496 pp.

[Cichon] Dieter J. Cichon, Thomas Kurner, "Propagation Prediction Models," Chapter 4, Available at www.it.lut.fi/kurssit/04-05/010651000/Luennot/Chapter4.pdf

Annex B: EESM PHY Abstraction

B.1 Objective

System level simulators have a challenge with simulating channels that are frequency selective where the signal is being impaired by interfering signals and thermal noise. Simulator must estimate the reception quality and rate (translated to BER BuER and PER) for each MS. Therefore it is infeasible and impractical to simulate the instantaneous performance of a wireless link in real-time. The following provides a method for approximating the reception quality through a PHY abstraction model for a short term.

B.2 Definition of PHY Abstraction

A PHY abstraction model is a mapping between a set of inputs from the system simulation, to a performance metric which is not dependent on the link.

- The computation load of this abstraction layer should be both feasible and independent of the actual channel observed.
- Typically, the performance metric is the signal-to-noise-ratio in AWGN conditions, for which modem performance curves can be easily generated.
- We shall refer to this performance metric by the name ‘effective SINR’, or ‘AWGN-equivalent SNR’.

$$P_{e,\,AWGN}(SIR_{\mathrm{eff}}) = P_{\mathrm{e,\,channel}}(SIR_{\mathrm{system}})$$

.

Many other technologies use a simple method based on averages for the PHY abstraction. This is not adequate for an OFDMA system. The average SINR observed by the receiver is not an adequate metric for performance evaluation for the following reasons

- FEC block bits are spread between sub carriers
- due to frequency selectivity of desired signal , each sub carriers observes different SNR
- Decoder behavior depends on the SINR fluctuations between FEC block bits, and not only on the average SINR
- Bursts that observe different channel and interference characteristics will display different BER / BLER results even though they may have the same average SINR.

PHY abstraction for OFDMA systems is a complicated mapping from a vector of the code-word’s per-subcarrier SINR levels to an ‘AWGN-equivalent SNR’ metric:

$$SINR_{eff} = f(SINR_1, SINR_2, ..., SINR_K)$$

Several mapping functions have been proposed over the years:

- Quasi-static method
- Convex method
- Shannon method
- EESM (**E**xponential **E**ffective **S**IR **M**apping)

EESM mapping is defined as (See reference [EESM]):

$$SINR_{eff} = -\beta \cdot \ln\left(\frac{1}{N}\sum_{i=1}^{N} e^{-\frac{SINR_i}{\beta}}\right)$$

where

- *SINRi* is the SINR of the *i*-th sub-carrier in the code-word,

b is an adjustment factor that depends on the FEC type , MCS and modem implementation specific .

B.3 Implementation of EESM (verify the steps)

1) **Calculate instantaneous per subcarrier SINR {SINR(i)}:**
   In OFDM based WiMAX SLS simulation, per subcarrier SINR is computed according to the location of BS and MSS, the channel gain, and per subcarrier interferes from neighboring cells.

2) **Calculate the instantaneous effective SINR ($SINR_{eff}$):**

$$SINR_{eff} = -\beta \cdot \ln\left(\frac{1}{N}\sum_{i=1}^{N} e^{-\frac{SINR_i}{\beta}}\right)$$

3) **Compute PER** from a table or a curve function calculated in advance (from LLS) for basic AWGN channel associated with given $SINR_{eff}$ from step two and the current used MCS:
   $PER = F_{AWGN}(SINR_{eff}, MCS)$

The methodology is based on the assumption that the actual PER for the instantaneous OFDM channel state {SINR(i)} can be approximated from the PER for a basic AWGN channel with the effective SINR for the specific MCS:

$$PER(\{SINR(k)\}, MCS) \approx F_{AWGN}(SINR_{eff}, MCS).$$

B.4 Beta (β) Training

See Annex E for b values to be used in simulation.

References

[1] "1xEV-DV Evaluation Methodology (V13), 3GPP2 / TSG-C.R1002.

[2] "MTMR Baseband Transceivers Needs for Intra- & Inter-system transparent UMTS/WLAN Operation", IST-2000-30116 FITNESS, D3.3.1/v1.0

[3] "OFDM EESM simulation Results for System-Level Performance Evaluations, and Text Proposal for Section A. 4.5 of TR 25.892", Nortel Networks, R1-04-0089, January, 2004

[4] "Feasibility Study for OFDM for UTRAN enhancement, Release 6", 3GPP TSG RAN, TR 25.892 v1.1.0, March 2004.

[5] "Link-system interface simulation methodologies", IEEE 802.20 Session #9, July 2004 [C802.20-04-67.ppt]

[6] "CINR measurements using the EESM method " IEEE C802.16e-05/141r1 ; 2005-03-09 ;

Annex C: MIC PHY Abstraction

C.1 Objective

For a network system-level simulator, to simulate the physical layer links between multiple BSs and SSs can be computationally prohibitive. Therefore, a PHY abstraction model is used to predict link layer performance in a computationally simply way.

C.2 Definition of PHY Abstraction

PHY abstraction is defined as a mapping function *L=F(C, MCS)* between *link layer performance indicator (L)* and *channel quality indicator (C)* for different modulation and coding schemes (MCS), which was obtained from PHY/Link Layer Simulation (LLS).

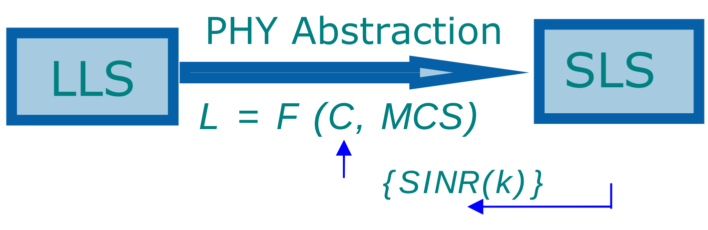


Figure C.2.1: PHY Abstraction

Choices of the indicators can be multiple. Especially for MIMO-OFDM wireless system which operates in frequency selective channel, the right choice is important and actually a matter of numerous researches.

Channel quality indicator *(C)* are derived from instantaneous channel state, such as the instantaneous SINR for each sub-carrier in OFDM as {SINR(k)}, (k=0,1,2…). The alternative metrics of *C* can be:

- Average SNR
- Mean instantaneous capacity (MIC)
- Effective SINR
- Exponential effective SINR and etc.

Link layer performance indicator (*L*) can be chosen according to the need of system level simulator:

- Packet (burst) error rate (PER)
- Block-error probability (BLEP)
- Bit error rate (BER)
- Symbol error rate (SER)
- Slot error rate (SLER) and etc.

C.3 Implementation of PHY Abstraction

There are several types of PHY abstraction models which are commonly used, depending on the indicators a model chooses:

- Average SNR:
- MIC
- Effective SINR Mapping (ESM)
- Exponential Effective SINR Mapping (EESM) etc.

Average SNR is known to be problematic for OFDM in broadband channels, because averaging the SINR on different tones which undergo different amounts of fading does not give a true picture of the channel.

In this document, the implementation of MIC and ESM are introduced. And we choose to use PER as the link layer performance indicator.

C.4 Implementation of MIC

4) **Calculate instantaneous per subcarrier SINR {SINR(k)}:**
   In OFDM based WiMAX SLS simulation, per subcarrier SINR is computed according to the location of BS and MSS, the channel gain, and per subcarrier interferes from neighboring cells.

5) **Calculate the MIC:**
   a. Compute the capacity of the *kth* subcarrier
   $$c(k) = \log_2(1 + SINR(k))$$
   b. Compute the MIC by averaging per subcarrier capacities across the subcarriers,
   $MIC = \frac{1}{N}\sum_{k=1}^{N} c(k)$, where N is the number of used subcarriers.

6) **Compute PER** from a table or a curve function calculated in advance (from LLS) for basic AWGN channel associated with given ***MIC*** from step two and the current used MCS:
   $$PER = F_{AWGN}(MIC, MCS)$$

   The methodology is based on the assumption that the actual PER for the instantaneous OFDM channel state {SINR(k)} can be approximated from the PER for a basic AWGN channel with the MIC for the specific MCS:

   $$PER(\{SINR(k)\}, MCS) \approx F_{AWGN}(MIC, MCS)$$

C.5 Implementation of ESM

7) **Calculate instantaneous per subcarrier SINR {SINR(k)}:** same as above.
8) **Calculate the instantaneous effective SINR (*$SINR_{eff}$*):**

a. Compute MIC: $MIC = \frac{1}{N}\sum_{k=1}^{N} c(k)$, same as above.

b. Compute effective SINR:

$$SINR_{eff} = 2^{MIC} - 1$$

9) **Compute PER** from a table or a curve function calculated in advance (from LLS) for basic AWGN channel associated with given ***SINR**$_{eff}$* from step two and the current used MCS:

$$PER = F_{AWGN}(SINR_{eff}, MCS)$$

The methodology is based on the assumption that the actual PER for the instantaneous OFDM channel state {SINR(k)} can be approximated from the PER for a basic AWGN channel with the effective SINR for the specific MCS:

$$PER(\{SINR(k)\}, MCS) \approx F_{AWGN}(SINR_{eff}, MCS).$$

References

[1] Shilpa Talwar, "System Level Simulations for 802.16e", Jan 2006.

[2] Alexander Maltserv, et al "802.16e System Level Simulations", Rev0.3, Mar 2005.

[3] Muthu Venkatachalam, "Some BKMs for AWG simlations", Jun 2006.

[4] Mustafa Demirhan, Ali Koc, Song Liu and Chunmei Liu, "Mobility Extensions for OPNET WiMAX Model", Apr 2006.

Annex D: MIM PHY Abstraction

D.1 Objective

Actual implementation of FECC coding/decoding in system level simulation is prohibitively complex. System Level simulations involve a large number of links between Sectors (~60) and Mobiles (10 to 100+/sector). Also it has to deal with many other functions that link level simulation do not (e.g., scheduler, upper layers, traffic models, network level issues such as queuing, etc). This indicates the need for abstraction models. Abstractions are based on short-term measurable data for a given code block such as computing an estimate of the probability of error. The second step is to flip a biased coin of the computed probability, to simulate channel performance. Link error prediction techniques require computation of an “effective SNR” and or an “effective code rate” used with either a single or a set of reference curves to yield an estimate of channel block error probability The following categories of error prediction methods have been applied for PHY abstraction

Effective SNR-based Prediction

–“Short Term” Method (STM)

–Non-linear Function Method

•Mutual Information Metric (MIM)

•Exponential Effective SIR Mapping (EESM)

Effective Code Rate Method (ECRM)

D.2 Comparison for various methods

Effective SNR is the actual instantaneous SNR averaged over the received code block. This method references curve that graphs a block error vs. instantaneous SNR averaged over the channel model in question. Its advantages are a) simple to compute and b) moderately accurate. This method has a few disadvantages. These are

– all the degrees of freedom are assumed to be captured by the mean and channel model type; this may not be adequate.

– distinct reference curves needed for each channel mode. Therefore it may be infeasible for spatial channel or other regimes

– need separate reference curves for each diversity scheme (Multiple antennas at Tx or Rx, HARQ)

Effective SNR in a non linear approach is computed by averaging a nonlinear function of the instantaneous SNR and using the inverse function of the average. That is by carrying out a summation over the current code block as follows

$$\gamma_{eff} = \mu^{-1}\left(\frac{1}{N}\sum_{1}^{N}\mu(\gamma_k)\right)$$

$\mu(\cdot)$ - Non-linear function

γ_k - SNR for the k[th] Symbol

γ_{eff} - effective SNR for the code block

Mutual Information Metric (MIM) is one type of a nonlinear function model. Reference 1 provides more details about MIM. "Exponential Effective SIR Mapping" is an alternative nonlinear function model (see reference 2) The exponential method uses the following formula with a free parameter β chosen for best fit

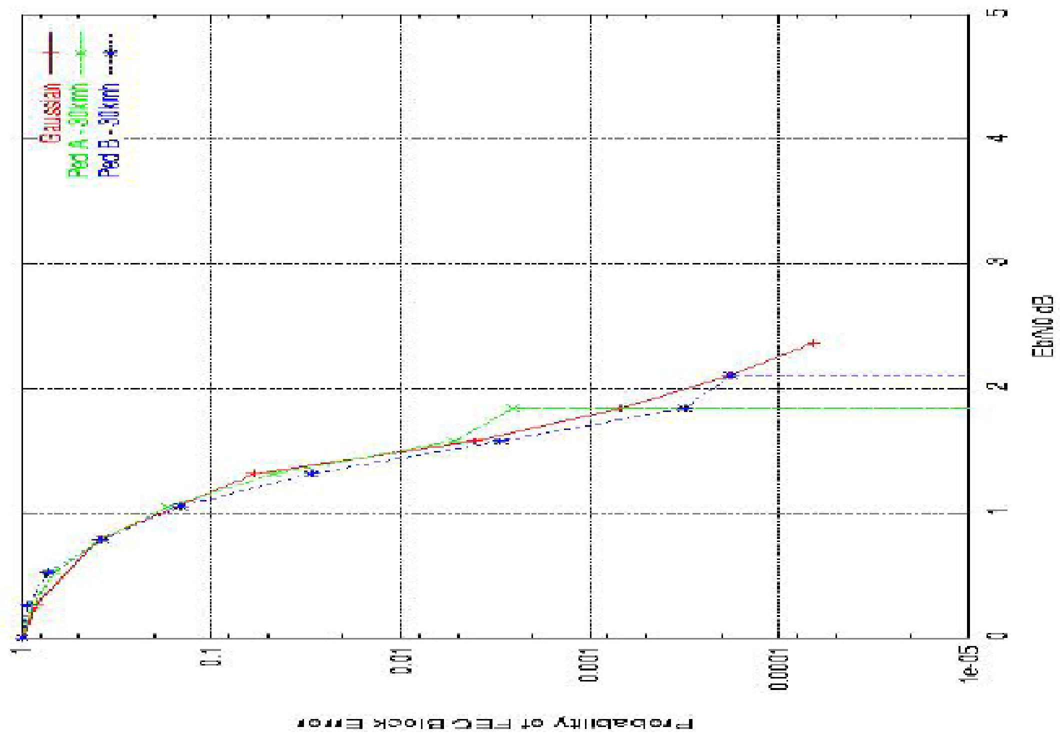
$$\mu(x) = e^{-\frac{x}{\beta}}$$

The advantage of the nonlinear methods is its robustness with respect to channel model. It can ideally refer back to a single reference curve (the basic Gaussian performance). But it might require a fudge factor

Effective Code Rate Method (ECRM) does not compute an effective SNR. It uses the standard short-term SNR and an effective code rate. (see Reference 3). The following formula is an example for calculating the effective rate based on the transmitted code rate R_0

$$R_{eff} = R_0 \frac{\frac{1}{N}\sum \gamma_k^2}{(\frac{1}{N}\sum \gamma_k)^2}$$

While ECRM provides computational simplicity its effectiveness has not been widely proven. However studies on MIM have shown that it is effective with OFDMA systems. The following graph shows the results from an experiment run on OFDMA 802.16 PUSC downlink with 60 byte blocks using Rate ½ mother code and QPSK. The study assumes ideal channel estimate, no HARQ, no power control. The study included a variety of 3GPP2 channel models



D.3 Summary

Effective SNR method using mutual information (MIM) appears to be an accurate and robust link-error-prediction method for OFDMA systems. Results have been carried out over a range of models, vehicle speed, modulation indices and code rates. The results from simulation agree with field measurements to within a fraction of a dB in almost every case. MIM has great potential to be used as the PHY model with OFDMA systems. Additional tests with MIM remains to be carried out for a) HARQ b) Non-ideal channel estimation c) Non-canonical and other channel models (e.g., SCM) b) Diversity (multi antenna)

References

[1] GPP2 TSG-C WG3, C30-20030414-034_LUC_Link Error Prediction_for_HARQ.doc

[2] IEEE C802.20-04-67_Link_System

[3] “Decoder” (C30-20051024-013-LU-Virtual-Decoder, 3GPP2 TSG-C WG3

Annex E: Reference EESM β values

The values presented in this annex are the recommended beta values for system level simulations. These values were computed using the following simulation conditions [BUPT07a and BUPT07b]:

- SISO
- PUSC mode
- SCM channel model with the velocity of 3Km/h (Ped-B) & 60Km/h (Veh-A)
- 100 independent channel realizations
- Ideal channel estimation is assumed.

Table E.1 Reference EESM Beta Values

Index	Mod	Rate	n=Rep	J	Data block size (bytes)	Encoded data block size (bytes) =48*n*M/8	Beta Ped-B (dB)	Beta Veh-A (dB)
1	QPSK	1/2	1	10	6	12	2.46	2.54
2	QPSK	1/2	2	10	12	24	2.28	2.26
3	QPSK	1/2	3	10	18	36	2.27	2.26
4	QPSK	1/2	4	10	24	48	2.18	2.12
5	QPSK	1/2	5	10	30	60	2.05	2.07
6	QPSK	1/2	6	10	36	72	2.00	2.06
7	QPSK	1/2	8	10	48	96	2.03	2.02
8	QPSK	1/2	9	10	54	108	2.04	2.01
9	QPSK	1/2	10	10	60	120	1.98	2.01
10	QPSK	3/4	1	6	9	12	2.56	2.50
11	QPSK	3/4	2	6	18	24	2.43	2.43
12	QPSK	3/4	3	6	27	36	2.46	2.44
13	QPSK	3/4	4	6	36	48	2.41	2.39
14	QPSK	3/4	5	6	45	60	2.41	2.41
15	QPSK	3/4	6	6	54	72	2.38	2.37

16	16-QAM	1/2	1	5	12	24	7.45	7.48
17	16-QAM	1/2	2	5	24	48	7.14	7.14
18	16-QAM	1/2	3	5	36	72	7.00	6.92
19	16-QAM	1/2	4	5	48	96	7.34	7.53
20	16-QAM	1/2	5	5	60	120	6.89	6.82
21	16-QAM	3/4	1	3	18	24	8.93	8.93
22	16-QAM	3/4	2	3	36	48	8.87	8.87
23	16-QAM	3/4	3	3	54	72	8.85	8.90
24	64-QAM	1/2	1	3	18	36	11.31	11.43
25	64-QAM	1/2	2	3	36	72	11.11	11.16
26	64-QAM	1/2	3	3	54	108	11.09	11.01
27	64-QAM	2/3	1	2	24	36	13.80	13.74
28	64-QAM	2/3	2	2	48	72	13.69	13.70
29	64-QAM	3/4	1	2	27	36	14.71	14.68
30	64-QAM	3/4	2	2	54	72	14.59	14.55
31	64-QAM	5/6	1	2	30	36	15.32	15.17
32	64-QAM	5/6	2	2	60	72	15.29	15.27

Note that beta values trained for Ped-B and Veh-A channel are quite similar for given format in most cases, coinciding with the theory that the beta training should be independent of channel realizations. There are some differences when the higher order modulations are used. Therefore, two beta tables are presented for different models in order to guarantee higher reliability of abstraction especially for higher order modulation. Beta values for QPSK modulation is around the theoretical value 2.0, especially when block length is small. The beta values presented could be used with the accuracy of 0.1, experiments have been made demonstrating the small changes of beta value do not influence the abstraction performance noticeably.

The abstraction performance is pretty good for QPSK modulation, also acceptable for 16QAM and 16QAM, although the performance degrades when higher order modulations are used. Note that EESM algorithm is derived for BPSK and QPSK modulation according to the Chernoff bounded pair-wise error probability (*PEP*). For higher-order modulations, such as 16QAM, it is not as straightforward to determine the exact expression for the exponential ESM in the same way. The reason is that different points in higher-order modulation constellation experience different channel attenuation and result in different error probabilities for each binary-symbol. The higher-order modulation in itself can be seen as a multi-state channel from a binary-symbol transmission point-of-view and its PEP can not be expressed explicitly. The formula used in the EESM derivation for BPSK and QPSK modulation is not suitable for 16QAM and 64QAM. They can not provide such tight error probability bounds for higher-order modulations as for QPSK modulation. Therefore, the abstraction performance degrades when higher-order modulation is adopted.

E.2 References:

[BUPT07a] Beijing University of Post and Telecommunications, “ESM Beta Training in 802.16e system 3.0,” WiMAX Forum, AATG Report, 2007, http://www.wimaxforum.org/apps/org/workgroup/aatg/download.php/24355/latest

[BUPT07b] Beijing University of Post and Telecommunications, “System Design and AWGN Results,” WiMAX Forum, AATG Report, 2007, http://www.wimaxforum.org/apps/org/workgroup/aatg/download.php/24354/latest

Annex F: Antenna Pattern and Orientation

F.1 Base station antenna pattern

All BS antenna elements have the beam pattern defined by 3GPP2. For angle of arrival θ (relative to the boresight direction), the antenna has gain (dB)

$$G(\theta) = G_{\max} + \max\left[-12\left(\frac{\theta}{\theta_{3dB}}\right)^2, -G_{FB}\right]$$

where θ_{3dB} is the 3dB beam width (degrees), $|\theta| \le 180$, $G_{\max}$ is the maximum antenna gain (boresight direction, dBi) and G_{FB} is the front-to-back ratio (dBi). An omni-directional antenna is created from this model by specifying $G_{FB} = 0$. Note that the antenna is typically oriented so that the boresight direction points to the center of the sector.

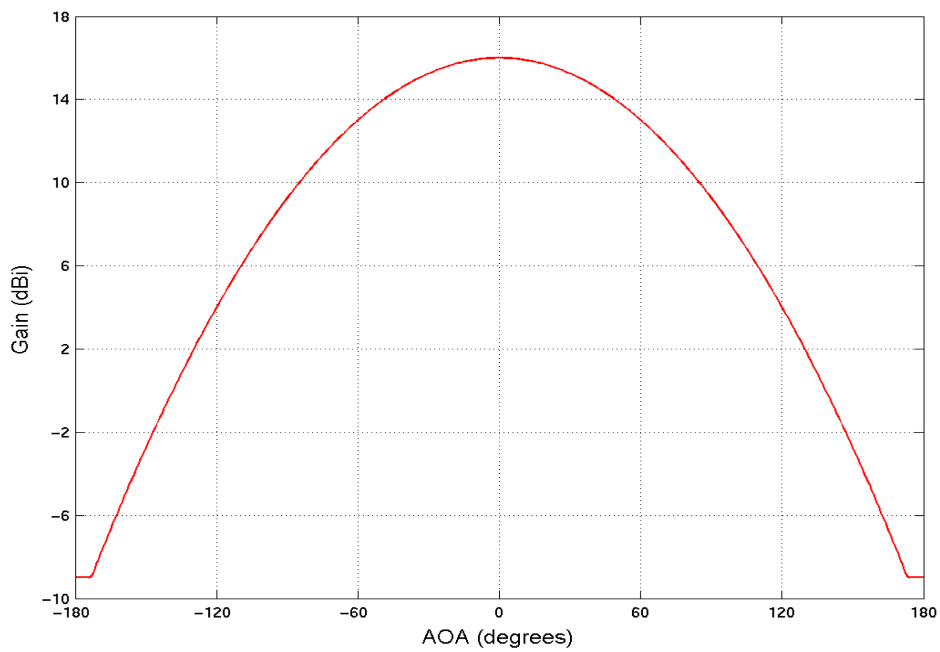


Figure F.1.1: Sector beam pattern for BS antenna

Annex G: Modelling PUSC in System Simulation

G.1 Introduction of PUSC

PUSC stands for partial usage of subchannels. And it is a way of subchannel allocation supported by WiMAX technology. There are two main categories of subchannel allocation scheme in WiMAX. *Contiguous subcarrier permutation* (Band AMC) allocates contiguous subcarriers into one subchannel. *Diversity subcarrier permutation* (FUSC, PUSC or TUSC) allocates non-contiguous subcarriers distributed over the entire bandwidth into subchannels through specific random permutation formulas.

The objective of diversity permutations is to achieve frequency diversity and inter-cell interference averaging. Furthermore, PUSC can be applied to allocate subset of subchannels (segments) to sectors of the same cell, if the cell is sectored.

G.2 Implementation of PUSC

The way PUSC permutation performs was standardized in [1, 2]. And UL PUSC differs from DL PUSC in subtle ways. For illustration purpose, we only summarize the implementation process for DL PUSC in **section 2.1**.

As we can see, the standard permutation scheme is fairly complex. If a system-level simulation is not going to perform frequency selective scheduling, a pseudorandom permutation can be implemented to approximate PUSC. The reason why we deem the approximation is acceptable and how to implement the pseudorandom permutation is explained in **section 2.2.**

G.2.1 Implementation of Standard DL PUSC [1,2]

DL PUSC permutation is comprised of an outer permutation and an inner permutation.

1) **Outer permutation** The outer permutation divides the subcarriers of one OFDMA symbol into 6 major groups of clusters using a renumbering sequence that is IDcell dependent.
   - **Dividing the subcarriers into (*Nclusters)* physical clusters** containing 14 adjacent subcarriers each (starting from carrier 0). The number of clusters, *Nclusters*, varies with FFT sizes.
   - **Renumbering the physical clusters into logical clusters**:
     - In the first PUSC zone of the downlink (first downlink zone) and in a PUSC zone defined by STC_DL_ZONE_IE() with 'use all SC indicator = 0', the default **renumbering sequence** is used for logical cluster definition.

       LogicalCluster = RenumberingSequence (PhysicalCluster)
     - For All other cases, DL_PermBase parameter (an integer ranging from 0 to 31) in the STC_DL_Zone_IE() or AAS_DL_IE() of DL-MAP messages shall be used:

       LogicalCluster =

       *RenumberingSequence((PhysicalCluster)+13*DL_PermBase) mod Nclusters*

     The renumbering sequence varies with FFT sizes.

- **Allocate logical clusters to six major groups**, according to FFT sizes. These groups may be allocated to segments, if a segment is being used, then at least one group shall be allocated to it (by default group 0 is allocated to sector 0, group 2 is allocated to sector 1, and group 4 to is allocated sector 2).

Table G.2.1: Subcarrier Groups

Clusters	***Nclusters***	**Group 0**	**Group 1**	**Group 2**	**Group 3**	**Group 4**	**Group 5**
FFT 2048	120	0-23	24-39	40-63	64-79	80-103	104-119
FFT 1024	60	0-11	12-19	20-31	32-39	40-51	52-59
FFT 512	30	0-9		10-19		20-29	
FFT 128	6	0-1		2-3		4-5	

2) **Inner permutation:** The inner permutation then operates separately on each major group, **allocating subcarriers to subchannel** within the group for each OFDMA symbol by

- first allocating the pilot carriers within each cluster (Figure 234 in [2])
- and then partitioning remaining data subcarriers into subchannels containing $N_{subcarriers}$=24 data subcarriers per subchannel. Each major group comprises several subchannels, and the size varies: even major groups are larger than odd major groups.
  - Partition the remaining data subcarriers into **24** groups of contiguous subcarriers.
  - Allocate a subchannel from one subcarrier from each of these groups according to ***permutation formula*** Equation (111) of 8.4.6.1.2.2.2. (Note: The subcarrier indexing within each group shall start from 0, where 0 is the lowest number subcarrier in the lowest numbered logical cluster belonging to the group.)

$$subcarrier(k,s) = N_{subchannels} \cdot n_k + \{p_s[n_k \bmod N_{subchannels}] + DL\_PermBase\} \bmod N_{subchannels} \qquad (111)$$

where,

- *subcarrier(k,s)* is the subcarrier index of subcarrier k in subchannel s.
- k is the subcarrier-in-subchannel index from $[0\ldots N_{subcarriers}-1]$.
- s is the index number of a subchannel, from $[0\ldots N_{subchannels}-1]$.
- $N_{subchannels}$ is the number of subchannels in current Major group.
- $N_{subcarriers}$ is number of data subcarriers allocated to a subchannel in each OFDMA symbol and equal to 24,
- $n_k = (k+13*s) \bmod N_{subcarriers}$
- *ps[j]* is the series obtained by rotating *basic permutation sequence* cyclically to the left s times.
- *DL_PermBase* is an integer ranging from 0 to 31, which identifies the particular BS segment and is specified by MAC layer. It is set to preamble IDCell in the first zone and determined by DL_PermBase parameter in the STC_DL_Zone_IE() or AAS_DL_IE() in the DL-MAP for other zones.

Table G.2.2: Number of Subchannels in Groups

)	**Even numbered major group**	**Odd numbered major group**
FFT 2048	$N_{subchannels}$ =12 Basic permutation sequence 12	$N_{subchannels}$ =8 Basic permutation sequence 8
FFT 1024	$N_{subchannels}$ =6 Basic permutation sequence 6	$N_{subchannels}$ =4 Basic permutation sequence 4
FFT 512	$N_{subchannels}$ =5 Basic permutation sequence 5	N/A
FFT 128	$N_{subchannels}$ =1 N/A	N/A

G.2.2 Implementation of PUSC approximation: pseudorandom permutation

G.2.2.1 Ideal Random Permutation

An ideal random permutation is defined as a permutation to randomly spread the subcarriers of the loaded subchannels throughout the entire channel bandwidth in uniform matter [3]. For two cells neighboring each other (assuming reuse of 1), the subcarrier overlapping for this ideal random permutation follows the well known binomial law ***Binomial(n, p).***

For example, for FFT size of 1024, there are total 30 channels (equals to 24*30 data subcarriers). If there are 2 subchannels to allocate, an ideal permutation will randomly choose 24*2 data subcarriers uniformly out of 720 data subcarriers and allocate them to 2 subchannels.

- The possibility of a subcarrier overlaps with a subcarrier in neighbor cell
  p = loading_factor = ***allocated_subchannels/total_subchannels***=2/30
- Repeat the experiment n times
  n=total_subcarriers=720
- The possibility of k subcarriers overlapping
  $C_n^k \, p^k \, (1-p)^{(n-k)}$

Comparing the subcarriers overlapping PDF of PUSC permutation to this ideal random permutation, the results taken from [3] (Figure G.2.1 and Figure G.2.2) show that PUSC performs very closely to the ideal random permutation.

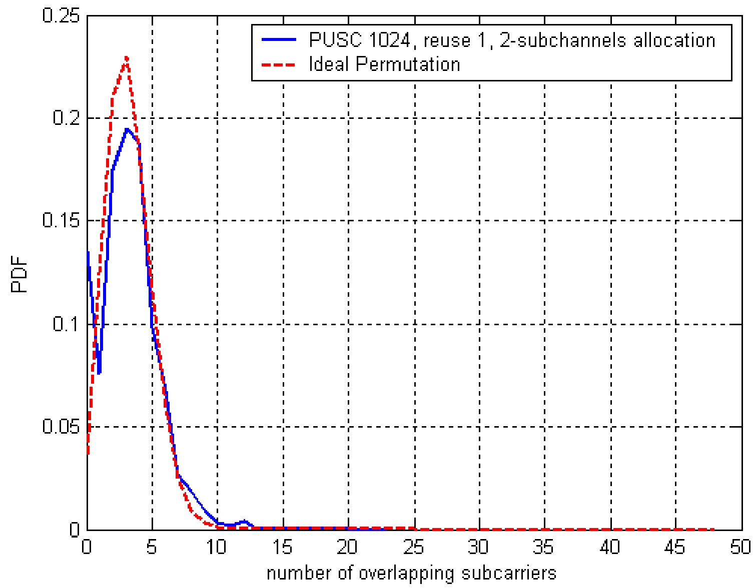


Figure G.2.1: subcarrier overlapping PDF for allocation size of 2 subchannels (6.7% loading).

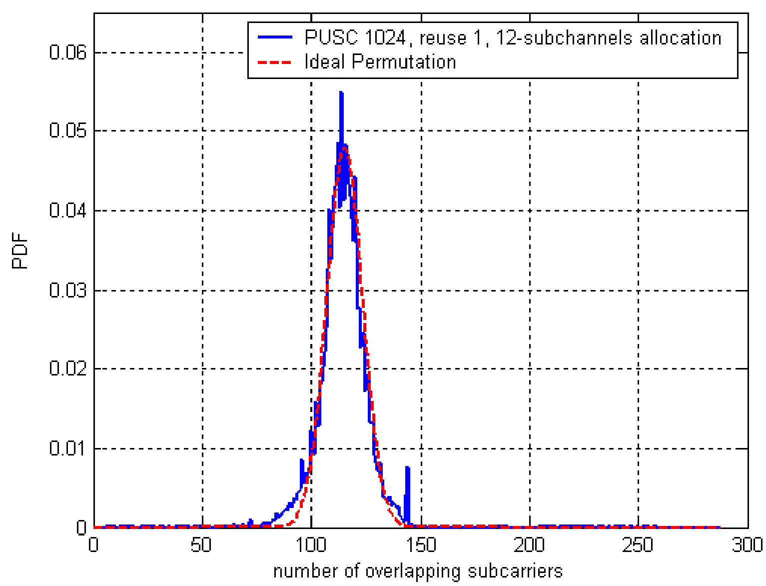


Figure G.2.2: subcarrier overlapping PDF for allocation size of 12 subchannels (40% loading).

G.2.2.2. Implementation of the pseudorandom permutation

To approximate the random permutation, a pseudorandom permutation can be implemented (pseudo code) to allocate the loaded subcarriers pseudo randomly:

```
        n = number_of_dataSubcarriers;
  f = randperm(n);  % randomly permutate the sequence [1, n]
  a = f(1:k);    % pick the first k numbers out of sequence f
```

The comparison between the subcarrier overlapping PDF of pseudorandom permutation and that of the ideal random permutation (Binomial distribution) in Figure 3 for subchannel=2 shows that the pseudo random permutation (implementation of ideal random permutation) can be used to approximate PUSC in system-level simulation for simplicity purpose.

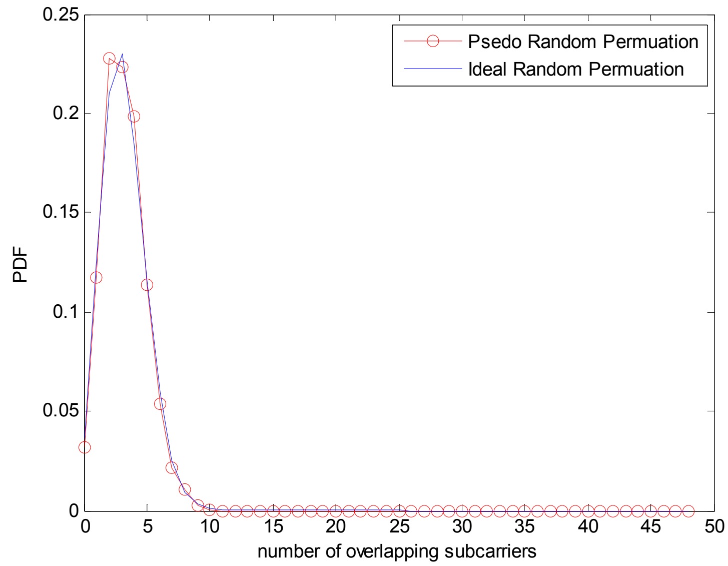


Figure G.2.3: subcarrier overlapping PDF for allocation size of 2 subchannels (6.7% loading).

Note: If the MAC scheduler will perform frequency selective scheduling, PUSC needs to be implemented without random approximation, because the approximation may lead to erroneous estimates of system performance.

References

[1] IEEE 802.16-2004

[2] IEEE 802.16e-2005

[3] Alexander Maltserv, et al "PUSC permutation analysis in interference enviroment", Rev0.4, April 2005.

[4] WiMAX Forum, "Mobile WiMAX – Part I: A technical Overview and Performance Evaluation", March 2006.

[5] Alvarion, "Hit ratio problems with the PUSC permutation", IEEE C802.16maint-5/083, Mar 2005.

[6] G. Nair, J. Chou, T. Madejski, K. Perycz, D. Putzolu and J. Sydir, "IEEE 802.16 Medium Access Control and Service Provisioning", Intel technology journal, Aug 2004

[Perf] WiMAX Forum, "Mobile WiMAX – Part I: A technical Overview and Performance Evaluation", March 2006.

[Profile] WiMAX Forum, "Mobile WiMAX System Profile", Feb 2006.

[EESM] IEEE EESM Reference, www.ieee802.org/16/tge/contrib/C80216e-05_141r3.pdf

Annex H: A Sample Link Budget Analysis

In this annex we provide link budget for a reference WiMAX system. The system parameters, OFDMA parameters, and propagation model for the reference system are listed in Table H.1, H.2, and H.3, respectively. Note that two sets of OFDMA parameters corresponding to 10 MHz channel bandwidth and 5 MHz channel bandwidth are provided in Table H.2. Only 10 MHz case is analyzed.

Table H.1: Mobile WiMAX System Parameters

Parameters	Value
Number of 3-Sector Cells	19
Operating Frequency	2500 MHz
Duplex	TDD
Channel Bandwidth	10 MHz
BS-to-BS Distance	2.8 km
Minimum Mobile-to-BS Distance	36 m
Antenna Pattern	70° (-3 dB) with 20 dB front-to-back ratio
BS Height	32 m
Mobile Terminal Height	1.5 m
BS Antenna Gain	15 dBi
SS Antenna Gain	-1 dBi
BS Maximum PA Power	43 dBm
Mobile Terminal Maximum PA Power	23 dBm
# of BS TX/RX Antenna	1/2/4
# of MT TX/RX Antenna	2
BS Noise Figure	4 dB
SS Noise Figure	7 dB

Table H.2: OFDMA Parameters

Parameters	Values Set 1	Values Set 2
System Channel Bandwidth (MHz)	10	5
Sampling Frequency (F_p in MHz)	11.2	5.6
FFT Size (N_{FFT})	1024	512
Sub-Carrier Frequency Spacing	10.9375 kHz	10.9375 kHz

Table H.2: OFDMA Parameters			
Parameters		Values Set 1	Values Set 2
Useful Symbol Time ($T_b = 1/f$)		91.4 us	91.4 us
Guard Time ($T_g = T_b/8$)		11.4 us	11.4 us
OFDMA Symbol Duration ($T_s = T_b + T_g$)		102.9 us	102.9 us
Frame duration		5 ms	5 ms
Number of OFDMA Symbols		48	48
DL PUSC	Null Sub-carriers	184	92
	Pilot Sub-carriers	120	60
	Data Sub-carriers	720	360
	Sub-channels	30	15
UL PUSC	Null Sub-carriers	184	92
	Pilot Sub-carriers	280	140
	Data Sub-carriers	560	280
	Sub-channels	35	17

Table H.3: Propagation Model	
Parameters	Value
Propagation Model	COST 231 Suburban
Log-Normal Shadowing SD (σs)	8 dB
BS shadowing correlation	0.5
Penetration Loss	10 dB

Tables H.4 and H.5 show the DL and UL link budgets for the above reference system. The value of 5.56 dB used for the Shadow Fade margin in the table assures a 75% coverage probability at the cell edge and 90% coverage probability over the entire area. The interference margin is 2 dB for DL and 3 dB for UL respectively assuming a frequency reuse 1/3/1. The interference margin can be reduced to 0.2 dB for a 1/3/3, reuse pattern but at the cost of reduced spectral efficiency. The Marco diversity gain is 4 dB assuming a 0.5 shadow fading correlation. The cell range can be estimated from the link budget using any one of a number of propagation models such as COST 231-Hata propagation model and Erceg-Greenstein model. The COST 231-Hata propagation model is based on empirical results in the 2 GHz band and tends to make very conservative prediction for 2.5 GHz. The Erceg-Greenstein model is another model commonly used in this frequency band and predicts ranges that are approximately 70% greater. Also note that the link budget corresponds to a DL cell-edge data rate of 4.32 Mbps and an UL cell-edge data rate of 102 Kbps; significantly higher values than the 3G systems. Higher cell-edge data and higher carrier frequency result in smaller cell size.

Table H.4: DL Link Budget for Mobile WiMAX

Mobile WiMAX Downlink				
Base Station Infrastructure	MAP	Traffic-PUSC		Units
Tx Power per Antenna Element	10.0	10.0	10.0	Watts
Number of Tx Antenna Element	2	2	2	
Cyclic Combining Gain	3.0	3.0	3.0	dB
Tx Antenna Gain	15.0	15.0	15.0	dBi
Pilot Power Boosting Gain	-0.7	-0.7	-0.7	dB
EIRP	57.3	57.3	57.3	dBm
Base Permutation Zone	PUSC	PUSC	PUSC	
Number of Occupied Sub-Carriers	840	840	840	
Power of Occupied Sub-Carrier	28.1	28.1	28.1	dBm
Mobile Unit, (Handset Indoor)				
Rx Antenna Gain	-1.0	-1.0	-1.0	dB
Rx Antenna Diversity Gain (2 Antenna)	3.0	3.0	3.0	dB
Rx Noise Figure	7.0	7.0	7.0	dB
Margins				
Log Normal Fade Margin	5.56	5.56	5.56	dB
Fast Fade Margin	6.0	6.0	6.0	dB
Interference Margin	2.0	2.0	2.0	dB
Penetration Loss	10.0	10.0	10.0	dB
Total Margin	23.56	23.56	23.56	dB
Mobile Rx Sensitivity				
Thermal Noise	-174	-174	-174	dBm/Htz
Sub-Carrier Spacing	10.94	10.94	10.94	kHz
Modulation	QPSK 1/8	QPSK 1/2	16QAM 1/2	
SNR Required	-3.31	3.49	8.93	dB
Delta from limiting cell range distance	0.82			
DL Traffic Data Rate		2.88	5.76	Mbps
Rx Sensitivity (per sub-carrier)	-129.9	-123.2	-117.7	dBm
Rx Sensitivity (composite)	-100.7	-93.9	-88.8	dBm
System Gain	-160.0	153.3	147.8	dB
Maximum Allowed Path Loss	136.4	133.7	128.2	dB

Table H.5: UL Link Budget for Mobile WiMAX

Mobile WiMAX Uplink				
Mobile Unit, (Handset Indoor)	Initial Ranging	FB Channel	Traffic Full Allocation	Units
Tx Power per Antenna Element	200	200	200	mw
Number of Tx Antenna Elements	1.0	1.0	1.0	
Tx Antenna Gain	-1.0	-1.0	-1.0	dBi

EIRP	22.0	22.0	22.0	dBm
Permutation Zone	**Ranging PUSC**	**FB Channel**	**PUSC**	
Available Sub-Carriers	144	70	840	
Allocated Sub-Channels		2.5	8	
Allocated Sub-Carriers	144	70	192	
Allocated Data Carriers	144	60	128	
Power per Occupied Sub-Carrier	0.43	3.56	-0.82	dBm
Base Station Infrastructure				
Rx Antenna Gain	15.0	15.0	15.0	dBi
Rx Antenna Diversity Gain	2.0	3.0	3.0	dBi
Rx Noise Figure	4.0	4.0	4.0	dB
Margins				
Log Normal Fade Margin	5.56	5.56	5.56	dB
Fast Fading Margin	4.0	4.0	2.0	dB
Interference Margin	3.0	3.0	3.0	dB
Penetration Loss	10.0	10.0	10.0	dB
Total Margin	**22.56**	**22.56**	**20.56**	**dB**
Base Station Rx Sensitivity				
Thermal Noise	-174	-174	-174	dBm/Hz
Sub-Carrier Spacing	10.94	10.94	10.94	kHz
Modulation Type	Ranging	Ranging	QPSK 1/8	
SNR Required	-6.0	-6.0	-2.5	dB
Delta from limiting cell range distance	0.17	0.66		
UL Traffic Data Rate			**102**	**kbps**
Rx Sensitivity (per sub-carrier)	-135.6	-135.6	-132.1	dBm
Rx Sensitivity (composite)	-114.0	-117.8	-111.1	dBm
System Gain	**153.1**	**157.2**	**149.3**	**dB**
Maximum Allowable Path Loss	**130.5**	**134.6**	**128.8**	**dB**

Annex I: NS2 Protocol Layer Modules

I.1 NS-2 Quick Overview

There From a user's perspective, NS is an Object-oriented Tcl (OTcl) interpreter that has a simulation event scheduler and network component object libraries, and network set up (plumbing) module libraries. To use NS, one programs in OTcl script language. To set up and run a simulation network, user should write an OTcl script that initiates an event scheduler, sets up the network topology using the network objects and the plumbing functions in the library, and tells traffic sources when to start and stop transmitting packets through the event scheduler. The term "plumbing" is used for a network set up, because setting up a network is plumbing possible data paths among network objects by setting the "neighbour" pointer of an object to the address of an appropriate object. When a user wants to make a new network object, he or she can easily make an object either by writing a new object or by making a compound object from the object library, and plumb the data path through the object.

I.2 Protocol Stack Modules available in NS-2

There are a number of protocol stack modules that are available in NS-2 that can be directly used for the WiMAX simulation effort. Since the development effort is primarily at the MAC and the PHY layers, the remaining parts of the protocol stack remain unchanged and can be directly over the developing lower layers. Some examples are given below. We note that these examples serve to give a quick overview of the broad simulation scenarios, are not comprehensive, and specific details are not delved into. Details on the implementation can be obtained from the NS-2 documentation and other online sources.

I.3 APPLICATION LAYER:

FTP: Inbuilt support exists in NS-2 to create and attach FTP agents to nodes (typically over TCP)

TELNET: inbuilt support in NS-2 to create and attach Telnet agents (typically over UDP)

CBR (Constant Bit Rate): Inbuilt support in NS-2 (typically used over UDP)

Exponential/Pareto ON-OFF/Poisson: Inbuilt support in NS-2 (typically over UDP)

Trace Driven: Support in NS-2 to specify a file in binary format, which specified inter-packet times, and packet size in bytes. Traffic corresponding to any desired type can be generated by using requisite models offline and then associating the binary file with the nodes in the simulation

We note that many additional application traffic sources have been developed for NS-2 over time and can be used in the WiMAX simulation effort, for example VoIP traffic models.

I.4 TRANSPORT LAYER

TCP: inbuilt support in NS-2 for TCP at the transport layer

UDP: Inbuilt support for NS-2 for UDP at the transport layer

I.5 NETWORK LAYER

There are a number of routing protocols available in NS-2, especially for routing in ad-hoc network scenarios, which can be directly used in WiMAX simulation scenarios. Examples include:

DSR: Dynamic Source Routing

AODV: Ad-Hoc On Demand Distance Vector Routing

DSDV: Destination Sequenced Distance Vector

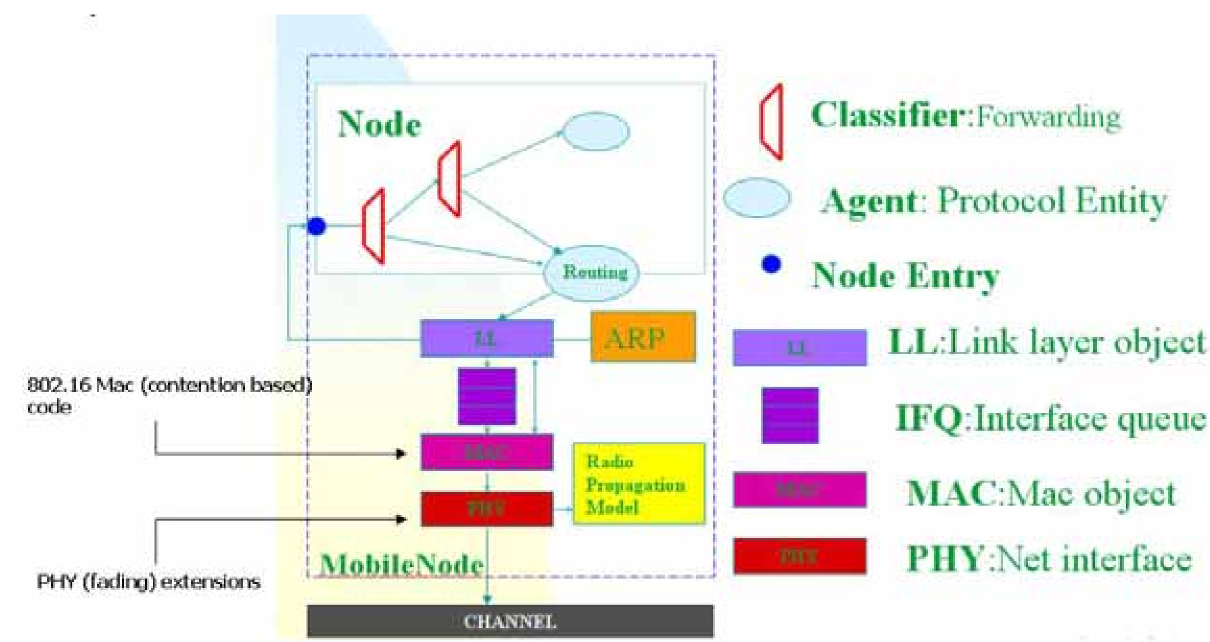
We do not focus much on the available routing protocols, since the current simulator emphasizes static scenarios, and as such, the routing layer can replaced by a trivial routing layer, with packet filtering occurring at the MAC layer. However, in the case of mobile simulation scenarios, it is likely that data sinks will be static, and some form of intelligent routing will be needed to get data from mobile sources of data to the sinks.

I.6 MAC LAYER

This is the most active area of development in the WiMAX simulation effort (together with some PHY layer abstractions to interface with BUPT). Like other MAC layer protocols available in NS-2 (for example, 802.11) the objective of the development at the MAC layer is to provide to the user the capability of using the 802.16 MAC using similar Otcl interfaces. Backward compatibility (for example, usage of 802.11 and such existing MACs in NS-2) may cease to be an option, due to additional features added at the PHY layer.

I.7 PHY LAYER

As mentioned earlier, the NS-2 PHY layer is being actively developed in concurrence with the MAC layer with the objective of creating a realistic PHY abstraction that can interface with BLER tables. More details on the specific PHY features and rationale are developed in the next section. This will serve as a description of the status quo in the NS-2 PHY.



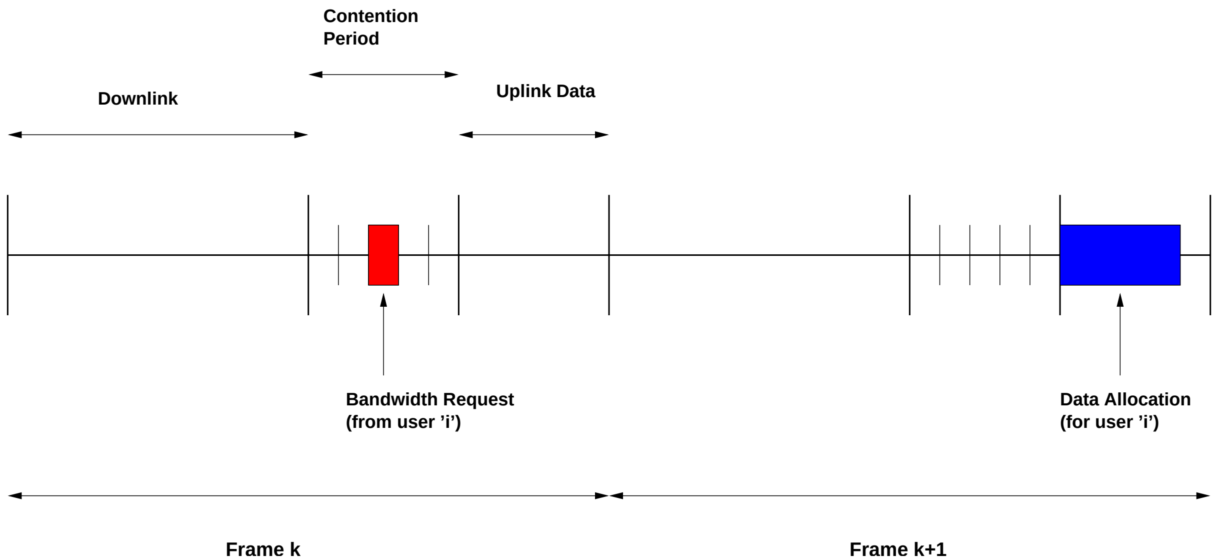
The picture above serves to illustrate the specific areas of development in the NS-2 mobile node structure.

I.8 NS2 FRAMWORK COMMON MODULES

This section describes the current status of the NS-2 WiMAX simulation effort. Broadly, the following features are available in NS-2. Note that the term feature is used flexibly and can refer to simulation scenarios as listed below.

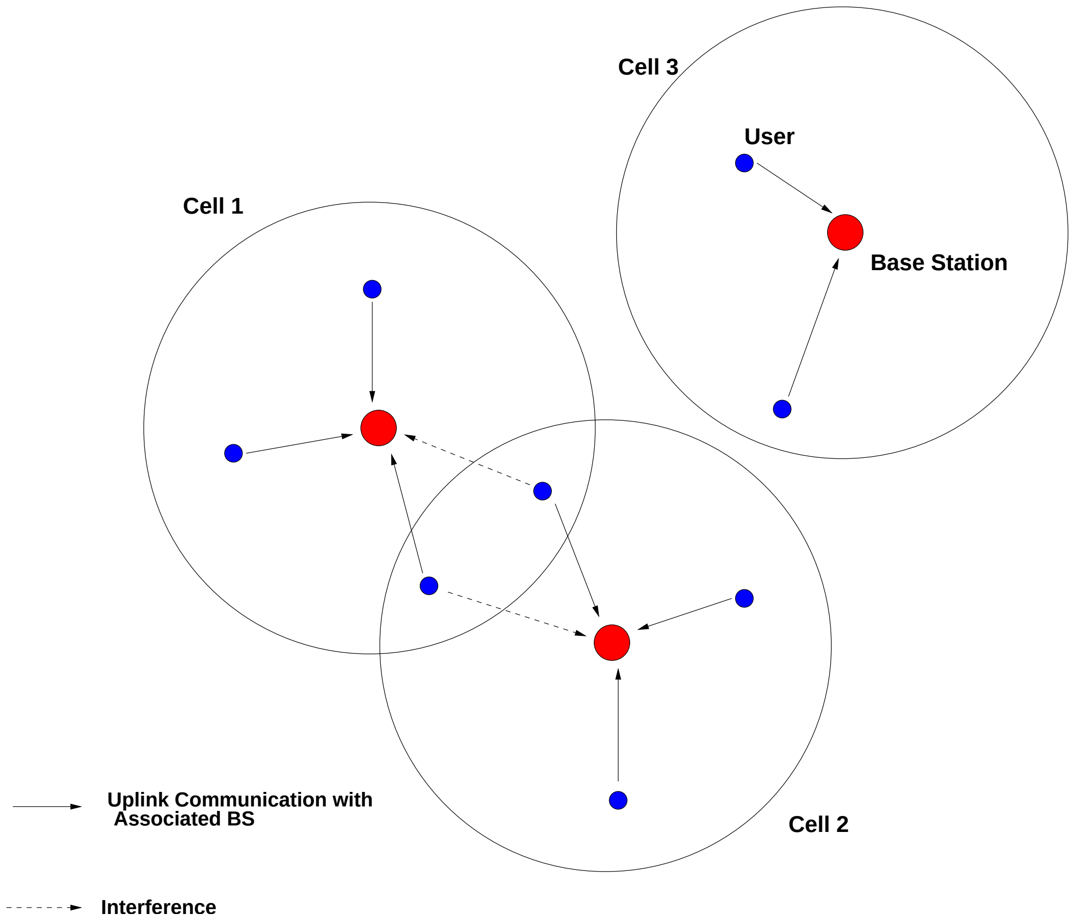
All of the current development is over a single channel, maintaining the inherent packet based simulation operation of NS-2.

Contention based operation (for Bandwidth Request)



Nodes can contend in a contention period defined by the Base Station during the Uplink Duration. The duration of contention is flexible on a frame-to-frame basis.

Multiple Cell simulation capability

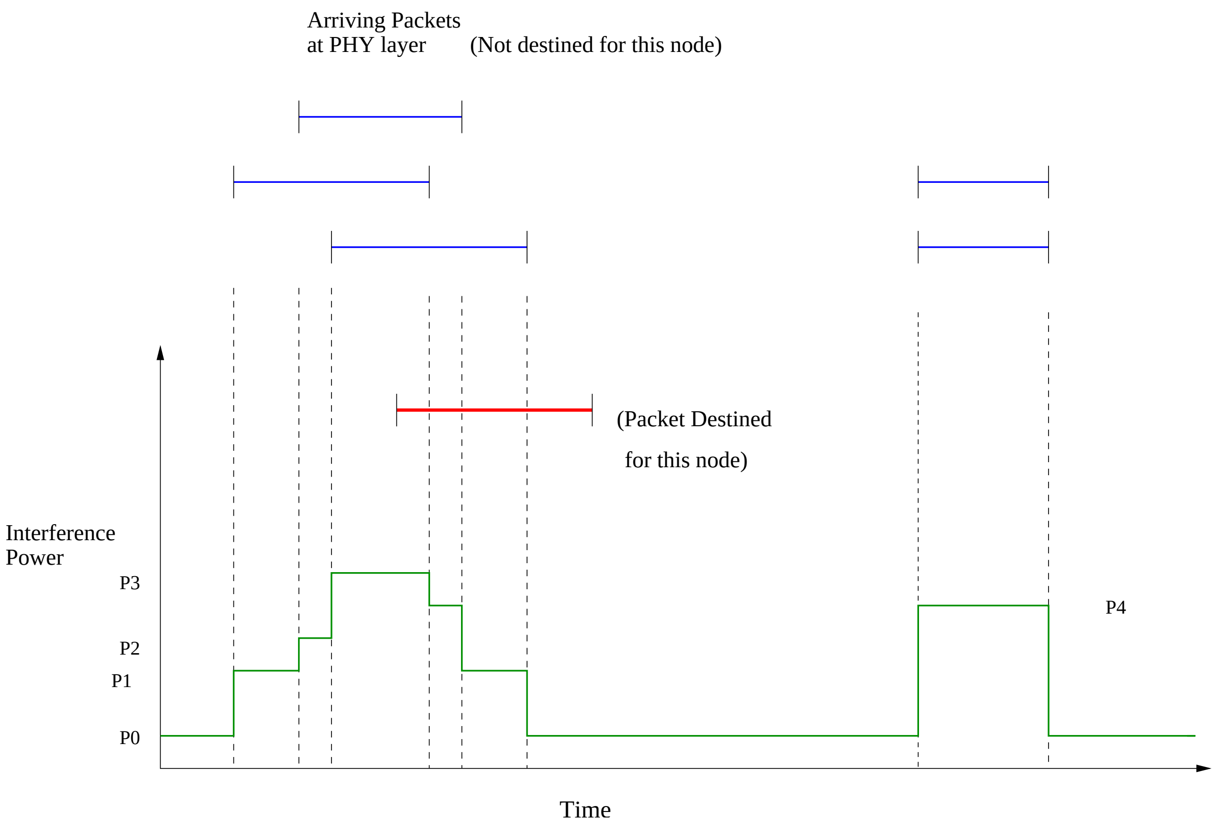


Support exits for the simulation of multiple cells, each with Base Station and associated users. This feature represents progress towards a 19-cell simulation scenario.

Combined Fast fading and shadowing extensions for radio propagation

Existing simplistic (deterministic) channel propagation model in NS-2 is extended to incorporate stochastic fading effects at small and large time scales.

Interference state maintenance for SINR based data decoding decisions



This feature represents progress towards simulation results that will incorporate explicit interference effects and their impact on system performance.

Given the current dependence on single channel and packet-based operation, the immediate focus is on relaxing these assumptions. Concurrently, code is also being re-organized to be aligned with the recommendations of the simulation software architecture document.

Annex J: List of Known Simulation Models of WiMAX Networks

WiMAX Forum is developing an NS-2 model of WiMAX networks. This model follows the methodology specified in this document and has been verified and discussed by various experts members of WiMAX member community. Release 1 of WiMAX Forum NS2 model is available to WiMAX Forum members. In addition, we include here a list of other known models.

J.1 NS2 Models

1 WiMAX Forum's NS2 Model, Release 1,
http://www.ecse.rpi.edu/Homepages/shivkuma/wimax/
(obtain password from Dr. Shyam Parekh, sparekh@Alcatel-Lucent.com)

2 Chang Gung University, Taiwan, WiMAX NS2 Simulator,
http://www.csie.cgu.edu.tw/~jhchen/

3 National Institute of Standards and Technologies (NIST), NIST IEEE 802.16 NS2 Simulator,
http://w3.antd.nist.gov/seamlessandsecure/files/80216/doc/INSTALL.TXT

J.2 Commercial Simulation Packages:

1. CoWare, Signal Processing Designer (SPD), www.coware.com
2. Opnet Technologies, www.opnet.com/wimax
3. Scalable Networks Technologies, Qualnet, www.qualnet.com

J.3 MATLAB Libraries

1 Mathworks, MATLAB Central File exchange,
http://www.mathworks.com/matlabcentral/fileexchange/loadCategory.do (Search for OFDM, MIMO, Raleigh, QAM, etc.)

J.4 FPGA Based Simulation Libraries

1 Altera, Scalable OFDMA Engine for Mobile WiMAX (MATLAB based Model),
http://www.altera.com/literature/an/an412.pdf

2 Cambridge Consultants, MATLAB-based WiMAX OFDM256 PHY reference design,
http://www.cambridgeconsultants.com/prop_ws_index.shtml

3 Xilinx, WiMAX OFDM Library Reference Design,
http://www.altera.com/literature/an/an412.pdf

Access Systems Corrigendum 2, Draft 4", May 2007

[Yagoobi04] Hassan Yagoobi, "Scalable OFDMA Physical Layer in IEEE 802.16 WirelessMAN", Intel Technology Journal, Vol 08, August 2004.

[WiMAX Arch2 2005] "WiMAX End-to-End Network Systems Architecture - Stage 2: Architecture Tenets, Reference Model and Reference Points," WiMAX Forum, December, 2005.

[WiMAX NWG Stage 3 Rel 1.0.0.] "WiMAX End-to-End Network Systems Architecture - Stage 3: Detailed Protocols and Procedures" WiMAX Forum, Mar 2007.

[WiMAX Comp2 2006] "Mobile WiMAX - Part II: Competitive Analysis", WiMAX Forum, February, 2006

[Cimini85] L.J. Cimini, "Analysis and Simulation of a Digital Mobile Channel Using Orthogonal Frequency Division Multiplexing," IEEE Trans. Comm., vol. COM-33, no. 7, pp 665-675, June 1985.

[VanNee00] Richard Van Nee and Ramjee Prasad, "OFDM for Wireless Multimedia Communications," Artech House, 2000.

[Xiao02] W. Xiao and R. Ratasuk, "Analysis of Hybrid ARQ with Link Adaptation", Proceedings of the Annual Allerton Conference on Communications, Control, and Computing, pp. 1618-1619, Oct 2002.

[Nair04] G. Nair, J. Chou, T. Madejski, K. Perycz, P. Putzolu and J. Sydir "IEEE 802.16 Medium Access Control and Service Provisioning", Intel Technology Journal, vol 08, August 2004.

[Wang05] F. Wang, A. Ghosh, R. Love, K. Stewart et.al., "IEEE 802.16e System Performance - Analysis and Simulation Results", Proc. of PIMRC, Berlin, Germany, Sept. 2005.

[DUplessis05] Philippe Duplessis, "HSOPA: Exploiting OFDM and MIMO to take UMTS beyond HSDPA/HSUPA" , Nortel Technical Journal, Issue 2, July 2005.

[Hoadley05] John Hoadley and Al Javed, "Overview: Technology Innovation for Wireless Broadband Access", Nortel Technical Journal, Issue 2, July 2005.

[Liva96] John Liva and Titus Kwok-Yeung Lo, "Digital Beamforming in Wireless Communications," Artech House Publishers, 1996.

[Alamouti98] S.M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," IEEE Journal on Selected Areas in Communications, vol. 16, pp 1451-1458, October 1998.

[Tarokh99] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "Space-time Block Codes from Orthogonal Designs," IEEE Transactions on Information Theory, vol. 45, pp. 1456-1467, July 1999.

[Foschini96] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multielement Antennas," Bell Labs Tech. J. pp. 41-59, Autumn 1996.

[Foschini99] G. J. Foschini, G.D. Golden, P.W. Wolniansky and R.A. Valenzuela, "Simplified Processing for Wireless Communication at High Spectral Efficiency," IEEE. Journal on Selected Areas in Communications, vol. 17, pp. 1841-1852, 1999.

[Salvekar04] A. Salvekar, S. Sandhu, Q. Li, M. Vuong and X. Qian "Multiple-Antenna Technology in WiMAX Systems," Intel Technology Journal, vol 08, Augus 2004.

[C.R1002-0] 3GPP2 C.R1002-0, CDMA2000 Evaluation Methodology, December 2004.

[3GPP R1-03-1303] 3GPP TSG-RAN-1, System-Level evaluation of OFDM - further Considerations", R1-03-1303, November 17-21, 2003.

[3GPP R1-03-1370] 3GPP TSG-RAN-1, "Effective SIR Computation for OFDM System-Level Simulations," R1-03-1370, Meeting #35, Lisbon, Portugal, November 2003.

[Yin06] Hujun Yin and Siavash Alamouti, "OFDMA - A Broadband Wireless Access Technology," IEEE Proc. of Sarnoff Symposium, March 2006.

[WiMAX App 2005] "Can WiMAX Address Your Applications?", Westech on Behalf of the WiMAX Forum, October 24, 2005

[Jain91] Raj Jain, "The Art of Computer Systems Performance Analysis:Techniques for Experimental Design, Measurement, Simulation, and Modeling," Wiley-Interscience, 1991.

[WiMAX Forum Mobile System Profile Rel 1.0] WiMAX Forum mobile system profile release 1.0 approved specification, revision 1.4.0:2007-05-02

[White Paper Part I] WiMAX Forum, "Mobile WiMAX – Part I: A technical Overview and Performance Evaluation", March 2006.