Shahid Mumtaz

Performance Analysis of 4G Wireless Networks Using System Level Simulator

Universidade de Aveiro Departamento de Electrónica, Telecomunicações e
Ano 2011 Informática

Shahid Mumtaz

Performance analysis of 4G wireless networks using system level simulator

To my Parents, my family and my son Gaberial Noor.

o júri

presidente
                                        Doutor Luís António Ferreira Martins Dias Carlos
Professor Catedrático da Universidade de Aveiro

Doutor Francisco António Bucho Cercas
Professor Associado com Agregação do ISCTE - Instituto Universitário de Lisboa

Doutor Manuel Alberto Pereira Ricardo
Professor Associado da Universidade do Porto

Doutor Atílio Manuel da Silva Gameiro
Professor Associado da Universidade de Aveiro

Doutor Fernando José da Silva Velez
Professor Auxiliar da Universidade da Beira Interior

Doutor Ramiro Samano-Robles
Investigador de Pós-Doutoramento do Instituto de Telecomunicações

Acknowledgement

Abstract

In the last decade, mobile wireless communications have witnessed an explosive growth in the user's penetration rate and their widespread deployment around the globe. In particular, a research topic of particular relevance in telecommunications nowadays is related to the design and implementation of mobile communication systems of 4th generation (4G). 4G networks will be characterized by the support of multiple radio access technologies in a core network fully compliant with the Internet Protocol (all IP paradigms). Such networks will sustain the stringent quality of service (QoS) requirements and the expected high data rates from the type of multimedia applications (i.e. YouTube and Skype) to be available in the near future. Therefore, 4G wireless communications system will be of paramount importance on the development of the information society in the near future.

As 4G wireless services will continue to increase, this will put more and more pressure on the spectrum availability. There is a worldwide recognition that methods of spectrum managements have reached their limit and are no longer optimal, therefore new paradigms must be sought. Studies show that most of the assigned spectrum is under-utilized, thus the problem in most cases is inefficient spectrum management rather spectrum shortage. There are currently trends towards a more liberalized approach of spectrum management, which are tightly linked to what is commonly termed as Cognitive Radio (CR).

Furthermore, conventional deployment of 4G wireless systems (one BS in cell and mobile deploy around it) are known to have problems in providing fairness (users closer to the BS are more benefited relatively to the cell edge users) and in covering some zones affected by shadowing, therefore the use of relays has been proposed as a solution.

To evaluate and analyse the performances of 4G wireless systems software tools are normally used. Software tools have become more and more mature in recent years and their need to provide a high level evaluation of proposed algorithms and protocols is now more important. The system level simulation (SLS) tools provide a fundamental and flexible way to test all the envisioned algorithms and protocols under realistic conditions, without the need to deal with the problems of live networks or reduced scope prototypes. Furthermore, the tools allow network designers a rapid collection of a wide range of performance metrics that are useful for the analysis and optimization of different algorithms.

This dissertation proposes the design and implementation of conventional system level simulator (SLS), which afterwards enhances for the 4G wireless technologies namely cognitive Radios (IEEE802.22) and Relays (IEEE802.16j). SLS is then used for the analysis of proposed algorithms and protocols.

Table of Contents

List of Figures

List of Tables

Abbreviations

16QAM	16 Quadrature Amplitude Modulation
1xEV-DO	1 x Evolution Data Only
2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
3GPP2	Third Generation Partnership Project 2
4G	Fourth Generation
64QAM	64 Quadrature Amplitude Modulation
AAS	Advanced Antenna Systems
ACK	Acknowledge
ACM	Adaptive Coding and Modulation
ADC	Additional Delay introduced by Cooperation
ADSL	Asynchronous Digital Subscriber Line
AES	Advanced Encryption Standard
AF	Amplify and Forward
AGC	Automatic Gain Control
AMC	Adaptive Modulation and Coding
AoA	Angle of Arrival
AoD	Angle of Departure
AP	Access Point
ARPU	Average Revenue Per User
ARQ	Automatic Repeat Request
AS	Angular Spread
ATM	Asynchronous Transfer Mode
AUT	Average User-Throughput
AUTC	Average User-Throughput by Cooperation
AvgCI	Average C/I
AWGN	Additive White Gaussian Noise
B3G	Beyond Third Generation
BE	Best Effort
BER	Bit Error Rate
BLER	Block Error Rate
BLERIF	BER Improvement Factor

BPSK	Binary Phase Shift Keying
BS	Base Station
BWA	Broadband Wireless Access
CAC	Connection Admission Control
CAPEX	Capital Expenditures
CAT	Cell Average Throughput
CBR	Constant Bit Rate
CC	Convolutional Encoder
CCCI	Combined Coverage and Capacity Indicator
CDMA	Code Division Multiple Access
CDMA 2000	Code Division Multiple Access 2000
CP	Cyclic Prefix
CPS	Common Part Sub-layer
CQI	Channel Quality Information
CQICH	Channel Quality Information Channel
EESM	Exponential Effective SINR Mapping
ELN	Explicit Loss Notification
ertPS	Enhanced Real Time Pooling Service
ETSI	European Telecommunications Standard Institute
EXP	Exponential scheduler
FCH	Frame Control Header
FER	Frame Erasure Rate
FFT	Fast Fourier Transform
FH	Frequency Hopping
FIFO	First In First Out
FTI	Fairness Throughput Indicator
FTIC	FTI by Cooperation
FTP	File Transfer Protocol
FUSC	Full Usage Sub-Carrier sub-channelization
HARQ	Hybrid Automatic Repeat Request
ICI	Inter-carrier interference
IE	Information Element
IETF	Internet Engineering Task Force
IFFT	Inverse Fast Fourier Transform
IP	Internet Protocol
IRR	Internal Rate of Return

IS-95	International Standard 95
ISD	Inter Site Distance
ISI	Inter Symbol Interference
ISO	International Standards Organization
ITU	International Telecommunications Union
LA	Link Adaptation
LAN	Local Area Network
LDCP	Low-Density Parity-Check
LUT	Look up Tables
MCS	Modulation and Coding Scheme
MIMO	Multiple Input Multiple Output
MISO	Multiple-Input Single Output
ML	Maximum Likelihood
MLD	Maximum Likelihood Decoder
M-LWDF	Modified Largest Weighted Delay First scheduler
MMR	Mobile Multi-hop Relay
MMSE	Minimum Mean Square Error
MPDU	Medium Access Control layer Protocol Data Unit
MRC	Maximum Ratio Combiner
MSC	Multi Source Cooperation
MU-MIMO	Multiple User Multiple Input Multiple Output
NACK	Negative Acknowledge
NAF	Non-Orthogonal Amplify-and-Forward
NCSW	Node Cooperative Stop and Wait
NGN	Next Generation Network
NLOS	Non Line of Sight
OFDMA	Orthogonal Frequency Division Multiple Access
OFDM	OFDM Multiple Access
OM	Opportunistic Multipath
OMC	Operations and Maintenance Centre
OPEX	Operational Expenditures
OS	Opportunistic Scheduling
OSI	Open Systems Interconnection
PAF	Packet Availability Factor
PAPR	Peak to Average Power Ratio
PDU	Packet Data Unit

PEP	pair-wise error probability
PER	Packet Error Rate
PERIF	PER Improvement Factor
PF	Proportional Fairness scheduler
QoS	Quality of Service
QPSK	Quaternary Phase Shift Keying
RA	Random Access
RAM	Resource Allocation Map
RAN	Radio Access Network
RAT	Radio Access Technology
RAU	Resource Allocation Unit
RB	Resource Block
R-BLER	Residual Block Error Rate
REC	Relay Enhanced Cell
RF	Radio Frequency
RN	Relay Node
RRM	Radio Resource Management
RT	Real Time
RTG	Receive to Transmit Gap
rtPS	Real Time Polling Service
RU	Resource Unit
S(I)NR	Signal to (Interference plus) Noise Ratio
SAF	Slotted Amplify-and-Forward
SAP	Service Access Point
SINR	Signal to Interference plus Noise Ratio
SISO	Single Input Single Output
SLS	System Level Simulator
SM	Spatial Multiplexing
SMTP	Simple Mail Transfer Protocol
SNR	Signal to Noise Ratio
SOFDMA	Scalable OFDMA
TDMA	Time Division Multiple Access
TD-SCDMA	Time Division Synchronous Code Division Multiple Access
TTG	Transition to Transmit Gap
TTI	Transmission Time Interval
UAT	User Average Throughput (provided by the direct source-destination

	link - no cooperation involved)
WCDMA	Wideband Code Division Multiple Access
WiFi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwaves Access
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Network
WP	Work Package
WSSUS	Wide-Sense Stationary Uncorrelated Scattering
WWW	World Wide Web
ZF	Zero Forcing

Chapter 1: Introduction

The purpose of this chapter is to provide the motivation behind this dissersation.We begin the chapter by summarizing the cellular background and introducing Cooperative and Cognitive wireless systems, which will be discussed throughout this dissertation. Next, we discuss the need for system level simulator and the motivation behind system level simulator. Afterward the objectives of this dissertation are presented. The chapter ends with the brief presentation of the organization of the dissertation and research contributions.

1.1   Background

Wireless networks can be classified into Wireless Local Area Networks (Wireless LANs, WLANs) with limited geographical coverage and cellular networks, which cover a large geographical area. For Wireless LANs, the Institute of Electrical and Electronics Engineers (IEEE) is the dominating standardization body [1], and the IEEE 802.11 standard family the prevailing technology. The most prominent active standard currently is IEEE 802.11g, which provides net data rates of up to 54 MBit/s with coverage of a few ten meters. Recent developments will push this to up to 248 MBit/s in the upcoming IEEE 802.11n standard with the help of Multiple Input Multiple Output (MIMO) techniques. Wireless LAN networks are designed to cover small areas also known as hotspots, such as local networks on private properties, on company premises, or in public places such as airports or cafes. This is further supported by a mobility mechanism on layer 2 that allows setting up a network consisting of multiple access points.

In contrast, cellular networks cover a large geographic area, such as a country. They divide the covered area into cells, each served by a base station transceiver. The size of a cell depends on the underlying technology, its surrounding terrain, and the expected traffic demand. In rural areas with a low traffic density, the diameter of a cell may be in the order of tens of kilometers. In urban areas with a high traffic density, the diameter may be as small as a few hundred meters. Small cells allow for higher data rates within the cells. Moreover, a large number of cells balance the offered traffic over many transceivers allowing for a higher network capacity compared to large cells. The base stations are part of the Radio Access Network (RAN), and they are connected to a wide area backbone network referred to as core network. Cellular networks usually offer a sophisticated mobility management, making them suitable for covering large areas. Many technologies co-exist in the area of cellular networks. The 2nd generation Global System for Mobile communications (GSM) is certainly the most successful system to date. In September 2007, there were over 2.5 billion GSM subscribers, which made up 85% of the world's mobile market. The GSM standards are currently maintained by the 3rd Generation Partnership Project (3GPP). Besides the 3GPP, other standardization bodies have been active in specifying cellular networks, such as the IEEE

with the IEEE 802.16e standard. Both the 3GPP and the IEEE are currently working on new standards based on Orthogonal Frequency Division Multiple Access (OFDMA), namely the 3GPP Long Term Evolution (LTE), and the IEEE 802.16m. The latter is the successor of IEEE 802.16e, another system that was already based on OFDMA. Recently, due to the explosion of interests in ad hoc networks and the research activity on the fourth generation (4G) wireless communication systems, the concept of cooperative communication brings to the mainstream of research in mobile communication academy and industry [2,3]. It is unlikely that 4G mobile systems can cover the same service area without additional infrastructure. The problem of path loss will prohibit the base station from communicating with users far away [4]. One way to overcome those problems is to use relays to extend the coverage area in cellular networks. Relays are introduced as a method to solve the problems.Cooperative communication involves two main ideas: (i) Use relays (or multi-hop) to provide spatial diversity in a fading environment, (ii) Envision a collaborative scheme where the relay also has its own information to send so both terminals help one another to communicate by acting as relays for each other (called "partners").

As 4G wireless services will continue to increase, this will put more and more pressure on the spectrum availability.There are currently trends towards a more liberalized approach of spectrum management, which are tightly linked to what is commonly termed as Cognitive Radio (CR). A cognitive radio is a radio that can change its transmitter parameters based on interaction with the environment where it operates [5], and additionally relevant here is the radio's ability to look for, and intelligently assign spectrum 'holes' on a dynamic basis from within primarily assigned spectral allocations. Spectrum is shared explicitly in one of three ways.

Opportunistic: where spectrum is used whenever the licensee does not use it.

Cooperative: where frequencies are allocated centrally based on real-time negotiation with the license.

Mixed: where sharing is cooperative when possible and opportunistically otherwise.

This is motivated by the fact that UMTS radio frequency spectrum has become, in a significant number of countries, a very expensive commodity, and therefore the opportunistic use of these bands could be one way for the owners of the licenses to make extra revenue.

1.2   Need for Simulations

The last decades have seen the proliferation of advances in software and computer processors that allow researchers to perform a wide variety of complex simulations on different types of systems. The approach consists of replicating all the relevant stochastic processes that occur in the system, using random number generators, and the management, processing or operations performed over such random variables. The results of such operations, processing and management algorithms are

collected over time and serve as performance metrics of the system. Due to this flexible approach, simulators are today a crucial element in research and industrial environments.

Computer simulations of cellular networks can be classified with respect to their level of abstraction and with respect to their simulation type [6,7]. Simulations in wireless and cellular networks can be classified in three categories with respect to the targeted layers in the protocol stack and with respect to the considered level of abstraction.

1.2.1   Physical / link level simulations

Physical level simulations, also known as link level simulations, are used for optimization and evaluation of physical layer properties and algorithms, such as symbol constellations or error correction codes. This requires a detailed modeling of the wireless link and a simulation at the symbol level. Their time scale of operation is therefore the baud rate of the system. Results are often Bit Error Ratios (BER) or Block Error Ratios (BLER), where the BLER may refer, for example to FEC or ARQ blocks. Due to the involved computational complexity, it is difficult to extend link level simulations to higher protocol layers, whose effects are on longer time scales.

1.2.2   Network level simulations

In order to simulate protocol layers above the physical layer, network level simulations usually use an abstract physical layer model. They often comprise the MAC layer and the IP layer, but may also extend to transport layer protocols or even applications. Subject to investigation may be ARQ performance, scheduling, queue management, protocol interactions, and the like. The physical layer abstraction makes necessary an appropriate interface. Effects from other cells are often neglected or represented by statistical models.

1.2.3   System level simulations

While physical and link level simulations commonly target only one or a few transmission links within a cell, system level simulations cover several cell sectors and transceivers within a certain geographic area. Classically, they use a simple model of the physical layer and abstract the effect of MAC protocols. System level simulations are used to evaluate system level aspects, such as network capacity and coverage, handover algorithms, admission control, and the like. The choice of the right simulator category depends on the investigated problem and the desired results. Usually, it is advantageous to decouple simulators working at different levels of abstraction in order to reduce complexity. In this case, appropriate interfaces are required for the interface between physical and network level simulations. However, some problems may require the combination of different simulation levels in one simulator. Recently, network and system level simulations have grown together, owing to the increasing complexity of MAC protocols and algorithms, which more and more interact with adjacent cells, or depend on adjacent cells in their

performance. Examples include interference coordination algorithms or certain aspects of scheduling algorithms. The basic differences between two simulators are shown in Table 1-1.

System level simulation (SLS) tool provide a virtual and flexible way to test all the envisioned algorithms under realistic conditions, without the need to deal with the problems of live networks or reduced scope prototypes. Additionally, the tools allow network designers a more or less rapid collection of a wide range of performance metrics that are useful for the analysis and optimization of different algorithms.
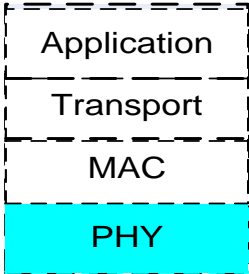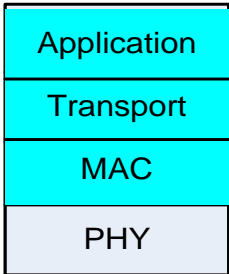
Link Level	System Level
Goal: Study different signal transmission and reception schemes	Goal: Application level performance.
Single Link	Multiple Users
Single Cell	Multiple Cells
Single Base Station	Multiple Base Stations
Emphasis on PHY	Emphasis on all Layers $\Rightarrow$ PHY abstracted

Table 1-1: Link and System Level

1.3 Motivation for System Level Simulator

Due to the complexity of wireless environments and modern radio access technologies, a SLS tool envisioned for their characterization must be extremely accurate as any slight deviation from real parameters might lead to inaccurate evaluation results. Therefore, accurate channel and propagation modeling tools, realistic traffic generators, tailored simulation and deployment scenarios, and calibrated results are required. Additionally, even when software tools have experienced a rapid development in recent years, certain processes of the network cannot be included in a dynamic fashion, but rather in a static and off-line manner. The reason behind this approach is that certain processes occur at completely different time scales, thereby being inefficient to simulate them in the same platform as this would lead to exaggerated simulation times. Therefore, the approach is to use alternative simulation tools that address processes at a very small time-scale, mainly at the

symbol and block level, and then import the results into the main system level simulator in a static way by means of a look up table or other relevant mathematical abstraction model.

We divide the SLS into three parts. According to the *Carrier's* view, capacity planning, performance optimization and operational guidelines are important factors in the modeling of SLS. From the *User's* prospective, operational guidelines are an important factor in modelling of SLS and according to *Vendor's* prospective, performance analysis of various features of SLS is a vital factor.

Typically, link level simulators (LLS) are used to model the link between the base and the mobile station, while SLS model the entire network. To most accurately predict the performance of cellular networks, a SLS, which includes the performance of link between each base and the mobile station, should be used. Unfortunately, the computational complexity of such simulators makes this impractical. An alternative technique consists of separately simulating the link and system components with an interface or mapping to combine the results. Therfore, the radio network simulations can be divided into two parts: LLS and SLS. The LLS is needed to build such a model for the system simulator, which is able to predict the receiver FER/BER performance, taking into account channel estimation, interleaving and decoding. The SLS is needed to model a system with a large number of mobile terminals, base stations, and algorithms operating in such a system. System performance evaluation of a given mobile wireless access technology requires simulations that carefully captures the dynamics of a multipath fading environment, the architecture of the air interface, mobile stations behaviour (in terms of mobility) and the types of applications used (properly defined traffic models). Also, since system level results depend intrinsically on the simulation scenario (propagation and interference environments, number and distribution of users within the cells), it is important that the assumptions and parameters used in the analysis are reported jointly with performance results. This procedure is fundamental for the correct validation of the results obtained and for benchmarking against other proposed scenarios for different wireless systems.

Cooperative diversity [8] schemes have demonstrated attractive features at the physical, link and medium access control layers. However, at such low layers, channel models and assumptions are, in general, simplistic, as the purpose is only to see whether the proposed schemes provide gains or not. Therefore, there is a large gap regarding the study of cooperative diversity schemes under more realistic scenarios.Other, issues such as estimation errors, capacity losses owing to signaling channels, and increased power and interference levels due to mobile and relay nodes, among many others should be considered. Additionally, modern networks cannot be considered as the simple collection of many independent links or users. On the contrary, real life networks, to which the

final algorithms are aimed at, consist of a large number of users, geographically dispersed base stations that interfere with each other, complex mobility patterns, different propagation models with complex probability distributions, and advanced radio resource management algorithms that take decisions about allocation based on measurements that are subject to estimation errors. Therefore, the final validation of the new cooperative diversity schemes should be made under much more realistic system level conditions, where it is probable that a very good algorithm at the physical layer does not have a good performance at the system level, or conversely, a physical layer scheme which is initially not very promising in combination with adequate management algorithms may provide attractive performance gains

There are currently trends towards a more liberalized approach of spectrum management, which are tightly linked to what is commonly termed as Cognitive Radio (CR). The development of such systems would include the sharing of the spectrum by several systems, which could for example include the coexistence of a cellular network with a CR. In order to propose such scenario, it is necessary to evaluate the implications of the coexistence on the performance of the different systems to ensure that the required QoS can still be met. A simulation tool is commonly used for the research and evaluation of CR networks. Today no simulation tool can readily handle the simulation of the coexistence of two or more systems. A brute-force solution, which simply combines two systems together, will not suffice due to computational complexity induced and difficulty to extend beyond two systems.

From another perspective, the last decades have seen the proliferation of advanced software and computer processors that allow researchers to perform a wide variety of complex simulations about different types of systems. The approach consists of replicating all the relevant stochastic processes that occur in the system, using random number generators, and the management, processing or operations performed over such random variables. The results of such operations, processing and management algorithms are collected over time and serve as performance metrics of the system. Due to this flexible approach, simulators are today a crucial element in research and industrial environments. Given the maturity of software related tools and the need to provide a high level evaluation of their proposed algorithms, wireless network designers commonly opt to use advanced software tools that aim to simulate in a very precise and accurate way all the relevant physical and radio resource management events of a network. These system level simulation tools, therefore, provide a virtual and flexible way to test all the envisioned algorithms under realistic conditions, without the need to deal with the problems of live networks or reduced scope prototypes. Additionally, the tools allow network designers a more or less rapid collection of a wide range of performance metrics that are useful for the analysis and optimization of different algorithms.

1.3.1 Objective

This dissertation has three main objectives:

1. Design of System Level Simulator (SLS).

2. Uses of System Level Simulator to analyses the algorithms and protocols.

3. Proposal of algorithms.

Traditionally, cellular radio access system is defined as the set of geographical areas called cells. In each cell there is one base station and a random number of mobiles, randomly located inside the cell. Mobiles access the network through Base station (BS's) as shown in Figure 1-1.To evaluate the performance of such system, a comprehensive system level tool is needed, which captures every aspect of the real cellular environment. To test the correctness of SLS, different algorithms and protocols are tested under various QoS parameters and their capacity/coverage for a given cellular environment is evaluated. Implementation of such a system level tool requires complex modeling of several touching issues. Main aspects to be considered in the modeling of the SLS are:

- Deployment scenarios
- Propagation scenarios
- Mobility scenarios
- Traffic scenarios
- Transmission/reception techniques
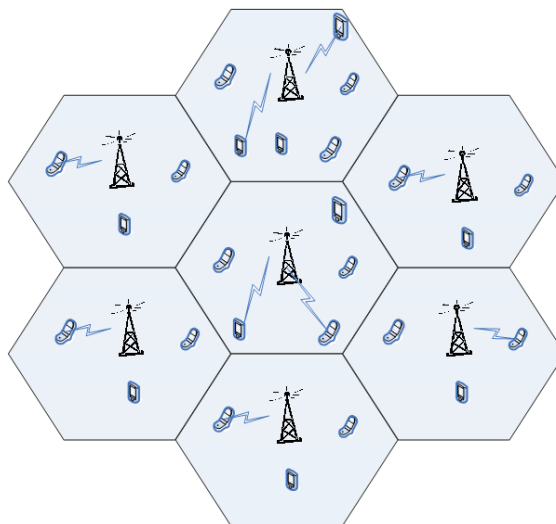
Figure 1-1: Conventional cellular system

1.4   Thesis Outline

The organization of dissertation is shown in Figure 1-2.

Chapter 2: describes the IEEE 802.16 standards for both Fixed WiMAX networks and Mobile WiMAX networks. The issues, which are of particular importance for the implementation of the System level simulator and the peculiarities of the PHY, and MAC layers, which are considered in the implementation of the System Level Simulator used for the realization of system level simulations, are provided. The overview of IEEE 802.16j and IEEE 802.22 are also given.

Chapter 3: describes the basics of system level simulation methodologies which are available in the literature, i.e. simulation type, cellular scenario, optimization techniques of cellular scenario and channel models

Chapter 4: provides all details followed into the design, modeling and implementation of the Link level and System Level Platform used in the realization of system level simulations for the Mobile WiMAX standard. Used traffic models for data generation as well as the channel models are provided. The cellular layout architecture used into the simulator is described. Of particular interest in the realization of system level simulations is the definition of the proper link to system level interface and the definition of the proper set of look up tables used in the mapping of physical layer performance. The procedure followed into the derivation of the Signal to Interference plus Noise Ratio (SINR) and the mapping function used to map the vector of SINRs into a single scalar to be inputted into the look-up tables is detailed.

Chapter 5: describes the enhanced version of SLS for the IEEE 802.22 and explains how opportunistic radios (ORs) can be used in an ad-hoc manner and further presents the new routing metrics which overall decrease the system interface from the base station.

Chapter 6: describes the enhanced version of SLS for the IEEE 802.16j.This relay technology is based on relay node deployment in a cellular system which overall increases the capacity of the system as compared to a conventional system. New relay selection algorithms are also defined.

Chapter 7: describes the conclusions and future work.
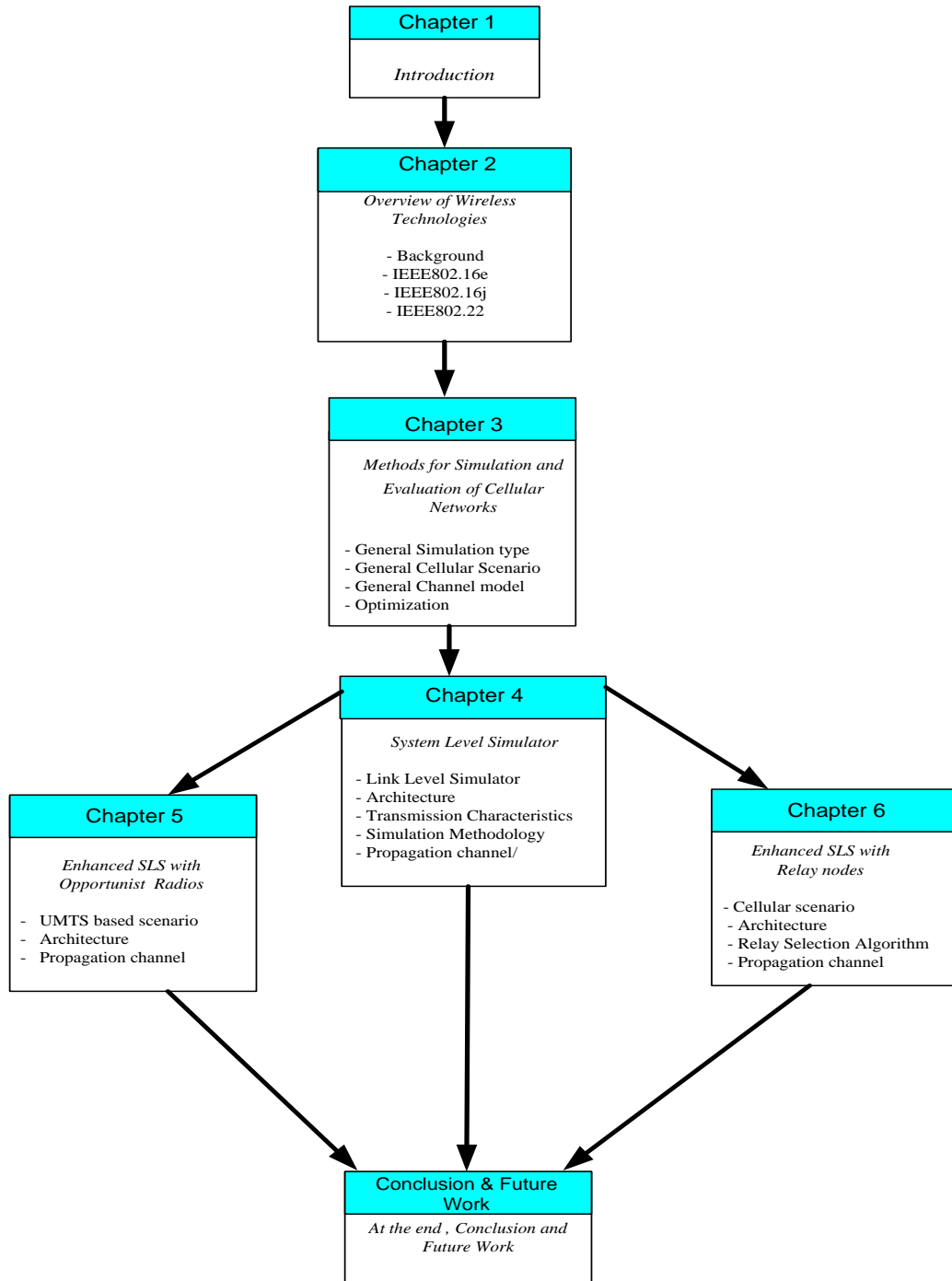
Figure 1-2: Dissertation Organization

1.5   Research Contribution

The following articles have been published in Conference and Journal during the PhD study.

1. **Book Chapters**

   - S.Mumtaz, Lee Than ,Rasool Sadeghi A.Gamerio, "Relay Implementation in WiMax System Level Simulator". Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering (LNICST),volume 2, 2009

2. **Journal Papers**

   - S. Mumtaz, A. Gameiro, and Kazi Saidul. " Calibration and Evaluation of Fixed and Mobile Relay-Based System Level Simulator ", Hindawi, Journal of Computer Systems, Networks, and Communications Volume 2010 (2010), Article ID 149257, 14 pages doi:10.1155/2010/1492572009.

   - S. Mumtaz, A. Gameiro, and Kazi Saidul. "Enhanced Algorithm for Link to System Interface Mapping ", International Journal of Computer Science and Information Security(IJSIS), an International Journal .USA, Vol. 3, July 2009.

   - S. Mumtaz, P. Marques, A. Gameiro, and J. Rodriguez. "Ad-hoc Behaviour in Opportunities Radio", Journal of Communications and Network(JCN), An International Journal Vol. 11, No.2, pag.186, April 2009.

   - A.Nascimento, J.Rodriguez, S.Mumtaz, , A.Gameiro. " Dynamic Resource Allocation For IEEE 802.16e ", in 'ACM/Springer Journal on Multimedia Systems', Mobile Networks and Applications, ACM, September, 2008 .

3. **Conference Papers**

   - S.Mumtaz, , A.Gameiro ,P Marques , "Clustering for Heterogeneous Wireless Ad-Hoc Networks" , IEEE, Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology(VITAE 09) , 2009. Aalborg/Denmark.

   - S.Mumtaz, , A.Gameiro ,P.Marques, "Application of Game Theory Ad-hoc Behaviour in Opportunities Radio" . IEEE, IARIA , The Eighth International Conference on Networks (ICN 2009) , Cancun, Mexico.

   - S.Mumtaz, A.Gameriro, "EESM for IEEE 802.16e: WiMaX", IEEE, ICIS, 2008 Portland, Oregon, USA.Page 4 of 4

   - Shahid Mumtaz, "Beamforming in Adaptive Array Antennas", IASTED International Conference on Wireless and optical Communication, 2007, Montreal, Canada.

1.6 References

[1] IEEE 802.16-2005: IEEE Standard for Local and Metropolitan Area Networks-Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access System—Amendment 2 : Physical Layer and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, February 2006.

[2] R. Pabst, B. Walke, D. C. Schultz, and et al, "Relay-based deployment concepts for wireless and mobile broadband radio," IEEE Communications Magazine, pp.80–89, Sep 2004.

[3] P. Gupta and P. R. Kumar, "The capacity of wireless networks,"*IEEE Transactions on Information Theory*, vol. 46, no. 2, pp.388–404, 2000.

[4] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—part I: system description," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1927–1938, 2003

[5] J. Mitola, "Cognitive radio for flexible multimedia communications," in *Proc. MoMuC*, 1999, p 3.

[6] Hämäläinen, S., P. Slanina, M. Hartman, A.Lappeteläinen, H. Holma, O.Salonaho, "A Novel Interface Between Link and System Level Simulations", Proceedings of ACTS Summit 1997, Aalborg, Denmark, October 1997, pp. 509 – 604.

[7] Modelling of Performance with Coloured Interference Using the EESM", Nortel Networks, 3GPP TSG-RAN WG1 #37, 2004.

[8] R. Pabst, B. H. Walke, D. C. Schultz et al., "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Communications Magazine*, vol. 42, no. 9, pp. 80– 89, 2004

Chapter 2: Overview of Broadband Wireless Technologies

The purpose of this chapter is to provide the executive summary of WiMaX standard. We begin the chapter by summarizing the activities of the IEEE 802.16, IEEE.16j and IEEE802.22. Next we discuss the salient features of WIMAX and briefly describe the physical and MAC –layer characteristics of WiMaX. Service aspects, such as quality of service, security, and mobility. The chapter ends with a brief discussion of the expected WIMAX performance.

2.1 Background

Broadband wireless networks attempt to provide broadband type of applications, initially envisioned for fixed networks, into the mobile environment. There are two different types of broadband wireless services: the first attempts to provide a set of services similar to that of the traditional fixed-line broadband but using wireless as the transmission medium. This is call *fixed wireless broadband*. The second is called *mobile wireless broadband* and it differs from the fixed one by offering the additional functionality of portability, nomadicity, and mobility. In this section, different proposed wireless systems intended for broadband wireless access are enumerated.

2.2 Broadband Wireless Technologies Overview

2.2.1 3G Cellular Systems

3G networks are currently being implemented by cellular network operators to deliver broadband applications over wireless. GSM operators evolved their second-generation (2G) networks by deploying UMTS High Speed Downlink Packet Access (HSDPA) and High Speed Uplink Packet Access (HSUPA) as part of the UMTS evolution. Traditional CDMA operators are deploying 1xEV-DO (1x Evolution Data Only) as part of their 3G solution. In China and other Asian countries, TD-SCDMA (Time Division-Synchronous CDMA) standard is also being implemented as the 3G solution for broadband wireless access networks.

2.2.2 High Speed Packet Access (HSPA)

HSDPA [1] is a downlink evolution of the UMTS standard defined in the Third Generation Partnership Project (3GPP) UMTS Release 5. It is capable of providing a peak user data rate of 14.4 Mbps, using a 5 MHz channel providing all 15 sub-channelization codes available in the radio frame which are assigned to one user only. HSDPA is based on a Time Division Duplex (TDD) frame with 15 slots and duration of 5 ms. Some enhancements such as spatial processing, diversity reception and multiuser detection are used to provide higher performance over basic systems. An uplink version, HSUPA, supports peak data rates up to 5.8 Mbps and is standardized as part of the 3GPP release 6 specifications.

2.2.3   1xEvolution Data Only (1xEV-DO)

1xEV-DO is a high-speed standard defined as an evolution to second-generation IS-95 CDMA systems by the 3GPP2 [2]. It supports a peak downlink data rate of 2.4 Mbps in a 1.25 MHz channel. The Revision A of this standard supports 3.1 Mbps and Revision B will support 4.9 Mbps. 3G systems are also evolving to support multimedia services.

2.2.4   1xEvolution Data Only Revision C (1xEV-DO Rev. C)

This is the major revision of the 1xEV-DO standard from 3GPP2 for longer-term evolution (LTE) to offer data rates of 70 Mbps to 200 Mbps in the downlink and 30 Mbps to 45 Mbs in the uplink, using 20 MHz channel bandwidth.

2.2.5   Long Term Evolution (LTE)

3GPP is developing the next revision of the 3G standards, named Long Term Evolution [3] (LTE). The objective of this new standard is to provide peak data rates up to 100 Mbps in the downlink and 50 Mbps in the uplink, with an average spectral efficiency of three to four times that of Release 6 HSPA. This will be achieved with the introduction of a completely new air interface based on OFDM/OFDMA (Orthogonal Frequency Division Multiplexing; OFDM Multiple Access), and MIMO.

2.2.6   Wireless Fidelity Systems (Wi-Fi)

Wi-Fi systems [4] are based on the IEEE 802.11 family of standards from 3GPP. It is primarily a local area networking (LAN) technology. Current Wi-Fi systems, based on IEEE 802.11a/g support a peak data rate of 54 Mbps with indoor coverage less than 100 m. Wi-Fi provides higher peak data rates than 3G systems thanks to the larger bandwidth used (over 20 MHz). However, capacity for outdoor scenarios is mainly reduced due to the inefficient multiple access scheme used, CSMA (Carrier Sense Multiple Access) along with the interference constraints of operating in the license-exempt band. Another limitation of Wi-Fi standard is that it was not designed to support user's mobility. New technologies are emerging with the revised IEEE 802.11n standard: multiple-antenna, spatial multiplexing and transmit diversity, which are envisioned to support a peak data rate of at least 100 Mbps.

2.2.7   IEEE802.16j

IEEE 802.16 has formed a task group to extend the IEEE 802.16e-2005 standard to include multi-hop communication, indicating that the field has reached a significant level of maturity. This amendment is called IEEE 802.16j [5], or as we sometimes refer to it, 16j. The history of 16j is quite recent; the task group was created in March 2006, and the first technical contributions were made in November of the same year. The purpose of IEEE 802.16j is not to standardize a new cellular network that includes multi-hop capability, but instead to expand previous single-hop

802.16 standards to include multi-hop capability. To expedite the standards process while still maximizing the likelihood that 16j will be useful in practice, designers limited their scope to the point-to-multipoint (PMP) OFDMA PHY mode of 802.16e-2005. Because companies are already well into the development stages of WiMAX-compatible mobile devices, 16j must be compatible with 16e MSs.

2.2.8   IEEE 802.22

It is expected that the demand for wireless services will continue to increase in the near and medium term, calling for more capacity and putting more and more pressure on the spectrum availability. While the use of advanced signal processing techniques may enable a very efficient usage of the spectrum even in the traditional framework of command and control spectrum policy, there is a worldwide recognition that these methods of spectrum management have reached their limit and are no longer optimal and new paradigms must be sought [6]. In fact, independent studies carried out in different places (e.g. [7]) have shown that most of the assigned spectrum is under-utilized. Thus, the problem is in most cases a problem of inefficient spectrum management rather than spectrum shortage.

The development of frequency agile terminals that can sense "holes" in the spectrum and adapt their transmission characteristics to use these "holes" may provide one tool to address and take advantage of this spectrum under-utilization. The evidence of the change and evolution in the approaches of spectrum management can already be seen in the development of the IEEE 802.22 cognitive radio based standard for fixed, point-to-multipoint, wireless regional area networks that operates on unused channels in the TV VHF/UHF bands between 54 and 862 MHz on a non-interfering basis [7]. The detecting of "holes" and the subsequent use of the unoccupied spectrum is referred to as opportunistic use of the spectrum. An Opportunistic Radio (OR) is the term used to describe a radio that is capable of such operation [8].

2.2.9   Wireless Interoperability for Microwave Access (WiMAX)

According to the WiMAX Forum, Mobile WiMAX is a broadband wireless solution that enables convergence of mobile and fixed broadband networks through a common wide area radio access technology and flexible network architecture [9-10]. In order to achieve high peak data rate transmissions in Non-Line of Sight (NLOS) scenarios, its air interface is based on the OFDMA multiple access for both the downlink as well as for the uplink connection. In order to address different channel bandwidths, ranging from 1.25 MHz to 20 MHz, according to regulator specifications from different markets, Mobile WiMAX implements scalable OFDMA (SOFDMA) [11,12]. This technique results in the modification of the size of the Fast Fourier Transform (FFT) used in the OFDM modulation, according to channel bandwidth, while keeping inter-carrier frequency separation constant, to provide orthogonality. WiMAX Medium Access Control (MAC)

layer is connection oriented for Quality of Service (QoS) support [10]. Some of the key features provided by Mobile WiMAX are enumerated as follows:

· OFDMA as multiple access technology for both downlink and uplink connections.

· Scalable OFDMA.

· Adaptive Modulation and Coding.

· Very high peak bit rates.

· Use of retransmission schemes at the link layer for fast and robust retransmissions.

· Use of flexible and dynamic per user resource allocation by means of OFDMA multiple access and sub-channelization schemes.

· Support for advanced antenna techniques such as MIMO and beamforming.

· Implementation of a connection oriented MAC layer for the provision of QoS.

· Robust security schemes and protocols.

· Support of mobility and nomadicity.

· All IP-based architecture for cost efficient and convergence with fixed IP networks.

IEEE 802.16e, also known as Mobile WiMAX [11] (Worldwide Interoperability for Microwave Access) is a broadband wireless solution that enables the convergence of mobile and fixed broadband networks through a common wide area broadband radio access technology (Orthogonal Frequency Division Multiplex Access OFDMA) and flexible network architecture. Mobile WiMAX is an extension to the previous IEEE 802.16d standard [11], also known as Fixed WiMAX, which is rapidly proving itself as a technology that will play a key role in the next generation of broadband wireless networks. One of the biggest potentials of Mobile WiMAX is that it offers scalability architecture in both radio access technology and network architecture which provides a great deal of flexibility in network deployment options. According to the predicted data rates and level of quality of service (QoS) to be expected from multimedia applications, and also to the throughput figures achieved under 3G cellular networks, it was realized since the beginning that the new standard should be built around the following key technical aspects:

· **Adequate Multiple Access Technology** – Mobile WiMAX supports OFDMA as a multiple access technology, whereby different users can be allocated different subsets of OFDM sub-carriers. OFDMA results in the exploitation of frequency diversity, besides the multiuser diversity achieved with opportunistic scheduling, under the normal OFDM air-interface in Fixed WiMAX. OFDM technology offers good resistance against multipath which suits WiMAX to operate in Non-Lien-of-Sight (NLOS) conditions.

· **Scalability for different Scenarios of Application** – Mobile WiMAX has a scalable PHY layer architecture that allows for data rates to scale with available channel bandwidth ranging

from 1.25 to 20 MHz [12]. This scalability is supported in the OFDMA mode, where the size of the Fast Fourier Transform (FFT) symbol may be changed according to the spectrum availability, while the inter-carrier frequency separation is kept constant.

- **Availability of High Data Rates -** Mobile WiMAX technology implements fast link adaptation schemes and scheduling over resources in time and frequency. With link adaptation, a number of modulations and forward error correction (FEC) coding schemes can be changed dynamically, according to the state of the channel reported on a frame basis. The scheduler, located in the base station, multiplexes the users in either frequency and/or time domains. The placement of the scheduler in the base station results in a faster resource allocation process.

- **Advanced antenna techniques** – WiMAX supports the use of multiple antenna techniques to increase overall system capacity and spectrum efficiency, such as space-time coding, spatial multiplexing and beamforming [11].

- **Provision of Quality of Service (QoS) demands** – One of the main achievements of the standard is the provision of QoS mechanisms designed to support a variety of applications including voice and multimedia services. The MAC layer has a connection-oriented architecture, capable of supporting multiple connections per each user terminal and is responsible for the assignment of specific service flow QoS parameters, according to the type of application. Bandwidth request mechanisms were developed for the provision of QoS under WiMAX standard.

- **Advanced Error Protection Mechanisms** – Mobile WiMAX supports a hybrid mechanism composed of automatic retransmission requests (ARQ) and FEC for connections requiring transmission feasibility (HARQ). Two types of ARQ mechanisms are considered in the standard: Chase Combining (CC) and Increment Redundancy (IC).

- **Robust Security** – Mobile WiMAX supports strong encryption, using Advanced Encryption Standard (AES) and has a robust privacy and key-management protocol. It also offers a flexible authentication architecture based on Extensible Authentication Protocol (EAP).

- **Mobility** – Mobile WiMAX has inherent features for the support of seamless handovers of real-time applications with very small latencies.

- **IP-based architecture** – The reference network is based on an all-IP platform, which relies on IP-based protocols for end-to-end transport, QoS, session management, security and mobility.

The IEEE 802.16 group was formed in 1998 to develop an air-interface standard for wireless broadband. The group's initial focus was the development of a LOS-based point-to-multipoint wireless broadband system for operation in the 10GHz–66GHz millimeter wave band. The

resulting standard—the original 802.16 standard, completed in December 2001—was based on a single-carrier physical (PHY) layer with a burst time division multiplexed (TDM) MAC layer. Many of the concepts related to the MAC layer were adapted for wireless from the popular cable modem DOCSIS (data over cable service interface specification) standard. The IEEE 802.16 group subsequently produced 802.16a, an amendment to the standard, to include NLOS applications in the 2GHz–11GHz band, using an *orthogonal frequency division multiplexing* (OFDM)-based physical layer. Additions to the MAC layer, such as support for *orthogonal frequency division multiple access* (OFDMA), were also included. Further revisions resulted in a new standard in 2004, called IEEE 802.16-2004, which replaced all prior versions and formed the basis for the first WiMAX solution. These early WiMAX solutions based on IEEE 802.16-2004 targeted fixed applications, and we will refer to these as fixed WiMAX [12]. In December 2005, the IEEE group completed and approved IFEEE 802.16e-2005, an amendment to the IEEE 802.16-2004 standard that added mobility support. The IEEE 802.16e-2005 forms the basis for the WiMAX solution for nomadic and mobile applications and is often referred to as mobile WiMAX. IEEE 802.16 standards are the typical BWA protocols, which are considered as an alternative solution to 3G and migration towards 4G [3]. Currently one of their workgroups is taking effort to draft an updating version based on the existent IEEE 802.16d/e standard. They are aiming at introducing relay into IEEE 802.16 networks to practice more economic deployment and more flexible radio resource management. The purpose of this amendment is to enhance coverage, throughput and system capacity of 802.16 networks by specifying 802.16j multi-hop relay [5] capabilities and functionalities of interoperable relay stations and base stations. The basic characteristics of the various IEEE 802.16 standards are summarized in Table 2-1 and Figure 2-1 shows the timeline of most relevant IEEE 802.16 standard documents and amendments. Note that these standards offer a variety of fundamentally different design options. For example, there are multiple physical-layer choices: a single-carrier-based physical layer called Wireless-MAN-SCa, an OFDM-based physical layer called WirelessMAN-OFDM, and an OFDMA based physical layer called Wireless-OFDMA. Similarly, there are multiple choices for MAC architecture, duplexing, frequency band of operation, etc. These standards were developed to suit a variety of applications and deployment scenarios, and hence offer a plethora of design choices for system developers. In fact, one could say that IEEE 802.16 is a collection of standards, not one single interoperable standard. For practical reasons of interoperability, the scope of the standard needs to be reduced, and a smaller set of design choices for implementation need to be defined. The WiMAX Forum does this by defining a limited number of system profiles and certification profiles. A *system profile* defines the subset of mandatory and optional physical- and MAC-layer features selected by the WiMAX Forum from the IEEE 802.16-2004 or IEEE 802.16e-2005 standard. It should be noted that the

mandatory and optional status of a particular feature within a WiMAX system profile might be different from what it is in the original IEEE standard.

	802.16-2004	802.16e-2005	802.16j	802.16m
Status	Completed June2004	Completed December 2005	Completed 2009	In Progress
Frequency Band	2GHz-11GHz	2GHz-11GHz for fixed 2GHz-6GHz for mobile	2GHz-11GHz for fixed 2GHz-6GHz for mobile	2.5GHz-11GHz for fixed 2GHz-6GHz for mobile
Application	Fixed NLOS	Fixed and Mobile NLOS	Fixed and Mobile Relay LOS/NLOS	Fixed and Mobile Relay LOS/NLOS
MAC architecture	Point-to-multipoint; Mesh	Point-to-multipoint; Mesh	Multi-Hop	Multi-Hop
Transmission scheme	Single carrier, 256 OFDM or 2,048 OFDM	Single carrier, 256 OFDM or scalable OFDM with 128, 512, 1024 or 2048 sub-carriers	OFDM with 128, 512, 1024 or 2048 sub-carriers	OFDM with 128, 512, 1024 or 2048 sub-carriers
Modulation	QPSK, 16QAM, 64QAM	QPSK, 16QAM, 64QAM	QPSK, 16QAM, 64QAM	QPSK, 16QAM, 64QAM
Data Rate	1Mbps-75Mbps	1Mbps-75Mbps	1Mbps-75Mbps (even more)	1Mbps-75Mbps (even more)
Multiplexing	TDM/TDMA/OFDMA	TDM/TDMA/OFDMA	TDM/TDMA/OFDMA	TDM/TDMA/OFDMA
Duplexing	TDD and FDD	TDD and FDD	TDD and FDD	TDD and FDD
Air-interface	WirelessMAN-SCa WirelessMAN-OFDM WirelessMAN-OFDMA WirelessMAN-HUMAN	WirelessMAN-SCa WirelessMAN-OFDM WirelessMAN-OFDMA WirelessMAN-HUMAN	WirelessMAN-OFDM WirelessMAN-OFDMA	WirelessMAN-OFDMA
WiMAX implementation	256-OFDM as Fixed WiMAX	Scalable OFDMA as Mobile WiMAX	Scalable OFDMA as Mobile WiMAX	Scalable OFDMA as Mobile WiMAX

Table 2-1: Basic Data on IEEE 802.16 Standards

Currently, the WiMAX Forum has two different system profiles: one based on IEEE 802.16-2004, OFDM PHY, called the fixed system profile; the other based on IEEE 802.16e-2005 scalable OFDMA PHY, called the mobility system profile. A *certification profile* is defined as a particular instantiation of a system profile where the operating frequency, channel bandwidth, and duplexing mode are also specified. WiMAX equipment is certified for interoperability against a particular certification profile.

With the completion of the IEEE 802.16e-2005 standard, interest within the WiMAX group has shifted sharply toward developing and certifying mobile WiMAX system profiles based on this newer standard like IEEE802.16j. All mobile WiMAX profiles use scalable OFDMA as the

physical layer. At least initially, all mobility profiles will use a point-to-multipoint MAC. It should also be noted that all the current candidate mobility certification profiles are TDD based. Although TDD is often preferred, FDD profiles may be needed for in the future to comply with regulatory pairing requirements in certain bands.
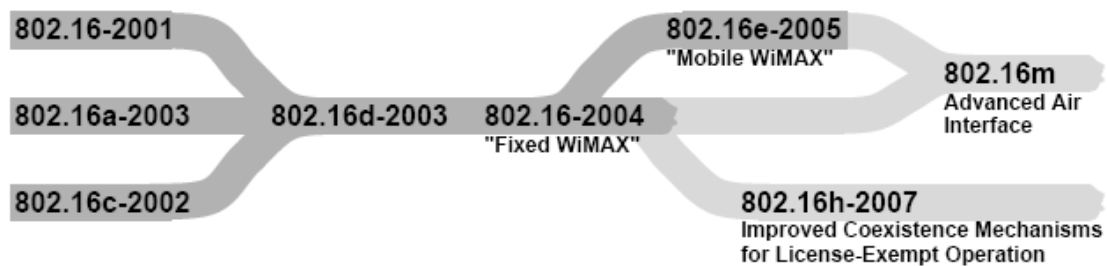
Figure 2-1: Timeline of most relevant IEEE 802.16 standard documents and amendments.

2.3   Salient Feature of IEEE802.16

WiMAX is a wireless broadband solution that offers a rich set of features with a lot of flexibility in terms of deployment options and potential service offerings. Some of the more salient features that deserve highlighting are as follows:

OFDM-based physical layer: The WiMAX physical layer (PHY) is based on orthogonal frequency division multiplexing, a scheme that offers good resistance to multipath, and allows WiMAX to operate in NLOS conditions. OFDM is now widely recognized as the method of choice for mitigating multipath for broadband wireless.

Very high peak data rates: WiMAX is capable of supporting very high peak data rates. In fact, the peak PHY data rate can be as high as 74Mbps when operating using a 20MHz wide spectrum. More typically, using a 10MHz spectrum operating using TDD scheme with a 3:1 downlink-to-uplink ratio, the peak PHY data rate is about 25Mbps and 6.7Mbps for the downlink and the uplink, respectively. These peak PHY data rates are achieved when using 64 QAM modulation with rate 5/6 error-correction coding. Under very good signal conditions, even higher peak rates may be achieved using multiple antennas and spatial multiplexing.

Scalable bandwidth and data rate support: WiMAX has a scalable physical-layer architecture that allows the data rate to scale easily with available channel bandwidth. This scalability is supported in the OFDMA mode, where the FFT (fast fourier transform) size may be scaled based on the available channel bandwidth. For example, a WiMAX system may use 128, 512, or 1,048-bit FFTs based on whether the channel bandwidth is 1.25MHz, 5MHz, or 10MHz,

respectively. This scaling may be done dynamically to support user roaming across different networks that may have different bandwidth allocations.

Adaptive modulation and coding (AMC): WiMAX supports a number of modulation and forward error correction (FEC) coding schemes and allows the scheme to be changed on a per user and per frame basis, based on channel conditions. AMC is an effective mechanism to maximize throughput in a time-varying channel. The adaptation algorithm typically calls for the use of the highest modulation and coding scheme that can be supported by the signal-to-noise and interference ratio at the receiver such that each user is provided with the highest possible data rate that can be supported in their respective links.

Link-layer retransmissions: For connections that require enhanced reliability, WiMAX supports automatic retransmission requests (ARQ) at the link layer. ARQ-enabled connections require each transmitted packet to be acknowledged by the receiver; unacknowledged packets are assumed to be lost and are retransmitted. WiMAX also optionally supports hybrid-ARQ, which is an effective hybrid between FEC and ARQ.

Support for TDD and FDD: IEEE 802.16-2004 and IEEE 802.16e-2005 supports both time division duplexing and frequency division duplexing, as well as a half-duplex FDD, which allows for a low-cost system implementation. TDD is favored by a majority of implementations because of its advantages: (1) flexibility in choosing uplink-to-downlink data rate ratios (2) ability to exploit channel reciprocity, (3) ability to implement in non-paired spectrum, and (4) less complex transceiver design. All the initial WiMAX profiles are based on TDD, except for two fixed WiMAX profiles in 3.5GHz.

Orthogonal frequency division multiple access (OFDMA): Mobile WiMAX uses OFDM as a multiple-access technique, whereby different users can be allocated different subsets of the OFDM tones.

Flexible and dynamic per user resource allocation: Both uplink and downlink resource allocation are controlled by a scheduler in the base station. Capacity is shared among multiple users on a demand basis, using a burst TDM scheme. When using the OFDMA-PHY mode, multiplexing is additionally done in the frequency dimension, by allocating different subsets of OFDM subcarriers to different users. Resources may be allocated in the spatial domain as well when using the optional advanced antenna systems (AAS). The standard allows for bandwidth resources to be allocated in time, frequency, and space and has a flexible mechanism to convey the resource allocation information on a frame-by-frame basis.

Support for advanced antenna techniques: The WiMAX solution has a number of hooks built into the physical-layer design, which allows for the use of multiple-antenna techniques, such as beamforming, space-time coding, and spatial multiplexing. These schemes can be used to

improve the overall system capacity and spectral efficiency by deploying multiple antennas at the transmitter and/or the receiver.

Quality-of-service support: The WiMAX MAC layer has a connection-oriented architecture that is designed to support a variety of applications, including voice and multimedia services. The system offers support for constant bit rate, variable bit rate, real-time, and non-real-time traffic flows, in addition to best-effort data traffic. WiMAX MAC is designed to support a large number of users, with multiple connections per terminal, each with its own QoS requirement.

Robust security: WiMAX supports strong encryption, using Advanced Encryption Standard (AES), and has a robust privacy and key-management protocol. The system also offers a very flexible authentication architecture based on Extensible Authentication Protocol (EAP), which allows for a variety of user credentials, including username/password, digital certificates, and smart cards.

Support for mobility: The mobile WiMAX variant of the system has mechanisms to support secure seamless handovers for delay-tolerant full-mobility applications, such as VoIP. The system also has built-in support for power-saving mechanisms that extend the battery life of handheld subscriber devices. Physical-layer enhancements, such as more frequent channel estimation, uplink sub-channelization, and power control, are also specified in support of mobile applications.

IP-based architecture: The WiMAX Forum has defined a reference network architecture based on an all-IP platform. All end-to-end services are delivered over an IP architecture relying on IP-based protocols for end-to-end transport, QoS, session management, security, and mobility. Reliance on IP allows WiMAX to ride the declining cost curves of IP processing, facilitate easy convergence with other networks, and exploit the rich ecosystem for application development that exists for IP.

2.4   WiMAX Physical Layer

This section details the implementation of the OFDMA-based PHY layer in Mobile WiMAX networks.

2.4.1   OFDM Basics

The WiMAX PHY layer is based on the OFDM technology, which is an efficient and low-cost scheme for high data rate transmission in a NLOS or multipath radio environment. OFDM is a spectrally efficient version of multicarrier modulation (MCM) where the sub-carriers are selected so that they are all orthogonal to one another over the symbol duration, thereby avoiding the need to have non-overlapping subcarrier channels to eliminate inter-carrier interference (ICI). It is extremely easy to implement OFDM modulators/demodulators in discrete time using Inverse Fast Fourier Transform (IFFT) and Fast Fourier Transform (FFT) chips, respectively. OFDM is very

efficient in eliminating signal distortion due to delay spread arising from multipath propagation because the stream of data is split among the orthogonal sub-carriers, resulting in an increase of the symbol time interval, which makes it more immune to Inter-Symbol-Interference (ISI). In addition, ISI can be completely eliminated with the insertion of guard intervals between OFDM symbols, larger than the expected multipath delay spread [12]. This guard interval is called Cyclic Prefix (CP) in OFDM. As long as the CP is longer than the channel delay spread, ISI is completely eliminated. The CP is a repetition of the last samples of the OFDM symbol that is appended to the beginning of the data payload. This mechanism makes the channel circular and enables the use of simple Maximum Ratio Combiners (MRC) as decoders in the receiver, instead of complex multi-user decoders such as the one used in CDMA for example. Figure 2-2 illustrates the creation of the CP.
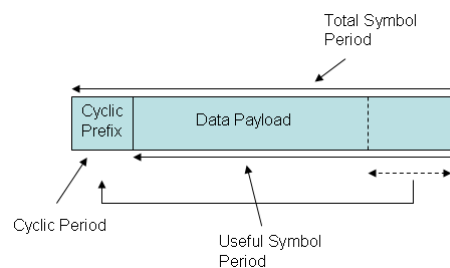
Figure 2-2: Data Symbol Structure and creation of cyclic prefix

The OFDMA symbol consists of three types of sub-carriers:

·   Data sub-carriers for data transmission.

·   Pilot sub-carriers for estimation and synchronization purposes.

·   Null sub-carriers for no transmission. These are used for guard bands and DC sub-carrier.
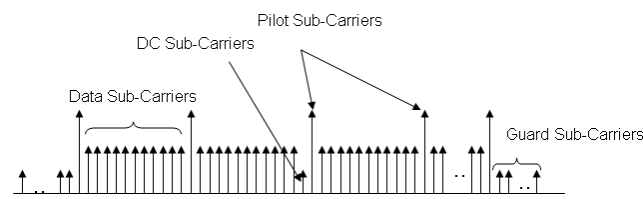
Figure 2-3: Data Symbol Structure and creation of cyclic prefix [5]

Active (data and pilot) sub-carriers are grouped into subsets of sub-carriers called sub-channels. Figure 2-3 illustrates the OFDMA sub-carrier structure implemented in Mobile WiMAX standard [12].

2.4.2   OFDMA Sub-Channelization in WiMAX

The WiMAX OFDMA-based multiple access supports sub-channelization (see ANNEX A) in both downlink and uplink. The minimum time-frequency unit of sub-channelization is one slot, which comprises 48 data sub-carriers. Fixed WiMAX PHY layer is based on OFDM technology and it allows a limited form of sub-channelization in the uplink direction: the multiple access scheme is OFDM/TDMA for downlink and OFDMA for uplink. In the downlink, only one mobile can transmit over all sub-carriers for the time interval corresponding to one OFDM symbol. In the uplink, 16 sub-channels are defined for allocation: 1, 2, 4, 8 or all 16 sub-channels can be assigned to a subscriber station (SS). Mobile WiMAX physical layer is based on the OFDMA multiple access technology. It allows the allocation of small sets of sub-carriers to different users in downlink and uplink directions. Sub-channels in Mobile WiMAX are comprised of sub-carriers which may be allocated contiguously or distributed according to a pseudo-random pattern, depending on the cell index, over the spectrum.

2.4.3   Diversity Permutation Sub-Carrier Sub-Channelization

The distributed sub-carrier permutation mode is a very useful scheme for averaging out inter-cell interference and avoiding deep fading, by allocating in each sub-channel sub-carriers in a pseudo-random way, provided different random sub-carrier distribution patterns are defined and assuming transmission is synchronized among neighbouring cells [9]. This mode is efficient in achieving frequency diversity, which makes it effective for scenarios in which it is difficult to track frequency selective channel variations, in order to allocate resources and/or select transmission parameters adaptively according to these variations. Therefore, it is expected to be suitable for users with high velocity and/or low signal-to-interference-plus-noise ratio (SINR). In this mode basic resource, units in frequency domain are called diversity sub-channels and result in one degree of freedom for resource allocation (time domain).In WiMAX there are several sub-channelization schemes based on the pseudo-random distribution of sub-carriers for both uplink and downlink. Two of the most important ones are Partial Usage of Sub-carriers (PUSC), which is mandatory for all WiMAX implementations and Full Usage of Sub-carriers (FUSC). PUSC enables the implementation of segmentation in the MAC layer. A segment is an individual instance of the MAC layer. Sub-carriers belonging to sub-channels from different segments do not collide, even if the same pattern is used in adjacent cells. This has to do with the way in which random allocation is performed for data and pilot sub-carriers. In FUSC mode all sub-carriers of each sub-channel are spread over the whole FFT spectrum, which forbids the implementation of segmentation for this channelization mode.

2.4.4   Band Adjacent Multi-Carrier (AMC) Sub-Channelization

In adjacent sub-carrier permutation mode, adjacent sub-carriers are grouped into clusters of contiguous sub-carriers. In this mode, the channel response can be seen as a flat fading channel. Thus, the frequency selectivity of the channel cannot be exploited, but the system can make better use of multiuser diversity over frequency domain, as long as the channel state does not change significantly during the scheduling period. This mode is suitable for scenarios with high SINR and/or with low mobile speeds, because it is more sensitive to inter-cell interference and to errors in the estimation of the channel quality [12].

The band AMC sub-channelization results in two degrees of freedom for resource allocation, as resources are available in both frequency and time domains. Band AMC allows system designers to exploit multiuser diversity, allocating sub-channels to users based on their frequency response over each sub-channel, assuming the state of each sub-channel is independent among different users (i.e., channel states are not correlated). Multiuser diversity over frequency domain provide significant gains in overall system capacity, provided the system assigns to each user a sub-channel that maximizes its received SINR.

2.5   Frame Structure of Mobile WiMAX

Mobile WiMAX systems can support time-division duplex (TDD) or frequency-division duplex (FDD) modes. For both FDD and TDD duplex modes the frame structure is the same, except that both uplink and downlink sub-frames are transmitted simultaneously over different frequency bands for FDD. For the first profile released for Mobile WiMAX the duplex mode to be used is TDD. In TDD mode, the frame is subdivided in two sub-frames separated by one guard interval and the downlink-to-uplink-sub-frame ratio may be varied to support different traffic profiles. The frame is composed of several zones that are divided according to sub-carrier allocation methods or Multiple-Input Multiple-Output (MIMO) modes. Figure 2-4 illustrates the structure of the TDD frame in Mobile WiMAX standard.The following fields are implemented:

Downlink Channels

·   **Preamble** – This is the first OFDM symbol in the frame and is used for time and frequency synchronization as well as for channel estimation.

·   **Frame Control Header (FCH)** – This field provides frame configuration information such as the Mobile Application Part (MAP) message length, the modulation and coding scheme (MCS) used in the DL-MAP and the used sub-carriers.

·   **DL-MAP and UL-MAP** – These fields convey the MAP messages (named as Information Elements – IEs) which indicate the starting times of bursts. IEs messages are used in the indication of the slots which are assigned to each burst in the data region. These fields are

broadcast after the FCH in the downlink sub-frame. Since both MAP fields contain critical information that needs to reach all users, they are often transmitted with the most robust modulation and coding scheme (BPSK with rate 1/2 coding and repetition coding). MAP messages could form a significant overhead in the frame, resulting in less efficiency for data allocation, particularly when there are a large number of users with small packets. To mitigate overhead, WiMAX can optionally use multiple sub-map messages where the dedicated control messages to different users are transmitted at higher rates, based on their individual SINR conditions. Each MAP message is composed of Information Elements (IE) containing information regarding each burst allocation in the TDD frame. Namely each IE contain:

- The description of the downlink burst profile (modulation and coding combination for each downlink burst).
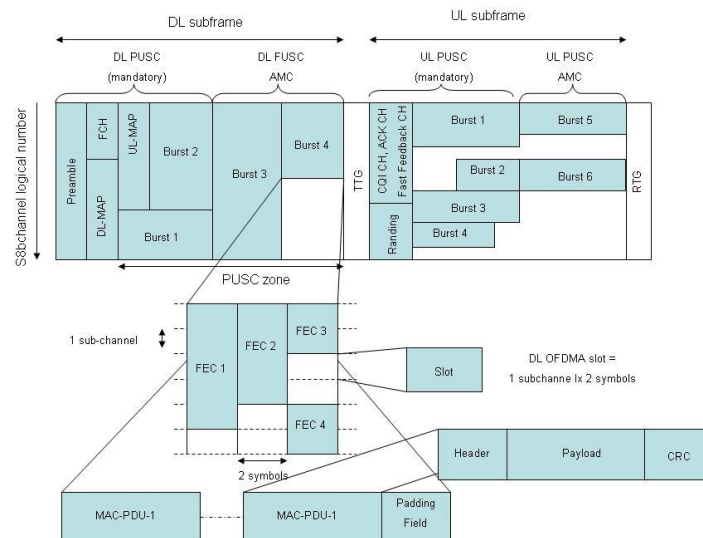
Figure 2-4: Frame Structure

- An optional field with the list of Connection Identifiers (CIDs) with packets mapped on the downlink burst to which the DL-MAP is referred to, including the number of connection identifiers in the list.
- The Burst Allocation information field:
  - · OFDMA symbol offset
  - · Sub-channel offset
  - · Number of sub-channels
  - · Number of OFDMA symbols
  - · Use of power boosting (+6dB to -9 dB)

· Indication on the use of Repetition Coding (1/2/4/6)

· **Data Bursts** – The minimum time-frequency resource which can be allocated for a given user is the slot. Each slot consists of one sub-channel over one, two or three OFDM symbols, depending on the sub-channelization scheme used. A contiguous sequence of slots assigned to a given user constitutes a burst. All slots in the same burst must transmit with the same MCS scheme. Bursts are allocated depending on user's traffic demand, QoS requirements and channel conditions.

Uplink Control Channels

Mobile WiMAX provides a number of control channels very useful for the exchange information such as channel quality information Acknowledge/Negative ACK (ACK/NACK) feedback for HARQ.

· **Ranging** – This is a region in the uplink sub-frame for contention-based access, which is used for a variety of purposes: it can be used to perform closed-loop frequency, time and power adjustments during network entry as well as periodically. It can also be used by the mobile station to make uplink bandwidth requests in contention mode.

· **Channel Quality Indication Channel (CQICH)** – This channel is used by the mobile station to feedback channel quality information that can be used in the base station scheduler for link adaptation purposes. This control channel is used to report the downlink SINR for either diversity sub-channels or band AMC sub-channels. This channel occupies one uplink slot in the fast-feedback region in the uplink sub-frame. For diversity sub-channels, the mobile terminal reports the average SINR of the preamble broadcast in the downlink sub-frame, from which the base station is able to determine the downlink MCS scheme level. For band AMC sub-channels, a mobile terminal can report the differential of SINR values of five selected frequency bands after reporting the SINR measurements of the five best bands.

· **Uplink ACK Channel** – This is a region allocated in the uplink sub-frame and is designed for the inclusion of one or more ACK channels for enabling HARQ in data transmission. Each uplink ACK channel occupies one half-slot in the HARQ ACK region and is implicitly assigned to each HARQ-enabled burst, according to the order of the HARQ-enabled downlink bursts in the DL-MAP. Thus, the mobile terminal can quickly transmit ACK or NACK feedback for downlink HARQ-enabled packet data using this uplink ACK channel.

· **Uplink Sounding** – Mobile WiMAX defines an uplink sounding zone in the uplink sub-frame for the definition of uplink sounding symbols, which are used in the support of MIMO and smart antenna beamforming. The base station measures the uplink channel response from uplink sounding waveforms transmitted by each mobile station and translates the measured

uplink channel response to an estimated downlink channel response, under the assumption of channel reciprocity for the TDD duplex mode.

Support of Advanced Antenna Technologies

Mobile WiMAX supports various multiple antenna technologies which are applied in different zones within a frame. For example, Adaptive Antenna Systems (AAS) is a kind of smart antenna processing which can be used for Space Division Multiple Access (SDMA), and allows the transmission of data bursts concurrently to spatially-separated mobile stations, by applying different beam patterns to them. Figure 2-5 shows a logical frame structure for AAS support.
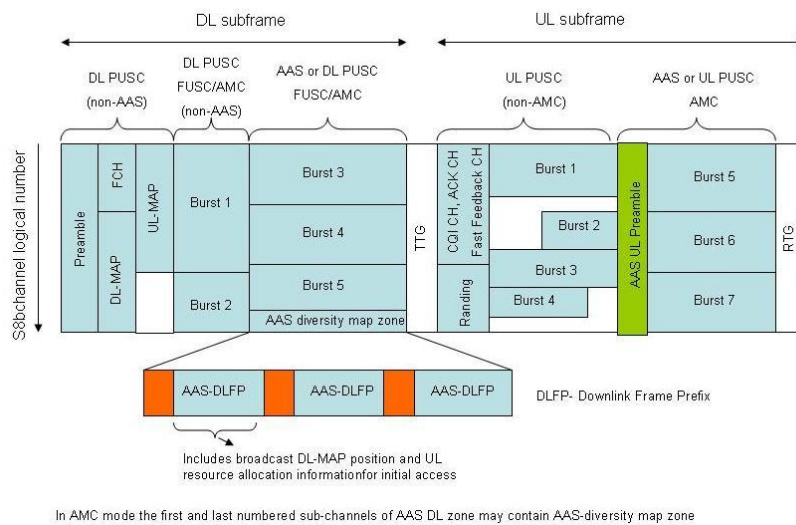
Figure 2-5: Frame structure for Mobile WiMAX using TDD duplex mode for AAS support

Downlink and uplink AAS zones are defined by a special broadcast MAP message and, as can be seen from the Figure 2.4, the downlink AAS zone includes an AAS diversity map zone which occupies two sub-channels. AAS-Downlink Frame Prefixes (AAS-DLPFs) in the DL AAS zone are transmitted using different beams from each other. Each mobile station, in AAS mode, scans these AAS-DLPFs using known AAS downlink preamble patterns which were previously reported and they choose the one with the best beam. Each AAS-DLFP includes the position of the broadcast DL-MAP which is beamformed. It can also be used to page a specific mobile that cannot receive the normal DL-MAP or to resource allocation information for uplink initial access. Once the mobile obtains the information regarding initial resource allocation through a broadcast DL-MAP pointed to by the AAS-DLFP, subsequent allocations can be managed with private DL-MAP and UL-MAP messages that are unicast and beam formed with high MCS levels.

2.6 WiMAX Medium Access Control Layer (MAC)

Mobile WiMAX standard was designed and developed from the outset for the delivery of broadband applications. Its MAC layer, in particular, has inherent features designed for the joint support of those burst data traffic applications with high peak rate demands and streaming and/or delay sensitive ones. The fine granularity and flexibility provided by the MAC layer in resource allocation, according to user's bandwidth needs, and the lower latency incurred in handling user's bandwidth requests and in making scheduling decisions, makes it possible to send data through the air-interface under the stringent QoS requirements of each type of service flow, and the efficient use of radio resources, with the consequent maximization of the achieved spectrum efficiency. Figure 2-6 illustrates the MAC layer architecture proposed for Mobile WiMAX.
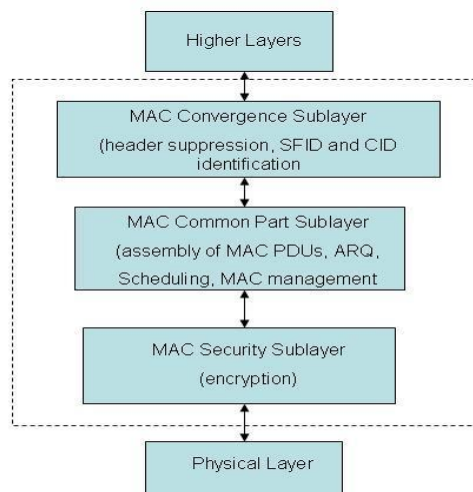
Figure 2-6: MAC Layer Architecture for Mobile WiMAX

2.6.1 MAC Layer Functional Blocks

The MAC layer is sub-divided in Convergence Sub-layer (CS) and Common Part Sub-layer (CPS).

·   The CS sub-layer is responsible for the interface between the MAC layer of WiMAX and other backhaul networks such as (Asynchronous Transfer Mode) ATM and IP-based networks. The CS sub-layer maps the QoS parameters from external networks to the set of QoS parameters used in the WiMAX standard. This sub-layer classifies Service Data Units (SDUs) to a proper connection in the CPS sub-layer with specific QoS parameters.

·   The CPS sub-layer is the most important element in the MAC layer. It is in the CPS sub-layer where the packet scheduler resides. Other MAC layer functionalities such as: packet

concatenation and/or fragmentation, packet error control through retransmissions and resource allocation are also performed inside the CPS sub-layer.

The MAC layer provides the interface between the PHY layer and upper transport layers. It takes advantage of different PHY layer services that address the needs of various mobile environments. The MAC layer performs several functions including: scheduling of bursts, link adaptation and error recovery, network entry procedures, and standard Packet Data Unit (PDU) creation tasks, such as fragmentation or packing, for each active connection. The MAC layer supports five QoS service classes that serve the data transfer needs of different real-time and non-real time media types.
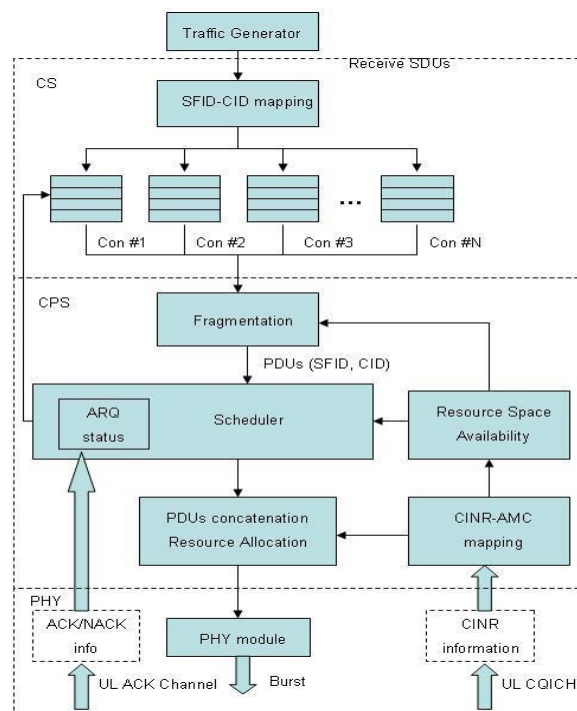
Figure 2-7: Mobile WiMAX Protocol layer

The MAC layer receives SDU packets from higher layers and organizes them into PDUs for transmission over the air. Multiple SDUs of same or different lengths may be aggregated into a single PDU (MPDU) in order to save MAC header overhead. In the same way, large SDUs may be fragmented so that they may be sent across frame boundaries. MPDUs may be of variable length and multiple MPDUs may be concatenated into a single burst to reduce MAP overhead. Each burst is transmitted using a single MCS that is signaled within the MAP message and may include MPDUs intended for one or more users. Each MPDU is segmented into forward error correction (FEC) blocks that are coded and interleaved within the burst. A number of contiguous OFDMA symbols using the same permutation formula to map sub-carriers to sub-channels is called a permutation zone.

In Mobile WiMAX the MAC layer is designed for handling applications with different QoS requirements. All services are connection-oriented, i.e., each service is mapped to one or multiple connections and is handled by the CS sub-layer and then the CPS. This is illustrated in Figure 2-7. As can be seen from the Figure 2-6, the CS classifies the SDUs to a proper connection with specific QoS parameters and, depending on the QoS requirements from each type of service data flow in the MAC layer, there are different mechanisms available for bandwidth allocation, which is the responsibility of the MAC layer.

2.7   Mechanisms for Quality of Service Support

Quality of service support is one of the essential features in WiMAX standards. Strong QoS control is achieved by using a connection-oriented MAC architecture where all downlink and uplink connections are controlled by the serving base station. Before any data transmission the base and mobile stations must establish a unidirectional logical link called a connection between both two MAC peers. Each connection is identified by a connection identifier (CID) which is a temporary address for data transmissions over the particular link. Fundamental to the provision of QoS in WiMAX is the implementation of service flows. A service flow is defined as a one-way flow of MSDUs on a connection associated with specific QoS parameters such as latency, jitter and throughput. These parameters are inputs to the scheduler in the MAC layer. Each service flow is identified by a service flow identifier (SFID). The base station is responsible for issuing the SFID and for mapping it to unique CIDs. The packet scheduler is placed inside the MAC layer in order to be fast enough and reduce the latency incurred by scheduling decisions, which would otherwise increase if it was placed inside the base station controller.

2.7.1   Service Classes in Mobile WiMAX

To support a wide variety of applications WiMAX defines five scheduling services:

Unsolicited Grant Service (UGS) – This is designed to support Real-Time (RT) service flows which periodically generate packets of fixed size. Service flows of type UGS are granted radio resources periodically without the need for the scheduler intervention and bandwidth request from mobiles. This results in a reduction of the associated signalling overflow and latency incurred in bandwidth requests. This type of service flow is particularly suitable for the support of applications with a constant bit rate (CBR applications). For uplink, this service offers fixed-size grants for data transport on a real-time periodic basis (implicit request). Connections configured with UGS are not allowed to utilize random access opportunities. Examples of the type of service implementing this type of scheduling service are T1/E1 and VoIP without silence suppression.

Real-Time Polling Service (rtPS) – This is designed to support delay sensitive real-time service flows with variable-size data packets on a periodic basis. This service offers periodic dedicated

request opportunities to meet real-time requirements, which results in more signalling overhead and latency than UGS. It is well suited for connections associated to service flows of type VoIP or video streaming services, such as Near Real Time Video (NRTV). For uplink this service offers periodic unicast request opportunities (piggyback request/unicast polling). Because the mobile station issues explicit requests, the protocol overhead and latency is increased, but this capacity is granted only according to the real need of the connection.

Extended Real-Time Polling Service (ertPS) – This is a mix of both UGS and rtPS service classes. It is designed to support real-time service flows that generate delay sensitive variable sized data packets on a periodic basis, as in rtPS, and the base station grants unicast grants in an unsolicited manner as in UGS. For uplink, this service offers a mechanism for periodic allocations, which may be used for requesting the bandwidth as well as for data transfer, considering the traffic characteristics of VoIP with silence suppression (piggyback request/unicast polling).

Non-Real-Time Polling Service (nrtPS) – This is designed to support delay-tolerant service flows consisting of variable-sized data packets for which a minimum data rate is required. The nrtPS particularly addresses Internet type of applications such as File Transfer Protocol (FTP) and web browsing. For uplink, this service offers unicast polls on a regular basis, in an interval on the order of 1s or less. For this service, connections may utilize random access transmit opportunities for sending bandwidth requests.

Best Effort (BE) – This is designed to support service flows that have no minimum service requirements and which are serviced on a resource-availability basis. BE service flows provide no guarantees either for minimum throughput assurance or maximum packet delay. For uplink, this service may offer contention request opportunities (contention based polling). The mobile sends requests for bandwidth in either random access slots or dedicated transmission opportunities. The occurrence of dedicated opportunities is subject to network load and the mobile cannot rely on their presence. Examples of BE service flows is the email application.

Table 2-2 illustrates the different types of data delivery services and respective QoS provisioning parameters in Mobile WiMAX systems. Table 2-3 provides the definitions for the QoS parameters. Finally, Table 2-4 describes the classification proposed for each one of the four traffic types implemented in this work, and the QoS parameters, which are considered in taking scheduling decisions. The Maximum Sustained Rate is actually not defined and/or used in the schedulers proposed in this work, because whenever a given user is scheduled for transmission it is assigned the amount of radio resources needed to empty its buffer.

Service Type	Application	QoS Parameters
UGS	T1/E1, VoIP without silence suppression	· Minimum reserved traffic rate · Maximum latency · Tolerated jitter
rtPS	Streaming audio or video	· Maximum reserved traffic rate · Minimum sustained traffic rate · Maximum latency · Traffic priority
ertPS	VoIP with silence suppression	· Maximum reserved traffic rate · Minimum sustained traffic rate · Maximum latency · Tolerated jitter · Traffic priority
nrtPS	File Transfer Protocol (FTP)	· Maximum sustained traffic rate · Minimum reserved traffic rate · Traffic priority
BE	Data Transfer, Web browsing	· Maximum sustained traffic rate · Traffic priority

Table 2-2: Types of data delivery services and their QoS requirements

Parameter	Definition
Maximum reserved traffic rate	Peak information rate of service
Maximum sustained traffic rate	Minimum amount of data to be transported when averaged over time
Maximum latency	Maximum allowable time between ingress of packet to convergence sub-layer and the forwarding of SDU to air interface
Tolerated jitter	Maximum delay variation for the connection
Traffic priority	Priority assigned to service flow

Table 2-3: QoS definitions

Application	QoS Category	QoS Specifications Used
VoIP	rtPS	Maximum Sustained Rate Maximum Latency Tolerance Priority
NRTV	rtPS	Maximum Sustained Rate Maximum Latency Tolerance Priority
FTP	nrtPS	Maximum Sustained Rate Minimum Reserved Rate Priority
WWW	BE	Maximum Sustained Rate Priority

Table 2-4: Specifications for the traffic types implemented in system level simulator

2.8   Bandwidth Request and Assignment in Mobile WiMAX

Whenever the mobile has multiple connections with the base station it has some control over bandwidth in order to share the resources among the connections. Scheduling in downlink and uplink is done by the base station. In the downlink the base station is the only one transmitting during the downlink sub-frame as it schedules all downlink connections. It allocates bandwidth

according to the needs of the incoming traffic in its buffers, without mobile station intervention. The bandwidth assignments are broadcast to all mobile stations in the DL-MAP and the mobiles will know exactly when to receive its own packets. In uplink the base station allocates bandwidth based on requests sent from each mobile station wishing to transmit and signals this allocation in the UL-MAP.

Figure 2-8: Example of service flow exchange in Mobile WiMAX

The mobile station has a plethora of ways to request bandwidth. Depending on the particular QoS and traffic parameters associated with a service type, the mobile station can combine the determinism of unicast polling with the responsiveness of contention-based requests and the efficiency of unsolicited bandwidth. A conventional way to request bandwidth from the mobile is for the base station to allocate dedicated or shared resources periodically to each mobile station, which can be used by these to send a bandwidth request PDU consisting of the bandwidth request header and the payload (polling). Polling can be done either individually (unicast) or in groups (multicast). When a mobile station is polled individually, it is allocated bandwidth to send bandwidth request messages. Multicast polling is a contention-based bandwidth request method used whenever there is insufficient bandwidth to individually poll many inactive mobiles. The allocation is multicast or broadcast to a group of mobile stations that have to contend for the opportunity to send bandwidth requests.

Figure 2-8 illustrates the mapping process from service applications into connection and service flows in the MAC layer and the association to the service classes described above [12].

2.8.1 Adaptive Modulation and Coding (AMC)

WiMAX supports a variety of MCS schemes for data transmission. The MCS scheme is selected according to the channel state and this selection can be performed on a burst-by-burst basis per link. On the downlink the base station uses the channel quality feedback indicator transmitted by the mobile station on the CQICH channel. On the uplink the base station estimates the channel quality based on the received signal. The MCS used is the one which maximizes the throughput while keeping the estimated block error rate (BLER) lower than the pre-defined threshold. This threshold depends on the type of service. AMC significantly increases the overall system capacity as it allows real-time trade-off between throughput and robustness on each link. The various modulation and coding schemes supported by WiMAX in downlink are: QPSK, 16 QAM and 64 QAM, which are mandatory for both Fixed and Mobile WiMAX. In uplink 64QAM is optional for Fixed WiMAX. FEC using convolutional coding is mandatory. The standard optionally supports turbo codes and low-density parity check (LDPC) codes at a variety of code rates as well.

2.9 Hybrid Automated Repeat Request (HARQ)

This is a combination of ARQ with FEC at the physical layer and provides for improved link performance over traditional ARQ, at the cost of implementation complexity. With HARQ retransmission is requested if the decoder is unable to correctly decode the received block. When a retransmitted coded block is received it is combined with the previously detected coded blocks and fed back to the input of the FEC decoder. The probability of success in the decoding of the data block is increased by combining the different replicas of the block. Two types of HARQ can be implemented: Chase Combining and Incremental Redundancy. Chase Combining was used in all simulations conducted in this work.

2.10 Fractional Frequency Reuse

WiMAX supports frequency reuse of one, i.e. all cells operate on the same frequency channel to maximize spectral efficiency. However, users on cell edge will suffer heavily degradation in connection quality due to inter-cell interference arising from co-channel use. In WiMAX users can operate on sub-channels which only occupy a small fraction of the whole channel bandwidth. This is achieved by sub-channel segmentation and the definition of permutation zones. A segment is a subdivision of the available OFDMA sub-channels and each segment is equal to a single instance of the MAC layer. The availability of sub-channelization schemes, such as PUSC, allows the coordination of sub-channel allocation to users at the cell edges in order to limit the inter-cell interference. This is achieved by grouping different sets of sub-channels in segments and by assigning each segment to a particular cell. In fractional frequency reuse, users with good channel conditions (measured as the SINR) have access to the full set of sub-channels, operating under a

frequency reuse of 1 and users with bad channel conditions will be allocated non-overlapping sets of sub-channels in order to preserve channel orthogonally, namely at the cell edge. This type of sub-channel allocation leads to the effective reuse factor taking fractional values greater than 1.In Figure 2-9 , F1, F2 and F3 represent different sets of sub-channels in the same frequency channel. The frequency reuse one is maintained for users in the centre of the cells, to maximize spectrum efficiency and fractional frequency reuse is implemented for users in the edge of the cells to assure edge-user connection quality.
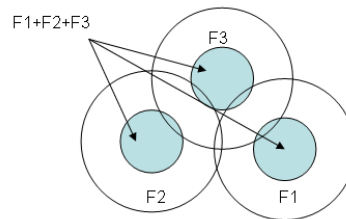
Figure 2-9: Fractional Frequency Reuse Implementation in Mobile WiMAX

2.11 Automatic Repeat Request (ARQ)

Hybrid-ARQ is an ARQ system that is implemented at the physical layer together with FEC, providing improved link performance over traditional ARQ at the cost of increased implementation complexity. The simplest version of H-ARQ is a simple combination of FEC and ARQ, where blocks of data, along with a CRC code are encoded using an FEC coder before transmission; retransmission is requested if the decoder is unable to correctly decode the received block. When a retransmitted coded block is received, it is combined with the previously detected coded block and fed to the input of the FEC decoder. Combining the two received versions of the code block improves the chances of correctly decoding. This type of H-ARQ is often called type I *chasecombining*. The WiMAX standard supports this by combining an *N*-channel *stop and wait ARQ* along with a variety of supported FEC codes. Doing multiple parallel channels of H-ARQ at a time can improve the throughput, since when one H-ARQ process is waiting for an acknowledgment, another process can use the channel to send some more data. WiMAX supports signaling mechanisms to allow asynchronous operation of H-ARQ and supports a dedicated acknowledgment channel in the uplink for ACK/NACK signaling. Asynchronous operations allow variable delay between retransmissions, which provides greater flexibility for the scheduler.

To further improve the reliability of retransmission, WiMAX also optionally supports type II H-ARQ, which is also called *incremental redundancy*. Here, unlike in type I H-ARQ, each (re)transmission is coded differently to gain improved performance. Typically, the code rate is

effectively decreased by every retransmission. That is, additional parity bits are sent every iteration, equivalent to coding across retransmissions.

2.12   Link Adaptation

The link adaptation allows a fair performance for the different applications and a good optimization of using the radio resources, realizing the QoS required for the transmission of the data streams. The link adaptation is an adaptive modification of the burst profile, mainly modulation and channel coding types, that take place in the physical link to adapt the traffic to a new radio channel condition. If the CINR decreases, change is made to a robust modulation and coding to improve the performance (data throughput); otherwise a less robust profile is picked up.

2.13   OFDM Pros and Cons

OFDM enjoys several advantages over other solutions for high-speed transmission.

• **Reduced computational complexity:** OFDM can be easily implemented using FFT/ IFFT, and the processing requirements grow only slightly faster than linearly with data rate or bandwidth. The computational complexity of OFDM can be shown to be $O(B \log BT_m)$, where B is the bandwidth and T_m is the delay spread. This complexity is much lower than that of a standard equalizer-based system, which has a complexity $O(B^2 T_m)$.

• **Graceful degradation of performance under excess delay:** The performance of an OFDM system degrades gracefully as the delay spread exceeds the value designed for. Greater coding and low constellation sizes can be used to provide fallback rates that are significantly more robust against delay spread. In other words, OFDM is well suited for adaptive modulation and coding, which allows the system to make the best of the available channel conditions. This contrasts with the abrupt degradation owing to error propagation that single-carrier systems experience as the delay spread exceeds the value for which the equalizer is designed.

• **Exploitation of frequency diversity:** OFDM facilitates coding and interleaving across subcarriers in the frequency domain, which can provide robustness against burst errors caused by portions of the transmitted spectrum undergoing deep fades. In fact, WiMAX defines subcarrier permutations that allow systems to exploit this.

• **Use as a multi-access scheme:** OFDM can be used as a multi-access scheme, where different tones are partitioned among multiple users. This scheme is referred to as OFDMA and is exploited in mobile WiMAX. This scheme also offers the ability to provide fine granularity in channel allocation. In relatively slow time-varying channels, it is possible to significantly enhance the capacity by adapting the data rate per subscriber according to the signal-to-noise ratio of that particular subcarrier.

• **Robust against narrowband interference:** OFDM is relatively robust against narrowband Interference, since such interference affects only a fraction of the subcarriers.

• **Suitable for coherent demodulation:** It is relatively easy to do pilot-based channel estimation in OFDM systems, which renders them suitable for coherent demodulation schemes that are more power efficient.

Despite these advantages, OFDM techniques also face several challenges. Firstly, there is the problem associated with OFDM signals´ having a high peak-to-average ratio that causes nonlinearities and clipping distortion. This can lead to power inefficiencies that need to be countered. Secondly, OFDM signals are very susceptible to phase noise and frequency dispersion, and the design must mitigate these imperfections. This also makes it critical to have accurate frequency synchronization.

2.14 Conclusions

This chapter presented an overview of Wireless Technologies especially IEEE802.16e, IEEE802.16j and IEEE802.22 and set the stage for more detailed exploration in subsequent chapters.

•WiMAX is based on a very flexible and robust air interface defined by the IEEE 802.16. IEEE 802.16 has formed a task group to extend the IEEE 802.16e-2005 standard to include multi-hop communication, indicating that the field has reached a significant level of maturity. This amendment is called IEEE 802.16j. The evidence of the change and evolution in the approaches of spectrum management can already be seen in the development of the IEEE 802.22 cognitive radio based standard for fixed, point-to-multipoint, wireless regional area networks.

• The WiMAX physical layer is based on OFDM, which is an elegant and effective technique for overcoming multipath distortion.

• The physical layer supports several advanced techniques for increasing the reliability of the link layer. These techniques include powerful error correction coding, including turbo coding and LDPC, hybrid-ARQ, and antenna arrays.

• WiMAX supports a number of advanced signal-processing techniques to improve overall system capacity. These techniques include adaptive modulation and coding, spatial multiplexing, and multiuser diversity.

• WiMAX has a very flexible MAC layer that can accommodate a variety of traffic types, including voice, video, and multimedia, and provide strong QoS.

•WiMAX has several features to enhance mobility-related functions such as seamless handover and low power consumption for portable devices.

•WiMAX offers very high spectral efficiency, particularly when using higher-order MIMO solutions.

2.15   References

[1]   IEEE 802.22 Working Group http://www.ieee802.org/22.R.

[2]   3GPP. Technical Specification Group Radio Access Network. High Speed Downlink Packet Access; Overall UTRAN Description. (3GPP TR 25.855 version 5.0.0).

[3]   3GPP TR 25.913. Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN). Available at http://www.3gpp.org

[4]   Wi-Fi Alliance (2005) Deploying WPA™ and WPA2™ in the Enterprise.

[5]   IEEE802.16j Relay's Task Group.2008; https://www.ieee802.org/16/relay/.

[6]   Diepenbeek C. Flexibility in frequency Management. 11th CEPT Conference, October

[7]   Shared   Spectrum   Company.   New   York   City   Spectrum   Occupancy MeasurementsSeptember2004,http://www.sharedspectrum.com/?section=presentations

[8]   Mitola J. Cognitive Radio for Flexible Multimedia Communications, MoMuC '99, pp. 3 –10, 1999.

[9]   Mobile WiMAX – Part I: A Technical Overview and Performance Evaluation, WiMAX Forum.

[10] Cicconetti, C.; Lenzini, L.; Mingozzi, E.; "Quality of Service support in IEEE 802.16 networks", IEEE Network Magazine, vol. 20, March 2006, pp. 50-55.

[11] IEEE. Standard 802.16e-2004.Part 16:Air interface for Fixed broadband wireless access systems . Amendment for physical and medium access control layers for combined fixed and mobile operation in licensed band. December 2004.

[12] IEEE. Standard 802.16-2004.Part 16:Air interface for Fixed broadband wireless access systems. October 2004.

Chapter 3: Methods for Simulation and Evaluation of Cellular Networks

This chapter gives the survey on the basics of simulation methodologies in wireless/cellular networks, type of simulation, simulation models and the requirement for modeling SLS available in the literature. Abstraction models for SLS will be presented and also explain that the implementation of a system level tool that requires complex modeling of several issues. To evaluate the performance of SLS a large number of statistics are collected for the computation of the metrics. These performance statistics are generated as outputs from the system level simulations and are used in the performance evaluation of the used scenarios and proposed algorithms.

3.1 Introduction

The high complexity of state-of-the-art and future cellular mobile networks makes their optimization and performance evaluation a challenging task. While some aspects in cellular networks may be evaluated analytically, testbeds or computer simulations are often the only way to go. Testbeds naturally deliver the most reliable results. However, they are difficult and expensive to install and maintain. Moreover, measurement trials are very time consuming, and it is very difficult to explore a large parameter space. Computer simulations are therefore a commonly applied method when it comes to evaluating the performance of cellular networks.

All relevant standardization bodies and industry consortia have specified reference scenarios, model parametrizations, and metrics in order to allow a comparison of simulation results generated by different research groups. These simulation specifications can be found in [1] for the 3GPP, in [2] for the WiMAX forum, and in [3] for the Next Generation Mobile Networks (NGMN) Alliance. The latter is an open alliance of leading network operators with the goal of influencing and pushing the development of next generation mobile networks.

The optimization of cellular networks can be done on many different levels and for a large number of different aspects. This includes static optimizations during the network planning process, as well as highly dynamic optimizations such as optimized queue management or scheduling decisions at the frame level. Many optimization problems that occur can be solved by means of simple problem specific heuristics. However, certain problems demand for a more generic solution approach.

Simulations in wireless and cellular networks can be classified in three categories (Link level, Network level, and System level) with respect to the targeted layers in the protocol stack and with respect to the considered level of abstraction as explained in chapter 1.

The choice of the right simulator category depends on the investigated problem and the desired results. Usually, it is advantageous to decouple simulators working at different levels of abstraction

in order to reduce complexity. In this case, appropriate interfaces are required between the simulators. However some problems may require the combination of different simulation levels in one simulator. Recently, network and system level simulations have grown together, owing to the increasing complexity of MAC protocols and algorithms, which more and more interact with adjacent cells, or depend on adjacent cells in their performance. Examples include interference coordination algorithms or certain aspects of scheduling algorithms.

Since SLS are often confused with system emulators and/or link layer simulators, these subsections aim to clarify their main differences, objectives and common functionalities. The first step in the design of a simulator or an emulator is to define the system to be analyzed. In this step it is necessary to determine the input data or initial conditions of the system, the system parameters or state variables and the performance metrics or output data of the system as shown in Figure 3-1.
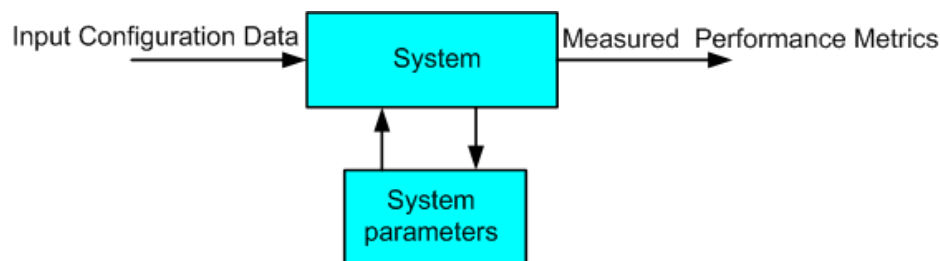
Figure 3-1: System to be simulated

In addition, their characteristics vary according to the type of system to be simulated or emulated. In general, a system simulator attempts to recreate the same system parameters by using abstract theoretical models. It also aims at collecting performance metrics to assess different properties of the system as show in Figure 3-2.

In contrast, an emulator does not aim at collecting performance metrics, but only at recreating system parameters. In general, an emulator dependency on abstract models is not as tight as that of a system simulator. In many cases an emulator is a piece of hardware that exactly reproduces all the desired system parameters as shown in Figure 3-3. For example the authors in [4] describe an emulator of WCDMA systems within the framework of the IST Wineglass project that is used to test RRM and scheduling algorithms.
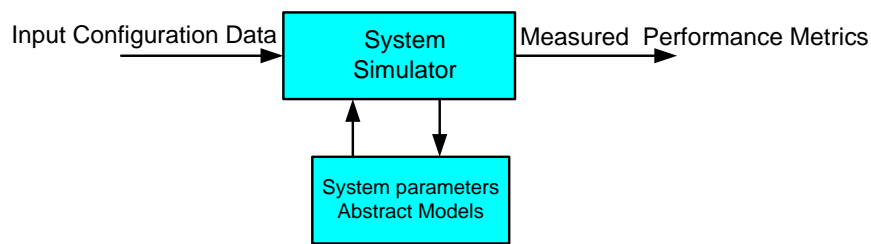
Figure 3-2: System Level Simulator
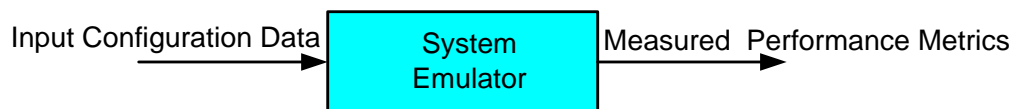
Figure 3-3: System Emulator

3.2   Simulation Terminologies

A simulation model is a particular type of mathematical model of a system. A mathematical model uses symbolic notations and mathematical equations to represent the system.
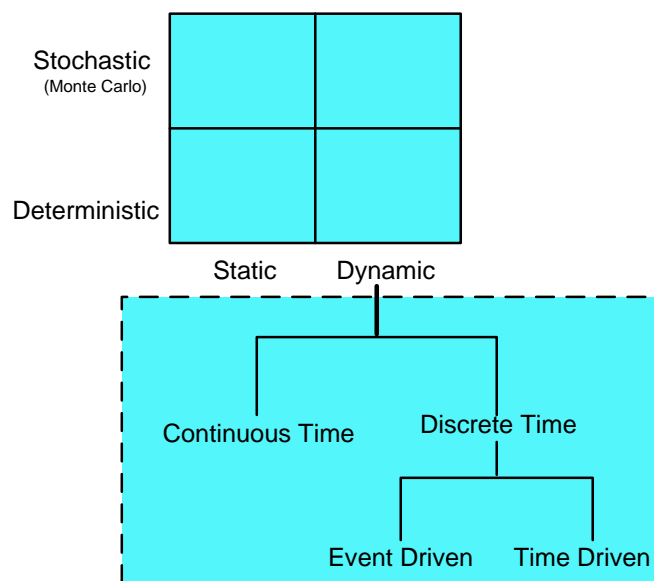
Figure 3-4: Classification of Simulations

Models can be classified in several ways as shown in Figure 3-4. The classifications are depicted along two main axes. Static and dynamic are on ´x´ axis, Stochastic and deterministic are on ´y´ axis of the model. The time evolution and the nature of the system inputs parameters describe the behavior of the above model. A model that attempts to describe the evolution over time is called

dynamic; otherwise, it is called static. The classical cellular planning tools, which are concerned with getting a static coverage map, are an example of static model.

If the nature of the system and /or its inputs is random, it is called stochastic otherwise, it is termed as deterministic. It is worth to mention that sometimes the word stochastic is not only used in connection with the nature of the system but also with the simulation procedure. Sometimes the simulation of the deterministic system may get so complex that it is better to simulate it stochastically. One example is to compute the decimals of the number ´π´. This is a deterministic problem in nature and ´π´ can be obtained through the appropriate series decomposition, but can also obtained stochastically, e.g. consider two random variables uniformly distributed over [0,1]. By generating very high number of pairs of variables and then verifying if the pairs fall inside or outside the unitary circle, the relative frequency of the point falling inside the circle tend to ´$\pi/4$´.

In the following, however, and for us, the meaning of stochastic will always be related to the nature of the system/inputs.Simulation models that contain no random variables are classified as *deterministic*. Deterministic Models have a known set of inputs, which will result in a unique set of outputs. A *stochastic* (Monte Carlo) simulation model has one or more random variables related to the nature of the system/inputs. Random inputs lead to random outputs. Since outputs are random, they can be considered only as estimates of the true characteristics of a model. In a stochastic simulation, the output measures must be treated as statistical estimates of the true characteristics of the system.

A *static* simulation model represents a system at a particular point in time. *Dynamic* simulation models represent systems as they change over time.

In the dynamic case we futher decompose the simulation as *Discrete-event systems simulation (time or event driven)* .The modeling of the systems in which the state variable changes only at a discrete set of points in time or in events. The simulation models are analyzed by numerical methods.

A *discrete model* is one in which the state variable(s) changes only at a discrete set of points in time. A *continuous model* is one in which the state variable(s) changes continuously over time.

3.2.1   Stochastic simulation (Monte Carlo simulation)

A *stochastic* (Monte Carlo) simulation model has one or more random variables related to the nature of the system/inputs. Typically, one is concerned in providing the statistics about the system behavior, although some specific techniques can eventually be of interest. These statistics are obtained by performing the averaging operation over large number of generated random triggers and inputs. The result is obtained by averaging over all results obtained from the generated random triggers and inputs, where all results must be obtained from uncorrelated sets of input values.

In mobile communication area, the Monte Carlo methods are used from Link to Network level simulations. One example is the deployment of mobile users in the cellular system. In this case, mobile terminals are randomly placed in the scenario (the mobiles are dropped onto the cells), and the resulting signal and interference situation are determined. The above-described paradigms (Figure 3.4) can be combined to hybrid approaches. A popular example is the combination of event-driven simulation with time driven simulation. In such an approach, an event-driven simulation is performed for every drop (TTI). For example, the mobile terminals could be repositioned for every drop, and a short event-driven simulation could be executed to obtain certain performance metrics including time-dependent processes such as ARQ. This simulation type will be used in this dissertation to produce simulation results and statistics.

3.3   System Level Simulator

Figure 3-5 show a system with a multiple cells and there is one base station and a random number of mobiles, randomly located inside each cell. Mobiles access the network through Base station (BS's). A ccomprehensive system level tool is needed to evaluate the performance of such system, which captures every aspects of the real cellular environment. This kind of tool is called System Level Simulator (SLS).A single simulator approach would be preferred, but the complexity of such simulator (including everything from transmitted waveforms to multi-cell network) is far too high with the required simulation resolutions and simulation times. In addition, domains of simulations of both simulators are very different, i.e. link level simulates a single link, one base station and one user whereas system level simulator simulates a multiple links as shown in Figure 3-5.

At link level, the granularity is in order of BER or if CDMA its chip error rate, which can be as low as nanoseconds, while at the network level it is at the level of packet durations, which is typically several orders of magnitude higher than the BER. Due to these computational complexities, it is not possible to simulate all networks in a one System level Simulator. Therefore, separate link and system level simulations are needed with an appropriate methodology to pass the results from the link level to the SLS.

This is the so-called Link-to-System (L2S) interface as shown in Figure 3-6,.which will be discussed in more detail in chapter 4. In practice, this interface is realized through a set of mapping tables known as Look up Tables (LUT). These mapping tables are constructed on the link level and they represent tabulated BER or FER function of instantaneous system level SINR.
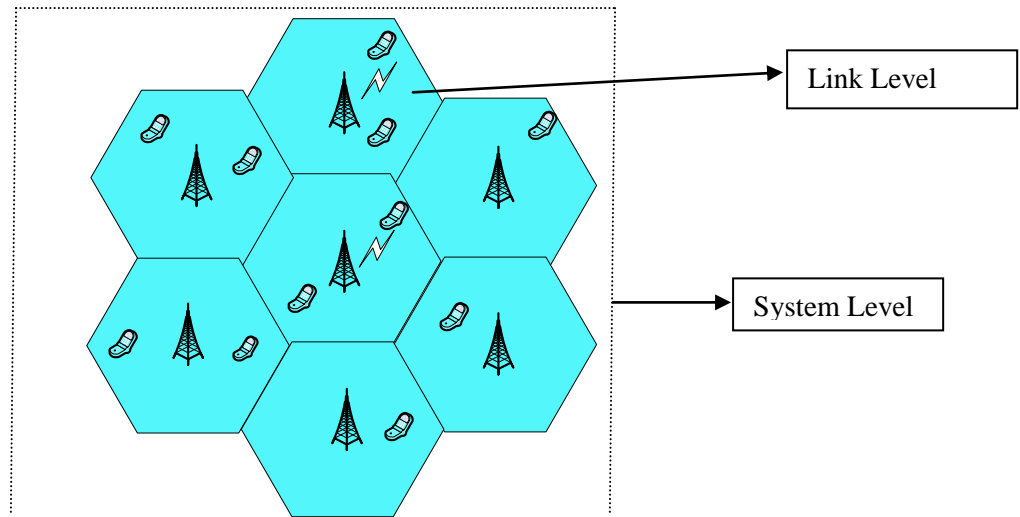
Figure 3-5: Wireless System Level Simulator
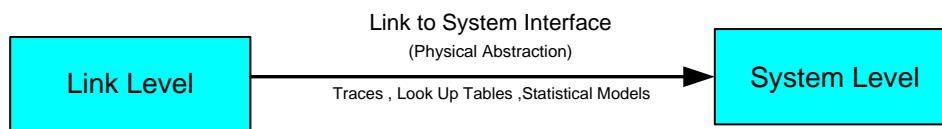
Figure 3-6: Interface between Link and System Level

3.3.1   Requirements for SLS Modeling

Implementation of a system level tool requires complex modeling of several issues. As shown in Figure 3.5, several aspects need accurate modeling. The minimum modeling involves three aspects:

- The users' behaviors;
- The traffics' behavior for the application considered;
- The radio aspects involved in the transmission of the signals from users to destination.

When considering the users' behavior, as their position can be random, there is need for a deployment model. Secondly, as they are mobile, there is need for a mobility model.Thirdly, as the user uses wireless communication to run application or services that lead to traffic pattern, there is need for a traffic model. Finally, the information is sent by radio waveform and there is need for the appropriate propagation models.

Therefore, the main modeling requirements for the SLS as shown in Figure 3-7 are:

- Deployment Modeling
- Mobility Modeling
- Traffic Modeling

- Propagation Modeling
- Transmission/reception techniques

This will be discussed in more detail in chapter 4.
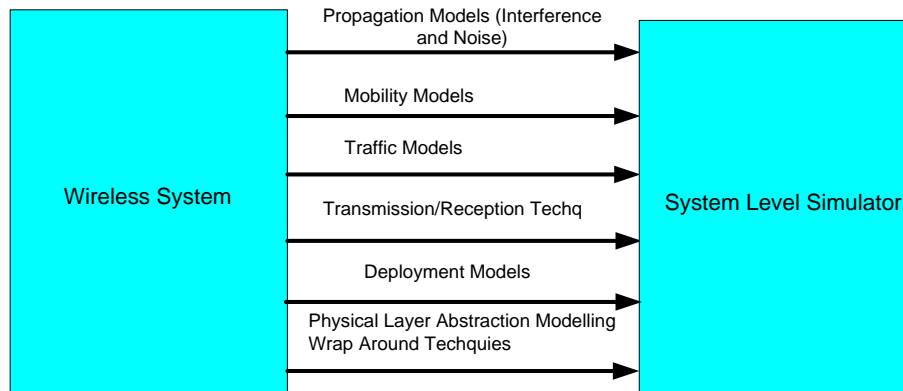
Figure 3-7: Modeling of Wireless communication System

3.3.2 Deployment Modeling

In order to determine the overall performance and capacity of a wireless network SLS that model the network with multiple BSs and MSs are required as shown in Figure 3.5.

We consider the hexagonal deployment. In each cell there is one base station and a random number of mobiles, randomly located inside the cell. Mobiles access the network through Base station (BS's). Deployment modeling will be discussed in more detail in chapter 4.

3.3.3 Mobility Models

Simulations with moving terminals require a suitable mobility model. For a scenario based on empirical channel models with no distinct terrain features, two popular mobility models are the random waypoint and the random direction mobility model. In the random waypoint mobility model illustrated in Figure 3-8(a), a target location within the movement plane is randomly determined. The terminal then moves to that position on a straight line and with a randomly chosen constant speed. The major drawback of this model is its non-uniform spatial terminal distribution, as terminals are more likely to be encountered in the center area [5]. In the random direction mobility model, the direction of movement and the length of the path segment are randomly chosen. The terminal moves with a randomly chosen constant speed until the end of the path segment is reached. Then a new direction is randomly chosen where a maximum turning angle can be set, as shown in Figure 3-8(b). The free path length is selected based on a given distribution function resulting in a particular mean free path length. Special actions have to be taken if the mobile terminal hits the scenario border. A uniform node distribution is achieved if the terminal

bounces off the scenario border like a billiard ball, or if a wrap-around technique is used [6]. In all models described above, a randomly chosen velocity may lead to problems, since slowly moving terminals take a much longer time to reach their destination compared to fast moving terminals. Therefore, the scenario will contain more slowly moving terminals than fast moving terminals once it has reached its steady state. If all terminals move at a constant speed, this problem naturally does not arise.

(a) Random waypoint   (b) Random direction
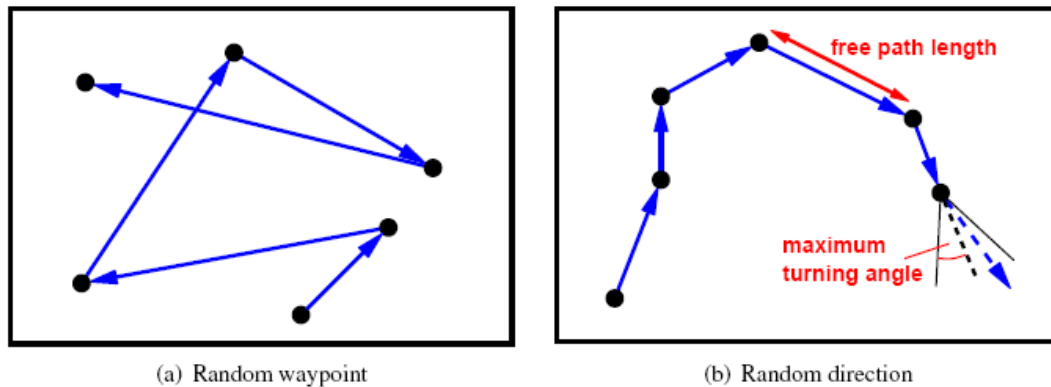
Figure 3-8: Illustration of commonly used terminal mobility models.

3.3.4 Traffic Models

The application traffic models are categorized into two types of traffic modeling: foreground and background traffic. The foreground traffic model represents a specific user behaviour or interaction with a device, whilst the background traffic is not directly related to a user interaction. Traffic models are represented for both user level and IP packet. Traffic modeling will be discussed in more detail in chapter 4.

3.3.5 Propagation Models

The term channel refers to the medium between the transmitting antenna and the receiving antenna. The characteristics of wireless signal changes as it travels from the transmitter antenna to the receiver antenna as shown in Figure 3-9. These characteristics depend upon the distance between the two antennas, the path(s) taken by the signal (i.e pathLoss), and the environment (buildings and other objects) around the path (i.e shadowing and fading). 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.This section gives the general overview of channel modeling in the system level simulator. Channel modeling will be discussed in more detail in chapter 4.

Wireless propagation models can be classified into two categories. *Deterministic models* usually make use of ray-tracing or related techniques [7]. They are based on a more or less precise

topographical representation of a certain area, which includes buildings and elevation information. The path loss is calculated by considering the physical paths along which the electromagnetic waves propagate, taking into account reflections, diffractions, and scattering. Deterministic models therefore deliver a very good map of the signal strength in a specific terrain, such as a particular city, which comes very close to reality. Usually, deterministic models include the effects of shadowing, but not the effects of fast fading. They are computationally very expensive and require a detailed database of the environment. They are commonly used for system level simulations and radio network planning, where exact results for a particular area are most important.

Empirical models are statistical models, which are generally based on results from measurement campaigns. In contrast to deterministic models, they do not take into account terrain features of a particular area. Empirical models have a much lower computational complexity compared to deterministic models and can be used for link, network, and SLS.
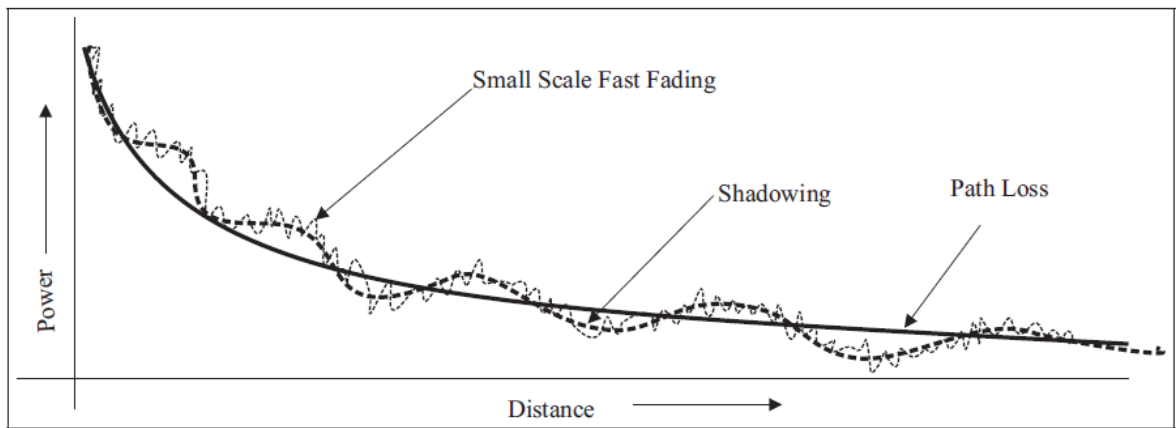
Figure 3-9: Wireless Propagation Channel Model

3.3.5.1 Path loss Models

Various empirical path loss models have been developed, usually by curve fitting of data obtained by measurements. A measurement-based approach implies that the model is valid for a certain parameter range, namely the range that the measurement data sets cover. The Okumura-Hata model is one popular empirical path loss model. In [8], Hata presented the following $R(\Delta x) = e^{-\frac{\Delta x}{d_{corr}} ln2} \sigma_2$ empirical formula for the path loss calculation as shown in equation 3.1:

$$L_{p\,(dB)} = A + B log_{10} d \qquad (3.1)$$

Hata obtained the parameters A and B from measurements performed by Okumura et al. in [9]. Influencing factors are the terrain category, the carrier frequency, and the antenna heights. The Okumura-Hata model can be used to model the path loss in macro cells with path lengths between 1 and 20 km and frequencies between 100 and 1500 MHz, though extrapolations to higher frequencies are possible. Mogensen et al. [10] proposed a modified Hata prediction model for the frequency range 1500−2000 MHz based on measurements in urban areas, which was further refined in [11]. Since this model was developed in the course of COST action 231, it is known as COST231-Hata model. COST 231 also proposed a combination of the Walfisch [12] and the Ikegami [13] models, known as COST231-Walfisch-Ikegami model. In contrast to the models before, it is suitable for smaller cells with a path loss distance in the range of 0.02 and 5 km. More recently, Erceg et al. [14] proposed a path loss model based on measurements in the 1.9 GHz band with possible path lengths between 0.1 and 8 km. It can be parametrized to model a variety of different scenarios. In particular, the authors also account for the variation of the path loss exponent between different cells. The path loss is given in an equation 3.2 [15]:

$$L_{p\,(dB)} = A + 10 \underbrace{\left(a - b h_b + \frac{c}{h_b} \right) log_{10} \left(\frac{d}{d_0} \right)}_{\text{Distance dependent Path Loss}} + \underbrace{10 x \sigma_\gamma \left(\frac{d}{d_0} \right)}_{\text{Cell individual variation}} + \underbrace{y \mu_\sigma + y z \sigma_\sigma}_{\text{shadowing}} \; ; d \geq d_0$$

$$\text{Distance dependent Path Loss} \qquad \text{Cell individual variation} \quad \text{shadowing} \qquad (3.2)$$

With

$$A = 20 log_{10}(4\pi \frac{d_0}{\lambda}) \qquad\qquad (3.3)$$

and $d_0 = 100$ m. $a, b, c, \sigma_\gamma, \mu_\sigma$ and σ_σ are model parameters which depend on the terrain category.

Three terrain categories are defined in [16], namely Hilly/Moderate-to-Heavy Tree Density (terrain category A), Hilly/Light Tree Density or Flat/Moderate-to-Heavy Tree Density (terrain category B), and Flat/Light Tree Density (terrain category C). $x, y,$ and z are independent zero-mean Gaussian variables with unit standard deviation. x and z vary from cell to cell, where y varies from location to location.

3.3.5.2 Shadowing Models

Gudmundson proposed a simple empirical model for temporally correlated shadowing in [17]. He used a simple correlation function, which exponentially decreases with the travelled distance. A model verification showed that the results obtained with this model are consistent with measurements obtained from urban and sub-urban environments. The correlation of the shadowing

values in the logarithmic domain can be expressed depending on the de-correlation length d_{corr} as follows [18]:

$$R(\Delta x) = e^{-\frac{\Delta x}{d_{corr}}ln2}$$

(3.4)

The de-correlation length d_{corr} is the distance, where the correlation has dropped to 1/e, i.e., $R(d_{corr})$=1/e. If a lognormal distribution of the shadowing values is desired, an auto-regressive filter of first degree can be used for their generation. If the filter operates in the logarithmic domain, its input are independent and identically distributed (i.i.d.) random variables, which are normal distributed with mean 0 and standard-deviation σ_1 .The output of the filter is also normal distributed with zero mean and standard deviation σ_2 . The multiplication factor R(Δx)determines the correlation between two subsequent samples at the filter output, where Δx is the distance between the two locations the samples are drawn for. In order to calculate the multiplication factor a, a further constraint has to be defined. It is thereby practical to require σ_1 to be identical to σ_2 .

3.3.5.3 Fast fading Models

Flat fading is the simplest form of fast fading. An appropriate fast fading model has to take into account the temporal correlation of the fading values. Two popular models for flat fading channels are filtered Gaussian noise and the sum of sinusoids method [19]. The sum of sinusoids method is very intuitive, as it constructs a complex fading envelope from the superposition of several complex sinusoids with different angular Doppler frequency and phase shift as shown in Figure 3-10.

In a frequency selective environment, the variation of the channel transfer function in frequency direction has to be taken into account. Frequency selective fast fading models are usually based on tapped delay lines, where the complex fading envelope results from a superposition of several taps representing signal components arriving at different times. Höher proposed a model where the complex envelope is a superposition of several complex sinusoids with different angular Doppler frequency, time shift, and phase shift [20]. The model proposed by the ITU-R in [21] constitutes of a superposition of a much smaller number of taps, where each tap undergoes Rayleigh fading [19]. For MIMO systems, the spatial correlation has to be considered as a further component. A popular model for this purpose proposed by the 3GPP and 3GPP2 is the Spatial Channel Model (SCM) [22]. SCM was developed for a carrier frequency of 2 GHz and a signal bandwidth of up to 5 MHz. The Spatial Channel Model Extended overcomes these limitations, with a carrier frequency range of 2−5 GHz and a bandwidth of up to 100 MHz [23], and it also adds further extensions. SCME was developed in the course of the European IST Winner project, which also developed two new

models in project phase 1 [24, 25] and phase 2 [26], respectively. A comparison of all four models can be found in [27].
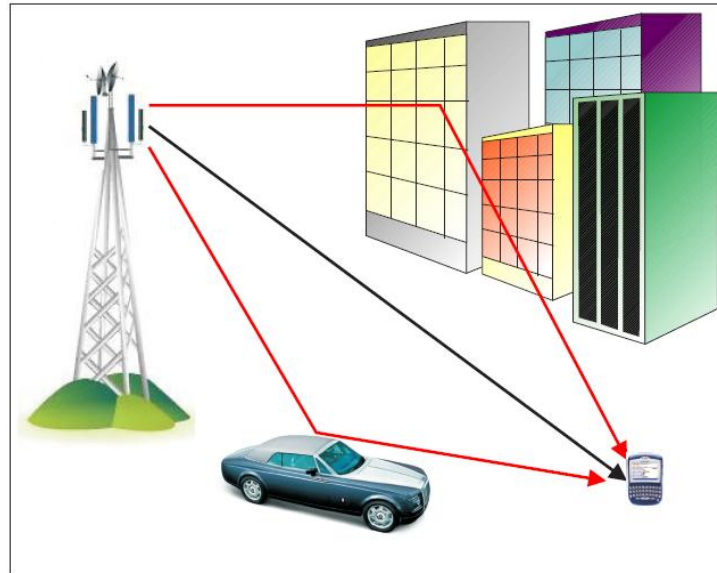
Figure 3-10: Multipath Propagation

3.3.5.4 Interference and Noise Modeling

Thermal noise is commonly modeled as Additive White Gaussian Noise (AWGN). This means the noise directly adds to the received signal's amplitude, its power spectral density is not frequency dependent, and the noise samples are i.i.d. and follow a Gaussian distribution. While this is quite simple, interference modeling is more difficult. There exist different types of interference. Intra-cell interference essentially adds a constant interference offset and is mostly independent of the applied scheduling and resource assignment strategy, except for MAI that occurs when several terminals are scheduled at the same time in a CDMA system. In contrast, inter-cell interference largely depends on the scheduling and resource assignment strategy, and thus also on possibly applied interference coordination schemes. It therefore has a large impact on the system performance. Consequently, in the remainder of this monograph, only inter-cell interference will explicitly be considered. The modeling of inter-cell interference depends on the air interface technology, and there mostly on how neighboring cells are separated. In CDMA systems like cdma2000 or UMTS, neighboring cells are separated by Pseudo Noise (PN) sequences, and the data signals are spread across the whole bandwidth. It is therefore valid to assume the inter-cell interference as additional AWGN. In FDMA/TDMA or OFDMA systems, this is more difficult. A commonly used approach to model inter-cell interference is to assume it as an additional source of additive Gaussian noise. The interference power is calculated using the same propagation models

that are used for the desired signals. This needs to be done separately for every time and frequency slot depending on the resource allocation in the interfering cells. The interference is therefore not white. Plass et al. have evaluated the accuracy of this approach in a 3GPP LTE environment [28] and found it to be a good approximation of the real interference situation. Exceptions are the application of advanced signal processing algorithms that rely on the exact interfering signal properties, such as interference cancelation or optimal combiners. Furthermore, the Gaussian approximation also does not hold if there is a highly dominating interferer. However, in such a situation, the SIR is relatively bad anyway, leading to a very low or even no throughput, and hence to a relatively low impact of this inaccuracy.

3.3.6   Wrap around technique

Cellular scenarios can be modeled in a number of different ways. The choice of the model depends on the problem under investigation and the required level of detail. The simplest possibility is a single-cell model in which one cell is modeled in detail and effects from other cells are taken into account by abstract models. For example, traces or statistical models can model other-cell interference. There may even be problems where effects of other cells only have an insignificant impact and can therefore be neglected. If a single-cell model is not sufficient, there are two different possibilities to model a scenario with multiple base stations.

First, a real environment with actual buildings, streets, etc. can be used. Base stations are placed at their real geographic coordinates, and mobile terminals move along distinct paths, such as streets. Such a model is typically used in conjunction with a deterministic propagation model. The computational complexity of this approach is relatively high. Alternatively, a more abstract model based on hexagonal cells and empirical propagation models can be used. An example is illustrated in Figure 3-11(a), which shows a scenario with 19 base stations, each serving three cell sectors. This results in a total of $N_{cells} = 57$ cell sectors. Even though such a hexagonal model does not represent a real world scenario, it allows for an easy modeling of a multi-cellular scenario. This makes it well suitable as a standard scenario, and it has been adopted as the basic scenario in all three mentioned bodies ([1], [2], [3]).

In the scenario of Figure 3-11(a), the border cells receive less interference than the center cells. Hence, the results will differ in each cell depending on its position. There are several possibilities ties to treat this problem. The easiest way is to evaluate only the center cells while the results of all other cells are ignored. This, however, wastes a lot of computing power, since only a fraction of the simulated cells is evaluated. A second possibility is to surround the scenario by interference generating cells, which generate interference based on a statistical model. In this case, all cells that are simulated in detail can be evaluated, even though border cells may still operate under different

conditions compared to center cells. A variant of such a method was presented by Berger et al. in [32], where other-cell interference within a cell was modeled based on the cell geometry factor and the terminal's line of sight angle to its serving base station. Last but not least, wrap around is an attractive technique.

(a) scenario without wrap-around   (b) identical layout with wrap-around
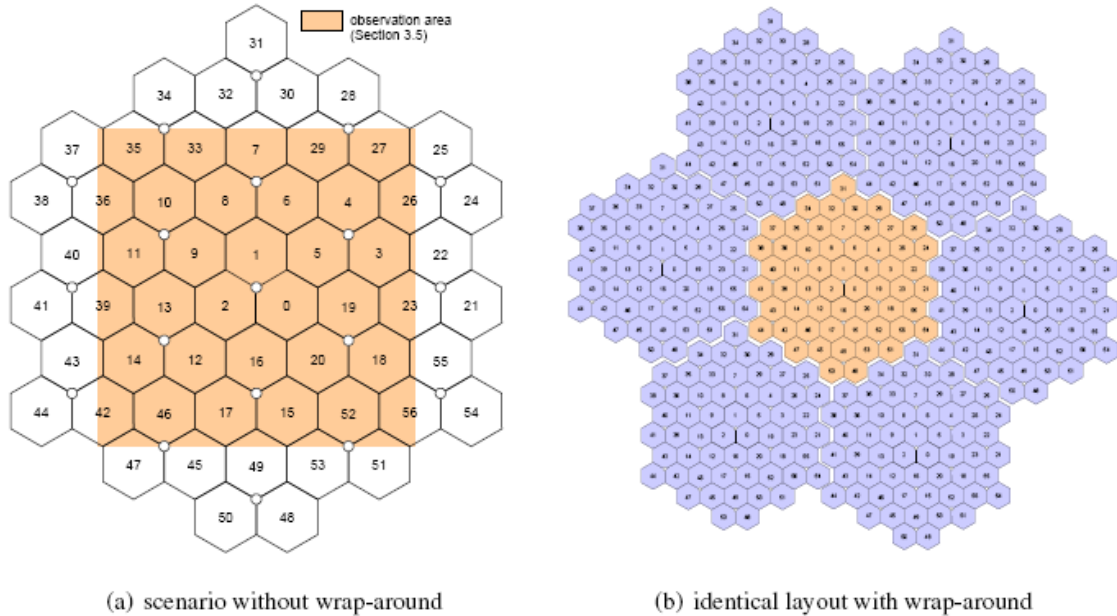
Figure 3-11: Cellular scenario with 19 base station sites and three sectors per base station.

Wrap around virtually repeats all base stations infinitely often in a regular grid, as it is illustrated in Figure 3-11(b). This corresponds to a mapping of the scenario in Figure 3-11(a) onto a torus. Mobile terminals and electromagnetic waves that leave the scenario re-enter it at the opposite side, where the impact of every transceiver is taken into account only once, namely by choosing the one closest to the point of interest. Hence, cell sectors number 27 and 47 now have a common borderline and cause interference on each other. With wrap around, all cells are equal, and measurements can be taken in every cell. This makes it a very efficient and widely used model.

3.3.7 Physical Layer Abstraction Models

On top of the model of a wireless channel and a cellular network, protocols and other building blocks need to be modeled in order to be able to evaluate the performance of an end-to-end connection in a network level simulation. This involves a large number of modeling decisions. First, the level of abstraction needs to be determined. Usually, network level simulations extend from the MAC layer to some higher protocol layer, making necessary an interface between the physical and the MAC layer. However, other levels of abstraction are possible. A higher level was chosen in [33] and [34], where the authors present models for the delay of IP packets in UMTS

networks. On the one hand, such models require more simplifications. On the other hand, they allow a fast and efficient simulation of higher layer protocols. If the model reflects all delay and loss effects that have an impact on a specific higher layer protocol, it can efficiently be used for their design and evaluation [35]. It is also possible to choose the abstraction below the MAC layer. In this case, certain parts of the physical layer would explicitly be simulated. Even though this brings along a high computational complexity, it may be necessary for systems where it is difficult to find suitable abstractions physical layer processes. MIMO systems are an example where the abstraction of physical layer issues becomes more and more difficult. Nevertheless, network level simulations starting at the MAC layer are very common and can efficiently be applied to many wireless technologies. A suitable interface between the physical and the MAC layer is required.

The interface between the physical and MAC-layer depends very much on the considered radio technology. Typically, an adequate model for the block error process on the MAC layer is developed, which is then parametrized by separate link level simulations. The model has to reflect all relevant characteristics of the block error process. This includes second-order statistics, such as the correlation of block errors. Models for block error processes can be put into one of the following three basic classes as shown in Figure 3.5.

3.3.7.1 Statistical models

The simplest statistical model is a random drop model with constant drop probability. Such a model is well suitable for certain systems, such as CDMA systems with perfect power control [36]. Even if it does not accurately reflect the link behavior in other systems, it may be a valid simplifying assumption for certain studies. More advanced statistical models were originally developed to represent bursty bit errors in link level simulations. The first model was presented by Gilbert in [37]. It is based on a Markov chain with two states, representing a good and a bad channel state, respectively. In the good channel state, no errors occur. In the bad channel state, symbol errors occur with a certain probability. This model captures the correlation and burstiness of symbol errors. Elliott [37] extended the model such that symbol errors may also occur in the good state with a certain probability, yielding the well-known Gilbert-Elliot channel model. The same approach can also be used to model the block error process on the MAC layer. For example, Wang and Moayeri [30] and Bergamo et al. [31] discuss finite-state models for the Rayleigh fading channel. Tralli and Zorzi developed models for the block error process within a WCDMA system based on N-state Markov chains that adequately capture the correlation between block errors.

Statistical models are easy to implement and have a low computational complexity. However, their development and parametrization is difficult. Moreover, they are restricted to relatively simple interfaces between the physical and MAC layer. Advanced features such as HARQ can be modeled only roughly, and it is time-consuming to model correlations among different connections.

3.3.7.2 Traces

In contrast to statistical models, traces constitute a direct interface between a link and a network level simulation. Typical traces obtained from link level simulations contain a block error probability for every time step. Trace-based models are easy to implement and have a low computational complexity. However, different traces must be generated for different scenarios, such as different terminal velocities, and trace files may be very large. It is also difficult to capture feedback effects from upper layers to the physical layer. In contrast to statistical models, it is easier to model advanced features such as HARQ with traces. On the other hand, traces also lack a detailed interference model, making it more difficult to model correlations among different connections.

3.3.7.3 Lookup tables

A very common approach to a link level interface is based on lookup tables, which map an SINR-value to a specific loss probability. The loss probability is usually given per FEC-block. If a data block on the MAC layer is composed of multiple FEC blocks, a separate calculation of the overall loss probability has to be done. Lookup-table based approaches require an SINR value for every FEC-block, which can either be obtained from traces, or by a direct calculation of the SINR in the simulation. Fast fading can be considered in two ways. Hämäläinen et al. [38] refer to these two possibilities as average value interface and actual value interface, respectively. The average value interface was originally discussed in [28] for system simulations of next generation TDMA systems. This interface type includes fast fading in the lookup tables, whose input is thus an average SINR value. In the actual value interface, instantaneous lookup tables are used, which do not average out the effects of fast fading. This requires the effects of fast fading to be included in the input SINR values.

While [32] and [31] consider flat fading channels, additional steps may be necessary when considering wide band frequency selective fading channels. If the data symbols of an FEC block experience different fading conditions, an effective SINR value for the table lookup has to be calculated. This may be the case if the FEC block is spread over a large enough frequency or time range. A commonly used approach is the Exponential Effective SINR Mapping (EESM) [39].

Lookup table based approaches are relatively accurate, and they allow modeling advanced features such as HARQ . They have a low computational complexity, despite the possibly time consuming calculation of (effective) SINR values. On the downside, link level simulations are necessary to obtain the lookup tables. Moreover, separate lookup tables have to be obtained for different modulation and coding schemes, for different FEC block sizes, and for different scenarios.

3.4   Evaluation and Metrics

To evaluate the performance of any system, outputs metrics and statistics are needed. To evaluate the peformace of SLS a large number of statistics are collected for the computation of the metrics. These performance statistics are generated as outputs from the system level simulations and are used in the performance evaluation of the used scenarios and proposed algorithms, which will be discussed in more detail in Chapter 4.

This section overviews relevant performance metrics, which are available in the literature and discusses their merits and limitations. The different metrics are grouped into two classes by the ITU-T and the IETF. The first class consists of objective performance metrics, referred to as Grade of Service by the ITU-T and Quality of Service (QoS) by the IETF. Objective metrics comprise all data that can be measured directly, such as packet delay and jitter, throughput, packet loss, and the like. Moreover, metrics that can be inferred from this data are considered objective metrics. In contrast, the second class of subjective performance metrics cannot be measured directly. Subjective metrics are collected under the terms Quality of Service (QoS) by the ITU-T and Quality of Experience by the IETF. The ITU-T defines subjective metrics as "the collective effect of service performances which determine the degree of satisfaction of a user of the service" [40]. Subjective metrics are always specific to a particular application or service. For some applications, it is possible to infer a subjective performance metric from measured objective metrics, such as for speech. However, most applications, like for example video applications, require extensive test trials with test users. In the following, only objective metrics will be considered.

Two important metrics have already been introduced, namely the overall spectral efficiency and the spectral efficiency at the cell edge. When multiplied with the available bandwidth, they yield the aggregate cell sector throughput and the cell edge throughput, respectively. While it is straightforward to determine the aggregate throughput in a cell sector, this is not the case for the performance at the cell edge. Clearly, the performance has to be measured for terminals that are close to the cell edge, but it is not obvious where exactly the cell edge area begins. One approach for this problem is to measure the cumulative distribution function (cdf) of all terminal throughput values in a cell sector, which basically shows how often a particular throughput result can be expected. In other words, if the cdf is obtained from a random and uniform user distribution, the cdf allows determining the probability that a randomly placed mobile terminal will achieve a certain throughput. [40]. Proposes to take the 5% quantile from this cdf as a scalar metric for the cell edge throughput. The 5% quantile indicates the throughput values of those mobile terminals with the lowest throughput in the scenario, which are in general located close to the cell edge. The NGMN Alliance [2] has also adopted this metric. In a Monte-Carlo simulation, the 5% quantile is

easy to determine, since every drop results in a set of throughput values for the randomly placed mobile terminals. This allows a simple calculation of the cdf. In an event-driven simulation with full terminal mobility, this is more difficult. In order to determine the quantile, the individual throughputs of the terminals within a short time period TSTP have to be measured. Subsequently, the cdf and the quantile of these short-term averages can be calculated. Obviously, the short-term period TSTP is an important parameter, which has a strong impact on the result. It must be long enough to even out effects of the MAC procedures, such as segmentation and retransmissions, but short enough to capture holes in the coverage.In addition to the just discussed throughput metrics, there are a number of indirect metrics that give an indication for the expected throughput without requiring a detailed system model.

Most important, the SIR and SINR directly relate to the maximum achievable spectral efficiency by the Shannon–Hartley theorem. Along with the SIR and SINR, it is important to study the occupied bandwidth in order to judge the resulting throughput. The occupied bandwidth is reflected in the resource utilization, where an occupation of the full available bandwidth corresponds to a resource utilization of 100%. Alike, a network with frequency reuse 3 has a resource utilization of up to only 1/3. It is obvious that an increase of the resource utilization counteracts an improvement of the SI(N)R, leading to a tradeoff between these two metrics. Note that the SI(N)R's mean value is a problematic metric, since it is dominated by high values close to the base station. If a single scalar value characterizing the SINR or SIR conditions is desired, it is better to consider the median, which is not susceptible to outliers of the SIR. Besides the above explained scalar metrics, it is also interesting to study position-dependent metrics. Position-dependent metrics illustrate the performance depending on the location of a mobile terminal. This is somewhat the case already for the throughput quantile, which depends on the position in the way that it corresponds to the throughput close to the cell edge. In order to obtain further insight into network coverage and position-dependent performance, it is convenient to plot for example the throughput or the SINR over the area.

3.5   Conclusions

In this chapter we explained all the necessary information for modeling SLS. Simulations in wireless and cellular networks can be classified in three categories (Link level, Network level, and System level) with respect to the targeted layers in the protocol stack and with respect to the considered level of abstraction and we also explained that the implementation of a system level tool requires complex modeling of several issues. The minimum modeling involves three aspects: the users' behaviors; the traffics' behavior for the application considered and the radio aspects involved in the transmission of the signals from users to destination.To evaluate the performance of SLS a large number of statistics are collected for the computation of the metrics. These performance statistics are generated as outputs from the system level simulations and are used in the performance evaluation of the used scenarios and proposed algorithms.

3.6   References

[1]   29 30 31   3GPP TS 25.814. Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA) (Release 7). 3rd Generation Partnership Project, June 2006.

[2]   WiMAX system evaluation methodology. Version 1.0, WiMAX Forum, January 2007.

[3]   NGMN performance evaluation methodology. Version 1.2, Next Generation Mobile Networks Alliance, June 2007.

[4]   T. Okumura, E. Ohmori, and K. Fukuda. Field strength and its variability in VHF and UHF land mobile services. Review of the Electrical Communication Laboratory, 16(9-10):825–873, September-October 1968.

[5]   J. F. Soufflet. The warehouse problem: A review. Regional and Urban Economics, 3(2):187–216, May 1973.

[6]   C. Boonthum, I. Levinstein, S. Olariu, E. Pigli, E. Shurkova, and A. Zomaya. Mobile computing: Opportunities for optimization research. Computer Communications,30(4):670–684, February 2007.

[7]   M. Hata. Empirical formula for propagation loss in land mobile radio services.IEEE Transactions on Vehicular Technology, 29(3):317–325, August 1980.

[8]   T. Okumura, E. Ohmori, and K. Fukuda. Field strength and its variability in VHF and UHF land mobile services. Review of the Electrical Communication Laboratory, 16(9-10):825–873, September-October 1968.

[9]   P. Mogensen, P. Eggers, C. Jensen, and J. Andersen. Urban area radio propagation measurements at 955 and 1845 MHz for small and micro cells. In Proc. IEEE Global Telecommunications Conference (GLOBECOM 1991),volume 2, pages 1297–1302, Phoenix, AZ, USA, December 1991.

[10] COST 231 Final Report: Digital Mobile Radio Towards Future Generation Systems. Office for Official Publications of the European Communities, Luxembourg,

1999.

[11] J. Walfisch and H. Bertoni. A theoretical model of UHF propagation in urban environments. IEEE Transactions on Antennas and Propagation,36(12):1788–1796, December 1988.

[12] F. Ikegami, S. Yoshida, T. Takeuchi, and M. Umehira. Propagation factors controlling mean field strength on urban streets. IEEE Transactions on Antennas and Propagation, 32(8):822–829, August 1984.

[13] V. Erceg, L. Greenstein, S. Tjandra, S. Parkoff, A. Gupta, B. Kulic, A. Julius, and R. Bianchi. An empirically based path loss model for wireless channels in suburban environments. IEEE Journal on Selected Areas in Communications, 17(7):1205–1211, July 1999.

[14] M. Gudmundson. Correlation model for shadow fading in mobile radio systems.

Electronics Letters, 27(23):2145–2146, November 1991.

[15] TR 101 112. Universal Mobile Telecommunications System (UMTS); Selection

procedures for the choice of radio transmission technologies of the UMTS. European Telecommunications Standards Institute, April 1998.

[16] G. L. Stüber. Principles of Mobile Communication. Kluwer Academic Publishers, 2nd edition, 2000.

[17] P. Hoeher. A statistical discrete-time model for the WSSUS multipath channel. IEEE Transactions on Vehicular Technology, 41(4):461–468, November1992.

[18] Recommendation ITU-R M.1225. Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000. International Telecommunication Union,February 1997.

[19] 3GPP TS 25.996. Spatial channel model for Multiple Input Multiple Output (MIMO) simulations (Release 6). 3rd Generation Partnership Project,September 2003.

[20] D. Baum, J. Hansen, and J. Salo. An interim channel model for beyond-3G systems: extending the 3GPP spatial channel model (SCM). In Proc. 61st IEEE Vehicular Technology Conference (VTC 2005-Spring), volume 5, pages3132–3136, Stockholm, Sweden, May 2005.

[21] H. El-Sallabi, D. Baum, P. Zetterberg, P. Kyosti, T. Rautiainen, and C. Schneider.Wideband spatial channel model for MIMO systems at 5 GHz in indoor and outdoor environments. In Proc. 63rd IEEE Vehicular Technology Conference (VTC 2006-Spring), volume 6, pages 2916–2921, Melbourne, Australia,May 2006.

[22] P. Kyösti et al. WINNER II interim channel model. Project Delivery IST-4- 027756 WINNER II D1.1.1 V1.1, November 2006.

[23] M. Narandzic, C. Schneider, R. Thomä, T. Jämsä, P. Kyösti, and X. Zhao.Comparison of SCM, SCME, and WINNER channel models. In Proc. 65th IEEE Vehicular Technology Conference (VTC 2007-Spring), pages 413–417,Dublin, Ireland, April 2007.

[24] L. T. Berger, T. E. Kolding, P. E. Mogensen, and L. Schumacher. Geometry based other-sector interference modelling for downlink system simulations. In Proc. 7th International Symposium on Wireless Personal Multimedia Communications (WPMC 2004), Abano Terme (Padova), Italy, September 2004.

[25] C. Bettstetter and C. Wagner. The spatial node distribution of the random waypoint mobility model. In Proc. German Workshop on Mobile Ad-Hoc Networks (WMAN), pages 41–58, Ulm, Germany, March 2002.

[26] C. Bettstetter. Mobility modeling in wireless networks: categorization, smooth movement, and border effects. ACM SIGMOBILE Mobile Computing and Communications Review, 5(3):55–66, July 2001.

[27] S. Plass, X. G. Doukopoulos, and R. Legouable. Investigations on link-level inter-cell interference in OFDMA systems. In Proc. Symposium on Communications and Vehicular Technology, pages 49–52, Liege, Belgium, November 2006.

[28] V. H. Mac Donald. Advanced mobile phone service: The cellular concept.The Bell System Technical Journal, 58:15–41, January 1979.

[29] M. C. Necker. A simple model for the IP packet service time in UMTS networks. In Proc. International Teletraffic Congress (ITC-19), pages 1475– 1485, Beijing, China, August 2005.

[30] M. C. Necker and S. Saur. Statistical properties of fading processes in WCDMA systems. In Proc. 2nd International Symposium on Wireless Communication Systems (ISWCS 2005), Siena, Italy, September 2005.

[31] S. Hämäläinen, P. Slanina, M. Hartman, A. Lappeteläinen, H. Holma, and O. Salonaho. A novel interface between link and system level simulations. In Proc. ACTS Mobile Communication Summit, pages 599–604, Aalborg, Denmark,October 1997.

[32] K. Littger. Optimierung: Eine Einführung in rechnergestützte Methoden und Anwendungen. Springer-Verlag, 1992.

[33] S. Hämäläinen, P. Slanina, M. Hartman, A. Lappeteläinen, H. Holma, and O. Salonaho. A novel interface between link and system level simulations. In Proc. ACTS Mobile Communication Summit, pages 599–604, Aalborg, Denmark, October 1997.

[34] M. C. Necker. A comparison of scheduling mechanisms for service class differentiation in HSDPA networks. International Journal of Electronics and Communications (AEÜ), 60(2):136–141, February 2006.

[35] 3GPP TSG RAN WG1#35 R1-031303. System-level evaluation of OFDM — further considerations. Technical document, Ericsson, Lisbon, Portugal, November 2003 .

[36] M. C. Necker and S. Saur. Statistical properties of fading processes in WCDMA systems. In Proc. 2nd International Symposium on Wireless Communication Systems (ISWCS 2005), Siena, Italy, September 2005.

[37] M. C. Necker and S. Saur. Statistical properties of fading processes in WCDMA systems. In Proc. 2nd International Symposium on Wireless Communication Systems (ISWCS 2005), Siena, Italy, September 2005.

[38] 3GPP TSG RAN WG1#35 R1-031303. System-level evaluation of OFDM — further considerations. Technical document, Ericsson, Lisbon, Portugal, November 2003 .

[39] Modelling of Performance with Coloured Interderence Using the EESM, Nortel Networks, 3GPP TSG-RAN WG# 37, 2004

[40] NGMN performance evaluation methodology. Version 1.2, Next Generation Mobile Networks Alliance, June 2007.

Chapter 4: System Level Simulator

This chapter provides all details followed into the design, modeling and implementation of the Link level and System Level Platform used in the realization of system level simulations for the Mobile WiMAX standard. Used traffic models for data generation as well as the channel models are provided. The cellular layout architecture used into the simulator is described. Of particular interest in the realization of system level simulations is the definition of the proper link to system level interface and the definition of the proper set of look up tables used in the mapping of physical layer performance. The procedure followed into the derivation of the Signal to Interference plus Noise Ratio (SINR) and the mapping function used to map the vector of SINRs into a single scalar to be inputted into the look-up tables is detailed.

4.1 Link Level Simulator

The development of SLS is presented in chapter 3 but due to computational complexities reasons link level is not included. Link level simulators try to model the behaviour of the system in a computationally efficient way. Performing link level simulations has a high computational cost. Due to this, these simulations are normally performed in advance, and the results obtained are stored. Then, those results can be easily used to model the PHY behaviour when other higher level issues want to be evaluated, avoiding lots of calculations.

As our system is based on OFDM, the Link level simulation is an OFDM chain with several options of coding, interleaving, punching and modulation. This simulator is targeted to the 1024-point FFT OFDM PHY layer as shown in Figure 4-1.Table 4-1summarize the simulation parameter for Link level simulation and calibration results are shown in Figure 4-2.
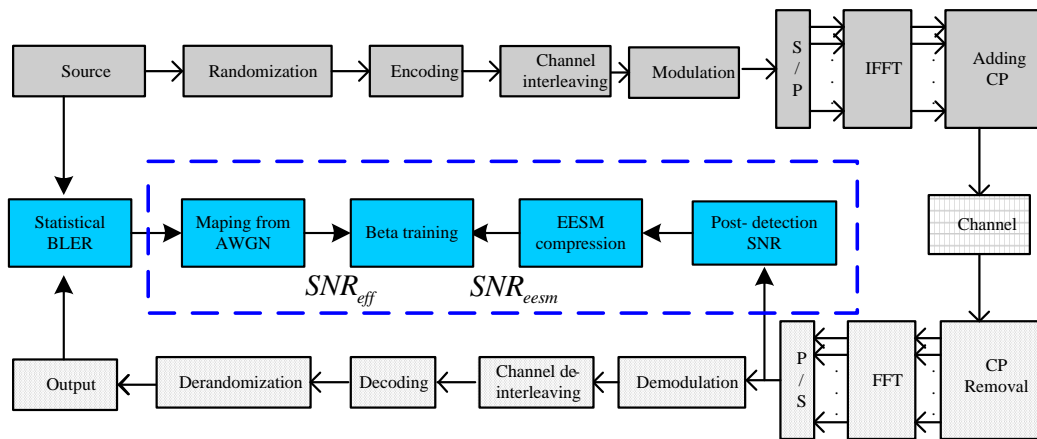
Figure 4-1: Link Level Simulator

Parameters	Values	
System Channel Bandwidth (MHz)	10	
Sampling Frequency (F_p in MHz)	11.2	
Subcarrier Frequency Spacing (f kHz)	10.94	
FFT Size (N_{FFT})	1024	
Coding	Convolutional (1/2,2/3,3/4,5/6) ,Turbo Coding(1/2,2/3,3/4,5/6) LDPC	
Modulation Schemes	QPSK,16QAM, 64QAM	
Channel	SISO/MIMO,3GPPchannel model(Ped B and VehA)	
	DL	UL
Null Subcarriers	184	184
Pilot Subcarriers	120	280
Data Subcarriers	720	576
Data Subcarriers per Subchannel	24	16
Number of Subchannels (N_S)	30	35
Useful Symbol Time ($T_b = 1/f$) in μs	91.4	
Guard Time ($T_g = T_b/8$) in μs	11.4	
OFDM Symbol Duration ($T_s = T_b + T_g$) in μs	102.9	
Number of OFDMA Symbols per frame (5ms)	48	
Data OFDM Symbols	44	

Table 4-1: Link Level Simulator
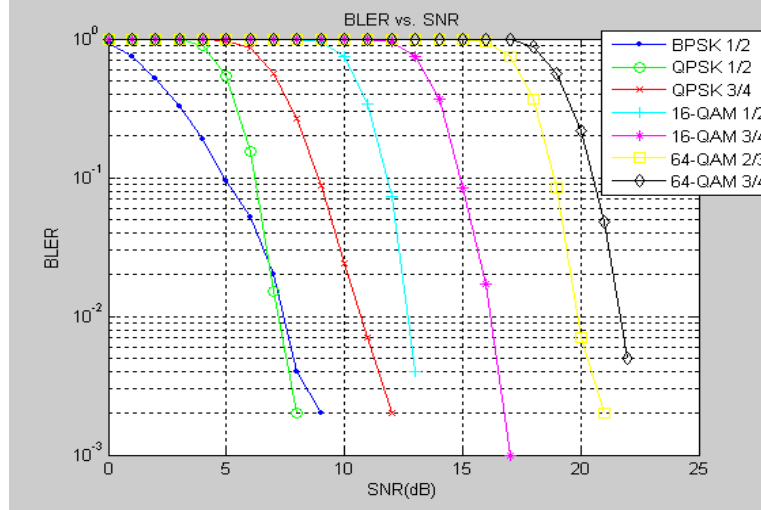
Figure 4-2: BLER vs. SNR plot for different modulation and coding profile on Veh A

4.2 Link layer Abstraction Procedure

The Block error rate (*BLER*) performance versus signal to inference and noise ratio (*SINR*) averaged over all channel realizations of one specific channel model has been widely used as the interface between the PHY- and system-level simulators. Two common interfaces are:

Average Value Interface – this type of interface reflects the radio link quality for a long time interval. This scenario is typical for mobile speeds corresponding to values of the coherence time smaller than the duration of a single transmission time interval, making it unrealistically to assume the channel constant along one or two radio frames. Only statistical channel behavior is assumed as channel state value is averaged over time. The average value interface is not accurate if there are fast changes in the interference due to, for example high bit rate packet users.

Actual Value Interface – this type of interface reflects the instantaneous value of the radio link.
It is suitable for scenarios of low mobility, resulting in a slow fading channel profile.

However, in many cases, the specific channel realization encountered may perform significantly different from the average performance. The abstraction method adopted in this dissertation is based on the EESM algorithm over link layer.The abstraction model (EESM) introduced is based on the concept of effective *SNR*. It takes the instantaneous channel and interference characteristics and other configurations into account and can provide the instantaneous performance of the real PHY layer system more accurately. The model consists of three main blocks [1] (see Figure 4-3).

- *Average SNR→Post-detection SNR*
- *Post-detection SNR→Effective SNR*
- *Effective SNR→BLER*

The first block covers the calculation of quality measures $SINR_p$ for all relevant Resource Elements. (p is the total number of resource).To facilitate further processing an appropriate transformation and compression of the $SINR_p$ to one single parameter SNR_{eff} is performed within the second block. The SNR_{eff} (or an equivalent metric BLER code word) is finally mapped to the Block Error Rate (*BLER*). The goal is to cover the resource- and power-allocation, the channel impacts and pre-/post-processing as well in the first part, i.e. by the extraction of the $SINR_p$. The following two steps have to account for the modulation and coding parameters. They should be designed in such a way that the final *BLER* mapping effectively models the performance of coding. In the following, the realization of the three parts will be explained in detail.
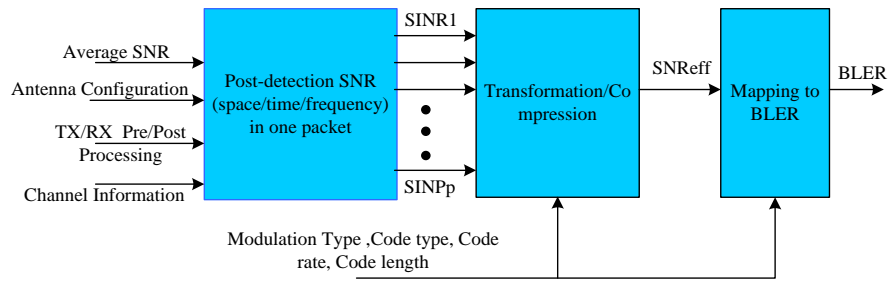
Figure 4-3: PHY-Abstraction Model

Block 1: *Average SNR→Post-detection SNR*

Basically the instantaneous SNR is chosen as a quality measure at System Level. In the first step, the computation has to cover at least a sufficient number of samples with respect to channel coherence conditions for the user of interest and its interferers as well. Therefore, the interface is capable to account for the instantaneous channel conditions (including fast fading). The SNR has to be understood as a post receiver Signal to Noise and Interference Ratio, i.e. after detection etc. In the following equation, P_{sig} stands for the power of the expected signals and P_{noise} stands for the power of white Gaussian noise involved in the system.

$$SINR_p = \frac{P_{sig}}{P_{noise}} \tag{4.1}$$

Post-detection SINR is calculated for each resource element, which means one for each time point and each frequency point. One packet is assumed to consist of p constellation points in the time/frequency space. All the SNR_p values belonging to a particular packet are "transformed/compressed/averaged" to a single effective *SNR*.

For a given channel allocation scheme one can reduce the number of SNR_p BLER code word according to the channel conditions, i.e. with respect to the variation of the fading in the time and frequency domain (coherence time and coherence bandwidth for frequency selective fading).

Block 2: *Post-detection SNR→Effective SNR*

The second step of this modeling approach is considered the most essential part. The concept of effective *SNR* is introduced in this part.

In this step, all the $SINR_p$ values belonging to a particular packet are "transformed / compressed / averaged" to a single effective *SINR*. One can say that it is another kind of averaging which is different from linear calculating.

For a given convolutional code and packet length *N,* the relationship between the packet error probability and signal-to-noise ratio can be determined for a conventional AWGN channel (with a constant *SNR,* via analysis or simulation. We assume that this relationship is known and represented as

$$PER_{AWGN} = h(SNR) \tag{4.2}$$

For a given L-dimensional vector SNR, which takes the instantaneous channel and interference characteristics and other configurations into account, we define the scalar effective SNR, as the SNR in an equivalent, constant signal-to-noise ratio AWGN channel which would yield the same packet error probability[2]. Thus, we rewrite the relationship as:

$$PER = f(SINR) = h(SNR_{eff})$$
$$SNR_{eff} = g(SINR) \tag{4.3}$$

The calculation of effective SNR is the most essential part, and it should cover such issues as current channel condition, the configuration of the antennas and the pre-and post- data processing in the real system so that the later processing of the model is independent of the channel and so on. The effective SNR is brought out as an ideal concept, there have been many ways to approximate the g function to get near to the concept, and EESM is one of them, which is explained in detail later.

Block 3: *Effective SNR→BLER*

As mentioned above, the effective *SNR* is defined as the *SNR* in an equivalent, constant signal-to-noise ratio AWGN channel, which would yield the same packet error probability. It is an ideal concept and the calculation approach described above is considered a good approximation for that. So, the mapping from mutual information to *SNR* should be simulated in the simple AWGN channel.

Since some adjustments have been done accordingly for every kind of modulation type in the second step explained above, the processing in the third step is expected to be independent of modulation scheme. That's to say, the mapping curves should be generated BLER "Code Triple"

(Coding Type, Code Rate and Code Word Length), without the consideration of modulation scheme. And also the results are kept with high resolution and can be sub-sampled according to the accuracy needed in the simulation. Following the three steps presented above, we can get the final block error rate for current configuration, channel condition, modulation and code scheme and so on. When HARQ with CC is introduced in the link layer, the post-detection SNR in the first step shall be calculated separately and combined first to get a new value as the input for the abstraction process. The other parts of the abstraction procedure remain the same.

4.3   EESM algorithm Introduction

The essential part of the link layer abstraction method is the nonlinear nature of the transformation results in an effective SNR. EESM algorithm adopted in step 2 for transforming process is introduced in detail.

4.3.1 Derivation of EESM

The exponential ESM is derived based on the Union-Chernoff bound of error probabilities [3]. The union bound for coded binary transmission and maximum-likelihood decoding is well known and given by equation (4.4)

$$P_e(\gamma) \leq \sum_{d=d_{min}}^{\infty} \alpha_d P_2(d, \gamma) \tag{4.4}$$

where

- γ is the channel symbol SNR,

- d_{min} is the minimum distance of the binary code,

- α_d is the number of code words with hamming weight d, and

- $P_2(d, \gamma)$ is the pair-wise error probability (*PEP*) assuming a certain Hamming distance d and a certain symbol SNR γ.

For BPSK transmission over an AWGN channel and using Chernoff-bounding techniques, the PEP can be upper bounded according to equation (4.5)

$$P_2(d, \gamma) = Q\left(\sqrt{2\gamma d}\right) \leq e^{-\gamma d} = P_{2,Chernoff}(d, \gamma) = \left[P_{2,Chernoff}(1, \gamma)\right]^d \tag{4.5}$$

The last part of the above expression implies that the Chernoff-bounded PEP is directly given by the Chernoff-bounded (uncoded) symbol-error probability. Thus, the Chernoff-bounded error probability $P_{e,Chernoff}(\gamma)$ only depends on the weight distribution of the code and the Chernoff-bounded symbol-error probability $P_{2,Chernoff}(1, \gamma)$ according to equation (4.6)

$$P_e(\gamma) \le \sum_{d=d_{\min}}^{\infty} \alpha_d P_2(d,\gamma) \le \sum_{d=d_{\min}}^{\infty} \alpha_d \left[P_{2,Chernoff}(1,\gamma) \right]^d = P_{e,Chernoff}(\gamma) \tag{4.6}$$

The basic principles for the Union Chernoff bound for a multi-state channel, i.e. a channel where different coded bits are subject to different SNR, is explained with the simple example of a 2-state channel. The principles are then straightforwardly extended to the general multi-state channel.

The 2-state channel is characterized by an SNR vector $\bar{\gamma} = [\gamma_1, \gamma_2]$ where, in general, the two states γ_1 and γ_2 occur with probability p_1 and p_2 respectively. Furthermore, the two SNR values are assumed to be independent from each other, which requires a corresponding interleaver in practice.

Let us now look at two arbitrary code words with Hamming distance d. The SNR value, either γ_1 or γ_2, associated with each of the d differing symbols depends on the respective symbol position. That means that the exact PEP for these two code words in case of a 2-state channel does not only depend on the distance d, but also on the position of the d differing symbols. Thus the union bound approach in the classical sense that all code-word pairs are compared would require detailed code knowledge about the bit positions. Instead, in this case the mean PEP, averaged over all possible positions of the d differing symbols, is used. This is equivalent to average over all possible cases how the SNR values γ_1 and γ_2 may be distributed among the d differing symbols. Hence, the Chernoff bounded PEP can be expressed as

$$P_{2,Chernoff}(d,[\gamma_1,\gamma_2]) = \sum_{i=0}^{d} \binom{d}{i} p_1^i p_2^{d-i} e^{-(i\gamma_1 + (d-i)\gamma_2)} = \left(p_1 e^{-\gamma_1} + p_2 e^{-\gamma_2} \right)^d \tag{4.7}$$

where the binomial theorem has been used to arrive at the final expression. To clarify the second expression, $p_1^i p_2^{d-i}$ represents the probability that i of the d differing symbols are associated with SNR γ_1 and the residual $(d\text{-}i)$ symbols are associated with SNR γ_2. There are $\binom{d}{i}$ such events and $e^{-(i\gamma_1 + (d-i)\gamma_2)}$ is the Chernoff-bounded PEP for such an event.

It can be noted that the term $p_1 e^{-\gamma_1} + p_2 e^{-\gamma_2}$ is the averaged Chernoff-bounded symbol-error probability for the 2-state channel. Therefore, the simple relationship found for the 1-state channels is also valid for the 2-state channel, i.e.

$$P_{2,Chernoff}(d,[\gamma_1,\gamma_2]) = \left[P_{2,Chernoff}(1,[\gamma_1,\gamma_2]) \right]^d \tag{4.8}$$

Moreover, from the polynomial theorem, it can be shown that the same is true for the general multi-state channel, characterized by a vector $\bar{\gamma} = [\gamma_1, \gamma_2, ..., \gamma_N]$, i.e.

$$P_{2,Chernoff}(d,\bar{\gamma}) = \left[P_{2,Chernoff}(1,\bar{\gamma}) \right]^d \tag{4.9}$$

This feature of the Chernoff-bounded PEP is now exploited to derive the exponential ESM.

The goal is to find an effective SNR value γ_{eff} of an equivalent 1-state channel such that the Chernoff-bounded error probability equals the Chernoff-bounded error probability on the multi-state channel, i.e.

$$P_{e,Chernoff}(\gamma_{eff}) = P_{e,Chernoff}(\bar{\gamma}) \tag{4.10}$$

Due to the feature stated above, this can be achieved by matching the respective Chernoff-bounded symbol-error probabilities

$$P_{2,Chernoff}(1, \gamma_{eff}) = P_{2,Chernoff}(1, \bar{\gamma}) \tag{4.11}$$

Inserting the Chernoff-bound expressions directly gives the exponential ESM:

$$\gamma_{eff} = -\ln\left(\sum_{k=1}^{N} p_k e^{-\gamma_k}\right) \tag{4.12}$$

or, for the case of OFDM with N carriers and different SNR γ_k on each carrier:

$$\gamma_{eff} = -\ln\left(\frac{1}{N}\sum_{k=1}^{N} e^{-\gamma_k}\right) \tag{4.13}$$

The above derivations have assumed binary transmission (BPSK). It is clear that, for QPSK modulation, the exponential ESM becomes

$$\gamma_{eff} = -2 \cdot \ln\left(\frac{1}{N}\sum_{k=1}^{N} e^{-\frac{\gamma_k}{2}}\right) \tag{4.14}$$

For higher-order modulation, such as 16QAM, it is not as straightforward to determine the exact expression for the exponential ESM. The reason is that higher-order-order modulation in itself can be seen as a multi-state channel from a binary-symbol transmission point-of-view. Instead, we simply state a generalized exponential ESM including a parameter β that can be adjusted to match the ESM to a specific modulation scheme or, in the general case, a specific combination of modulation scheme and coding rate. A suitable value for the parameter β for each modulation scheme and/or coding rate of interest can then be found from link-level simulations.

$$SNR_{eesm} = -\beta \ln\left(\frac{1}{N}\sum_{i=1}^{N} e^{-\frac{SNR_i}{\beta}}\right) \tag{4.15}$$

Where, N is the total number of subcarriers, $SINR_i$ is a vector $[SINR_1, SINR_2, ..., SINR_N]$ of the per-subcarrier SINR values, which are typically different in a frequency selective channel. β is the parameter to be determined for each Modulation Coding Scheme (MCS) level, and this value is used to adjust EESM function to compensate the difference between the actual BLER and the predicted BLER.

To obtain β value, several realizations of the channel have to be conducted using a given channel model (e.g., Ped B and Veh A). Then BLER for each channel realization is determined using the

simulation. Using the AWGN reference curves generated previously for each MCS level, BLER values of each MCS is mapped to an AWGN equivalent SINR. These *AWGN SINRs* for '*n*' realizations can be represented by an '*n*' element vector $SINR_{eff}$. Using a particular β value and the vector $SINR_i$ of subcarrier *SINRs*, an effective *SINR* is computed for each realization. For '*n*' realizations, we get a vector of computed effective *SINRs* denoted by $SINR_{eesm}$. The goal is to find the best possible β value that minimizes the difference between computed and actual effective *SINRs* as shown in equation 4.16.

$$\beta = arg\ \min_{\beta}\left\| SINR_{eff} - SINR_{eesm}(\beta)\right\|$$

(4.16)

4.3.1.1 Simulation Condition

- SISO
- PUSC mode
- 3GPP channel model with the velocity of 3Km/h & 60Km/h
- 100 independent channel realizations
- CTC with MCS formats in the following Table 4.3.
- Ideal channel estimated is assumed.

Beta values of different format are trained on PB and VA channel respectively through adequate link layer simulation of 802.16e system. The obtained beta values for look up are shown in the following Table 4-2 and Table 4-3.

format	QPSK 1/2	QPSK 3/4	16QAM 1/2	16QAM 3/4	64QAM 1/2	16QAM 3/4
Beta(dB)	2.18	2.38	7.34	8.85	11.09	14.59

Table 4-2: Beta values for PB channel (3Km/h)

format	QPSK 1/2	QPSK 3/4	16QAM 1/2	16QAM 3/4	64QAM 1/2	16QAM 3/4
Beta(dB)	2.12	2.37	7.53	8.90	11.01	14.55

Table 4-3: Beta values for VA channel (60Km/h)

There are also alternative methods to present the verification of link layer abstraction, such as

- Effective SNR from EESM *Vs.* Effective SNR from AWGN for given channel realization
- Predicted BLER *Vs.* Simulated BLER for given channel realization
- Predicted BLER *Vs.* Simulated BLER for given effective SNR

All the verification methods above demonstrate the abstraction performance equivalently, which are presented in the following Figure 4-4, Figure 4-5 and Figure 4-6 taking VA channel (QSPK 3/4) for example. In the three figures, the simulation results used for abstraction are the same and the samples are obtained from 100 independent channel realizations. In Figure 4-4, the distance from the red star which stands for the estimated SNR from EESM algorithm to the blue circle which stands for the ideal effective SNR from performance curve on AWGN channel of the same channel realization shows the predication accuracy. In Figure 4-5, the red stars stand for the *BLERs* obtained from link layer simulation while the blue circles stand for the *BLERs* obtained from abstraction procedure using EESM algorithm with the optimized beta value. In Figure 4-6, effective SNR is presented and the *x* axis. The blue line stands for the BLER performance on AWGN channel, which means the effective SNR of blue line is the ideal value in theory, while, the red stars are drawn using the estimated effective *SNR* through EESM algorithm as the x coordinate. In reverse, the blue performance curve on AWGN curve are used for look up during the abstraction procedure, so the *BLER* value on the blue curve stands for the prediction value and the BLER value of the red star of the same effective SNR stands for the realistic statistical *BLER* value from simulation. The extent of the red stars scatter from the blue line shows the predication accuracy. All the three figures above demonstrate the abstraction performance equivalently in different aspects. Obviously, the third method is the most clear and direct way and is widely adopted in abstraction verification.
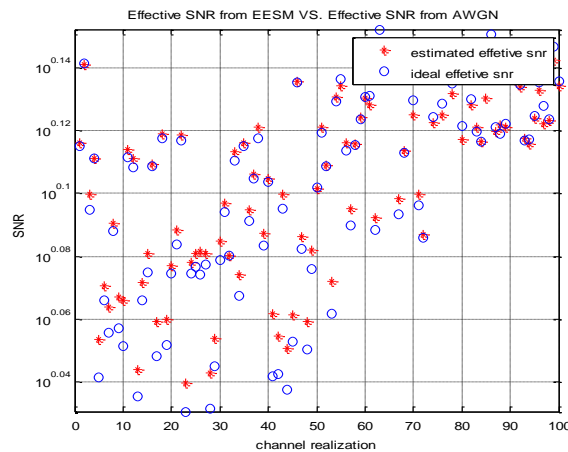
Figure 4-4 : Effective SNR from EESM VS. Effective SNR from AWGN
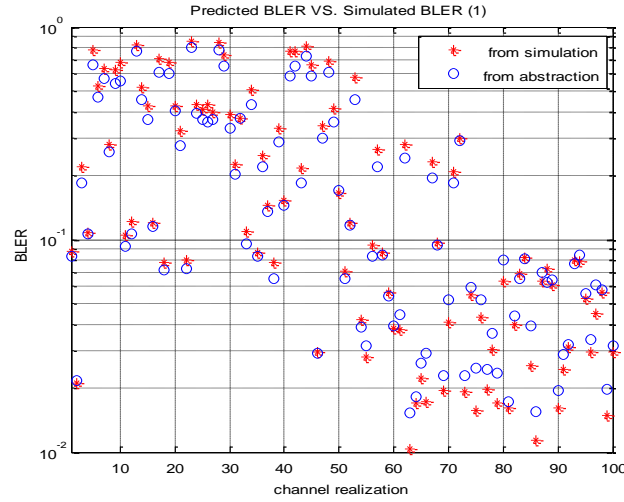
Figure 4-5: Effective SNR from EESM VS. Effective SNR from AWGN
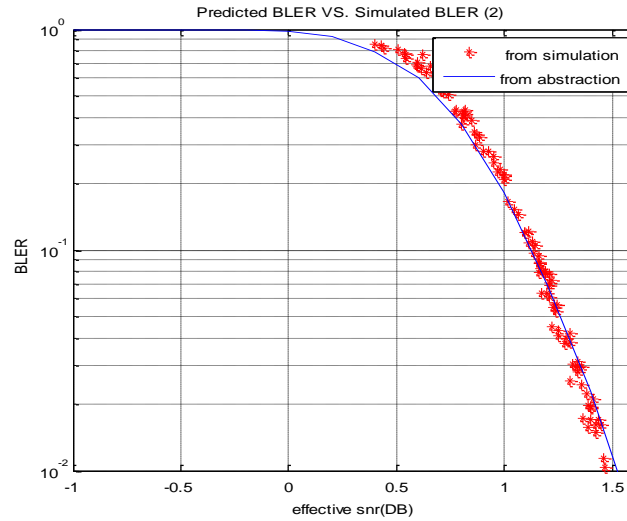
Figure 4-6 : Effective SNR from EESM VS. Effective SNR from AWGN

Beta values trained for PB and VA 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 modulation is adopted, 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 cut with the accuracy of 0.1, experiments have been made demonstrating the small change of beta value does not influence the abstraction performance noticeably. For QPSK modulation, the predicted effective SNR from EESM algorithm from different independent channel realizations (red dots in the figure) scatters closely around the effective SNR concept in theory (blue curve). It shows high reliability of EESM

algorithm in the process of physical layer abstraction. The abstraction performance is adequate for QPSK modulation, also acceptable for 16QAM and 16QAM, although the performance degrades when higher order modulation is adopted. Note that EESM algorithm is derivation from BPSK and QPSK modulation according to the Chernoff bounded pair-wise error probability (*PEP*). For higher-order modulation, 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 results 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 cannot be expressed explicitly. The formula used in the EESM derivation for BPSK and QPSK modulation is not suitable for 16QAM and 64QAM. They cannot 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.

Figure 4-7: Beta value for PB according to Table 4-2
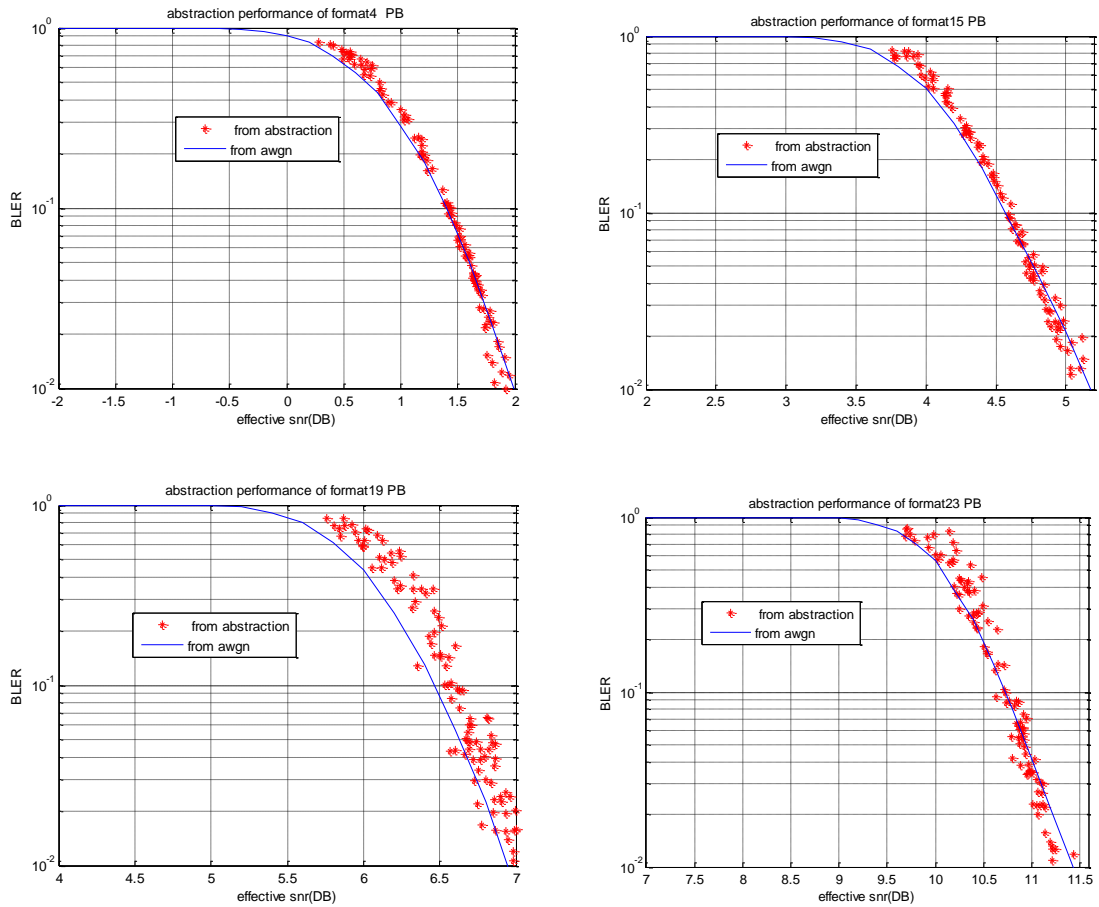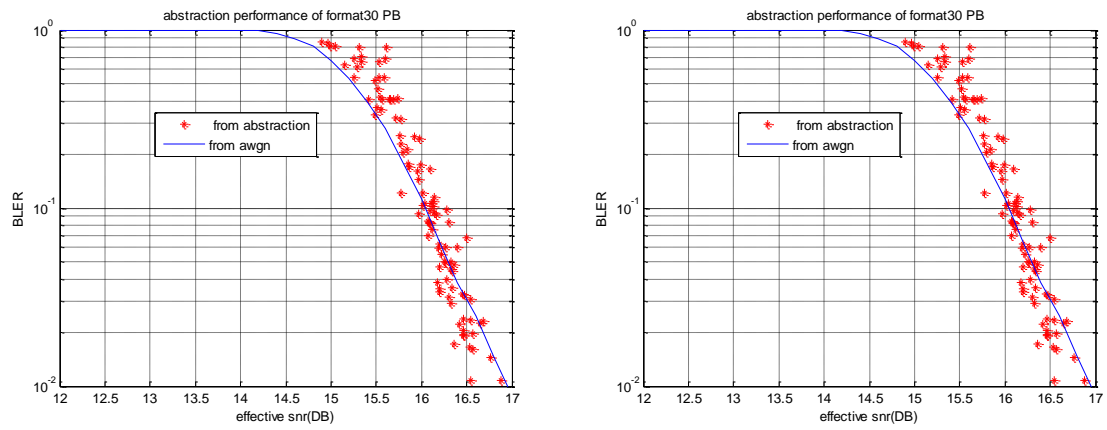
Figure 4-8: Beta value for PB according to Table 4-3
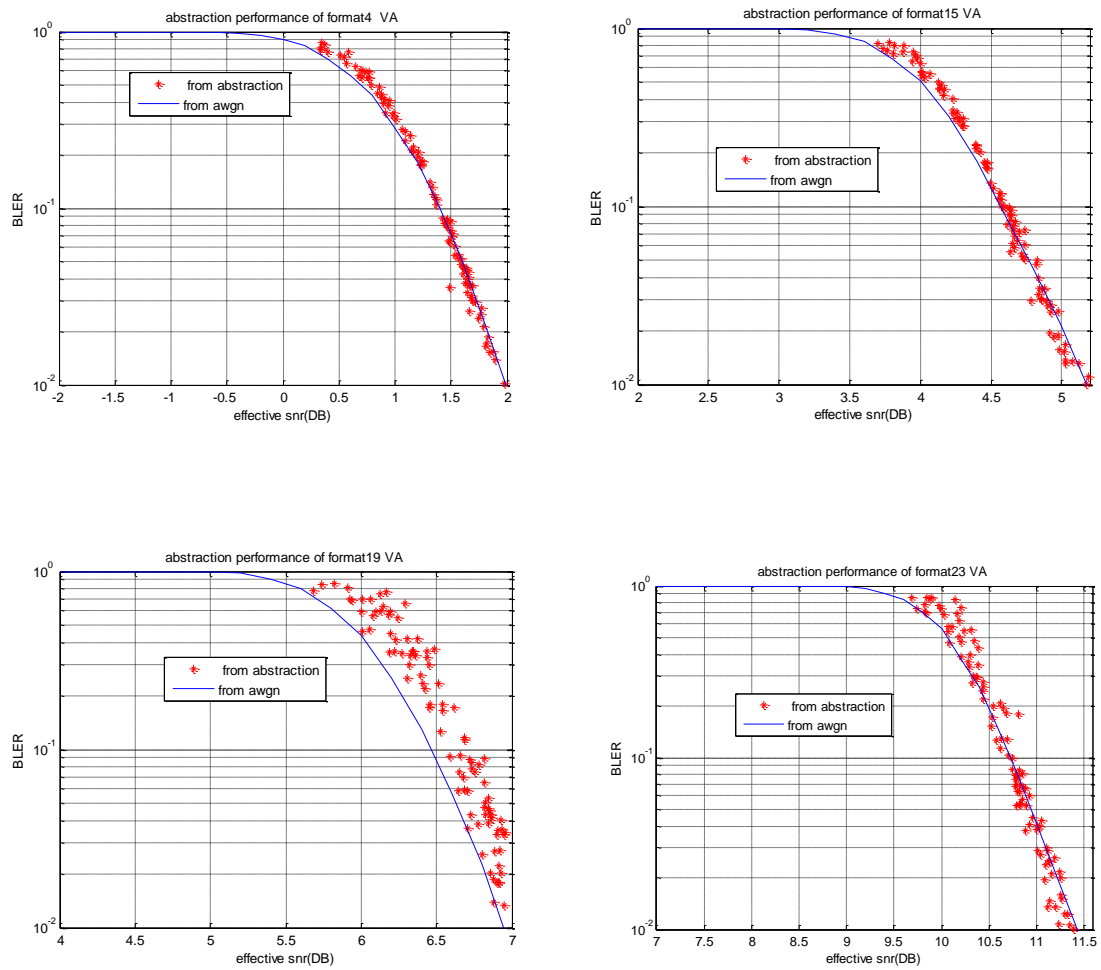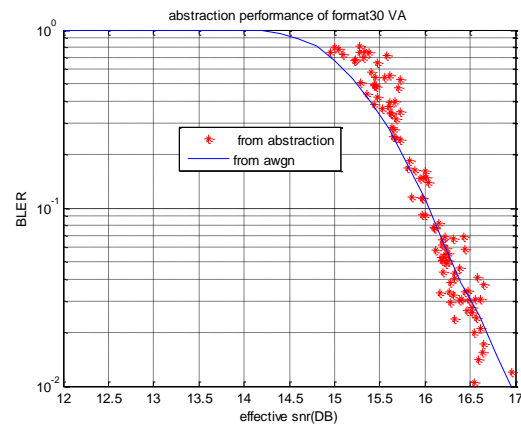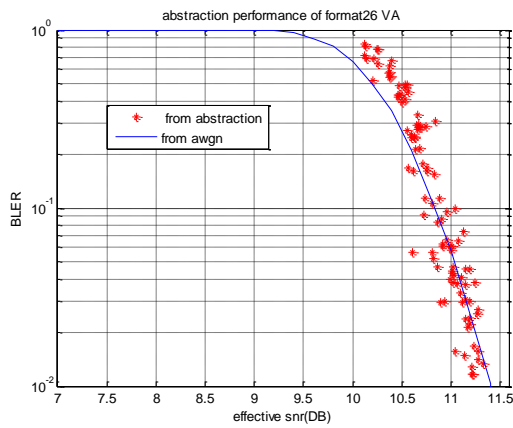
Block Error Rate or Frame Error Rate

The performance of the physical layer is modeled by means of Look-Up Tables (LUT) in which the performance of the radio link is encapsulated. This performance is measured as the variation of the Frame Error Rate (FER) or Block Error Rate (BLER) with the Signal-to-Interference plus Noise Ratio (SINR) along with parameter beta and it is averaged over many channel realizations of one specific channel model used.

4.4   System Level Simulation Methodology

The performance evaluation of a broadband wireless system such as Mobile WiMAX, for scenarios of application as close as possible to reality, must be conducted on system-level simulation platforms, with realistic models for traffic and signal propagation phenomena such as: path-loss, shadowing and fast fading; mobility patterns and traffic generation for supported users; inter-cell influence, etc. [4]. The more accuracy achieved in the implementation of these models in the simulation platform, the closer the simulator outputs are to reality. A properly designed system-level simulation platform suits the derivation of the performance figures needed in the evaluation of the impact and satisfaction of the standard, in terms of system requirements such as: spectrum efficiency, system capacity, quality of service support, end-user satisfaction and cost-efficiency. If the results obtained from system-level simulations are satisfactory, then hardware can be designed and manufactured.

The methodology followed in system level simulations depends on a different set of assumptions regarding: type of wireless system simulated, air interface technology, simulations complexity and time resolution, interface with other layers of the protocol stack, such as the physical layer, network layout, channel and interference modelling and application traffic models.

In particular, the following aspects must be thoroughly considered with care in developing a system level tool and on performing system level simulations:

Network Scenario

- This is related to the type of environment considered in the simulations: urban, rural, vehicular or indoor.

Network Layout

- Amount of tiers and number of base stations simulated.

- Type of cells: one omnidirectional cell or three, six sectored cells, for example.

- Number of mobile stations and their distribution over the network coverage area.

Radio Resource Management

- Enable/disable power control.

- Enable/disable user mobility and handover.

- Definition of the radio resources according to the type of air interface and medium access layer.

Physical Layer Modelling and Abstraction

- Definition of the metrics used to map physical layer performance to higher layers of the protocol stack.

- Definition of the types of interfaces used in the interaction between system and physical layers.

Propagation and channel modelling

- Path loss propagation.

- Slow fading (shadowing) propagation.

- Fast fading channel modelling.

Interference modelling

- Intra-cell, inter-cell and inter-system interference.

Implemented radio access system

- Multiple access to radio resources, circuit switch/packet switch.

Traffic models for application services

- Choice of traffic models: emulation by using pre-defined traffic models or use of real traces from real networks.

Performance metrics

- Metrics for network evaluation performance.

- Metrics for user satisfaction evaluation.

Simulation complexity and time resolution

There is a trade-off between accuracy and simulation execution time. The correct balance must be found. The main simulation components of a complete simulation tool are illustrated in Figure 4-9, according to simulation procedures elaborated in [4].
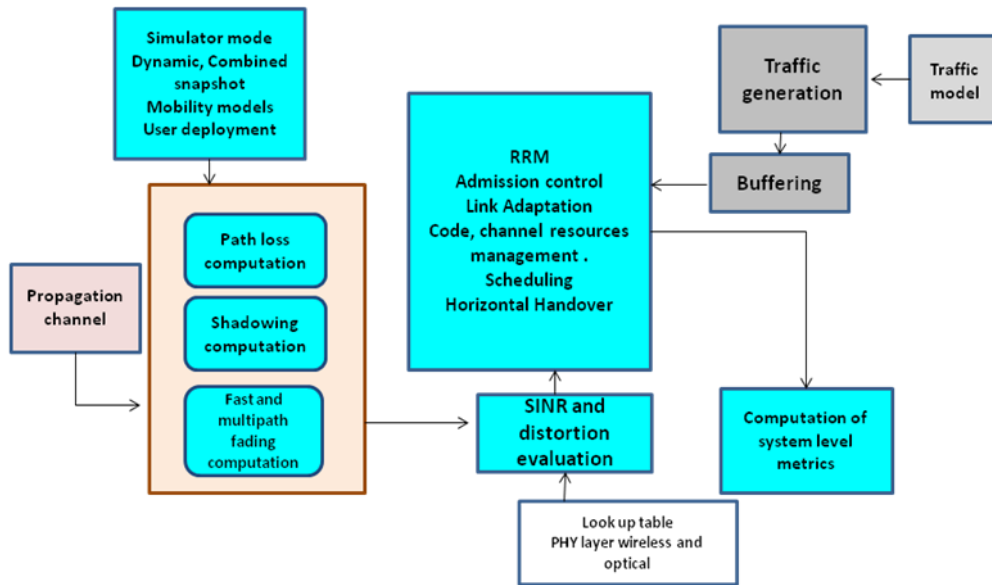
Figure 4-9: Simulation Components

4.5   Logical Architecture and Simulation Mode

4.5.1   Functionalities, input parameters and output metrics

MOTION is the system level simulation tool that was initially designed for the simulation of a beyond 3G mobile communication system. However, the design has evolved to include the possibility of simulating any type of wireless communication network. This is achieved by using the modular approach of C++ object oriented design and analysis. Hence, every object contains all information and functions that correspond to its functionality, and which also maps to real-life objects found in wireless systems. The system level tool has been designed having in mind extendibility to model heterogeneous cellular systems such as HSDPA, WiMAX, etc.

Figure 4-10 illustrates a functional block diagram of the MOTION simulator. The inputs and configuration files are shown as arrows entering the main block. The calculated metrics and output files are shown as arrows exiting from the main block.

4.6   Simulator modes: Dynamic mode and combined snapshot

Two different types of simulations can be performed: using a Combined Snapshot-Dynamic mode or a Dynamic mode.

Dynamic mode: mobility is enabled as mobiles travel along the network coverage area performing handovers. Mobiles are dropped in the network in the beginning of the simulation run and remain active since the instant of activation, which can be coincident with the beginning of the simulation

run or defined by some random distribution. Only one simulation run is performed and mobiles are removed at the end of the simulation. Statistics are collected as mobiles travel through the network coverage area. Path-loss, shadowing and fast fading propagation components are re-computed at every Transmission Time Interval (TTI, equal to the frame transmission period). The new position of the mobile station in the next transmission time interval is also computed according to the chosen mobility model. Figure 4-11 illustrates the concept of fully dynamic mode, where users are free to move across the entire coverage area and where metrics are calculated for all the cells in the system

.

Figure 4-10 : Funcationl Block Diagram of MOTION SLS

Combined snapshot-dynamic mode: mobility and handovers are disabled. A given number of simulation runs are performed. Mobile stations are drawn on the network in the beginning of each simulation run and are removed at their end. They remain active since the instant of activation, which can be coincident with the beginning of the simulation run or be defined by some random distribution. In this mode, path-loss and shadowing are computed at the beginning of each simulation run and remain constant until the end of the run. Fast fading is re-computed at every transmission time interval. This mode increases simulation speed as the different simulation runs (snapshots) can be performed in parallel. Graphic representations of the combined snapshot mode with non-central and central option are displayed in Figure 4-12 and Figure 4-13, respectively. Table 4-4 summarizes the features of each one of the simulation modes.
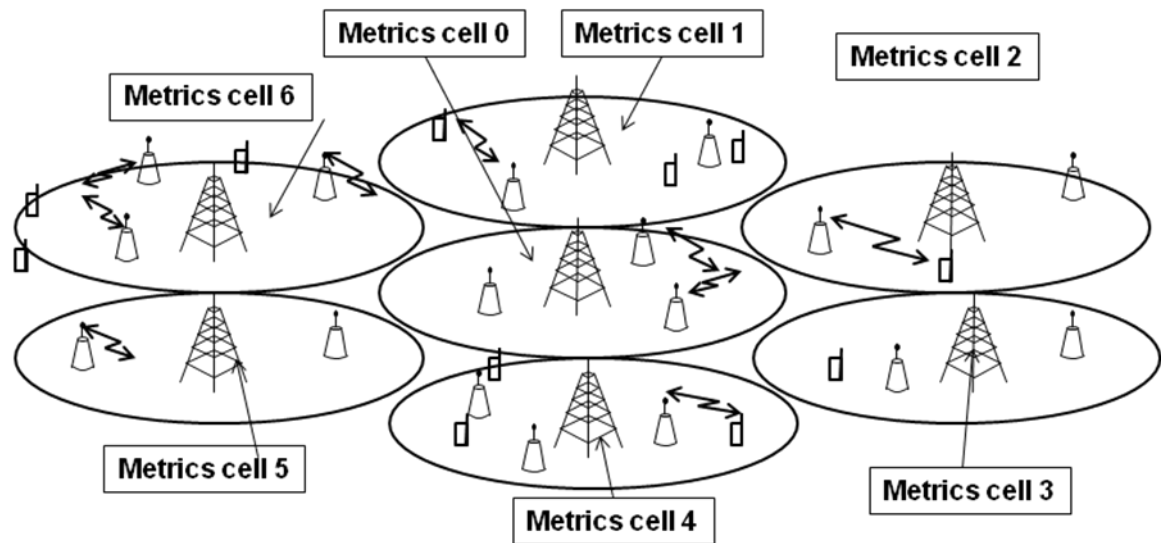
Figure 4-11: Graphic representation of the fully dynamic mode.
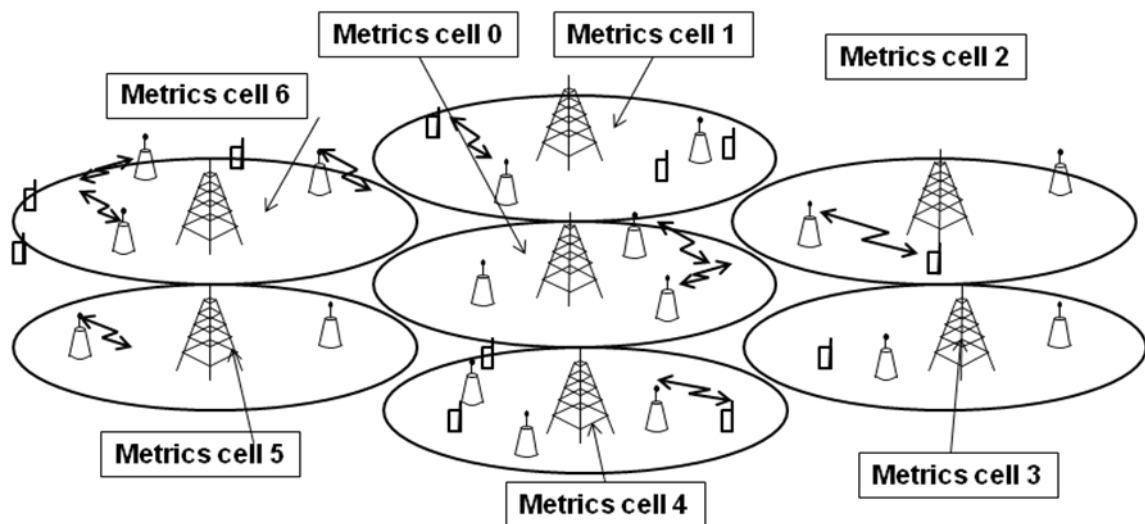
Figure 4-12: Graphic representation of the combined-snap shot mode with non-central option.
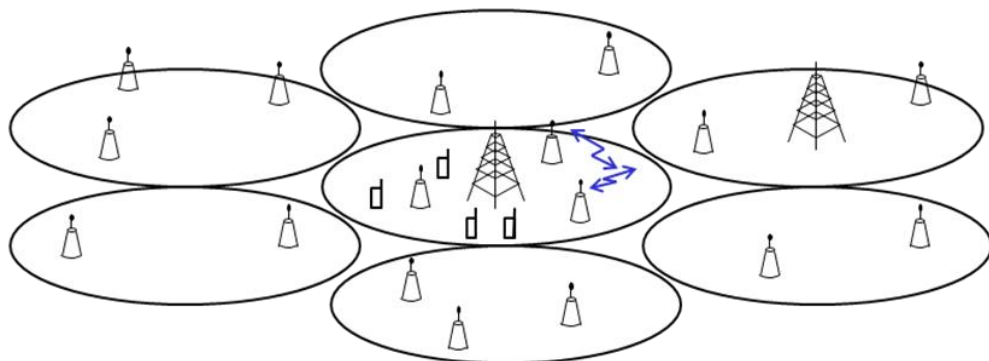
Figure 4-13: Graphic representation of the combined-snap shot mode with central option.

Mode Option	Central or Non-Central Options	Path Loss & Shadowing	Mobility	Fast and Multipath fading
Combined Snapshot (CSM)	Central and non-central	Calculated only once at the beginning of simulation run	Disabled Random positions	Recomputed every TTI
Fully Dynamic (FDM)	Non-central only	Recomputed every TTI	Mobility and handover enabled	Recomputed every TTI

Table 4-4: Options and features for the different simulation modes

4.7 Deployment scenarios

4.7.1 Hexagonal multi-cell deployment

Each base station can be configured with one sector (omni-directional antenna pattern, see Figure 4-14) or with three sectors/cells (directional antenna pattern). The number of tiers included in the simulation can also be adjusted as well as the frequency reuse pattern (see Figure 4-15). Table 4-5 summarizes the user input parameters for the configuration of the omni- and sectorized hexagonal cell configurations.

In both modes, mobiles are randomly uniformly distributed over a hexagonal network of base stations. Each base station can be configured with one sector (omni-directional antenna pattern) or with three sectors/cells (directional antenna pattern). All system level simulations conducted in this work were performed assuming an urban environment model. The simulated network is constituted of 57 sectors (19 base stations with 3 sectors each), composing a 3 tier hexagonal cellular network layout.
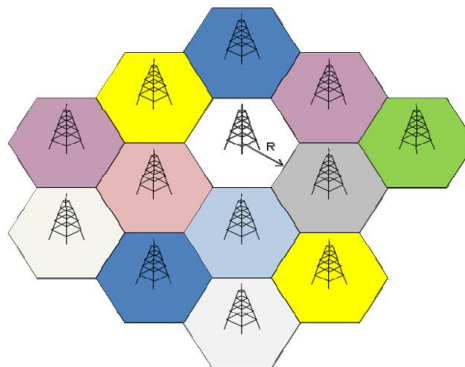
Figure 4-14 : Hexagonal Multi-cell Deployment Scenario
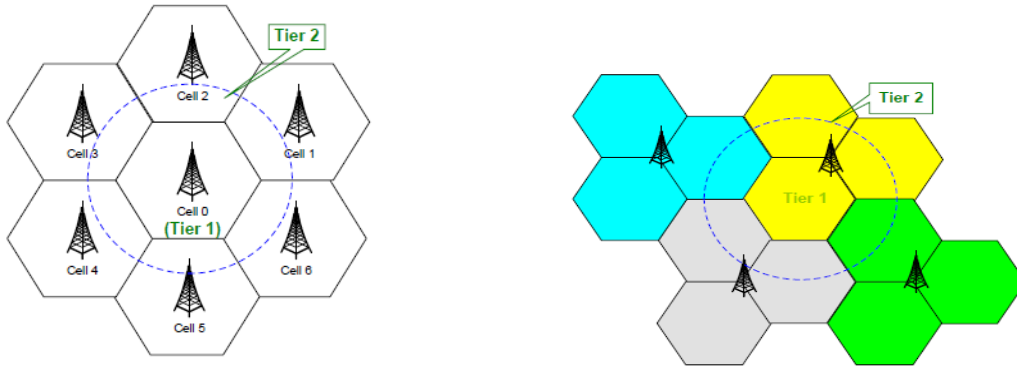
Figure 4-15: Omni vs. Sectorized Hexagonal cell multi-tier Configurations used in the Simulator

Parameters for the hexagonal deployment
Number of BSs
Cell Radius
Number of Tiers
Number of Sectors
Frequency Reuse
BS/MS height
Noise Destiny
Receiver Sensitivity
BS Tx Power

Table 4-5: Main Parameter for Hexagonal cell Deployment Scenario

4.7.2 Antenna radiation pattern

The 3-sector antenna pattern used for each sector is plotted in Figure 4-16. The antenna pattern used for the sectored antenna deployment only considers the horizontal pattern corresponding to a main sector of 70 degrees. According to the model for the typical antenna pattern proposed in [5], power attenuation is computed as a function of the angle between the antenna pointing direction and the mobile to base station direction, as given by equation 4.17.

$$A(\theta) = - \min\left[12\left(\frac{\theta}{\theta_{3dB}}\right), A_m\right]$$
(4.17)

Where:

· $-180 < \theta < 180$ is the angle between the antenna's pointing direction and the mobile to base station line-of-sight direction in degrees,

· $\theta_{3dB} = 70^{\circ}$ is the beam width at 3dB

· $A_m = 20dB$ is the maximum attenuation.
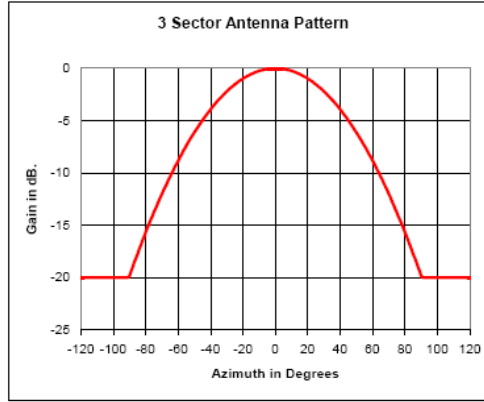
Figure 4-16: Antenna pattern for 3 sectors

4.7.3   Simulation Steps

As explained in the previous section, two types of cell configurations can be defined for simulations: central-cell and non-central cell approach. In the central-cell approach mobiles are dropped along the coverage of the central base station and statistics are collected only for the cells of this base station. Naturally the central cell approach simulation method can be enabled only in conjunction with the combined snapshot-dynamic mode, as mobility modelling is disabled. The cells in the remaining tiers are assumed as fully loaded, i.e., transmitting with maximum power and contributing to interference only.

For each frame interval the following events are generated:

·   Packets are generated according to the traffic model.

·   The fast fading channel is updated.

·   Dynamic resource allocation is executed.

·   Packet quality detection is performed.

The BLER resulting from decoding the information transmitted along a single resource unit (RU) is denoted by $BLER_{RU}(SINR_{RU})$ and is obtained from the link-to-system interface, using as input the Signal to Interference plus Noise Ratio $SINR_{RU}$. Then a random variable, uniformly distributed between 0 and 1 is drawn as shown in Figure 4-17. If the random value is less than $BLER_{RU}(SINR_{RU})$, the block is considered as erroneous and a Negative Acknowledge (NACK) message is sent back to the base station on the associated signalling channel. Otherwise, the block is deemed as error free and an Acknowledge (ACK) message is transmitted. The success or failure

in the decoding of the transmitted block of information is computed from decoding each individual resource into which the data block is mapped into. Assuming that a total amount of N_{res} radio resources are used in the transmission and that the decoding is an independent and identically distributed random process, the BLER for the whole radio block is given by equation (4.18).

$$BLER_{RB} = 1 - \left[1 - BLER_{RU}\left(SINR_{RU}\right)\right]^{N_{res}} \tag{4.18}$$
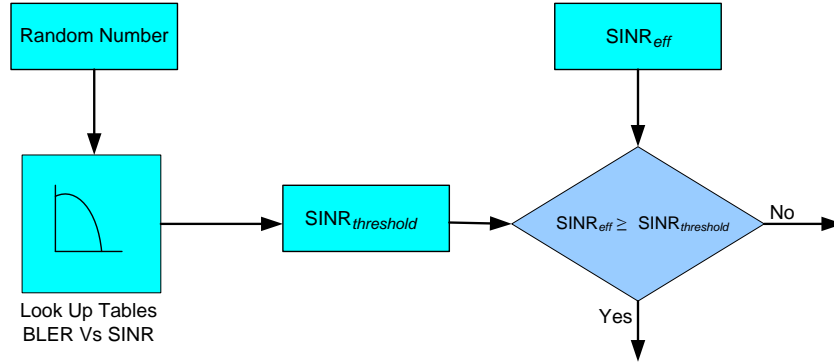
Figure 4-17: Packet decoding process

In each simulation run (snapshot) the following steps are followed as shown in Figure 4-18:

1. Mobile stations are dropped independently with uniform distribution throughout the system. Each mobile corresponds to an active user session that runs for the whole duration of the drop.

2. Mobiles are assigned channel models. This can be in support of a channel mix or separate statistical realizations of a single type of channel model.

3. Mobiles are assigned a traffic model.

4. Cell assignment to a mobile station is based on the received power at the mobile station from all potential serving cells. The cell with the best path to the mobile station, taking into account slow fading, path-loss and antenna gains, is chosen as the serving sector.

5. For simulations that do not involve handover performance, evaluation of the location of each mobile station remains unchanged during a drop and the mobile's speed is used only to determine the Doppler effect of fast fading. The mobile station is assumed to remain attached to the same base station for the duration of the drop.

6. Fast fading is computed for each mobile station in each TTI. Slow fading and path loss are assumed as constant during the whole simulation run.

7. Packets are withdrawn from the buffers of the traffic models. Packets are not blocked as the queues are assumed to be infinite. Start times for each traffic type for each user should be randomized.

8. Packets are scheduled with a packet scheduler using the required metric. Packet, decoding errors result in packet retransmissions. In the Dynamic Resource Allocation (DRA) module a Hybrid Automatic Repeat Request (HARQ) process is modelled by explicitly rescheduling a packet as part of the current packet call and after a specified feedback delay period.

9. For a given drop the simulation is run for the pre-defined duration and then the process is repeated with the mobile stations being dropped at new random locations.

10. Performance statistics are collected for mobile stations in all cells.
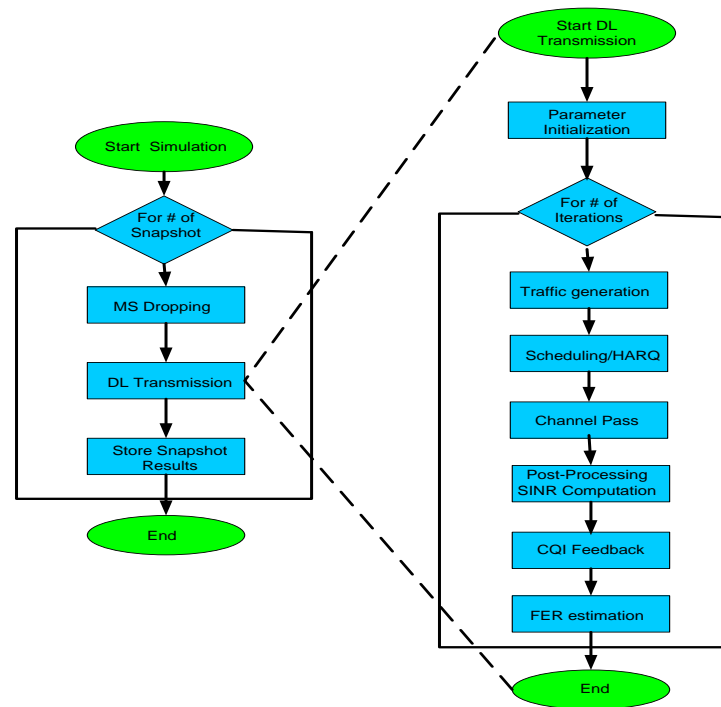
Figure 4-18: Simulation Step

4.8 Propagation Channels

The radio propagation is divided into three distinct components, namely path loss, slow fading (shadowing) and fast fading which already explained in chapter 3.The decrease of the transmitted radio signal impinging on the receiver antennas is the result of their contribution. Accurate modelling of each one of these three-radio propagation components depends on the simulation scenario envisaged for the system-level simulations. Namely, the simulation scenario can be described according to the following characteristics:

· Type of environments: indoor, urban, suburban and rural.

· Mobile speed: pedestrian, vehicular, train.

· Type of receiver used in the signal processing at the receiving end.

· Antenna radiation pattern.

· Antenna configuration used in the communication between the transmitter and the receiver (SISO, SIMO, MISO, MIMO).

· Radio transmission parameters: carrier frequency, system bandwidth, etc.

4.8.1 Fast Fading Model

This section explains the fast fading model used in system level simulator .The fast fading component of the signal is generated according to a modified Jake's model, for the fast generation of independent Rayleigh faders according to the method proposed in [6-7]. In order to speed-up simulations, the multi-path channel model is used for the serving cell, while a flat fading channel model (with only one tap) is assumed for neighbouring cells as shown in Figure 4-19
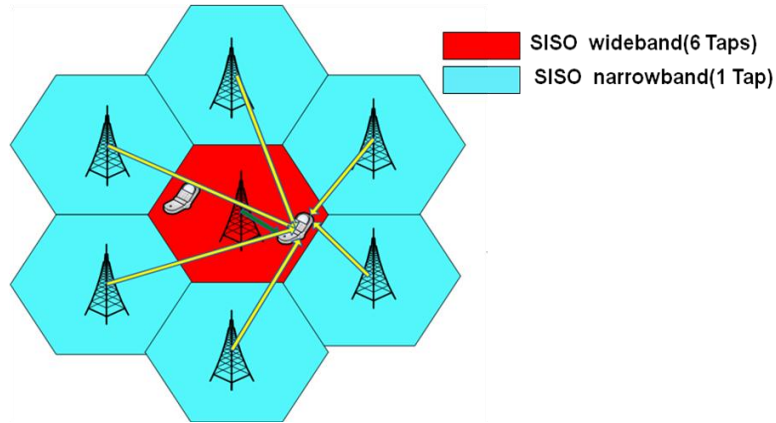
Figure 4-19: Fast fading modeling

The mobile speed and carrier frequency are the parameters considered in the generation of the fading statistics. In this context a channel model corresponds to a specific number of paths, a power profile giving the relative powers of these multiple paths and Doppler frequencies to specify the fade rate. ITU multi-path channel models for narrowband SISO are proposed in [8,9]. These models are based on a discrete version of the scattering function of the propagation channel and are designated as tapped delay line models. Each tap is characterized by an attenuation A_i, a corresponding delay τ_i a Doppler frequency υ and a Doppler Power Spectrum (DPS) $P_s(\upsilon, \tau_i)$ at the ith tap.

Table 4-6 and Table 4-7 detail the parameters used in the definition of each type of channel model proposed by ITU [10]. The channel models assigned to a specific user remains fixed over the duration of a simulation run.

Channel Model	Multi-path Model	Number of Paths	Speed (Km/h)	Fading
Model 1	Ch-100	1	30	Jakes
Model 2	Ch-100	1	120	Jakes
Model 3	Ch-104	6	30	Jakes
Model 4	Ch-104	6	120	Jakes
Model 5	Ch-102	4	3	Jakes
Model 6	Ch-103	6	3	Jakes

Table 4-6: Parameters for the different types of fast fading channel models for SISO

Channel Model		Path 1	Path 2	Path 3	Path 4	Path 5	Path 6
Flat Fading Ch-100	Path Power (dB)	0	-	-	-	-	-
ITU Ped. A Ch-102	Path Power (dB)	-0.51	-10.21	-19.71	-23.31	-	-
	Delay (ns)	0	110	190	410	-	-
ITU Ped. B Ch-103	Path Power (dB)	-3.92	-4.82	-8.82	-11.92	-11.72	-27.82
	Delay (ns)	0	200	800	1200	2300	3700
ITU Veh. A Ch-104	Path Power (dB)	-3.14	-4.14	-12.14	-13.14	-18.14	-23.14
	Delay (ns)	0	310	710	1090	1730	2510

Table 4-7: Multi-Path Channel Models for Performance Simulation

A separate link level simulation must be performed (see section 4.2) for each specific channel model and mobile stations' velocity combination.

In [11] System Level Simulations are conducted to validate Mobile WiMAX standard using the ITU's normalized power profiles for channel models such as ITU Vehicular A: Ch-104 and ITU Pedestrian-B: Ch-103 [12-13]. These models are illustrated in Table 4.8. The absolute power values are normalized so that they sum to zero dB (unit energy) for each given channel.

4.8.2 Signal to Interference plus Noise Ratio (SINR) Modelling

Mobile WiMAX standard is an OFDM-based technology. If one designates the set of sub-carriers available for data transmission in each OFDM symbol as N_{data}, the power available at the base station for data transmission (not considering the power boost used in the transmission of pilot sub-carriers used in channel estimation) by P_{data}, and if one splits this power uniformly over the set of data sub-carriers, the power assigned for the transmission of data sub-carrier $n, n \in [0,...,N_{data}-1]$ is given by equation (4.19):

$$P_{data}^{(n)} = \frac{P_{data}}{N_{data}} \qquad (4.19)$$

The transmitted baseband signal for user k is given by equation (4.20)

$$x_k(t) = \sum_{n=0}^{N-1} \sqrt{p_k^{(n)}} \, d_{n,k} \exp(j2\pi nt/T), \quad 0 \le t \le T, \quad n=1,...,N; \quad k=1,...,K \qquad (4.20)$$

Where:

· K is the number of active users in the cell

· N is the number of sub-carriers composing one OFDM symbol.

· $p_k^{(n)}$ denote the power allocated to user k in sub-carrier n

· $d_{n,k}$ denotes the complex symbol at the output of the OFDM modulator for the nth sub-carrier of user k with $b_{n,k}$ bits.

The power transmitted over each sub-carrier must follow the constraint of the total power available for data transmission in the cell, according to equation (4.21)

$$\sum_{k=1}^{K} \sum_{n \in S_k} p_k^{(n)} \leq P \tag{4.21}$$

Where S_k is the set of sub-carriers assigned to user k.

The time-invariant channel (for each frame period) between the base station and user k has impulse response given by equation (4.22)

$$h_k(t) = \sum_{l=1}^{L} \beta_{k,l} \delta(t - \tau_{k,l}) \tag{4.22}$$

Where:

· L is the number of paths in the multipath channel.

· $\beta_{k,l}$ is the gain of the lth path of user k, including path loss and shadowing.

· $\tau_{k,l}$ is the path delay of the lth path of user k, including path loss and shadowing.

In the receiver the nth symbol of user k is given by equation (4.23)

$$y_{n,k} = d_{n,k} \sqrt{p_k^{(n)}} \left(\sum_{l=1}^{L} \xi_{k,l}(n) \right) + z_{n,k} = d_{n,k} \sqrt{p_k^{(n)}} g_{n,k} + z_{n,k}, \quad n = 1,...,N \tag{4.23}$$

Where: $\xi_{k,l}(n) = \beta_{k,l} \exp[-j2\pi(f_c + n/T \tau_{k,l})]$ captures the impact of propagation characteristics of path l of the multipath channel for user k at sub-carrier n. The term $z_{n,k}$ is the Gaussian noise at the receiver after sampling.

In System-Level Simulations mobile stations are randomly dropped along the network coverage area being simulated. When the mobile station becomes active the best serving cell (according to signal strength) is selected and it is assumed this is the cell where the mobile station camps on. As mentioned in the previous sections, the signal coming from the best cell (serving cell) is modelled as a frequency selective fading channel, whereas the signal coming from neighbouring ones is

modelled according to a flat frequency fading channel. Assume then a given mobile station MS_i is camping in the coverage area of cell $Cell_j$. Assume also a SISO channel. The power received from serving base station $BS_{serving}$ for data sub-carrier $i, i \in [0,...,N_{data}-1]$, on mobile station MS_i in the nth frame interval is given by equation (4.24).

$$P_{BS_{serving}}^{(i)}(n) = \frac{P_{data}^{(i)} \cdot \left| H_{BS_{serving}}^{(i)}(n) \right|^2 G_{BS_{serving}} G_{MS_i}}{PL_{MS_i BS_{serbing}} SH_{MS_i B_{serving}} L_{loss}}$$ (4.24)

Where:

· $\left| H_{BS_{serving}}^{(i)}(n) \right|^2$ is the instantaneous power from the serving base station $BS_{serving}$ at the ith data sub-carrier at the nth frame interval.

· G_{MS_i} is the gain of the antenna at the mobile station MS_i.

· $G_{BS_{serving}}$ is the gain of the antenna at the serving base station $B_{serving}$.

· $PL_{MS_i BS_{serving}}$ is the path-loss between serving base station $BS_{serving}$ and mobile station MS_i.

· $SH_{MS_i BS_{serving}}$ is the shadowing loss between serving base station $BS_{serving}$ and mobile station MS_i

· L_{loss} encompasses the other losses in the transmission (cable losses, body loss,…).

As sub-carriers are mutually exclusively assigned inside each cell there is no intra-cell interference. Only inter-cell interference must be considered. The interfering power arriving at mobile station MS_i from neighbouring cells is given in equation (4.25)

$$P_{Inter}^{(i)}(n) = \sum_{BS_j \in \{BS_{Inter}\}} \frac{P_{data,BS_j}^{(i)} \cdot \left| H_{BS_j}^{(i)}(n) \right|^2 G_{BS_j} G_{MS_i}}{PL_{MS_i BS_j} SH_{MS_i BS_j} L_{loss}}$$ (4.25)

Where:

· $\{B_{Inter}\}$ is the set of interfering base stations and $BS_j \in \{B_{Inter}\}$

· $\left| H_{BS_j}^{(i)}(n) \right|^2$ is the instantaneous power from the interfering base station BS_j at the ith data sub-carrier at the nth frame interval.

· G_{BS_j} is the gain of the antenna at the interfering base station BS_j.

· $PL_{MS_i BS_j}$ is the path-loss between interfering base station BS_j and mobile station MS_j.

· $SH_{MS_i BS_j}$ is the shadowing loss between interfering base station BS_j and mobile station MS_i

· L_{loss} encompasses the other losses in the transmission (cable losses, body loss,…).

According to equations (4.24) and (4.25) the SINR at sub-carrier i and for the nth frame interval is given by equation (4.26).

$$SINR^{(i)}(n) = \frac{P^{(i)}_{BS_{serving}}(n)}{P^{(i)}_{Inter}(n) + N_0 W_i F_{MS_i}}$$ (4.26)

Where:

· N_0 is the received noise spectral density.

· W_i is the sub-carrier bandwidth.

· F_{MS_i} is the noise figure at the mobile station.

Although the method followed in equations (4.24-4.26) for the derivation of the SINR is perfectly general, in the system level simulator platform, the SINR for each data sub-carrier k is computed according to the method derived in [14-15], which are in accordance to the previous description, and is given by equation (4.27).

$$SINR^{(k)}(n) = H^{(k)}(n).\overline{G}.\left(\frac{N}{N+N_p}\right).\frac{R_D}{N_{SD}/N_{ST}}$$ (4.27)

Where:

· $H^{(k)}(n)$ is the channel's current frequency response for the serving cell (propagation from interfering cells is modelled as flat fading).

· $\overline{G}$ if the Geometric Factor between the mobile station and its serving and interfering cells and is given by equation (4.28).

· N is the FFT size, including pilot, data and guard sub-carriers.

· N_p is the cyclic prefix length.

· R_D is the percentage of maximum total available transmission power allocated to data sub-carriers.

· N_{SD} is the amount of data sub-carriers per each OFDM symbol.

· N_{ST} is the amount of useful (pilot plus data) sub-carriers per OFDM symbol.

The Geometric Factor is defined by equation (4.28)

$$\overline{G} = \frac{G(Cell_0, MS) x \dfrac{1}{PL(Cell_0, MS) x SH(Cell_0, MS)}}{\sum\limits_{k=1}^{N} G(Cell_k, MS) x \dfrac{1}{PL(Cell_k, MS) x SH(Cell_k, MS)} + N_0 WF_{MS}} \tag{4.28}$$

In equation (4.28) $Cell_0$ is the serving cell and N is the total amount of interfering base stations.

Assuming that multi-path fading magnitudes and phases, respectively $M_p(n)$ and $\theta_p(n)$, are constant over the frame interval for each path p of the tapped delay channel filter, the frequency-selective fading power for the kth sub-carrier is given by equation (4.29) [14].

$$H^{(k)}(n) = \left| \sum_{p=1}^{N_{paths}} M_p A_p \exp\left(j\theta_p\right) \exp\left(-j2\pi f_k T_p\right) \right|^2 \tag{4.29}$$

Where:

· p is the tap index (from 1 to 6) of the tapped delay model.

· A_p is the amplitude value of the long-term average power for the pth tap of the tapped delay filter.

· T_p is the relative time delay of the pth tap of the tapped delay filter.

· f_k Is the relative frequency offset of the kth sub-carrier within the spectrum of the OFDM symbol

Parameters A_p and T_p depend on the type of ITU channel used in the modelling of multi-path channel propagation.

4.8.3 MIMO channel Modeling

There are different methods for modeling the MIMO channel at the system level. These methods are grouped into two different categories:

Ray-based: the channel coefficient between each transmit and receive antenna pair is the summation of all rays at each tap of the multi-path filter at each time instant, according to the antenna configuration, gain pattern, angle of arrival (AoA) and angle of departure (AoD) of each ray. The temporal channel variation depends on the travelling speed and direction relative to the AoA/AoD of each ray.

Correlation based The MIMO channel coefficients at each tap are mathematically generated. According to independent and identically distributed Gaussian, random variables, according to the antenna correlation and the temporal correlation, correspond to a particular Doppler spectrum.

3GPP Spatial Channel Model In this model the channel gain between each pair of antennas, at both ends of the communication link, results from the superposition of the contributions from each individual path of the tapped delay line model. All simulations are conducted by using 3GPP Spatial Model. More details about this channel model can be found in [15]

The derivation of SINR for the MIMO channel is performing in two separate parts, detailed in [16]:

a) Computation of the desired user radio signal arriving from the mobile station-serving cell.

b) Computation of the interfering user radio signal arriving from neighbouring cells.

Three types of neighbouring cells are shown in Figure 4-20.

1. Near strong cells, whose interference is modeled by a MIMO channel

2. Near weak cells, whose interference is modeled as wideband (frequency-selective) channel

3. Far Cells whose interference are modeled as narrowband (flat) channel
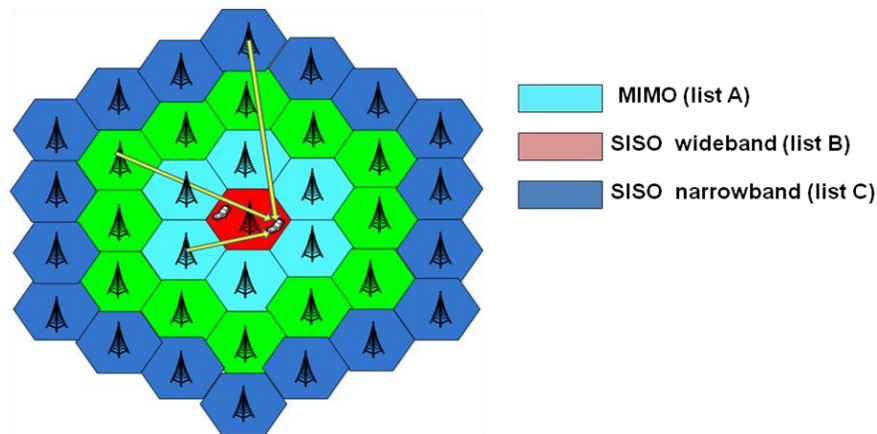
Figure 4-20: MIMO modelling

4.8.4   Link Level Interface Modelling for Mobile WiMAX System

The general principals of the Link-To-System Interface (LLI), integrating the physical and system level performance results are applied to the specific scenario of system level simulations under the Mobile WiMAX standard. Figure 4.21 illustrates the methodology followed in system level simulations and performance evaluations for the OFDM radio technology, in which Mobile WiMAX is based [14].

·   On system level a set of mobile stations are randomly dropped over the network deployment area. Depending on its physical location each mobile station is characterized by the

Geometric Factor $\overline{G}$, which depends on distance-dependent path-loss between the serving and neighbouring cells, shadowing value and thermal noise power. The Geometric Factor is derived by averaging-out the influence of the fast fading component on the propagated signal.

In each TTI the instantaneous value of the radio channel between each base station and the mobile station is derived, assuming only the fast fading component. As the Mobile WiMAX standard is OFDM-based it is preferred to express the channel in the frequency domain, by its frequency response $H(f)$, which depends on the type of radio scenario being simulated: pedestrian, vehicular, etc. For low/medium Doppler, i.e. low/medium channel coherence time, the channel is assumed constant during a TTI and its instantaneous value is considered in the computation of the SINR used in the LUTs.
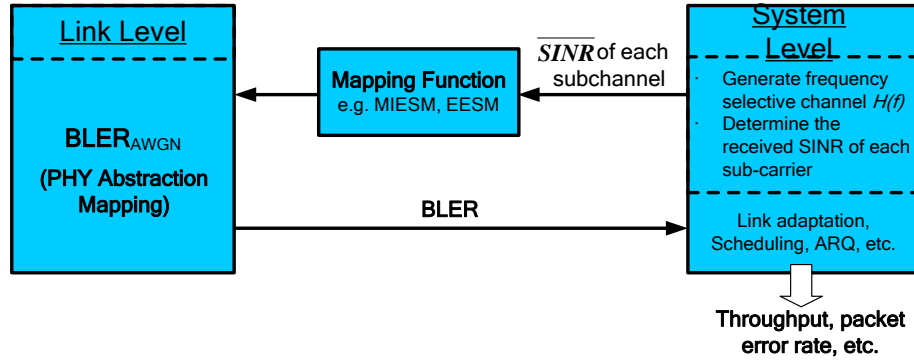
Figure 4-21 : Schematic view of system level methodology

· Although OFDM uses cyclic prefixes for the limitation of Inter-Symbol Interference (ISI) at each sub-carrier, the performance (SINR) over the whole set of sub-carriers encompassing each OFDM symbol will change due to the influence of the frequency selectivity of the mobile radio channel. The output from the System Level Simulation Module is a vector of frequency response channel gains $H(f)$ and a scalar Geometric Factor $\overline{G}$. A mapping function $SIR_{eff}(\overline{G}; H(f))$ is used in order to map the geometry and the frequency response to an effective SINR value. The interference part of the effective SIR includes both the inter-cell interference and noise, as the intra-cell interference is not considered due to the orthogonality among adjacent sub-carriers.

· The effective SINR is used to find the BLER probability from the basic Additive White Gaussian Noise (AWGN) performance curves which emulate the link level performance.

The only task that remains now is the definition of the suitable mapping function, $SIR_{eff}(\overline{G}; H(f))$. The mapping function should depend on the exact modulation and coding scheme (MCS) used as well as on other transmission formats (SISO, MIMO…), but not on the channel model implemented on the simulations. Any mapping function designed to be used in a system level evaluation must be thoroughly verified by means of link level modulations. The obtained mapping function must be verified for different MCS schemes/transmission modes. Different types of SINR mapping functions were proposed in the literature. The role of an SINR-based mapping function is to provide the expected BLER/FER of a coded block of information as a function of the SINRs of its data symbols, which can be transmitted through different types of resources (time, frequency, codes or spatial beams), depending on the type of multiple access implemented in the wireless system. This calculation may involve the computation of the weights of both pre and post processing matrices, at the transmitter and the receiver, as for example whenever beamforming or spatial multiplexing is used in connection with MIMO. The vector of instantaneous SINR values associated to the resources assigned to the transmission of a data block is mapped into a given set of data sub-carriers and is computed from the corresponding fading amplitudes of the different sub-carriers at the receiver. Due to multi-path propagation the SINR values in this vector are received with different levels of quality and should be considered separately by the LUTs for the generation of the BLER/FER/PER/BER metrics. However, the amount of resource elements in all domains (power, time, frequency, code, space) is too high for them to be considered individually by the LUTs. As a consequence the vector of SINRs must be compressed to a lower dimension order, preferably to a one or two-dimensional vector, before being inputted to the LUT.

Assume an OFDMA-based multiple access schemes with SISO and inter-cell interference, modeled as AWGN, accumulated in the thermal noise component of the SINR. Assume also that at the nth frame interval the radio block data symbols are mapped into a set of N data sub-carriers of the OFDM symbol. Equation (4.30) computes the compressed SINR by means of an arithmetic mean over the set N of data sub-carriers.

$$SINR_{eff}(n) = \overline{G}\frac{\left(\sum_{k=0}^{N-1}\frac{|h_k(n)|^2}{\sigma_n}\right)}{N} \tag{4.30}$$

Where:

·	$\overline{G}$ is the actual geometry factor between the user, its serving and neighboring cells, given by equation (4.28).

- $\left|h_k(n)\right|^2$ is the instantaneous power in the kth data sub-carrier of the sub-carrier set to which the data is mapped into in the nth frame interval.

- N is the size of the sub-carrier set.

- σ_n is the noise power, which includes interference from other cells (modeled according to Gaussian distribution).

Figure 4.22 illustrates this compression principle. The arithmetic mean-based compression scheme is the simplest and less accurate effective SINR mapping. It is not accurate because it averages-out the variations of the channel along the set of sub-carriers, i.e., it underestimates some samples and overestimates others.
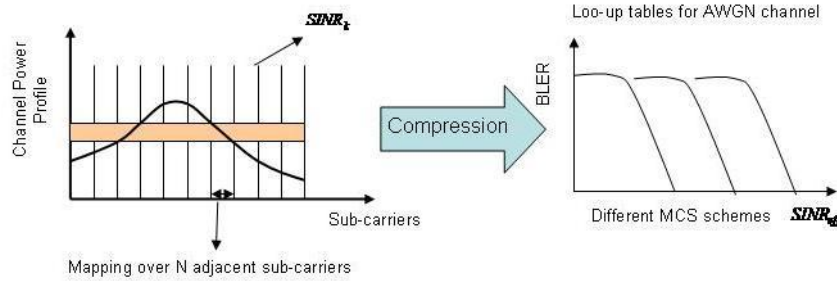
Figure 4-22: Illustration of SINR compression and mapping

Different methods for compressing the vector of SINRs into a single scalar SINR are proposed in the literature. For this dissertation we used EESM which has been explained in section 4.1.1. See annex B for other compressing methods.

4.9   Traffic Models

The application traffic models are categorized into two types of traffic modelling: foreground and background traffic. The foreground traffic model represents a specific user behaviour or interaction with a device, whilst the background traffic is not directly related to a user interaction. Traffic models are represented for both user level and IP packet level. The usage of these two levels of model is the following:

User level traffic model:

- This type of traffic model models the user behaviour interaction in an application. It is used with simulations which include detailed application layer, transport layer and IP layer model on top of the physical layer and MAC layer models.

· Application performance metrics specific to the application can be evaluated and a scheduling mechanism for the application of QoS in the MAC layer can be evaluated.

IP packet level traffic model:

· This type of traffic model is generally obtained from a network traffic measurement and is represented as statistical packet distributions, such as packet size distribution and packet inter-arrival time distribution at the IP layer. This model can be used with a simulation which does not include detailed protocol layers above MAC layer.

· The IP packet level traffic model can be directly applied to the MAC and PHY models. With this type of traffic mode, evaluation of both the application performance and the scheduling mechanism at the MAC layer is not an easy task.

· This type of traffic model is used in generally in system level simulations for the air link resource allocation.

These are traffic models at the IP packet level:

· Full Queue (FQ) traffic model in which it is assumed that there is an infinite amount of data bits waiting in the queue of each active user in the system. This traffic model is particularly interesting in accessing the maximum capacity of the network.

· Voice over IP traffic model.

· Near Real Time Video with an average source bit rate of 32Kbps, 2Mbps and 10Mbps.

· World Wide Web (WWW) traffic model with a source bit rate of 64Kbps, 2Mbps and 10Mbps.

· File Transfer Protocol (FTP) traffic model with a source bits rate of 64 and 384Kbps.
  Traffic models, which are used in MOTION SLS, are shown in Figure 4-23.

4.10 Performance Metrics

Along the execution of each TTI a number of statistics are collected for the computation of the metrics used in the evaluation of the performance of the system level simulation platform. These performance statistics are generated as outputs from the system level simulations and are used in the performance evaluation of the used scenarios and proposed algorithms. The following parameters are used as inputs for the computation of performance metrics:

· Simulation time per run: T_{sim}

· Number of simulation runs: D

· Total number of cells being simulated: N_{cells}

· Total number of users in cells of interest (cells being simulated): N_{users}

- Number of packet calls for user u : p_u

- Number of packets in i^{th} packet call of user u : $q_{i,u}$
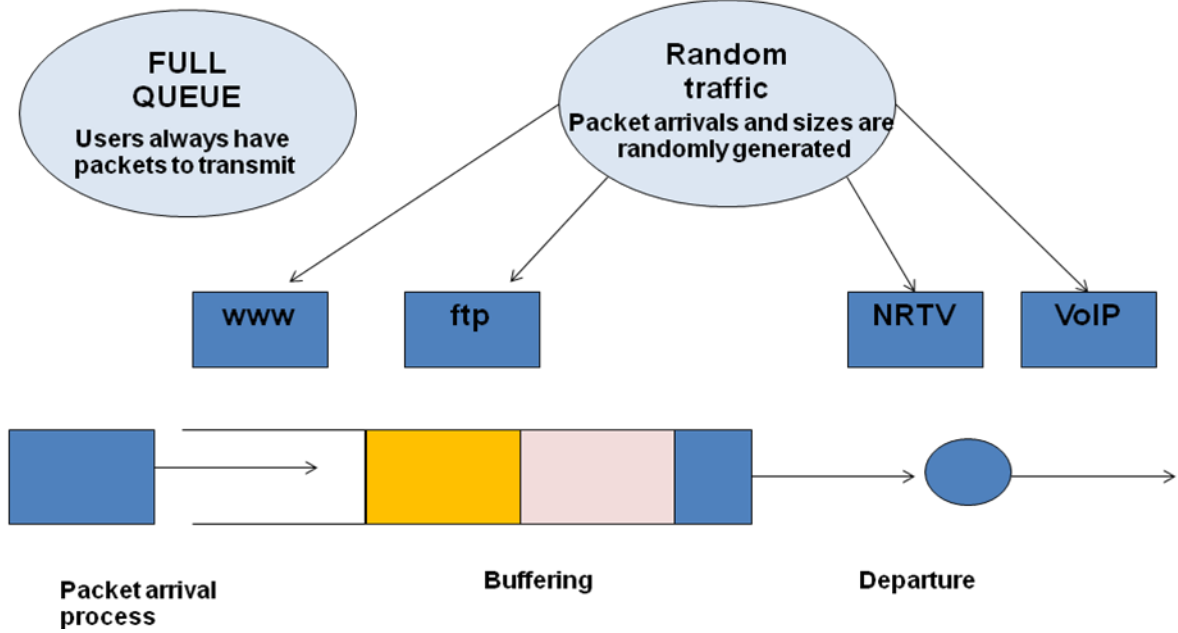
Figure 4-23 : Traffic Models

4.10.1 Throughput Performance Metrics

Average Service Throughput per-Cell

The average service throughput per cell is defined as the sum of the total amount of bits successfully received by all active users in the system, divided by the product of the number of cells simulated and the simulation duration as shown in the following equation.

$$R_{service}^{DL(UL)} = \frac{\sum_{u=1}^{N_k^{usersDL(UL)}} \sum_{i=1}^{p_{u,k}^{DL(UL)}} \sum_{j=1}^{q_{i,u,k}^{DL(UL)}} b_{j,i,u}}{N_{cells}T_{Sim}}$$

Where $N_k^{users,DL(UL)}$ is the number of users transmitting in DL(UL) in the kth cell, $p_{u,k}^{DL(UL)}$ is the number of packet calls for user u in cell k , $q_{i,u,k}^{DL(UL)}$ is the number of packets for the ith packet call for user u in cell k and $b_{j,i,u}$ is the number of bits received with success in the jth packet of packet call i for user u in cell k .

Average Over-The-Air (OTA) Cell Throughput (kbps/cell) (3GPP Definition)

The average OTA throughput per cell is defined as the sum of the total amount of bits being successfully received by all active users in the system divided by the product of the number of cells being simulated in the system and the total amount of time spent in the transmission of these packets as shown in the following equation.

$$R_{OTA}^{DL(UL)} = \frac{\sum_{u=1}^{N_k^{users,DL(UL)}} \sum_{i=1}^{P_{u,k}^{DL(UL)}} \sum_{j=1}^{q_{i,u,k}^{DL(UL)}} b_{j,i,u}}{N_{cells}.T_{Trans}}$$

Where T_{trans} is the time required to transmit these packets.

Average Over-The-Air (OTA) Cell Throughput (kbps/cell) (Peak Bit Rate Definition)

This metric is very similar to the OTA throughput. But here all bits (correct and erroneous) are considered in its computation as shown in the following equation.

$$R_{OTA\_PBR}^{DL(UL)} = \frac{\sum_{u=1}^{N_k^{users,DL(UL)}} \sum_{i=1}^{P_{u,k}^{DL(UL)}} \sum_{j=1}^{q_{i,u,k}^{DL(UL)}} b_{j,i,u}^{Trans}}{N_{cells}.T_{Trans}}$$

Where $b_{j,i,u}^{Trans}$ is the number of bits received (with error or success) in the jth packet of packet call i for user u in cell k

Offered Cell Load (kbps/cell) (3GPP Definition)

This metric is used in the evaluation of the data load (in kbps) withdrawn from the base station's buffers for transmission, i.e., the influence of the channel in the transmission of the data is not being considered as shown in the following equation.

$$R_{OL3GPP}^{DL(UL)} = \frac{b_{Sent}^{DL(UL)}}{N_{Cells}T_{Sim}}$$

Where $b_{Sent}^{DL(UL)}$ is the total amount of data bits that have been withdrawn from the base station's queues and sent over the air interface for DL(UL) connection, for all mobile stations being simulated over the whole simulation run.

Offered Cell Load (kbps/cell) (Network Definition)

This metric measures the offered load from the core network to the base station for all mobile stations being simulated in the system over the whole simulation run as shown in the following equation.

$$R_{OLNetwork}^{DL(UL)} = \frac{b_{Network_{net}}^{DL(UL)}}{N_{Cells}T_{Sim}}$$

Where $b_{Network}^{DL(UL)}$ is the total amount of data bits that have arrived to the base station's queues from the core network over the whole simulation run. The following Figure illustrates the relation among these different metrics.
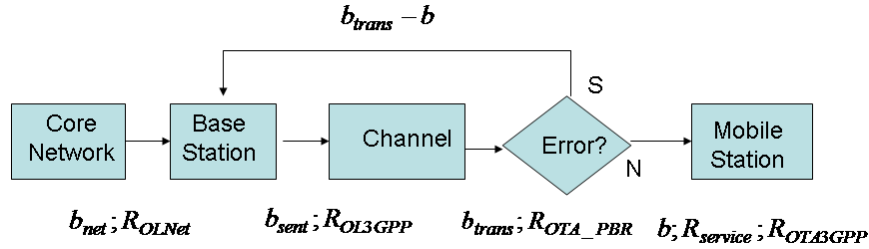
User Average Peak Bit Rate at a Given Distance (kbps)

This metric gives the average peak bit rate of a given user at a given distance d, in steps of 10m, from the base station. For one user i the average peak bit rate is defined as shown in the following equation.

$$R_{PBR}(i) = \frac{\sum_{k=1}^{N(i)} R_k(i)}{N(i).T}$$

Where $N(i)$ is the total amount of frames received by the mobile station i (both transmitted and retransmitted frames are taken into account) and $R_k(i)$ is the total amount of bits in the kth frame received by mobile station i.

Per User Service Data Throughput

The user's service data throughput is defined as the ratio of the number of information bits successfully received by the user and the total simulation run time. If user u has $p_u^{DL(UL)}$ downlink (uplink) packet calls with $q_{i,u}^{DL(UL)}$ packets for the ith downlink (uplink) packet call and $b_{j,i,u}$ bits in the jth packet the average user throughput for user u is given by the following equation.

$$R_u^{DL(UL)} = \frac{\sum_{i=1}^{p_u^{DL(UL)}} \sum_{j=1}^{q_{i,i}^{DL(UL)}} b_{j,i,u}}{T._{Sim}}$$

Per-User Average Service Throughput

The average per-user service throughput is defined as the sum of the user service throughput of each user divided by the total number of users in the system. as shown in the following equation.

$$\overline{R_u^{DL(UL)}} = \frac{\sum_{u=1}^{N_{users}} R_u^{DL(UL)}}{N_{\cdot users}}$$

Average Packet Call Throughput for a User

If there are N_{users} in the cell of interest and $R_{k,u}^{DL(UL)}$ is the service throughput for the nth user in the cell, the DL or UL service throughput for the cell is given by the following equation

$$R^{DL(UL)} = \sum_{u=1}^{N_{users}} R_u^{DL(UL)}$$

The packet call throughput is equal to the total amount of bits per packet call received with success divided by the duration of the packet call. If user u has $p_u^{DL(UL)}$ downlink (uplink) packet calls with $q_{i,u}^{DL(UL)}$ packets for the ith downlink (uplink) packet call and $b_{j,i,u}$ j^{th} packet call then the average packet call throughput is given by the following equation

$$R_u^{pc,DL(UL)} = \frac{1}{p_u^{DL(UL)}} \left(\sum_{i=1}^{p_u^{DL(UL)}} \frac{\sum_{j=1}^{q_{i,u}^{DL(UL)}} b_{j,i,u}}{\left(T_{i,u}^{end,DL(UL)} - T_{i,u}^{start,DL(UL)}\right)} \right)$$

Where $T_{i,u}^{start,DL(UL)}$ is the time instant at which the transmission of the first packet of the ith DL (UL) packet call for user u starts and $T_{i,u}^{end,DL(UL)}$ defines the time instant at which the last packet of the ith DL(UL) packet call for user u is received with success. For uncompleted packet calls this parameter is set to the simulation end time.

Average per-User Packet Call Throughput

The average per-user packet call throughput is defined as the sum of the average packet call throughput of each user divided by the total number of users in the system. as shown in the following equation.

$$\overline{R_u^{pc,DL(UL)}} = \frac{\sum_{u=1}^{N_{users}} R_u^{pc,DL(UL)}}{N_{\cdot users}}$$

Throughput Outage

The throughput outage is defined as the percentage of users with service data rate $R_u^{DL(UL)}$ less than a pre-defined minimum rate $R_{\min}$.

Cell Edge User Throughput

The cell edge user throughput is defined as the 5th percentile point of the CDF of user's average packet call throughput.

4.10.2 Performance Metrics for Delay Sensitive Applications

Packet Delay

For an individual packet the delay is defined as the time elapsed between the instant when the packet enters the queue at transmitter and the time when the packet is received successfully by the mobile station. If a packet is not successfully delivered by the end of a run its ending time is the end of the run. Assuming the jth packet of the ith packet call destined for user u arrives at the base station (mobile station) at time $T_{j,i,u}^{arr,DL(UL)}$ and is delivered with success to the mobile station (base station) at time $T_{j,i,u}^{dep,DL(UL)}$, the packet delay is defined as in the following equation .

$$Delay_{j,i,u}^{DL(UL)} = T_{j,i,u}^{dep,DL(UL)} T_{j,i,u}^{arr,DL(UL)}$$

Use Average Packet Delay

The average packet delay is defined as the average interval between packets originated at the source station (mobile or base station) and received at the destination station (base or mobile station) in a system for a given packet call duration. The average packet delay for user u is given by the following equation .

$$D_u^{avg,DL(UL)} = \frac{\sum_{i=1}^{p_u}\sum_{j=1}^{q_{i,u}}\left(T_{j,i,u}^{dep,DL(UL)} - T_{j,i,u}^{arr,DL(UL)}\right)}{\sum_{i=1}^{p_u} q_{i,u}}$$

Residual Frame Erasure Rate (FER)

This metric is computed for each user and for each packet service session. A packet service session contains one or several packet calls depending on the application. A packet service session starts when the first packet of the first packet call of a given service begins and ends when the last packet of the last packet call of the same service has been transmitted. One packet call contains one or several packets. The Residual FER is given by the following equation.

$$FER_{residual} = \frac{\eta_{dropped\_packets}}{\eta_{packets}}$$

Where $\eta_{dropped\_packets}$ is the total amount of dropped packets in the packet service session and $\eta_{packets}$ is the total amount of packets in the packet session. A dropped packet is the one in which the maximum number of transmission attempts has been achieved without the packet being successfully decoded.

Packet Loss Ratio

The packet loss ratio is computed for each user and for each packet service session and is defined as in the following equation.

$$PDR = \frac{\eta_{discarted\_packets}}{\eta_{packets}}$$

Where $\eta_{discarted\_packets}$ is the total amount of packets discarded due to time-out (delay bound violation and maximum number of transmission attempts achieved).

Spectral Efficiency (bps/Hz)

This is the ratio of correctly transmitted bits over the radio resources to the total amount of available bandwidth. The average cell spectral efficiency is defined as in the following equation.

$$SE = \frac{R}{BW_{eff}}$$

Where R is the aggregate cell throughput, BW_{eff} is the effective channel bandwidth, defined as $BW_{eff} = BW * TR$, where BW is the used channel bandwidth and TR is the time ratio of the link. For example for TDD with DL:UL=2:1, TR = 2/3 for DL and 1/3 for UL.

System Outage

A user is said to be in outage if more than a given percentage of packets experience a delay greater than a certain time. The system is said to be in outage if any individual users are in outage.

System Capacity

System capacity is defined as the maximum number of users that can be serviced without making the system exceed the maximum allowed outage probability.

Fairness Criteria
It may be an objective to have uniform service coverage resulting in a fair service offering for best effort traffic. A measure of fairness under the best effort assumption is important in assessing how well the proposed solution performs.

The fairness is evaluated by determining the normalized cumulative distribution function (CDF) of the per user throughput. The CDF is to be tested against a predetermined fairness criterion under several specific traffic conditions.

Let $T_{put}(k)$ be the throughput for user k. The normalized throughput with respect to the average user throughput for user k is given by the following equation.

$$\tilde{T}_{put}(k) = \frac{T_{put}(k)}{\underset{i}{avg}\, T_{put}(i)}$$

Moderately Fair Criteria

The CDF of the normalized throughput with respect to the average user throughput for all users is determined. This DCF shall lie to the right of the curve given by the three points in the following Table.

Normalized Throughput w.r.t average user throughput	CDF
0.1	0.1
0.3	0.2
0.5	0.5

Moderately fair criterion CDF

Short Term Fairness Indication

During the simulation, the following short-term fairness indicator should be computed and recorded every τ ms (τ is suggested to be 20 or 40) as shown in the following equation.

$$F(t) = \frac{\left| \sum_{i \in A} \hat{T}_i(t) \right|^2}{|A| \sum_{\substack{i \in A \\ i}} \hat{T}_i^2(t)}$$

Where $\hat{T}_i(t)$ is the amount of service received by the ith user in time interval $[t, t+\tau)$. A is the set of users with nonzero buffers in $[t, t+\tau)$ and $|A|$ is the cardinality of A. The minimum of $F(t)$ during the simulation time, defined as $F_{\min} = \underset{t \in \{0, \tau, 2\tau, 3\tau, \ldots T_{slot}\}}{\min} F(t)$ can serve as an indication of how much fairness is maintained all the time.

4.11   System Level Simulator Architecture

This section presents system level simulator architecture:

• TaskManager **manager1** (filename_config,name1,dirname_results): task manager constructor is called which intialises the simulation object platform. Given a directory name for the simulation results, and the input configuration file (handled thereafter by the Read_file object), it sets the mode of operation of the simulator as shown in Figure 4-24. The structure chart of the simulator after initialization is given by Figure 4-25. The chart below shows the interconnections between the core classes where the direction of information flow is indicated by the arrow heads.
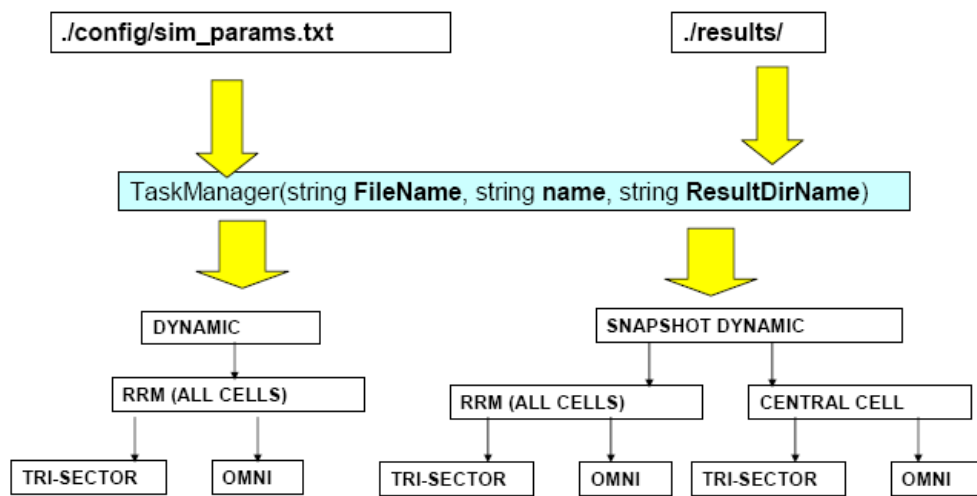
Figure 4-24: Simulator Mode of operation

- manager1.PrintInfo ():prints general information about the simulation configuration and prints this out to the console; furthermore the outputfilenames for the results are defined.

• manager1.Run():

update_timers: updates all the relevant timers of the mobile ARQ processes; associated with the dynamic resource allocation protocol and independent of the system architecture

o updateResourceAllocation: updates the resource allocation map with the current mobile id; ready for transmission. This can be effectively perceived as a FIFO buffer, that contains the resource allocation mappings for the current mobile to be transmitted, the next mobile to be transmitted on the next TTI, and the "to be scheduled mobile". Although the latter is set to default values in this function.

o updateQueueStatus: updates the queue status of each active mobile; thus it is responsible for transferring the packets from the *downlinkqueueinfo* object at time "t" to the user queue for each user on the network. The simulator always assumes, regardless of the simulator mode of operation, that all the user traffic throughout all the simulation lifetime has been pre-generated, either at the simulator initialization stage or when a new user is added to the network.

o Mobility: mobility function invoked only when the simulator is operating in dynamic mode.

o set_resources(): this function will compute the updated channel and shadowing values for each mobile on the network. (For dynamic snapshot mode, the pathloss, and the shadowing values are precalculated either at the inialisation stage, or when a new packet user is added onto the system but remain fixed throughout each simulation run). The channel values include the effect of path loss, shadowing and fast fading. (Channel value calculated herein refers to the channel coupling gain). The shadowing values are either generated based on the 2D correlated shadowing model, or using the traditional 1D uncorrelated shadowing modeling approach

o handoverCPICH: this function function invoked only when the simulator is operating in dynamic mode (hard handover is implemented).

o computeResourceAllocation(): the core of the dynamic resource allocation scheme is carried out here. The new resources are acomputed according to the following methodology:

  ▪ manage ARQ messages: will update the timers and ARQ priorities according to the ARQ messages.

  ▪ updateCellPriorities: the high and low priority lists are generated per cell; a service which is then exploited by the scheduler to choose the ARQ process for transmission based on maxCI policy.

  ▪ SchedulingV1():carries out scheduling, link adaptation, and updates the "to be scheduled mobile" data attributes (resource allocation) in each cell object.

• Packet transmission(): simulates packet transmission, and updates the ARQ process accordingly.

SIR updates on CPICH(): updates SIR measurement taking into account the delay between measurement and actually using the SIR value.

PacketQuality(): packet transmitted is rececived and quality tested to determine whether a block
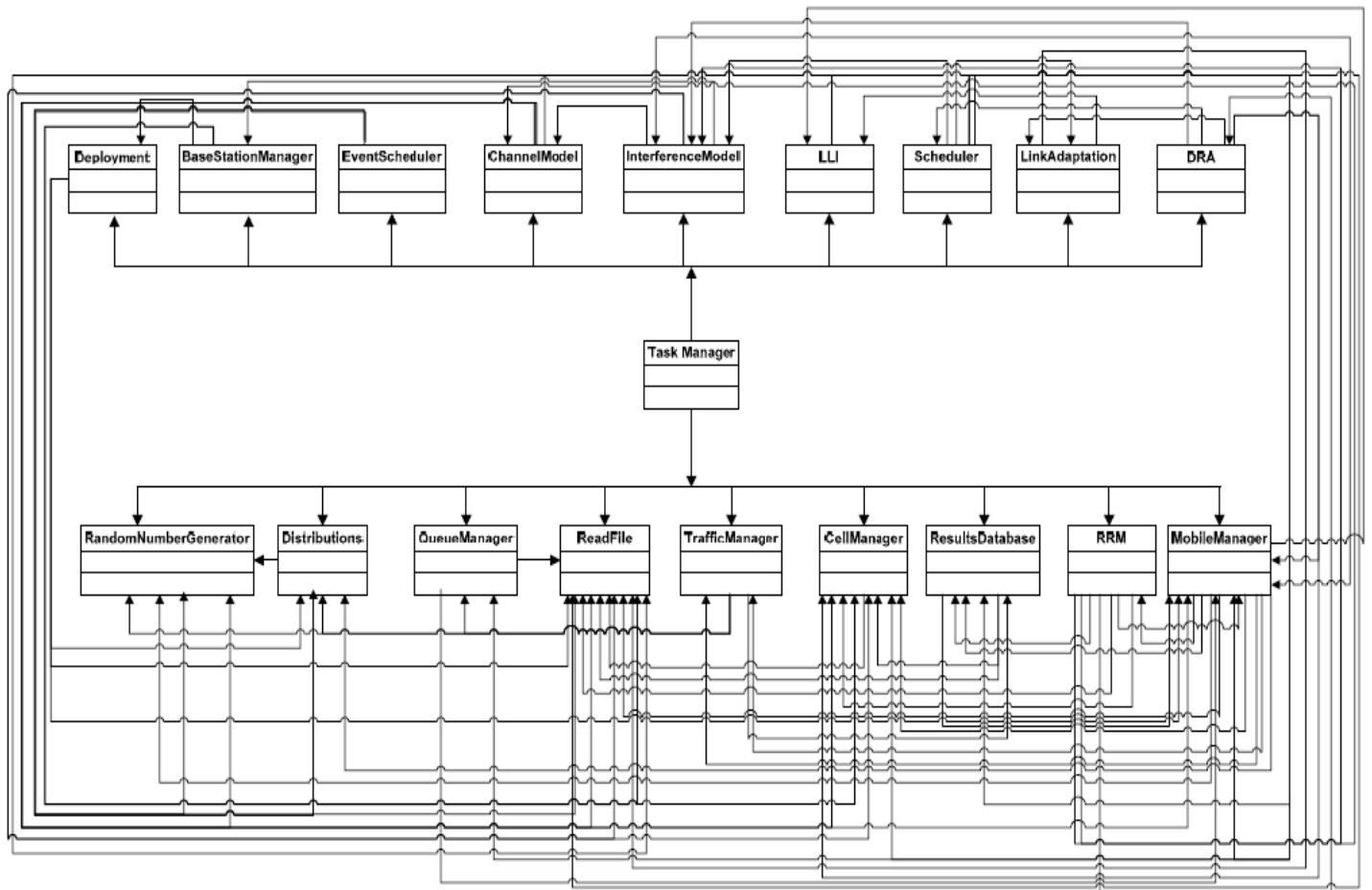
has been received in error. ARQ messages are generated accordingly

Figure 4-25: Simulator Structure Chart

- GetFirstEventAtTime: if the simulator is configured to assume a rand arrival process, then on each TTI, the task manager samples the event scheduler to asses whether an event has occurred. An event is defined to be either a *service start event*, or *service end event* and reflects a user either entering the network, or terminating the connection.

    - GetEventType: the event type will either be a *service start event*, or *service end event* and will also encompass information related to the type of service and the connection length.

- AddNewPacketUser: Every time a user enters the system, a new mobile object is created and thus this function will also invoke method functions within the mobile to discover its serving cell, channel values, position, and trigger constructor functions to create the mobility data object, channel quality indicator, ARQ processes and the associated traffic queues

- revemovePacketUser: this function will destroy the mobile object and dependant objects, and update the results database accordingly.

- RemoveEvent: will remove the event that triggered the termination of the mobile object from the event scheduler.

- measureUser: will measure the current number of users in the network.

4.12   System Profile

A TDD frame with duration of 5 ms (corresponding to the transmission time interval - TTI) and a Fast Fourier Transform (FFT) size of 1024 sub-carriers (corresponding to a channel bandwidth of 10MHz) are considered. In the TDD frame resources are available in two domains: frequency (sub-channels) and time (OFDM symbols). Each sub-channel comprises a number of sub-carriers arranged in such a way according to the type of sub-channelization used and direction of the connection. For the different types of sub-channelization schemes implemented in the Mobile WiMAX standard, see appendix A.

Two basic types of sub-channelization schemes for the diversity sub-carrier permutation mode are:

· Partial Usage Sub-Carrier Channelization (PUSC) - 720 data sub-carriers are organized into three segments of 240 data sub-carriers each. Each sub-channel comprises 30 data sub-carriers which amount to 30 data sub-channels (10 sub-channels per segment). The sub-carriers in the preamble are divided into 3 segments and the CQI is estimated from a sample of sub-carriers inside each segment. Three different measures are used to estimate the CQI, one for each segment, and all channels inside each segment are described by the same CQI associated to the same segment.

· Full Usage of Sub-Carrier Channelization (FUSC) - 768 data sub-carriers are organized into 48 adjacent groups of 16 data sub-carriers each. Each sub-channel is formed by taking one sub-carrier from each one of the 48 groups and in total there are 16 sub-channels available for data transmission. As each one of the 48 data sub-carriers are spread along the whole spectrum in the same way it is reasonable to assume that all channels will result in similar SINR values. Segmentation is not implemented and the CQI is estimated from a set of

sample sub-carriers taken over the whole spectrum of the OFDM symbol. FUSC mode needs only one CQI value to describe the sate of each one of its data sub-channels.

From a resource managing point of view it is possible to use all sub-channels available in the frame within neighbouring cells, amounting to a 1/1/1 frequency reuse factor from an interference perspective. Alternatively, sub-channel segmentation may be employed to divide the available sub-channels into three segments, each allocated to one of three cells within a base station, amounting to a 1/3/1 reuse factor from an interference perspective.

The map of resources associated to the PHY radio frame comprises a number of slots. Each slot is the smallest resource granularity which can be allocated for data transmission. A group of slots using the same MCS scheme constitutes a burst. A burst can be assigned to more than one user provided the same MCS scheme is followed in the transmission of all packets mapped into the slots of the burst.

This work follows the guidelines from the WIMAX Forum system profile which determines the implementation of the DL PUSC sub-channelization mode. With DL-PUSC configuration each sub-channel comprises 24 data sub-carriers and each slot comprises a resource space equal to two OFDM symbols by one sub-channel. Table 4.8 lists the system parameters regarding the implementation of the DL-PUSCin the basic system level simulator platform.

Parameters	Values	
System Channel Bandwidth (MHz)	10	
Sampling Frequency (F_p in MHz)	11.2	
Subcarrier Frequency Spacing (f kHz)	10.94	
FFT Size (N_{FFT})	1024	
	DL	UL
Null Subcarriers	184	184
Pilot Subcarriers	120	280
Data Subcarriers	720	576
Data Subcarriers per Subchannel	24	16
Number of Subchannels (N_S)	30	35
Useful Symbol Time ($T_b = 1/f$) in μs	91.4	
Guard Time ($T_g = T_b/8$) in μs	11.4	
OFDM Symbol Duration ($T_s = T_b + T_g$) in μs	102.9	
Number of OFDMA Symbols per frame (5ms)	48	
Data OFDM Symbols	44	

Table 4-8: WiMAX PHY Information – UL/DL PUSC sub-channels

Table 4-9 lists the size of the transport block carrier in each slot for each one of the MCS schemes implemented in the SLS. These are the basic constituents of each resource addressed in the MAC layer for data transmission.

Not all OFDM symbols in the frame are available for data transportation. Some of these symbols are used for conveying signaling and control information and also for channel estimation. This constitutes an overhead, whose size depends on the number of users scheduled in the frame and on the size of each burst in slots [18].

MCS Level	Modulation	Coding Rate	Symbol size (bits)	FEC block size (in bits
0	QPSK	½	24	48
1	QPSK	2/3	32	64
2	QPSK	¾	36	72
3	16QAM	½	48	96
4	16QAM	2/3	64	128
5	16QAM	¾	72	144
6	64QAM	½	72	144
7	64QAM	2/3	96	192
8	64QAM	¾	108	216

Table 4-9: Data rates for MCS level

In particular, the total amount of users which can be scheduled per frame depends on the type of MCS scheme used in the transmission of the downlink and uplink Mobile Application Part (MAP) control regions, defined after the preamble in the downlink sub-frame. As MAP messages are broadcasted to all mobiles in the cell and they contain vital information for the decoding of the data bursts, they must be properly decoded even for those users in the cell edge, which are more affected from inter-cell interference. As a consequence they must be transmitted with the most robust MCS scheme (rate 1/2 convolution coding (CC) and modulated with Quaternary Phase Shift Keying (QPSK)) and with a higher degree of protection by means of repetition coding (coded symbols are repeated one, two, four or six times so that mobile stations in the cell edge can successfully decode them). Therefore, the number of OFDM symbols available for data transmission decreases as the number of users scheduled in the frame increases and/or more robustness is used in transmission of DL/UP-MAP signaling.

4.12.1 Link Adaptation

At the beginning of each frame period the choice of the proper transmission mode (i.e. MCS scheme) used is defined by the link adaptation module. In the simulations 9 MCS schemes, encompassing QPSK, 16QAM and 64QAM and the convolution/Turbo encoder are used, according to the profiles envisioned by the WiMAX forum. For each burst created in the resource allocation space the MCS scheme to be used is chosen according to the decision rule defined in equation (4.31):

$$i = \arg \max_{i \in MCS_{set}} \left[R_i (1 - BLER_i) \right] \tag{4.31}$$

Where MCS_{set} represents the set of modulation and coding schemes, R_i is the throughput achieved for the selected MCS scheme and $BLER_i$ is the predicted Block Error Rate (BLER) for the MCS scheme used, which is a function of the predicted CQI reported from the mobile station ($\gamma_{k,CQI}$) and the accumulated SINR value on the Chase Combiner from previous transmission attempts ($\gamma_k^L = \sum_{t=1}^{L-1} \gamma_{k,CQI}(n-t)$) as defined in equation (4 32):

$$BLER_k^{(i)}(n) = f\left(\gamma_{k,CQI}(n) + \gamma_k^L, MCS_k(n) \right) \tag{4.32}$$

The BLER is a threshold which depends on the type of service being processed. For voice services a lower delay per packet is preferred in detriment of a higher BLER. For data services a higher delay can be supported but with a lower BLER. The scheduler selects the most spectrally efficient MCS, i.e., the one that maximizes the achievable bit rate and at the same time complies with the expected BLER under the desired threshold level. The channel quality is tracked by the CQI parameter for each channel used in the mobile station. For the DL-PUSC mode, and according to the inherent frequency diversity achieved with sub-carrier pseudo-random allocation, it is assumed that all sub-channels have the same state. The CQI is then updated with a period of T_{CQI} frames, which is a multiple of the frame period, combining past information and information provided by measurements, according to the time-smoothing formula expressed in equation (4.33):

$$CQI_k^{(i)}(n) = 0.7x\gamma_k^{(i)}(n) + 0.3xCQI_k^{(i)}(n-1) \tag{4.33}$$

Where:

· $CQI_k^{(i)}(n)$ is the value of the CQI for data sub-carrier k of user i for time period $\left(lT_{CQI}, (l+1)T_{CQI} \right)$.

· $\gamma_k^{(i)}$ is the reported CQI from measurements performed for data sub-carrier k of user i.

The CQI is estimated by each mobile station attached to the cell by means of the SINR value of the OFDM symbol carried in the frame's preamble. The CQI is reported back to the base station via the Channel Quality Indicator control Channel (CQICH) on the uplink sub-frame, on a frame-by-frame basis. The resulting CQI is a scalar value computed by the Exponential Effective SINR Method (EESM) mapping rule applied to the vector of CQI values according to equation (4.34):

$$SINR_{eff} = -\beta \ln\left(\frac{1}{N} \sum_{k=1}^{N} e^{-\frac{SINR_k}{\beta}} \right) \tag{4.34}$$

Where:

· β is the correction parameter used to adapt the formula to the different types of scenarios used in the simulations.

· N is the number of sub-carriers in the vector of CQI values.

4.12.2 Asynchronous Hybrid Automatic Repeat Request (HARQ)

The number of slots forming the data burst used in the transmission of a given MPDU depends on its size and on the selected MCS scheme. HARQ procedures for error recovery, where soft combining of information associated with new and previous erroneous transmissions, is carried out in order to minimize the amount of redundant information and power transmitted over the air interface [19,20]. According to this mechanism, each MPDU being transmitted for the first time is then mapped into one of the available HARQ processes, for simultaneous transmissions from the same user. Each HARQ process is in charge of the transmission and re-transmissions of a single MPDU until it is successfully received, and is associated to one buffer in the mobile station to store the result of the combination of successive versions of the same MPDU. On the reception of each version of the MPDU, the mobile station combines this current version with previous ones of the same MPDU using Chase Combining [20]. Retransmissions of the same MPDU keep the original MCS scheme used in the first transmission attempt.

Once an HARQ process has been selected for transmission, the scheduler must wait for an ACK/NACK message from the mobile station before selecting the HARQ process again. HARQ buffers are freed when the radio block is successfully received or when the maximum number of allowed transmission attempts $N_{attempts}$ has been achieved. Due to the time required for signaling feedback and processing of the information at both ends of the transmission chain, the minimum time interval between two successive transmissions of a particular HARQ process is equal to two frame periods.In the uplink sub-frame there is an HARQ Acknowledge (HARQ – ACK) channel region for the inclusion of one or more ACK channels(s) for HARQ support. This UL-ACK channel is implicitly assigned to each HARQ-enabled DL burst according to its order in the DL-MAP. Thus, the user can quickly transmit ACK or NACK feedback messages for DL HARQ-enabled bursts using this UL ACK channel.

4.12.3 Scheduler

Packet schedulers must be designed properly to be reactive to changes in the channel and traffic patterns, in order to respond fast to deviations from the requested QoS of even the most delay sensitive applications. The scheduler is located inside each base station to enable rapid response to traffic requirements and channel conditions. As data packets are associated to service flows with

well defined QoS requirements, the scheduler can correctly determine the packet transmission ordering through the air interface. Packets must be given priority according to the set of QoS metrics which have been negotiated between the network service provider and the end user. At each frame period the scheduler provides transmission opportunities to eligible mobiles with data to send, starting with the highest ranked user and then proceeding to lower ranked ones in sequence. The CQI reports are obtained from every user on a frame-by-frame basis and the scheduler re-computes the mobile's access priority at every frame period.

4.12.4 Type of Scheduler

Scheduling process is divided into two main groups:

(i) **Packet scheduling** methods that base the scheduling decisions on recent UE channel quality measurements (i.e. executed on a TTI basis) that allow tracking the instantaneous variations of the user's supportable data rate. These algorithms have to be executed in the Base station in order to acquire the recent channel quality information.

These methods can exploit the multiuser selection diversity, which can provide a significant capacity gain when the number of time multiplexed users is sufficient [21].

(ii) **Packet scheduling** methods that base their scheduling decisions on the average user's signal quality (or that do not use any user's performance metric at all).

4.12.4.1   Fast Scheduling Methods

• *Maximum C/I (Max. CI)*: This scheduling algorithm serves in every TTI the user with largest instantaneous supportable data rate. This serving principle has obvious benefits in terms of cell throughput, although it is at the cost of lacking throughput fairness because users under worse average radio conditions are allocated lower amount of radio resources. Nonetheless, since the fast fading dynamics have a larger range than the average radio propagation conditions, users with poor average radio conditions can still access the channel.

Proportional Fair (PF): This algorithm was firstly described in [22] and further analysed in [23]. According to [24], the Proportional Fair scheduler serves the user with largest relative channel quality:

$$P_i = \frac{R_i(t)}{\lambda_i(t)} \quad i = 1,...,N$$

where $P_i(t)$ denotes the user priority, $R_i(t)$ is the instantaneous data rate experienced by user i if it is served by the Packet Scheduler, and λ_i is the user throughput. This algorithm intends to serve users under very favourable instantaneous radio channel conditions relative to their average ones, thus

taking advantage of the temporal variations of the fast fading channel. In [24], Holtzman demonstrated that the Proportional Fair algorithm asymptotically allocates the same amount of power and time resources to all users if their fast fading are iid (identically and independently distributed) and the rate $R_i(t)$ is linear with the instantaneous EsNo. the AMC functionality. The Proportional Fair method has another interesting property. According to [25], the scheduler policy provides the so-called *proportional fairness*. This criterion has been defined in [25]: a set of user throughputs $\{\lambda_i\}, i = 1,...,N$ is proportionally fair if it is feasible and if for any other feasible vector of throughputs $\{\lambda_i^*\}$ the aggregate of proportional changes is zero or negative:

$$\sum_{"i} \frac{\lambda_i^* - \lambda}{\lambda_i} \leq 0 \ \ i = 1,...,N$$

This fairness criterion favours poor channel quality flows less emphatically than the max-min fairness criterion, which allows no increase in any λ_i, regardless of how large the increase is, if it is at the expenses of the throughput decrease of any other user j such $\lambda_j < \lambda_i$, no matter how small the decrease is.

Fast Fair Throughput (FFTH):

This method aims at providing a fair throughput distribution among the users in the cell (in a max-min manner), while taking advantage of the short term fading variations of the radio channel. In [29], Barriac proposes an elegant modification of the Proportional Fair algorithm to equalize the user throughput:

$$P_i = \frac{R_i(t)}{\lambda_i(t)} \cdot \left[\frac{max_j\left\{\overline{R_j(t)}\right\}}{\overline{R_i(t)}} \right]$$

where P_i describes the priority of user i, $R_i(t)$ represents the supportable data rate of user i at instant t, $\overline{R_i(t)}$ is the average supportable data rate of user i, $max_j\left\{\overline{R_j(t)}\right\}$ is a constant that indicates the maximum average supportable data rate from all the users, $\lambda_i(t)$ represents the throughput of user i up to instant t. Note that the $R_i(t)$ term in the denominator compensates the priority of less favourable users, and distributes evenly the cell throughput to all users if their fast fading are iid and the rate $R_i(t)$ is linear with the instantaneous EsNo.

4.12.4.2   Slow Scheduling Methods

• *Average C/I (Avg. CI)*: This scheduling algorithm serves in every TTI the user with largest average C/I with backlogged data to be transmitted.

• *Round Robin (RR)*: In this scheme, the users are served in a cyclic order ignoring the channel quality conditions. This method outstands due to its simplicity, and ensures a fair resource distribution among the users in the cell. It is interesting to observe that the Round Robin scheduling method satisfies the proportional fairness criterion described by Kelly in [25].

• *Fair Throughput (FTH)*: There are various options to implement a fair throughput scheduler without exploiting any a priori information of the channel quality status [23]. Users are served in every TTI with lowest average throughput. This method is considered as a slow scheduling one because it does not require any instantaneous information of the channel quality.

From an implementation point of view, the slow scheduling methods have a lower degree of complexity than fast scheduling ones, because it require the information of channel quality measurements for all the users in the cell.

The Proportional Fair is a good choice to be implemented because it provides a competitive performance as well as fairness. In addition it takes into account channel quality as well as previously achieved throughput. In contrast, MAX C/I schedulers are better from a cell aggregate throughput perspective. The reason is that MAX C/I take into account the user channel quality and tries to allocate the resources to the user with the best channel quality leading to high data rates for that user.

The Round Robin scheduler is better from the user perspective since it is a fair scheduler that tries to distribute the resources equally among the users. So each user will get some resources to transmit with, whereas in MAX C/I some users with bad channel conditions might never get the chance to transmit anything, which is called as the user starvation problem. It is expected that distributed antenna systems will improve the fairness metrics of MAX C/I schedulers as, on average, users across the cell have a higher probability of getting a good channel with respect to any of the distributed nodes. Table 4-10[23] summarizes the Packet Scheduling methods.

Figure 4.26 shows the entire resource allocation process. At the beginning of each frame the scheduler allocates resources and builds a list of requests; required information, i.e. users' MCSs and the amount of data in each queue, are retrieved by the scheduler from the queues' manager and the link adaptation module.

Requests are passed to the frame mapper for creating rectangular data regions in the frame and obtaining the final frame structure. It is important to remark that due to the rectangular shape constraint, the percentage of the frame actually mapped depends on the particular request list and can vary in each frame [28,29]. Therefore a mapping algorithm able to map on average up to 90% of the frame has been implemented and a feedback on actually allocated slots is provided by the frame mapper when the frame structure is completed in order to update users' average throughput.

Finally, just before transmission, packets are drawn from the queues and MAC PDUs are formed adding MAC overheads and applying fragmentation and packing if needed [30].

The Link Adaptation module is responsible for the adaptation of MCSs according to the CQI measured and reported by each SS. In this case the CQI corresponds to the SINR calculated through the so-called EESM link to system interface.

| PS Methods	Scheduling Rate	Serve Order	Radio Resource Fairness
Round Robin (RR)	Slow($\approx$ 100ms)	Round robin in cyclic order	Proportional throughput fairness & Same amount of average radio resources
Fair Thought (FTH)	Slow($\approx$ 100ms)	Served user with lowest average throughput	Max-min throughput fairness
C/I based(C/I)	Slow($\approx$ 100ms)	Served according to highest slow-averaged channel Quality	Unfair distribution of radio resources in favour of high G Factor users.
Proportional Fair(PF)	Fast ($\approx$ Per TTI basis)	Served according to highest relative instantaneous channel quality	Proportional throughput fairness & Same amount of average radio resources under certain assumptions
Fast Fair Throughput (FFTH)	Fast ($\approx$ Per TTI basis)	Served according to highest equalized relative instantaneous channel quality	Max-min throughput fairness under certain assumptions
Maximum C/I	Fast ($\approx$ Per TTI basis)	Served according to highest instantaneous channel quality	Unfair distribution of radio resources in favour of high G Factor users

Table: 4-10 Summary of Packet Scheduling Methods
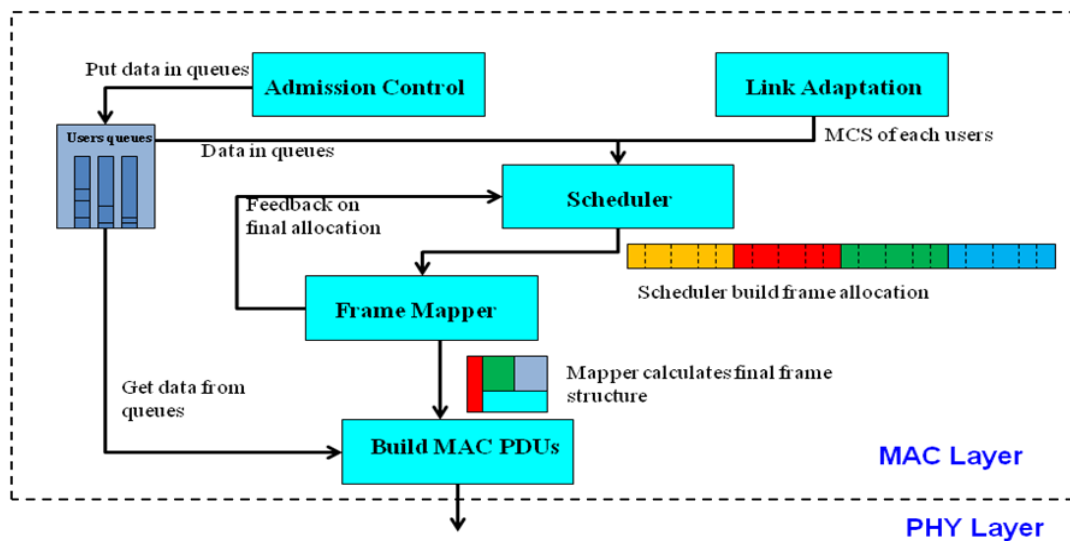
Figure 4-26 : MAC functional block

4.12.5 Resource Manager

Once the list of scheduled users is composed the resource manager starts with the resource allocation process. The resource manager is responsible for the allocation of sub-channels and power over each sub-channel/resource unit. The allocation of the proper sub-channels should be able to exploit physical layer information such as SINR, MCS level and velocity of the user, attempting to maximize resource efficiency and constrained to the power available for data transmission at the base station. The velocity is very important because it determines the proper type of sub-channelization mode used (adjacent AMC or diversity permutation). The sub-channel allocation algorithm should also satisfy the applications' QoS requirements. But, as the scheduler is normally assigned the task of selecting the users and dimensioning the amount of resources to be allocated in each frame, the resource manager does not take into account applications requirements.

4.12.5.1 Resource Map Definition

Figure 4-27 illustrates the IEEE802.16e TDD frame structure. Although mobile WiMAX standard support both frequency division duplexing (FDD) and time division duplexing (TDD), only the TDD mode is being supported by the system profiles designed by the WiMAX forum for equipment compliance and interoperability. Also TDD offers some advantages over the FDD mode such as the support of asymmetrical data rates in UL and DL and also the fast estimation of the downlink channel in the UL transmission [31,32].

In the frame, resources are available in two domains: frequency (sub-carriers) and time (OFDM symbols). The OFDM symbols available for data transmission and the subcarriers constitute distinct logical slots. Different types of logical channels are available, depending of the type of channel model used. The standard defines two basic types of generic channel modes: channel modes that are intrinsically diverse in frequency called diversity sub-carrier permutation modes and the channel modes that make use of the frequency selective nature of the radio channel.

We focus primarily on the so-called Partially Usage Subcarrier Channelization (PUSC) channel mode. In PUSC mode, every sub-channel is available for allocation in the cell. The data subcarriers are randomly allocated along the whole spectrum of the FFT. In the PUSC mode the small unit of allocation is the slot. A group of slots defines a burst in the radio frame. A burst can be assigned to more than one user provided the same modulation and coding scheme is followed in the transmission of all packets allocated to it. PUSC schematic description of the Resource Allocation Map (RA) in the System Level Simulator is given by Figure 4-28.

Total number of symbol available in one OFDM frame are 48, as we consider only downlink transmission.Therefore, our frame contains 30 OFDM symbols (35-5=30 , 5 symbols for control information e.g. Preamble, FHC the and remaining symbols are used for uplink transmission.)

According to PUSC configuration, 1 slot contains 2 OFDMA symbols in time domain and 720 data subcarriers in frequency domain. So, we have a total of 30 logical sub-channels that contain 24 subcarriers each (30*24=720) in frequency domain.

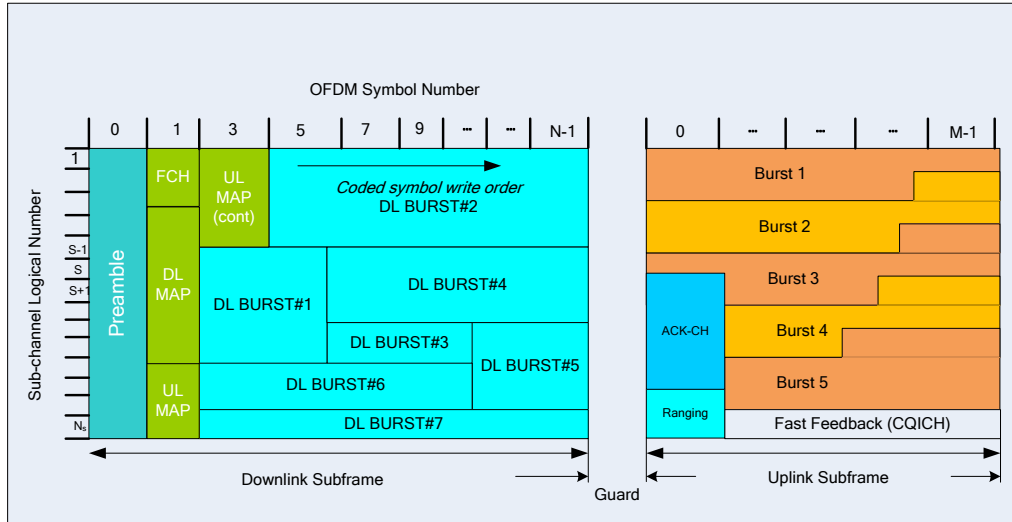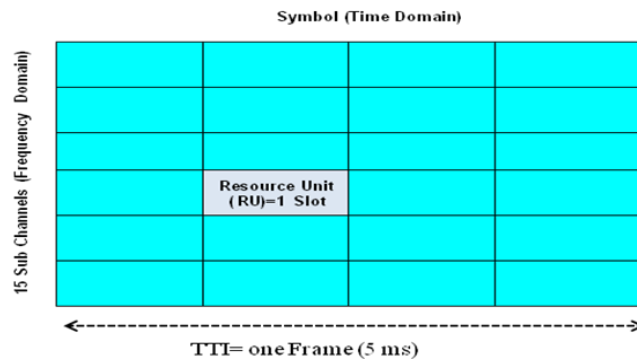**Figure 4-27 : WIMAX Frame Structure**

.

Figure 4-28 : Resource Allocation Map (RAM-PUSC)

To calculate the theoretical throughput of each coding and modulation schemes using PUSC configure, we illustrate the following numerical example.

Consider we have,

      1 slot = 2 OFDM symbols

      1 sub-channel = 24 data subcarriers

      Downlink symbol= 30 symbol

      Frame length= 5ms

Then, total number of slot in one OFDM frame is (30/2)*30 = 450 slots, and 48(2*24) sub-carriers. For example, each slot is modulated with 64QAM and 3/2 coding rate then, 48*$\log_2$ (64)*3/4=216 bit.Peak data rate=216*450(slots)/5ms = 19440000 bit per second or 19.44 Mbps as shown in Table 4-10.

MCS	Theoretical Throughput(Mbps)
QPSK ½	4.32
QPSK ¾	6.48
16 QAM ½	8.64
16 QAM ¾	12.96
64 QAM ½	12.96
64 QAM ¾	19.44

Table 4-10: Theoretical throughputs

It is worth mentioning that a smaller RAU size will result in higher overhead because more users can be assigned resources in the frame (depending on the offered load per user). Meanwhile, a larger RAU size will result in higher error rates because the BLER must be computed for all RAUs used in the transmission of the MPDU, assuming a large burst of data is transmitted. The size of the RAU is a trade-off between efficiency in data allocation per resource, overhead (in terms of DL/UL-MAP size and padding bits to fill each RAU) and the probability of error in the decoding of each resource. The adequate size depends on the type of traffic model being supported. For example: web traffic model packets have much larger size than VoIP ones. It results then that it is more efficient to map web packets into resources of larger size than VoIP packets and vice-versa for VoIP packets.

If fractional frequency reuse is used users in the edge of the cell (with a fraction of the bandwidth available) will have a smaller set of RAUs for data allocation. DL-PUSC mode allows segmentation of the resource space into 3 sets of orthogonal instances of the MAC layer. Therefore the resource space for these users is made up of 5 RAUs only.

Whenever a given user is selected for transmission, either for a new or for an old MPDU assigned to an active HARQ process, its packets are concatenated and/or segmented in order to fill the amount of RAUs available for allocation in the frame. Padding bits are added if necessary. The set of RAUs assigned to the user constitutes a given burst in the frame and each burst is composed of a group of contiguous RAUs.

Each burst is individually assigned to a single user and all slots in the same burst are transmitted with the same MCS scheme. Subsequent to scheduling and mapping of the IP packets to the available RAUs, the base station broadcasts the map with the RAM to the network via DL-MAP control fields in the DL sub-frame. Each mobile station utilizes this signaling information to anticipate packet transmission and to perform the demodulation and decoding of the transmitted data using the appropriate MCS scheme. The Resource Allocation Map is updated every frame by the base station.

In the beginning of each TDD frame the scheduler assigns free RAUs from the resource map according to the individual priorities assigned to each user by the scheduler. Whenever a user is selected by the scheduler, the scheduler withdraws the selected IP packets from the respective queue or from the respective HARQ process in case of retransmission. The amount of RAUs requested for transmission depends on the size (in bits) of the amount of packets in the queue. The size is computed by the link adaptation module. The whole process elapses along a cycle equivalent to two frames and is illustrated in Figure 4-29.

· **Phase A1:** The base station computes the RAM to be used for data transmission during the Phase C1 that follows.

· **Phase B1:** The base station broadcasts in the DL-MAP sub-field the RAM determined during Phase A1, so that all users in the cell have the information required to demodulate the packets they will receive during Phase C1.

· **Phase C1:** The base station sends data according to the RAM determined in Phase A1.

· **Phase D1:** The base station receives ACK messages from users which have received their packets successfully during Phase C1, and likewise receives NACK messages from users which have received their packets un-successfully during Phase C1. This signaling is sent in the UL-ACK region in the uplink sub-frame.

· **Phase A2:** The mobile station receives the broadcasted RAM from the base station.

Phase B2: The mobile station receives the MPDU mapped onto a set of RAUs, according to the RAM received in Phase A2.

· **Phase C2:** The mobile station processes the received MPDU. If it is only another version of a first MPDU transmission the mobile station performs Chase Combining on the different replicas of the same packet.

Phase D2: After decoding the received MPDU the mobile station sends an ACK/NACK message to the base station if it has received the MPDU with or without error respectively during Phase B2. This is to inform the base station about the status of the decoding process
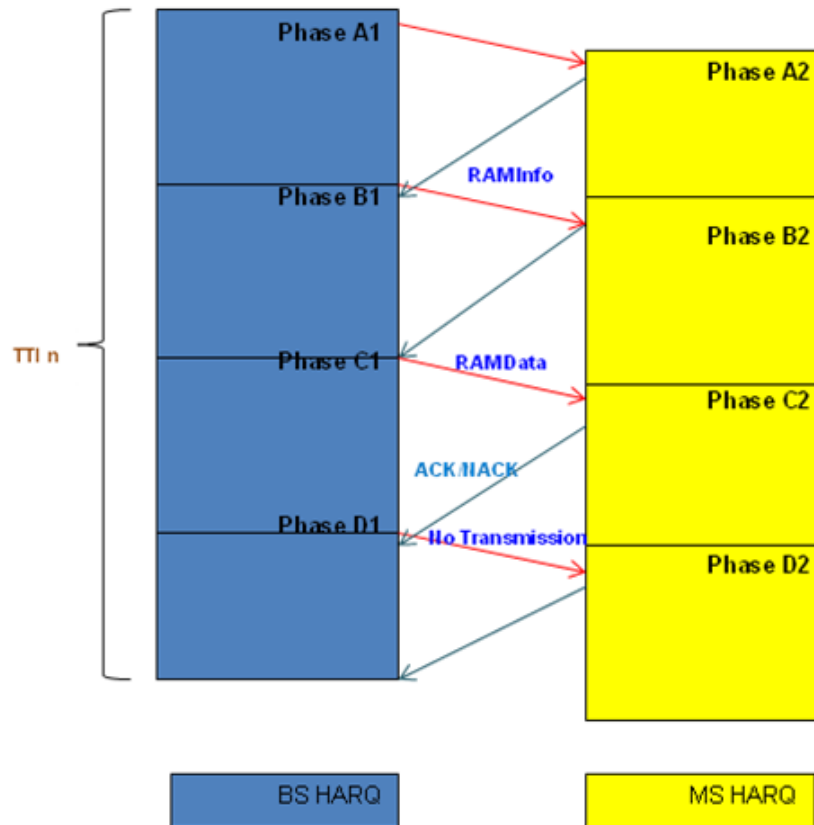
Figure 4-29: WiMAX DRA cycle

Each HARQ process of each mobile station handles only one MPDU mapped onto a given amount of RAUs of the MAC frame. The same amount of RAUs is used in the retransmission of the MPDU, with the same MCS scheme. In each frame period each HARQ process is in one of two states:

· **Active State:** The HARQ process number n of mobile station k is busy handling a MPDU. Different versions of the same MPDU may be sent over the air-interface if the HARQ process is retransmitting the MPDU. The initial version of the MPDU is stored in the buffer dedicated to the HARQ. The result from the combination of successive received versions of the same MPDU is stored in the buffer dedicated to HARQ process number n of mobile station k.

· **Inactive State:** The HARQ process number n of mobile station k is idle. In the base station the buffer dedicated to the HARQ process number n is empty or being used by another HARQ process number n from another mobile station different from mobile station k.

4.12.5.2 Resource Allocation Procedures

The description of the resource allocation procedures encompasses the methodologies for packet scheduling, resource map allocation and link adaptation. The resource allocation procedure is sub-divided in four steps:

Step 1: Determining HARQ processes lists

In the beginning of each scheduling period two lists of HARQ processes are computed: the Low Priority list and the High Priority list.

· **High Priority List** – This list comprises only HARQ processes waiting for another transmission attempt. After the first transmission attempt, and if there is an error, a timer (Timer Priority) is activated. If the HARQ process is not scheduled again before the expiration of this timer, it will be inserted in the High Priority list. At the same time another timer (Timer Discard) is activated. If the HARQ process is not released before the expiration of this timer, the HARQ process is initialized and the corresponding information stored in its buffer is lost. Figure 4.30 illustrates the flow chart describing the creation of the two types of priority lists and the activation/deactivation of both timers. The delay threshold is service specific: for real-time services such as VoIP or near-real time video (NRTV), the delay threshold is zero, which means that the HARQ process will be inserted into the High Priority list the next scheduling period. For non-real time services such as the WWW the delay threshold can be larger.

· **Low Priority List** – For a mobile with new packets stored in their buffer, an inactive HARQ process, which can be assigned for the first transmission attempt, is searched for in the set of HARQ processes assigned to the mobile. If there is any it is inserted into the low priority list. This process is repeated for all active users in the cell and with new packets. The HARQ processes waiting for another transmission opportunity and with the Priority Timer not set are also inserted into this list.

It is assumed that the base station hosts a waiting queue for each mobile station. Both lists are then sorted in descending order of the priority, computed for each HARQ process according to the scheduling algorithm. The base station selects the HARQ processes with highest priority in each list. The principle behind the definition of the Priority Timer is to increase the probability of serving a given user, with an active HARQ process in retransmission, before the delay bound is achieved for the packets stored in its buffer. Of course this mechanism is not fully QoS-compliant because new packets can remain in buffer until they are dropped without service, as they are no put into the High Priority List.
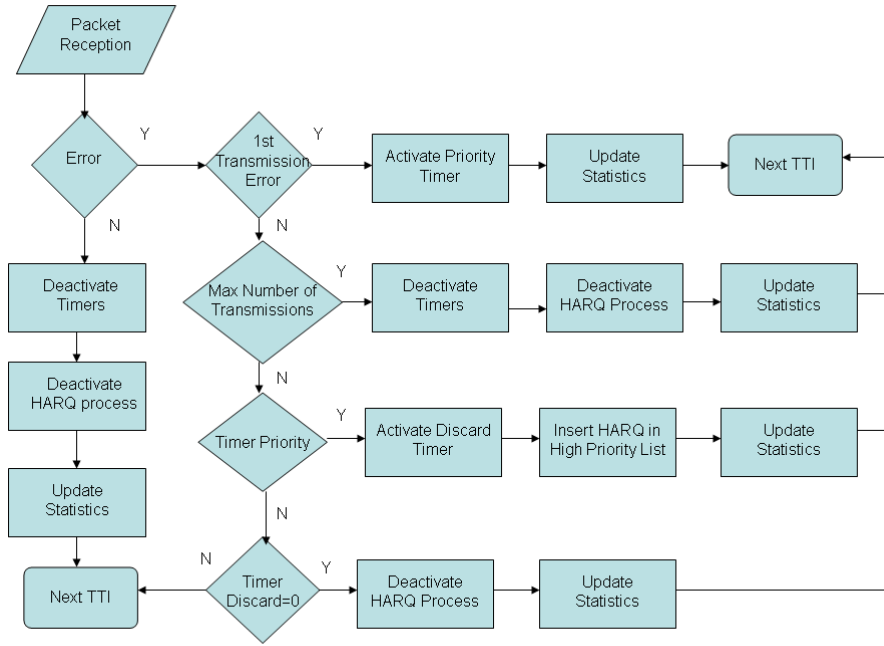
Figure 4-30: Flowchart with the creation of priority lists and activation/deactivation of timers

The Discard Timer avoids an HARQ process being indefinitely in active state, which could result in a situation where a given mobile has no HARQ processes for transmission of new packets arriving to its buffer. The definition of the proper number of HARQ process to assign to each user depends on the maximum delay bound of the service, the value assigned to the Discard Timer and on the complexity pretended for the simulation.

Step 2: Scheduling HARQ processes in High Priority list

All HARQ processes in High Priority list are attempting a retransmission. Therefore, the same MCS scheme, same power per resource and same amount of resources are re-used in the retransmission. The resource allocation module iterates the High Priority list, starting with the process with highest priority. Then the same RAUs in the map of resources are allocated, with the same power, data and MCS scheme used in previous transmission, to the HARQ process. If there are any, RAUs remaining unallocated after all HARQ processes in the list are served, are deactivated and put into the state of availability in order to be assigned to the HARQ processes in the Low Priority list.

The remaining power for data transmission, P_{budget}, is computed by adding the power assigned to each RAU already occupied in the map of resources, P_{data}, and subtracting the result from the maximum power available for data transmission in the cell, $P_{\max}$. It is given by equation (4.35).

$$P_{budget} = P_{\max} - P_{data} \tag{4.35}$$

Step 3: Scheduling HARQ processes in Low Priority list

It there are RAUs available for data transportation after the HARQ processes in the High Priority list are serviced then the resource allocation module iterates the Low Priority list. The HARQ processes in retransmission are scheduled only if there are enough RAUs for allocation as the MPDU stored in its buffer cannot be fragmented. The same MCS scheme is selected for the transmission of a new version of the MPDU. The power per RAU is computed according to subsections below.

For HARQ schemes performing a first transmission attempt the link adaptation module selects the most appropriate MCS scheme, $SINR_{selected}$ according to the predicted SINR, $SINR_{pred}$, based on the frame's preamble and reported from each mobile. Mobile stations with a CQI value lower then a given admission threshold are not considered by the scheduler even if they have information ready for transmission. Mobile stations with a CQI which do not correspond to a transmission with quality good enough to result in a high probability of success in the decoding of the packet are transmitted with the most robust MCS scheme.

The following steps are performed until one of the following three conditions is not satisfied anymore: (i) the list of priorities is empty, (ii) there is no more information to be transmitted (iii) or there are no more resource units for allocation.

Step 3.1: - The base station selects the mobile with highest priority in each list. The mobiles from the list of higher priority have preference in the resource allocation.

Step 3.2: The base station computes the current remaining transmit power per resource unit, $P_{t,RU}$, in function of the remaining transmit power, P_{budget}, and the number of remaining resource units, N_{budget}, in the resources map. It is given by equation (4.36).

$$P_{t,RU} = P_{budget} / N_{budget} \tag{4.36}$$

Step 3.3: The best MCS scheme, $MCS_{selected}$, is selected according to the predicted CQI, CQI_{pred}, following the link adaptation mechanism. The chosen criteria is the maximization of the expected throughput, $R(1 - BLER_{pred})$, while keeping the expected BLER, $BLER_{pred}$, lower than the threshold value, $BLER_{thres}$. This threshold depends on the type of service. The predicted SINR is computed as in equation (4.37).

$$SINR_{pred} = CQIxP_{t,RU} \tag{4.37}$$

Step 3.4: The power per resource, $P_{t,RU}$, is computed according to one of two different strategies:

Uniform power distribution per resource

The power available for data in the cell is uniformly distributed per each resource unit in the map of resources, according to equation (4.38).

$$P_{t,RU} = P_{dat} / N_{RU} \tag{4.38}$$

Non-uniform power distribution per resource

The power assigned to each resource depends on the reported CQI. If the selected MCS scheme is the most robust one and the expected BLER is larger than the threshold one for this service $\left(BLER_{pred} > BLER_{thres}\right)$ then a larger transmit power may be allocated to each resource unit assigned to the mobile. The new power per resource unit must be such that the expected SINR, $SINR_{pred}$, equals the threshold, $SINR_{thres}$. It is computed according to equation (4.39).

$$P_{t,RU} = \min\left(SINR_{thresh} / CQI, P_{t,RU,\max}\right) \tag{4.39}$$

Where $P_{t,RU,\max} = P_{budget}$. The value of the threshold for the SINR, $SINR_{thresh}$, is computed from the LUT in order to make the predicted BLER to be equal to the expected threshold for the service. The new value of the power to be assigned to each resource unit cannot be greater than the remaining power, P_{budget}.

The maximum number of resource units, N_{RU}, available for allocation for the computed power per resource unit, $P_{t,RU}$, is then given by equation (4.40).

$$N_{RU} = P_{budget} / P_{t,RU} \tag{4.40}$$

Step 3.5: The number of bits carried in each resource unit is deduced from the selected MCS scheme and the available number of resource units. It is given by equation (4.41).

$$N_{transmitable} = N_{RU} \, xSize_{RU} \tag{4.41}$$

Step 3.6: The number of bits that will be transmitted is a function of the number of bits waiting in the base station queue, N_{queue}. It is computed according to equation (4.42).

$$N_{transmit} = \min(N_{queue}, N_{transmitable}) \tag{4.42}$$

Step 3.7: The number of resource units to be allocated in the map of resources is computed by equation (4.43)

$$N_{RU,transmit} = \left\lceil \frac{N_{transmit}}{Size_{RU}} \right\rceil \qquad (4.43)$$

Step 3.8: The Resource Allocation Map is updated by allocating $N_{RU,transmit}$ to the given mobile with the selected MCS scheme $MCS_{selected}$ and the computed transmit power per resource unit: $P_{t,RU}$.

Step 3.9: The base station updates the mobiles list as well as the remaining power available for transmission, P_{budget}, and the amount of resource units still available for allocation, N_{budget}.

Step 4: Interaction with the physical layer

Finally the Resource Allocation Map is broadcasted by the base station, providing the respective mobiles with data mapped in the resources map with control information regarding the allocation of bursts, the MCS scheme used in the transmission and respective transmission power assigned to each resource unit.

In Summary the main characteristics of the scheduler are the following:

· It handles mixed traffic (in particular non-real time services and real-time services).

· HARQ processes are organized in two priority lists. Higher priority is given for HARQ processes active and waiting for a retransmission after a given period of time defined by the priority timer.

· In the lower priority list the priority of the new transmissions and re-transmissions is defined by the scheduler.

· Users in the edge of the cell may be assigned a higher power per resource unit if there is power available for allocation.

· Re-transmissions are performed with the same amount of resources and the same MCS scheme selected form the first transmission attempt.

· Priority computations are based on the type of service, waiting queue status and channel current conditions.

4.13 Conclusions

This chapter presents in a detailed way the design, modeling and implementation of the Link level and System Level simulators. Traffic models and channel models for SLS were also discussed in this chapter. The cellular layout architecture for SLS was fully described. Of particular interest in the realization of system level simulations is the definition of the proper link to system level interface and the definition of the proper set of look up tables used in the mapping of physical layer performance. The procedure followed into the derivation of the Signal to Interference plus Noise Ratio (SINR) and the mapping function used to map the vector of SINRs into a single scalar to be inputted into the look-up tables is detailed. This chapter also describes in detail all steps followed in each frame period by the communication protocol, between the base and the mobile stations, for the transmission of the information associated to the scheduled users.

4.14 References

[1]   IEEE Standard 802.16e-2005, Amendment to IEEE Standard for Local and Metropolitan Area Networks -Part 16: Air Interface for Fixed Broadband Wireless Access Systems-Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands..

[2]   Effective SINR approach of link to system mapping in ofdm/multi-carrier mobile network", Esa Tuomaala, 2005

[3]   "System-level evaluation of OFDM – further considerations", R1-0313103, RAN WG1 #35..

[4]   ETSI, "Universal Mobile Telecommunications System (UMTS); Selection procedures for the choice of the radio transmission technologies of the UMTS (UMTS 30.03 version 3.2.0)", TR 101 112 v3.2.0, April 1998.

[5]   Gudmundson, M.; "Correlation Model for Shadow Fading in Mobile Radio Systems", Electronics Letters, vol. 27, pp. 2145-2146, Nov. 1991.

[6]   W.C. Jakes, Microwave Mobile Communications, Wiley, New York, 1974

[7]   Yunxin Li; Xiaojing Huang; "The generation of independent Rayleigh faders", Proceedings of the IEEE International Conference on Communications (ICC 2000), vol. 1, June 2000, pp. 18-22.

[8]   Yunxin Li; Xiaojing Huang; "The generation of independent Rayleigh faders", Proceedings of the IEEE International Conference on Communications (ICC 2000), vol. 1, June 2000, pp. 18-22.

[9]   ITU, "Guidelines for evaluation of radio transmission technologies for IMT-2000" Recommendations ITU-R M.1225, 1997.

[10]   3GPP R1-030224, Nortel Networks, "Update of OFDM SI simulation methodology".

[11]   Mobile WiMAX – Part I: A Technical Overview and Performance Evaluation, WiMAX Forum.

[12]   3GPP R1-030224, Nortel Networks, "Update of OFDM SI simulation methodology".

[13]   Medbo J.; Anderson H.; Schramm, P.; Asplund, H.; "Channel models for HIPERLAN/2 in different indoor scenarios", COST259 TD(98), Bradford, UK, April 23-24 1998.

[14]   3GPP TR 25.892 (2004) "Feasibility study for OFDM for UTRAN enhancement, V1.1.0.

[15]   Wang, F.; Ghosh, A.; Love, R.; Stewart, K.; Ratasuk, R.; Bachu, R.; Sun, Y.; Zhao, Q.; "IEEE 802.16e system performance: analysis and simulations", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2005, (PIMRC 2005), vol. 2. pp. 900-904.

[16]   3GPP TSG-RAN1#48 R1-070674, LTE physical layer framework for performance verification, Feb. 2007.

[17]   ETSI SMG2 Universal Mobile Telecomunnications System (UMTS); "Selection procedures for the choice of radio transmission technologies of the UMTS", TR 101 112.

[18] Gao, Y.; Zhang, X.; Jiang, Y.; "Performance Evaluation of Mobile WiMAX with Dynamic Overhead", EEE Vehicular Technology Conference, 2008, (VTC 2008), Fall, Sept. 2008, pp. 1-5.

[19] Wang, F.; Ghosh, A.; Sankaran, C.; Fleming, P.; Hsieh, F.; Benes, S.; "Mobile WiMAX systems: performance and evolution", IEEE Comm. Magazine, vol. 46, no. 10, Oct. 2008, pp. 41-49.

[20] Chase D.; "Code Combining – A Maximum Likelihood Decoding approach for Combining an Arbitrary Number of Noisy Packets," IEEE Transactions on Communications, vol.33, pp.385 – 393, May 1985.

[21] Knopp, R.P.; Humblet, P.A.; "Information Capacity and Power Control in Single-Cell Multiuser Communications", Proceedings of the International Conference on Communications, (ICC), June 1995, Seattle, USA.

[22] Bender P.; Black P.; Grob M.; Padovani R.; Sindhushayana N Viterbi S, "CDMA/HDR: a bandwidth efficient high-speed wireless data service for nomadic users", IEEE Communications Magazine, vol. 38, no. 7, Jul. 2000, pp. 70-77.

[23] Kolding T. et al. *Performance Aspects of WCDMA Systems with High Speed Downlink Packet Access (HSDPA).* Vehicular Technology Conference, 2002. VTC 2002 Fall. Volume 1. pp. 477-481.

[24] 24 Holtzman J.M. *CDMA Forward Link Waterfilling Power Control.* Vehicular Technology Conference, 2000. VTC 2000 Spring. Volume 3. pp. 1663-1667.

[25] Jalali A. et al. *Data Throughput of CDMA-HDR a High Efficiency-High Data Rate Personal Communication Wireless System.* Vehicular Technology Conference, 2000. VTC 2000 Spring. Volume 3. pp. 1854-1858.

[26] Elliott R.C. et al. *Scheduling Algorithms for the CDMA2000 Packet Data Evolution.* Vehicular Technology Conference, 2002. VTC 2002 Fall. Volume 1. pp. 304-310

[27] Berger L. et al. *Interaction of Transmit Diversity and Proportional Fair Scheduling.* Vehicular Technology Conference, 2003. VTC 2003 Spring. Volume 4. pp. 2423-2427

[28] Holtzman J.M. *Asymptotic Analysis of the Proportional Fair Algorithm.* Personal, Indoor and Mobile Radio Communications, 2001 12th IEEE International Symposium on , Volume: 2, Sep 2001. pp. F-33-F-37.

[29] Barriac G., et al. *Introducing Delay Sensitivity into the Proportional Fair Algorithm for CDMA Downlink Scheduling.* Spread Spectrum Techniques And Applications, 2002 IEEE Seventh International Symposioum on, Volume 3, 2002. pp. 652-656

[30] Jain, R.; Chakchai, So-In; Al Tamimi, A-k; "System-level modeling of IEEE 802.16e mobile WIMAX networks: key issues", IEEE Wireless Communications Magazine, vol. 15, no.5, Oct. 2008, pp73-79.

[31] Spatial channel model for multiple-input multiple-output simulations (Release 6), 3GPP TR 25.996, 2003-05

[32] "Multilayer Optimization in Radio Resource Allocation for the Packet Transmission in Wireless Networks" Alberto de Jesus Nascimento,PhD Thesis 2009,University of Aveiro

Chapter 5: Enhanced System Level Simulator with Opportunist Radios

The UMTS radio frequency spectrum has become, in a significant number of countries, a very expensive commodity, and therefore the opportunistic use of these bands could be one way for the owners of the licenses to make extra revenue. This chapter explain how we utilize these frequency bands by using opportunistic radio and further elaborate how opportunistic radio works in an ad-hoc manner.

5.1   Introduction

In the previous chapters, we focused on the development of system level simulator for conventional cellular systems and respective performance evaluation. There are currently trends towards a more liberalized approach of spectrum management, which are tightly linked to what is commonly termed as Cognitive Radio (CR). The development of such systems would include the sharing of the spectrum by several systems, which could for example include the coexistence of a cellular network with a cognitive radio one. In order to propose such scenario it is necessary to evaluate the implications of the coexistence on the performance of the different systems to ensure that the required QoS can still be met. In this chapter, we make use of the system level simulator developed for conventional cellular and extend it to assess how a CR can use the UMTS bands without causing interference. In the following, we present the main motivation behind the usage of CR.

It is likely that the demand for wireless services will continue to increase in the near and medium term, calling for more capacity and putting more and more pressure on the spectrum availability. While the use of advanced signal processing techniques may enable a very efficient usage of the spectrum even in the traditional framework of command and control spectrum policy, there is a worldwide recognition that these methods of spectrum management have reached their limit and are no longer optimal and new paradigms must be sought [1]. In fact, independent studies carried out in different places (e.g. [2]) have shown that most of the assigned spectrum is under-utilized. Thus, the problem is in most cases a problem of inefficient spectrum management rather than spectrum shortage as shown in Figure 5-1.

The development of frequency agile terminals that can sense "holes"(A spectrum hole is a band of frequencies assigned to a primary user, but, at a particular time and specific geographic location, the band is not being utilized by that user.") in the spectrum (Figure 5-2) , adapt their transmission characteristics to use these "holes" may provide one tool to address and take advantage of this spectrum under-utilization.
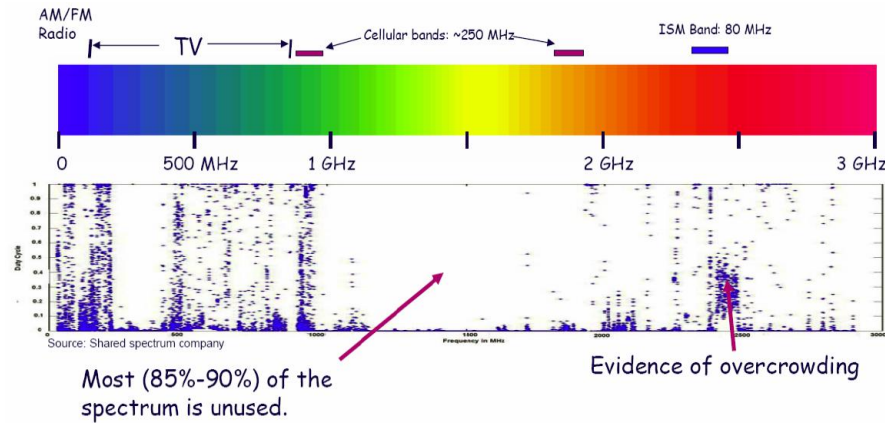
Figure 5-1: Spectrum Utility

The most general scenario of CRs distinguishes two types of users which share a common spectrum portion with different rules: Primary (or licensed) Users (PUs) have priority in spectrum utilization within the band they have licensed, and Secondary Users (SUs) must access the spectrum in a non-intrusive manner as shown in Figure 5-3. Primary Users use traditional wireless communication systems with static spectrum allocation. Secondary Users are equipped with CRs and exploit Spectrum Opportunities (also known as Opportunistic Radio (ORs)) to sustain their communication activities without interfering with PU transmissions.
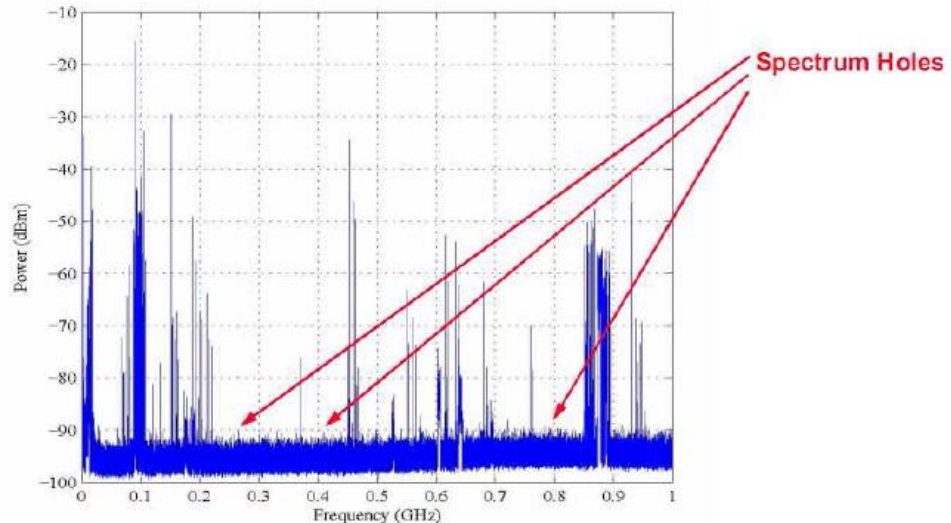
Figure 5-2: Spectrum Holes for Primary users

Devices with cognitive capabilities can be networked to create Cognitive Radio Networks (CRNs), which are recently gaining momentum as viable architectural solutions to address the limited spectrum availability and the inefficiency in the spectrum usage [3].CRNs is a network with a cognitive process that can perceive current network conditions, and then plan, decide, and act on those conditions [4].
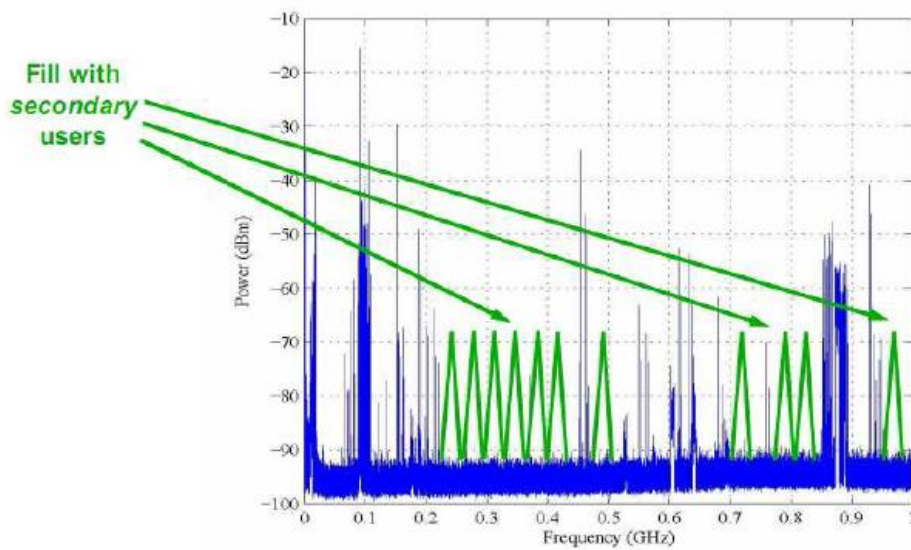
Figure 5-3: Spectrum Holes Fill with Secondary users

For the routing mechanism to be 'cognitive', it must have three elements of processes: observing, reasoning and acting as shown in Figure 5-4 ([5], [6]).
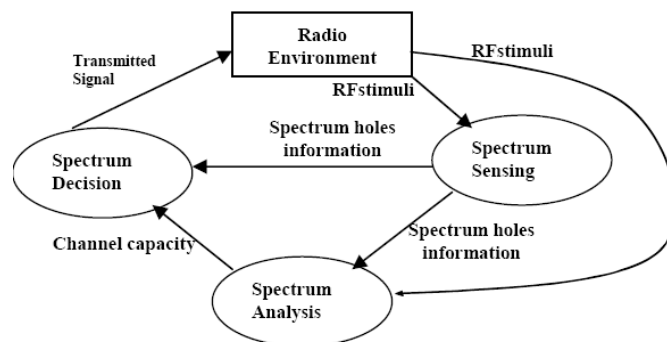

Figure 5-4: Cognitive Cycle

The observing process refers to how necessary information is gathered for implementation. The reasoning process is where the cognitive entity, e.g. a node (ORs), considers (orient, plan, decide, and learn) the way to behave for its various goals based on the information gathered from observing. The acting process is about how to implement the decision made by reasoning. We focus on the reasoning process mainly, assuming necessary information for each node is available through observing and different adjustment can be carried out in terms of acting.

In order to serve a goal of the network level (e.g. finding the shortest path), there must be a certain mechanism to comprehensively link the nodes in the system and enable them to function in a collective way.

Most of the research on CRNs to date has focused on single-hop scenarios, tackling Physical (PHY) layer and/or Medium Access Control (MAC) layer issues, including the definition of

effective spectrum sensing, spectrum decision and spectrum sharing techniques [7, 8]. Only very recently, the research community has started realizing the potentials of multi-hop CRNs which can open up new and unexplored service possibilities enabling a wide range of pervasive communication applications. Indeed, the cognitive paradigm can be applied to different scenarios of multi- hop wireless networks including Cognitive Wireless Mesh Networks featuring a semi-static network infrastructure [9], and Cognitive radio Ad Hoc Networks (CRAHNs) characterized by a completely self-configuring architecture, composed of CR users which communicate with each other in a peer to peer fashion through ad hoc connections[10].

Conventional routing algorithms normally find the route with shortest path to improve efficiency [11]. However, in CRAHNs, shortest path routing is not necessarily the best solution. In such scenario, other criteria, such as interference, capacity, etc., should be considered while making the routing decisions. Over the course of the last two decades, many approaches have been proposed to address the issue of dynamic spectrum allocation. In the mid-to-late 1990s, some of the initial foundational work on spectrum sharing was done in [12], [13], and [14]. It was shown that from a queuing theory perspective, independent networks could achieve an overall better capacity through cooperation. A survey on distributed spectrum management functionalities is given in [15] with spectrum sensing, sharing, decision and mobility as key components. A common control channel is introduced to enable a distributed coordination between cognitive radio users. In [16] and [17] spectrum pooling is proposed as an innovative strategy to enhance spectrum efficiency, by overlaying a new mobile radio system on an existing one without requiring any changes to the actual licensed system. In [18] the performance of spectrum sharing is investigated for the universal mobile telecommunications system (UMTS) frequency division duplex (FDD) downlink.

Digital Enhanced Cordless Telecommunications (DECT) can be regarded as a first implementation of spectrum sensing. A DECT telephone selects a frequency channel based on sensing of the channels available for DECT. However, DECT uses an exclusively assigned frequency band; sensing is only used to determine the best available channel within the band.

The first application that senses the available channels to detect and avoid other users is Unlicensed Radio LAN in the 5 GHz band. The Radio LAN uses a subset of OSA, which is called Dynamic Channel Selection (DFS). DFS is used to prevent a device from accessing a specific frequency channel if it is in use by a primary user, notably radar systems. The difference between OSA and DFS is that DFS is not used to seek spectrum access, but to prevent spectrum access if co-channel interference might occur. Close cooperation between regulators and industry was needed to define and standardize DFS in such a way that it can detect all different radar systems that are active in the bands involved.

Cognitive radio based standard for fixed, point-to-multipoint, wireless regional area networks that operates on unused channels in the TV VHF/UHF bands between 54 and 862 MHz on a non-interfering basis [19].The first application of Cognitive Radio is foreseen in the "white spots" of the TV-bands based on opportunistic spectrum access. The Federal Communications Commission (FCC) of the USA has already published a notice of proposed rulemaking to permit unlicensed opportunistic access to white spaces in the TV bands (FCC 2004). In response to this notice, the IEEE has created a working group (IEEE 802.22) which aims to develop a standard based on opportunistic spectrum access of the television bands to provide fixed wireless broadband access in rural and remote areas.

5.1.1   Objective

A simulation tool is commonly used for the research and evaluation of CR networks. Today no simulation tool can readily handle the simulation of the coexistence of two or more systems. A brute-force solution, which simply combines two systems together, will not suffice due to computational complexity induced and difficulty to extend beyond two systems. This chapter covers the simulation scenarios, which are implemented in order to observe the system performance as, OR terminals are introduced in the system level simulator. For such scenario and independently of the specific algorithms or protocols used, resorting to simulations is indispensable for the capacity evaluation of the OR network and its impact on the licensed one. In fact, such an evaluation in realistic scenarios is far too complex to be handled by simple models amenable to theoretical analysis. Therefore, we enhanced System level simulator that allows us to assess the coexistence of OR as a secondary system together with primary (victim device) UMTS FDD cellular systems. The OR network has Cognitive radio Ad Hoc Networks (CRAHNs) topology and exploits opportunities from a spectrum pool of several UMTS frequencies. Different routing strategies for this CRAHNs network are investigated and compared.

5.2   Platform for UMTS Scenario

5.2.1   UMTS scenario Overview

The UMTS radio frequency spectrum has become, in a significant number of countries, a very expensive commodity, and therefore the opportunistic use of these bands could be one way for the owners of the licenses to make extra revenue. However, UMTS UL bands capacity has been mainly under-utilized due the typical internet traffic asymmetry. In the Internet based applications the traffic patterns are asymmetric with much lower usage of the uplink band due to a negligible amount of control traffic in comparison to a large amount of data downloading. Recently spectrum occupancy measurements performed in Europe pointed out a 20 dB difference between the power spectrum density (PSD) measured at UL lower than the measured in DL bands [19]. Typically the

UMTS cell planning studies includes an activity factor ratio 1:10 for uplink to downlink traffic, as a consequence the UMTS FDD is usually a downlink capacity-limited system [20]. On the other hand, exploiting UL bands the victim device is the UMTS base station, likely far from the opportunistic radio, which creates local opportunities due the path loss and shadowing between the OR transmitter and the UMTS base station.

5.2.2 Interference Temperature

Interference is typically regulated in a transmitter-centric way, which means interference can be controlled at the transmitter through the radiated power, the out-of-band emissions and location of individual transmitters. However, interference actually takes place at the receivers. To be both highly reliable and efficient in spectrum usage, cognitive radio systems must be aware of other radio channels that may interfere with its receivers. It must also be aware of the vulnerabilities of other radio systems to allow for maximum usage of the shared spectrum. To allow for the interoperability of different radio users in the same spectrum space, interference temperature was proposed as a metric to limit the power and bandwidth available to new systems without degrading existing wireless systems [21].

The interference temperature model is an entirely new concept for dynamic spectrum access. Higher Interference yields lower SIR, which means lower capacity is achievable for a particular signal bandwidth. OR nodes search for gaps in frequency and time where the measured interference is low enough to achieve communication at a target capacity, subject to overall interference constraints defined by the interference temperature model. In order to constraint absolute interference we can employ spectrum shaping techniques where the OR senses the shape of the interference power spectrum and designs a waveform whose shape is its converse. For instance, using OFDM the OR could create a notched power spectrum that can take advantage of non-contiguous spectrum segments, increasing the OR's signal bandwidth. This spectrum shaping technique is called "interference fitting" by [22]. A major distinction is that all other proposed schemes avoid licensed signals, while the interference temperature model tries to coexist with them. This metric raises new challenges, a comprehensible overview of this topic can be found at [23] [24].

5.2.2.1 Interference Temperature Model

A new model for measuring interference, referred to as interference temperature shown in Figure 5.5 was introduced by the FCC [21]. The model shows the signal of a radio station designed to operate in a range at which the received power approaches the level of the noise floor. As additional interfering signals appear, the noise floor increases at various points within the service area, as indicated by the peaks above the original noise floor. Unlike the traditional transmitter-centric approach, the interference temperature model manages interference at the receiver through

the interference temperature limit, which is represented by the amount of new interference that the receiver could tolerate. In other words, the interference temperature model accounts for the cumulative RF energy from multiple transmissions and sets a maximum cap on their aggregate level. As long as OR users do not exceed this limit by their transmissions, they can use this spectrum band. This means more flexibility in dynamic spectrum access, and provides the opportunity for increased overall capacity. Based on sensing information and assuming a worst case approach, e.g. a free space propagation link between the OR transmitter and the victim device, the OR should compute the current interference temperature in a specific band. The difference between the interference temperature limit and the current interference temperature, indicates the spectrum opportunities for OR operation in that band and a particular location.

More specifically, defining a licensed band as Δf, the thermal background noise is given by:

$\sigma_n^2 = kT_n\Delta f$, where the Boltzmann's constant $k = 1.38 \times 10^{-23} WHz^{-1}K^{-1}$, and T_n is the noise temperature (measured in *Kelvin*).
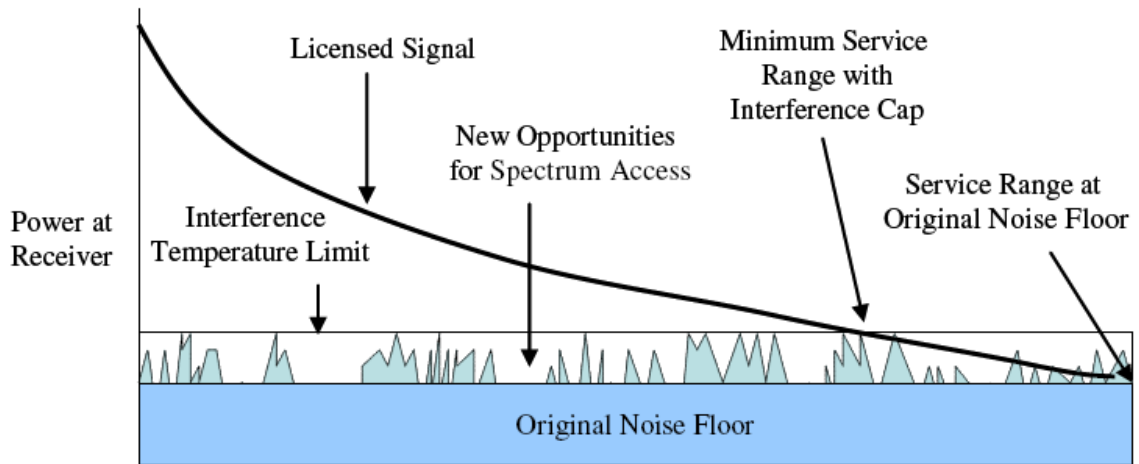

Figure 5-5 : Interference Temperature Model [6]

The total noise and interference power at the licensed receiver antenna is given by equation (5.1),

$$\sigma_{total}^2 = \sigma_n^2 + \sigma_{Inter}^2$$
(5.1)

Where the interference term, σ_{Inter}^2 , is due to the contribution of *M* opportunistic radios using this band, as shown in equation (5.2)

$$\sigma^2_{Inter} = \sum_{m=1}^{M} \int P_m \left| H_m(f) \right|^2 df \qquad (5.2)$$

P_m is the transmitted power by the OR terminal m and $H_m(f)$ is the channel frequency response between the OR transmitter m and the licensed receiver antenna. Thus, the total interference temperature is given by, equation (5.3)

$$T_{total} = \sigma^2_{total} / k\Delta f \qquad (5.3)$$

Finally, if for this particular band, the defined interference temperature limit is T_{max}, the opportunities in this band are given by the amount of extra interface that the PU receiver can support, that is, $\Delta T = T_{max} - T_{total}$,T_{max} is set by a policy decision for each particular licensed system and T_{total} is an output of the sensing algorithm. Based on ΔT, the OR system should compute the maximum power that is allowed to transmit in each licensed band. Figure 5.6 illustrates ΔT for two licensed receivers.
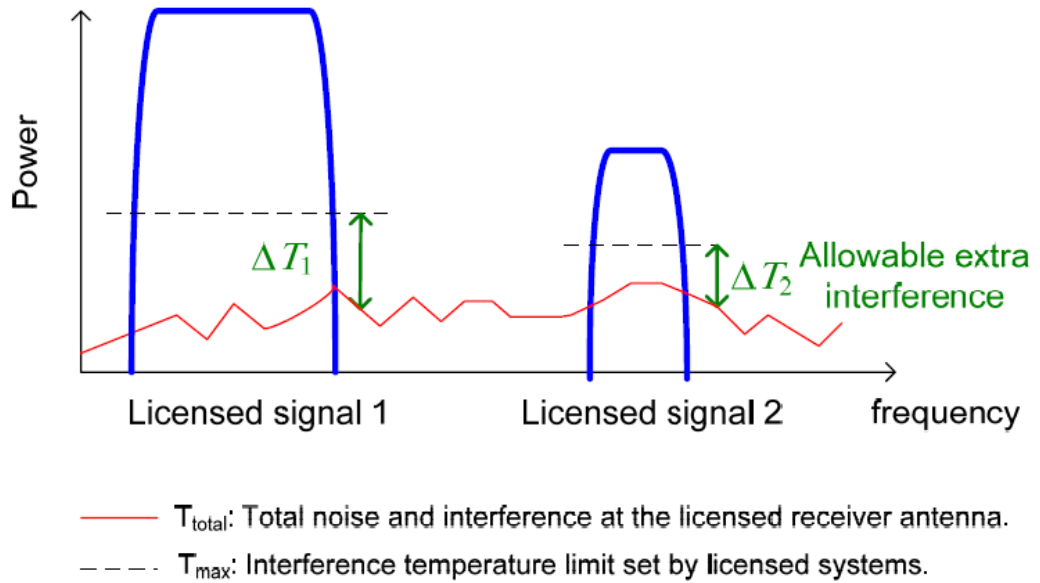

Figure 5-6 : Illustration of ΔT: amount of extra interface that a licensed receiver can support.

There are two main challenges in implementing the interference temperature model. The first involves identifying licensed signals which should be solved by reliable sensing and classification algorithms; this topic was deeply addressed by [26]. The second problem involves measuring interference temperature in the presence of a licensed signal. The OR transceiver must measure the interference floor underneath the licensed signal. This can be relatively easy if it has knowledge of the licensed waveform's structure. For example, perhaps it can measure during an interval when the signal is not present, as often occurs in bursty, time-multiplexed signals. In addition, if the ORs

have precise knowledge of the signal's bandwidth and centre frequency, they can estimate the interference based on measurements performed in neighbor's guard band. Moreover, in a real scenario, sensing information is collected from several OR locations, around the PU location. How to extrapolate the sensing data to estimate the interference level at the licensed receiver (T_{total}) is a very challenging problem.

The interference temperature metric could be used to set maximum acceptable levels of interference, thus establishing a 'worst-case' environment in which a receiver would operate. Interference temperature limits could thus be used, where appropriate, to define interference protection rights. Threshold levels could be set for different bands, geographic regions or services. These levels could serve as benchmarks to guide engineering trade-offs for opportunistic radio equipment design. In next section the interference temperature limit will be computed for UMTS scenarios.

5.2.2.2 Opportunities in UMTS FDD bands

In this scenario we consider an OR system that operates over UMTS UL bands in an opportunistic way. The OR terminal sense the path loss between its location and the UMTS base station. With this information, given by the sensing algorithm, the ORs adapt its power to limit the interference with the UMTS system within certain tolerable value. Figure 5.7 illustrates this scenario which exploits the Interference temperature model explained in Section 5.3.1.
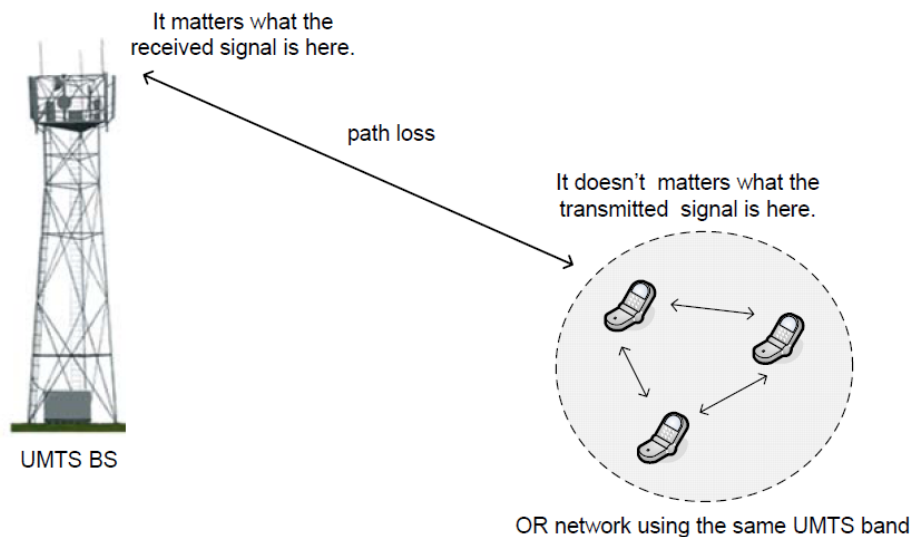

Figure 5-7 : Interference management differentiation between the transmitter and the receiver.

As the UMTS is a DS-CDMA system, thus all users transmit the information spreaded over the 5 MHz bandwidth at the same time and therefore users interfere with one another. Figure 5.8 shows a typical UMTS FDD paired frequencies used in Europe. *SF* is the length of the spreading code. The typical asymmetric load between uplink and downlink creates spectrum opportunities in UL bands since the interference temperature (amount of new interference that the UMTS BS can tolerate) is not reached.
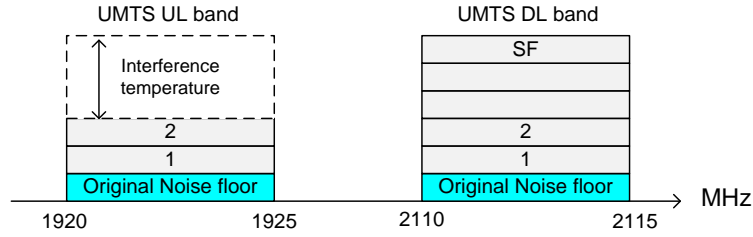

Figure 5-8: Example of UMTS FDD spectrum bands with asymmetric load

In order to fully exploit the unused radio resources in UMTS UL bands, the OR network should be able to detect the vacant channelization codes using a classification technique, see for instance [27]. Thus, the OR network could communicate using the remaining spreading codes which are orthogonal to those used by the UMTS terminals. However, classification and identification of specific spreading codes is a very challenging problem, especially for real time applications. In addition ,synchronization between UMTS UL signals and the OR signals, mandatory to keep the orthogonality between codes at UMTS BS, will be a difficult problem without cooperation between the two networks (UMTS and OR).

To apply the interference temperature model in this scenario, the OR should measure the current interference temperature at UMTS BS location (ΔT), which is infeasible, because no cooperation between the UMTS and the OR is assumed, instead we assume that by a policy decision, the OR network is allowed to fill part of the available interference temperature in UMTS UL bands with a certain amount of extra interference (e.g., $\mu=1$ dB rise above the noise floor). A similar approach is assumed by [28] recommendation, where the acceptable interference in the victim receiver is defined as the interference level that causes 1 dB degradation in receiver sensitivity. For simplicity we consider that the aggregated signal coming from the OR network is like-AWGN and causes a maximum noise rise equal to μ dB, as shown in Figure.5.8.

5.3   System Model Description

Figure 5.8 illustrates the scenario where an opportunistic radio network operates within an UMTS cellular system. We consider CRAHNs (OR) network of *M* nodes operating overlapped to the UMTS FDD cell. The OR network acts as a secondary system that exploit opportunities in UMTS

UL bands.The OR network has an opportunity management entity which computes the maximum allowable transmit power for each OR node in order to not disturb the UMTS BS.
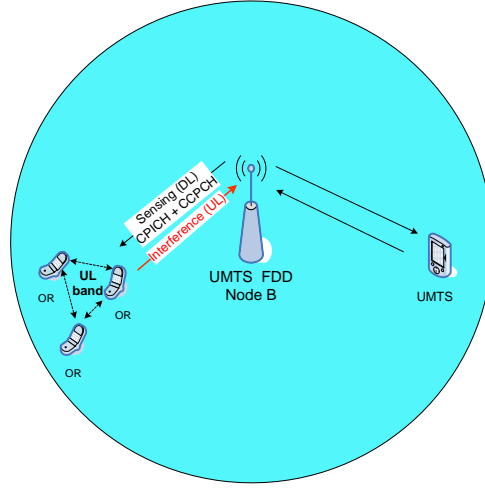

Figure 5-9: Ad-hoc Network operating in a licensed UMTS UL Band

Although the OR network use an UMTS UL band, sensing is performed in UMTS DL band. This is because it is easier to detect UMTS DL channels, since there are always active DL channels. In addition, no idle time is required for sensing because OR communication and sensing are performed in different bands with a frequency separation of 190 MHz. The price to be paid for sense and transmit in different bands is that because of mismatch between UL and DL shadowing, sensing algorithm will introduce some errors in the maximum power that the OR can transmit which could origin interference at UMTS BS.

The maximum transmit power allowed to a particular OR node (*POR*) is computed using a non-interference rule that takes into account the aggregated interference of the entire OR network, (5.4)

$$10\log\left(\sum_{k=1}^{K} 10^{\frac{P_{OR}(k)+G_{OR}+G_{BS}-\hat{L}p(k)}{10}}\right) \leq 10\log\left(10^{\frac{Nth+\mu}{10}} - 10^{\frac{Nth}{10}}\right) - \Gamma.$$ (5.4)

Where G_{OR} is the OR antenna gain, G_{BS} the UMTS BS antenna gain, Lp the estimated path loss between the OR node and the UMTS BS, K is the Number of ORs, performed by a sensing algorithm, and *Nth* is the thermal noise floor. μ is a margin of tolerable extra interference that, by a policy decision, the UMTS BS can bear. Finally, Γ is a safety factor to compensate shadow fading and sensing impairments. Notice if the margin of tolerable interference is set $\mu=0$ the OR network must be silent.

It is straightforward to extend this scenario to a 3G multi-operator case where several UMTS UL frequencies cover the same region. In this case, the OR can exploit a spectrum pool of some UMTS

UL carriers. Figure 5.10 shows a spectrum pool mechanism. As basic principle, the OR node rent the most appropriate UMTS UL band to transmit the required power (*POR_target*) to meet a QoS target. Whenever the P_{OR} is sufficient, the OR signal is formatted by the spectrum shaping module, for instance, using an OFDM modulator. If during OR transmission the allowed *POR* becomes lower than the target, the OR leaves that frequency and switches to another one.
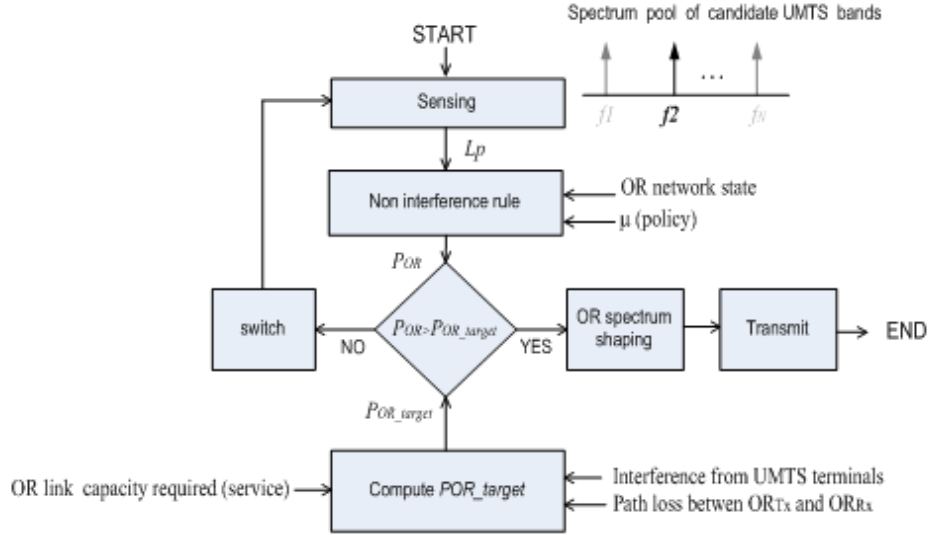

Figure 5-10 : UMTS spectrum pool mechanism.

Sensing is done on a periodic basis to follow the OR's node movement and the correspondent path-loss change. It is assumed that the OR doesn't know the position of the UMTS BS, thus is not possible to use apropagation model to estimate the path loss $\hat{L}_p$. The path loss is estimated based on the estimated received power, given by the sensing algorithm (energy detector, cyclostationary detector [28]). However, the problem of estimate the path loss between the UMTS BS and the OR location is that the output power level of the traffic channels varies according the traffic load. We use a conservative approach, assuming the minimum power transmitted by an UMTS BS, i.e., the power due the control and signaling channels that there is always active in DL bands (P-CPICH). Therefore the path loss can be estimated by (5.5)

$$\hat{Lp} = (P\min_{Tx\_UMTS\_BS} - P_{Rx}) \tag{5.5}$$

5.3.1 ORs with Ad-Hoc Technology

According to the network architecture, cognitive radio (CR) networks can be classified as the infrastructure-based CR network and the CR in Ad-Hoc Networks (CRAHNs) [29]. The infrastructure-based CR network has a central network entity such as a base station in cellular

networks or an access point in wireless local area networks (LANs). On the other hand, the CRAHN does not have any infrastructure backbone. Thus, a CR user can communicate with other CR users through ad hoc connection on both licensed and unlicensed spectrum bands. The changing spectrum environment and the importance of protecting the transmission of the licensed users of the spectrum are mainly important in CRAHNs. In following, we describe the unique features of CRAHNs.

Choice of transmission spectrum: In CRAHNs, the available spectrum bands are distributed over a wide frequency range, which vary over time and space. Thus, each user shows different spectrum availability according to the primary user (PU) activity.

Topology control: In CRAHNs, as the licensed spectrum opportunity exists over large range of frequencies, sending beacons over all the possible channels is not feasible. Thus, CRAHNs are highly probable to have incomplete topology information, which leads in an increase in collisions among CR users as well as interference to the PUs.

Multi-hop/multi-spectrum transmission: The end-to-end route in the CRAHN consists of multiple hops having different channels according to the spectrum availability. Thus, CRAHNs require collaboration between routing and spectrum allocation in establishing these routes. Moreover, the spectrum switches on the links are frequent based on PU arrivals.

Distinguishing mobility from PU activity: In CRAHNs, a node may not be able to transmit immediately if it detects the presence of a PU on the spectrum, even in the absence of mobility. Thus, correctly inferring mobility conditions and initiating the appropriate recovery mechanism in CRAHNs necessitate a different approach.

As we are using CRAHN technology in ORs nodes as shown in Figure 5.9, now we extend the coexistence analysis to the case of an ad hoc OR network with several nodes, multiple hop communication and routing mechanisms.

Routing by definition is a process of selecting paths. For a routing problem, the route discovery algorithm is required to find the ideal route in the system. The classic Dijkstra's algorithm [31] is chosen for this work for the route discovery problem. It is originally derived for solving the single-source shortest path problem for a graph with non-negative path costs. The functionality of Dijkstra's original algorithm can be extended for various purposes. For example, OSPF (open shortest path first) protocol is an implementation of Dijkstra's algorithm for Internet routing [31]. Dijkstra's algorithm can effectively choose the route with lowest accumulative cost. In order to achieve higher level goals in the system, we can exploit the optimization function of Dijkstra's algorithm by redefining the cost. That is to say, the definition of cost is manipulated in order to serve a different purpose other than finding the 'shortest' path.

In the following section we propose routing metrics which are based on hop count and capacity information of OR node which operated over UMTS FDD band. Here, we discriminate between some important capacity related terminologies used in this section. ´Originating capacity´ (C_o) of OR node is the goodput generated from the node. `Relaying capacity`(C_r) of the node refers to the relaying burden the node needs to carry for communications initiated by others OR nodes in the system.´Requested capacity´ of a node is the level of capacity that is able to satisfy all the requested traffic that needs to be transmitted by the node, includeing both originating traffic and relaying traffic as shown in equation (5.6)

$$C_{req} = C_o + C_r \tag{5.6}$$

where C_{req}, C_0 and C_r are the requested, originating and relaying capacities. ´Maximum capacity´ of the node denotes the maximum capacity level that can be transmitted by the node regarding the capabilities constraint of the node.This capability constraint of the node refers to the ability to find and utilize the UMTS FDD spectrum resource, the power level it uses to tranmit to it recervier(s) and interference. ´Available capacity´ is the capacity achieved by request under the maximum capacity constraint as shown in equation (5.7)

$$C_a = \min(C_{req}, C_{max}) \tag{5.7}$$

where C_a is the available capacity, C_{max} is the maximum capacity.The last term introduced here is ´virtual capacity´(C).It is defined as the capacity allocated to the node for the specific routing task.It is a portion of the available capacity on the node ($C \in C_a$)

5.3.2 Hop based Routing

To find the shortest path for a routing mission in terms of hop count, the cost can be defined as in equation (5.8):

$$w_{OR\_Tx,OR\_Rx} = 1 \ , \ \ \forall_{OR\_Tx,OR\_Rx} \in N \tag{5.8}$$

Where w_{OR_Tx, OR_Rx} denotes the cost to transmit from link *OR_Tx' to 'OR_Rx'* and *N is* the total number of ORs in the system. This routing scheme is often known as the 'hop count' or 'shortest path routing', and is the most commonly used routing metric in existing routing protocols.

5.3.3 Capacity-based Routing

With an ad hoc wireless network in our UMTS system, the scalability of the network is a major constraint due to the relaying burden each OR node has to carry. OR nodes in the system can

potentially improve the scalability of the network if the higher capacity OR nodes are placed in the right positions where more traffic has to be relayed. Intuitively, it is desirable to divert relaying traffic away from the OR nodes which have limited capacity to ones which are more capable to relay. For this purpose, we can define the cost as follows: in equation (5.9)

$$w_{OR\_Tx,OR\_Rx} = \frac{1}{C_{OR\_Rx}}, \quad \forall_{OR\_Tx,OR\_Rx} \in N \quad \ldots \tag{5.9}$$

C_{OR_Rx} denotes the virtual capacity of the receiving node OR_Rx a node with higher capacity will be more likely to be chosen for the object route because the cost is less compared with a lower capacity node.

We consider that the transmission power of each OR node is identical, then C $_{OR\_Rx}$ is determined by two factors:

- The interference suffered by the UMTS base station.
- The spectrum resource that can be utilized by the OR.

Assuming power levels and available bandwidth at each OR node, we let both of them equal to 1 as shown in equation (5.10). I_{OR_Rx} represents the number of interference sources from the UMTS base station which node OR_Rx suffers.

$$w_{OR\_Tx,OR\_Rx} = \frac{1}{\dfrac{1}{C_{OR\_Rx}}} < -116 \; dbm, \quad \forall_{OR\_Tx,OR\_Rx} \in N \tag{5.10}$$

$$I_{OR\_RX} = 10log\left(\sum_{k=1}^{M} 10^{\frac{P_{OR}(k)+G_{OR}+G_{BS}-\hat{L}p(k)}{10}} \right) \tag{5.11}$$

In a cognitive radio context, having more available spectrum for the OR node implies that the OR node is capable of finding a larger spectrum hole (or more spectrum holes) and exploiting it for transmission. If we only consider a system without bandwidth constraint at each node as a special case, interference is then the only concern. It is a worst-case scenario, in which the interference problem is dealt with in the most conservative way and has the most severe impact on capacity.

5.3.4 System Level simulation platform for UMTS

This section discuss the UMTS based System level simulator. Full dynamic system level simulator for opportunistic use of 3G licensed bands, based on the UMTS air interface specifications, has

been developed. The simulator with a time-driven simulation engine in it, recalculates mobile positions, path-losses, generates call events and determines the SINR, at regular time intervals. Seven macrocells, including the given cell in the centre and six adjacent cells in the first tier are considered, with a wrap around effect to keep the density of the users fixed. Based on mobility models, the direction of the mobile stations can be changed after the mobile station travels a minimum distance, the decorrelation length, according to a given probability. The speed of mobile stations is assumed constant. Mobile stations are also assumed to be uniformly distributed on the cell layout and their direction is randomly chosen at initialization.

The structure of the UMTS system level simulator is shown in Figure 5.11. The system level simulator interfaces with the link level simulator through look up tables (LUTs) and therefore can be used with a large number of different PHY layers, provided the appropriate LUTs are present as input to the simulator. Several propagation and traffic models are available, and the simulator computes the entire channel losses (fast and slow fading), thereby ensuring accuracy in the system level parameters computed. The outputs are the parameters that usually characterize packet transmissions: BER, FER, packet loss, packet delay, etc, and global system and individual statistics can be calculated using these output parameters. The system level simulation platform is constituted by two main parts, the UMTS FDD network and the opportunistic network. We focus the opportunistic network blocks that have to be designed and implemented.

Figure 5-11: System level simulation platform for coexistence analysis.

5.3.4.1 The system level UMTS simulator

The UMTS FDD system level simulator interfaces with the link level simulator through look-up tables (LUTs) as input to the simulator. Several propagation and traffic models are available, and the simulator computes the entire channel losses (slow and fast fading), thereby ensuring accuracy in the system level parameters computed. The outputs are the parameters that usually characterize packet transmissions: Throughput, BLER and Packet Delay.

The UMTS traffic generation block contains real and non-real time service traffic models. Handover block, as the name suggests, includes the UMTS handover algorithm. The radio resource management block comprises a call admission control algorithm, to regulate the operation of the network; a link adaptation algorithm, to select the appropriate parameters in function of the current radio conditions and a scheduler that decides how to allocate the appropriate resources, based on the service type, the amount of data, the current load in the cell, etc.

The Power control block contains mechanisms to provide similar service quality to all communication links despite the variations in the channel conditions. The interference block determines the average interference power received by central base station, i.e. intra-cell and inter-cell interference. The SINR evaluation block measures, at the UMTS base station, the ratio between the bit energy, and the interference level that includes the extra interference caused by the opportunistic network. The statistics from link level LUT block correspond to the link level chain simulation results, which are stored in SIR vs. PER tables for specific simulation environments. Finally, the computation of UMTS system level metrics block returns the UMTS FDD network results such as Service Throughput, Block Error Rate and Packet Delay.

Notice that the mobile model and the propagation model blocks are common to both networks. The mobility block models mobile movement in indoor, urban, and rural environments. Parameters associated with mobility include speed, probability to change speed at position update, probability to change direction, and the de-correlation length. The propagation block models path loss, slow fading and fast fading. Channel models for indoor environments, outdoor urban and rural environments are available.

5.3.4.2 Opportunistic Network

The opportunistic system level simulator will interface with the link level simulator through look-up tables (LUTs). The propagation and traffic models developed for the UMTS FDD network will be used, and the entire channel losses (slow and fast fading) computed. The outputs will be the parameters that usually characterize packet transmissions: Throughput, BLER and Packet Delay.

The LUT sensing algorithm characterization block contains the cyclostationary detector's performance, i.e. the output detection statistic, d, as a function of the SNR measured at the sensing antenna for different observation times [26]. The sensing OR-UMTS path loss block estimates the path loss between UMTS BS and the OR location through the difference between the transmitted power and the estimated power given by cyclostationary detector (LUT sensing algorithm characterization block output). The OR traffic generation block contains real and non-real time service traffic models.

OR QoS block defines the minimum data rate, the maximum bit error rate and the maximum transmission delay for each service class. The non-interference rule block compute the maximum allowable transmit power without disturbing the UMTS BS applying a simple non-interference rule (according to policy requirements).

5.3.4.3 Coexistence Analysis

In order to evaluate the coexistence between UMTS FDD networks and opportunistic networks that share the UL FDD bands, have a lower priority and are not allowed degrading the performance of the UMTS system.

Task Manger is responsible for overall coordination of the system. The scheduler mechanism will generate the arrival process of the users, according to a Poisson arrival process. The objective of Call Admission Control is to regulate the operation of a network in such a way that ensures the uninterrupted service provision to the already existing connections and at the same time accommodates in an optimum way the new connection request. The scheduler decides how to allocate the appropriate resources, based on the service type, the amount of data, the load on the common and shared channels, the current loading in the cell and the radio performance of each type of transport channel. Link adaptation can be considered as a component of Dynamic Channel Assignment. Power control mechanism and similar service quality are provided to all communication links despite the variations in the channel conditions, which means larger proportion of the total available power is consumed for the bad channel conditions. Handover is common to all dynamic system level simulators, and required to maintain link quality at the cell boundaries. Simulation Map description of the cellular map, which includes the cell descriptions, base station locations, and the manner in which it will model mobile movement at the system boundaries. Simple ARQ is employed for non-real time services. Dynamic Channel Allocation algorithm provides extra performance, but is not a pivotal element in the simulator. To provide an adaptive solution, the UMTS FDD system level platform must be integrated into the Link Level platform. The mobility block models mobile movement in indoor, urban, and rural environments.

The Results Database is responsible for storing, managing and analysing the results that are obtained during the course of the simulations.

In the following, we briefly explain the opportunistic network blocks that have to be designed and implemented, using a C++ design methodology approach.

First of all, we assume that the OR knows a priori the UMTS carrier frequencies and bandwidths, which have been isolated and brought to the baseband. In order to get the maximum allowable power for OR communications the OR nodes need to estimate the path loss from its location to the UMTS BS, i.e., the victim device. The opportunistic user is interested in predefined services which should be available everytime. This motivates the proposal of defining a set of usable radio front end parameters in order to support the demanded services classes under different channel conditions. Basically, at the beginning of each time step the opportunistic radio requires certain QoS guarantees including certain rate, delay and minimum interference to the primary user (non interference rule policy). The opportunistic network has an opportunity management entity which computes the maximum allowable transmit power for each opportunistic node in order the aggregated interference does not disturb the UMTS BS.

Employing scheduling algorithms, we can provide a good tradeoff between maximizing capacity, satisfying delay constraint, achieving fairness and mitigating interference to the primary user. In order to satisfy the individual QoS constraints of the opportunistic radios, scheduling algorithms that allow the best user to access the channel based on the individual priorities of the opportunistic radios, including interference mitigation, have to be considered.

The objective of the scheduling rules is to achieve the following goals:

- Maximize the capacity;

- Satisfy the time delay guarantees;

- Achieve fairness;

- Minimize the interference caused by the opportunistic radios to the primary user.

A power control solution is required to maximize the energy efficiency of the opportunistic radio network, which operates simultaneously in the same frequency band with an UMTS system. Power control is only applied to address the non-intrusion to the services of the primary users [33], but not the QoS of the opportunistic users.

A distributed power control implementation which only uses local information to make a control decision is of our particular interest. Note that each opportunistic user only needs to know its own received SINR at its designated receiver to update its transmission power. The fundamental concept

of the interference temperature model is to avoid raising the average interference power for some frequency range over some limit. However, if either the current interference environment or the transmitted underlay signal is particularly non uniform, the maximum interference power could be particularly high.

5.4 OR procedure modes

In order to share spatially/temporally the spectrum between an opportunistic radio and K primary users, some specific modifications/extensions have been done in the system level simulator. Before analysis these extensions, we present the two opportunistic radio procedure modes implemented in the SLS platform, static mode (Combined Snapshot mode) and dynamic (Fully Dynamic) modes.

5.4.1 Static Mode

In the static mode, the cell coverage area is considered as a matrix m x n, as shown in Figure 5.12. For a given frequency f_n and instant of time t, we introduce a "virtual opportunistic radio" in the system which run the entire cell/matrix, OR (f_n, t, $x_1...x_n$, $y_1...y_m$).

For each of the cell positions, the channel losses between the opportunistic radio and the primary user i, $L_o(f_n,t,1,i)$, are calculated and the signal to interference ratio in the primary user i determined. Notice that the SIR threshold, the antenna gain and the power transmitted by the OR are obtained from the configuration file.
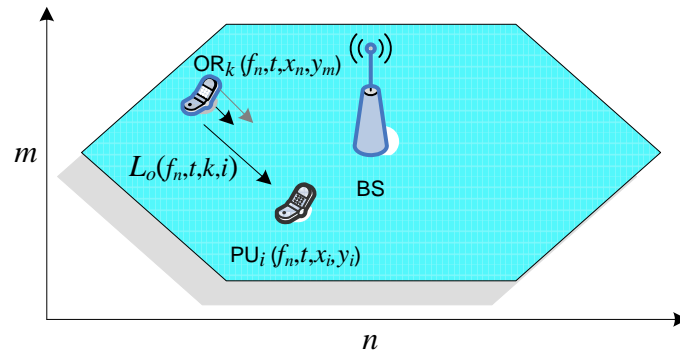
Figure 5-12 : Static mode

5.4.2 Dynamic mode

In the dynamic mode, we consider a communication with duration T_{cd}. During the call the active users into the cell coverage area, opportunistic radio and primary users (K_{all}), cover a route $ro(1,T_{cd})$ e $r(f_1...f_N$, $1...K$, $T_{cd})$, respectively, as illustrated in   Figure 5.13. The mobility model implemented on the system level simulator gives these routes; the mobility model applied is the

same for both user types, i.e. opportunistic radio and primary users. At the instant of time t_1, the opportunistic radio tries to start communicating on frequency f_n; the channel losses between the opportunistic radio and the K active primary users in the frequency f_n, $L_o(f_n,t_1,1,1...K)$ are calculated as well as each primary user SIR.

If the condition $SIR(i) \geq \gamma(i)$ is satisfied for $i=1...K$, the opportunistic radio transmits in the frequency f_n, at the instant of time t_1. Otherwise, the opportunistic radio tries to start communicating on frequency f_{n+1}, f_{n+2}, and so on. This frequency search stops when the opportunistic radio finds a communication opportunity or when no more frequencies are available.
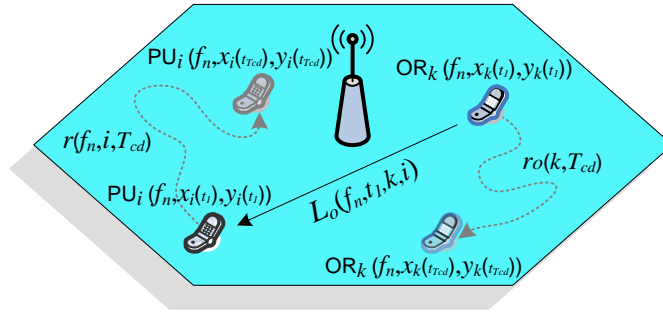
Figure 5-13: Dynamic mode

5.4.3   Channel losses calculation

The interference caused by the opportunistic radio k on the primary user i, depends on a set of parameters (power transmitted, antenna gain, etc...), however, only one is unknown and variable, the channel losses L_o. Thus, we present the approach considered to calculate the channel losses.

The opportunistic radio transmission power, P_o, is defined at the beginning of each simulation and stays constant during all the simulation time, as well as the opportunistic radio antenna gain, G_{RO}. On the other hand, the primary users losses and antenna gains are defined as, $L_{UP}(1...K)=L_{UP}$=constant and $G_{UP}(1...K)=G_{UP}$=constant, respectively.So, we need to calculate the channel losses L_o in order to estimate the interference caused by the opportunistic radio on the primary user i, as shows Figure 5-14.

The channel losses between the opportunistic radio and the primary user i, are given by the propagation losses PL_o, by the shadowing SH_o and by the fast fading FF. (see chapter 4)

5.4.3.1 Propagation losses

The propagation losses are obtained through the COST-Walfisch-Ikegami model, considering an urban environment and frequency 2000 MHz. The propagation losses PL_o (dB), in function of the

distance d between the opportunistic radio and the primary user i, are given by expression. Therefore, the distance d between the two terminals is obtained through the Pythagoras theorem and consequently the propagation losses are determined by (5.12).
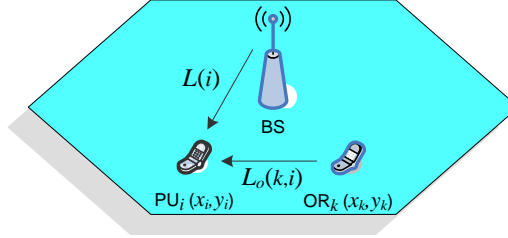
$$d = \sqrt{\left(x_i - x\right)^2 + \left(y_i - y\right)^2}$$

(5.12)

Figure 5-14: Channel losses L_o

5.4.3.2 Shadowing

The bi-dimensional shadowing, between the base station and the primary user i, is generated at the beginning of each simulation (t_o). For example, at instant t_1, the shadowing between the base station and the primary user i, $SH(i)$, is calculated/updated from the shadowing value generated at the instant t_o, $SH_{old}(i)$, as shown in Figure 5-15.

The simulation level simulator calculates the covered distance by the primary user i, during the time transmission interval t_{ti}, i.e. the distance d between the PU$_i$ position at the instant t_o and the PU$_i$ position at instant t_1.
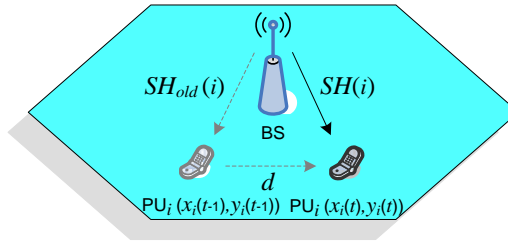
Figure: 5-15: Shadowing calculation between the base station and the

primary user i

After that, through the distance d and the de-correlation length, the system level simulator determines the shadowing spatial correlation, between the position ($x_i(t_0)$, $y_i(t_0)$) and ($x_i(t_1)$, $y_i(t_1)$) of the cell. Finally, the shadowing between the base station and the primary user i, at instant t_1, is obtained by equation (5.13)

$$SH(i) = \sqrt{1 - R^2} \times X + R \times SH_{old}(i)$$

(5.13)

where *SH* corresponds to the shadowing at instant *t* and SH_{old} to the shadowing at instant *t-1*. *X* defines a Gaussian distribution variable $N(\mu,\sigma)$, with mean μ (zero) and standard deviation σ (shadowing standard deviation).

In order to calculate the shadowing values between the opportunistic radio and the primary user *i*, we assume an approach based on the shadowing calculation process implemented on the system level simulator. On the one hand, we consider that the primary user *i* position, at instant *t*, corresponds to the opportunistic radio position, as shows the Figure 5.16. On the other hand, we consider that the primary user *i* position, at instant *t-1*, corresponds to the base station position, BS(0,0).
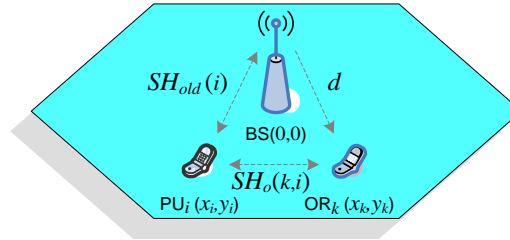
Figure 5-16: Shadowing calculation approach between the opportunistic radio and the primary user

Thus, we determine the distance *d* between the base station and the opportunistic radio as well as the shadowing spatial correlation, *R*, between these two points. Although the frequencies UL and DL are separated, we assume that the shadowing $SH_{old}(i)$ is symmetric, i.e. the downlink shadowing is equal to the uplink shadowing. Once know the shadowing on the base station position and the shadowing spatial correlation, we can calculate the shadowing between the opportunistic radio and the primary user *i*, $SH_o(i)$.

5.5 Simulation Metrics

The basic intention for using ORs in cellular networks is the exploitation of the licensed spectrum opportunities. Notice that the ORs can use these opportunities, if they do not cause harmful interference to the owners of the licensed bands, i.e. the cellular operators. The metrics defined in this section are intended to evaluate the spectrum/communication opportunities.The Figure 5-17 gives an overview of all the metrics. There are different system inputs parameters (depicted with yellow boxes in Figure 5-17) that have more or less influences the measured metrics.

Probability of Communication

Probability of Communication (*Pc*) means the probability that an opportunistic radio dropped in a random position of the cell does not disturb the existing communication in primary network. As a

result, probability of communication gives an idea of the OR communication opportunities, in a given cell and at a given time.
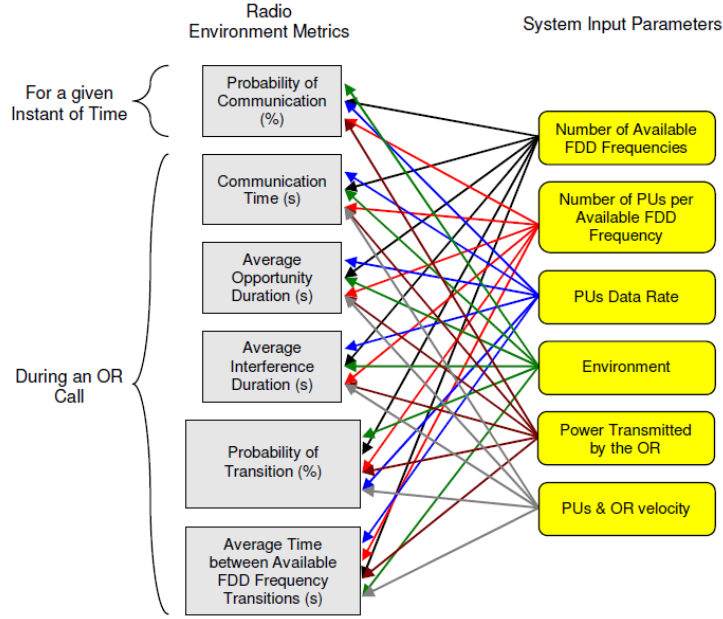

Figure 5-17: Overviews of radio environment metrics

This metric is strongly dependent on the number of available FDD frequencies in the cell coverage area, the number of PUs active in each available FDD frequency, the type of environment (related to the channel losses) and the power transmitted by the OR (interference caused on the PUs). The PUs data rate (related to the PUs SIR requirement) has a lesser impact on the probability of communication.

Probability of Communication can be calculated by computing, for a given instant t_q, the allowable/forbidden communication regions (Interference Maps) for each available FDD frequency, i.e. positions in the cell where PUs SIR is greater/lesser, respectively, than a given requirement γ. The union $\Omega_A(f_1....f_N,t_q)$ over the whole set of frequencies, of the allowable communication positions at a given instant t_q is the union of the allowable communication regions for each of the frequencies, and therefore the probability that a given instant t_q a TDD OR located at position x is allowed to start communication is given by, (5.14)

$$P_c = Pr(x \in \Omega_A(f_1...f_N,t_q))$$
(5.14)

Therefore assuming ergodicity we get, (5.15)

$$P_c = \overline{\Omega} / A_{cel}$$
(5.15)

Where $\overline{\Omega}$ is the average of the allowable regions and A_{cel} the area of a single cell.

Communication Time

Communication Time (T_c) corresponds to the time, during a call with duration T_{cd}, when the OR does not interferes harmfully with the primary users, i.e. when the OR transmission is allowed. This metric directly reflects the OR communication opportunities in terms of time and will be measured in seconds or milliseconds. The number of available FDD frequencies in the cell coverage area, the number of PUs active in each available FDD frequency, the type of environment and the power transmitted by the OR directly reflect the performance of this metric. The PUs data rate and the mobile terminals

velocity have a lesser impact on the communication time metric. The Communication Time is calculated based on an UMTS FDD OR call, i.e. the OR tries to start communicating on frequency f_1. If the condition $SINR \geq \gamma(n)$ is satisfied for $n=1\ldots K$, the OR transmits in that frequency.

Otherwise, the OR tries to communicate on frequency f_2, f_3 and so on. This frequency search stops when the OR finds a communication/spectrum opportunity or when no more frequencies are available. This process is repeated for each time transmission interval (T_{ti}) of the OR call. So, the T_c in seconds, is given by (5.16)

$$T_c = \sum_{t_q}^{N_{Tti}} T_{ti}(t_q), when\ P_o(t_q) = 0$$

(5.17)

where N_{tti} corresponds to the number of time transmission interval, during the FDD OR call. T_c can also be obtained through the follow expression, (5.18)

$$T_c = T_{cd} - T_{off}$$

(5.18)

where T_{cd} corresponds to the FDD OR duration call and T_{off} to the time without OR transmission during the call, and is given by (5.19)

$$T_c = \sum_{t_q}^{N_{Tti}} T_{ti}(t_q), when\ P_o(t_q) = 0$$

(5.19)

Average Opportunity Duration

Opportunity Duration (Δt_0) returns the OR communication duration, in a given cell and at a given FDD frequency (or more than one), without any interruption; the interruptions are derived from the restrictions of the primary network. Thus, the opportunity duration simply means the vacancy of a certain available FDD frequency (or more than one), in a certain geographical location during a specific period of time and will be measured in milliseconds or seconds. This metric will give an

indication of the communication time interval before breaking and may therefore be used to match traffic sessions to this parameter.

Figure 5-14 shows how the opportunity duration Δt_0 is determined, considering N=2. Figure 5-18 illustrates the OR carrier frequency status for a set of time transmission intervals (notice that a 90 seconds OR call corresponds to 45000 T_{ti}). The legend of the figure is the following:

• Frequency 0: No opportunities are available in the FDD systems and the FDD OR terminal is OFF.

• Frequency 1&2: The FDD OR communicates through the FDD frequency 1 and 2, respectively (ON).
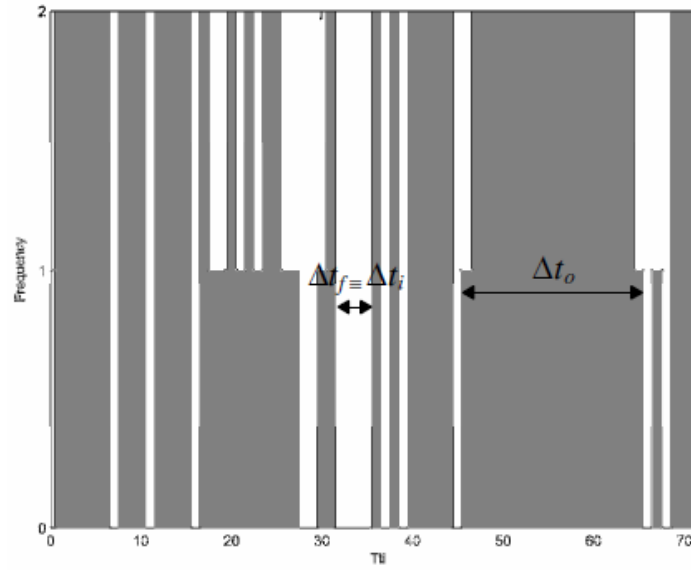
Figure 5-18: Time transmission intervals with and without FDD OR communication (N=2)

Average Interference Duration

According to the example of Figure 5-18, the time interval in which the OR is not allowed to transmit can be defined as Forbidden Communication Duration (Δt_f), i.e. consecutive time in Frequency 0. However, if we assume that the FDD OR may transmit without taking into account the $SINR \geq \gamma(n)$ condition for n=1…K, the time interval in which the OR is not allowed to transmit is simply equal to the time interval in which the OR will interfere harmfully with the PUs, i.e. Interference Duration. Interference Duration (Δt_i) will be measured in milliseconds or seconds.

This metric will give an indication of the average *interference duration* Δt_i. Both metrics, Δt_o and Δt_f, are represented in Figure 5-13.time interval that an OR communication can cause on the PU systems. Finally, we define Both metrics, Δt_o and Δt_i are strongly dependent on the number of

available FDD frequencies in the cell coverage area, the number of PUs active in each available FDD frequency, the type of environment and the power transmitted by the OR. The PUs data rate and the mobile terminals velocity have a lesser impact on these metrics.

5.5.1 Simulation Results

Simulations were carried out to compute the statistics defined in the previous section related to the OR communication opportunities. To compute these statistics the interference maps are generated and are assumed available to the OR. The main parameters used for the simulations are summarized in Table 5.1. We consider an omni-directional cell with a radius of 577 meters. Each available frequency, in a maximum of 12, contains 64 primary users. Each of these primary users receives the same power from the base station (perfect power control). We assume the primary users data rate equal to 12.2 kbps (voice call); the E_b/N_o target for 12.2 kbps is 9 dB. Thus, and since the UMTS receiver bandwidth is 3840 kHz, the signal to interference ratio required for the primary users is sensibly -16 dB.There is (minimum one) opportunistic radio in the cell coverage area, which has a transmitted power range from -44 to 10 dBm . The opportunistic radio duration call is equal to 90 seconds, i.e. 45000 time transmission intervals. We furthermore consider load characteristics identical in every UMTS cellular system and the frequencies are close enough so that the same statistical models apply.

Parameter Name	Value
UMTS DL system	
Time transmission interval (T_{ti})	2 ms
Cell type	Omni
Cell radius	577 m
Radio Resource Management	
Nominal bandwidth (W)	5 MHz
Maximum number of available frequencies ($N_{[max]}$)	12
Data rate (R_b)	12.2 kbps
E_b/N_o target	9 dB
SIR target (γ)	-16 dB
Spreading factor	16
Spectral noise density (N_o)	-174 dBm/Hz

Step size PC	Perf. power ctrl
Channel Model	Urban
Carrier frequency	2 GHz
Shadowing standard deviation (σ)	8 dB
Decorrelation length (*D*)	50 m
Channel model	ITU vehicular
Mobile terminals velocity	30 km/h
Primary User (PU)	
Number of primary user(s) per cell/frequency (*K*)	64
Sensibility/Power received	- 117 dBm
Antenna gain	0 dBi
Noise figure	9 dB
Primary user losses	0 dB
Orthogonality factor	0
Opportunistic Radio (OR)	
Number of opportunistic radio(s) in the cell coverage area (*M*)	4
Maximum/Minimum power transmitted ($P_{o\ [max/min]}$)	10/-44 dBm [16]
Antenna gain	0 dBi
Duration call	90 s

Table 5-1: Main parameters used for the simulations

5.5.2 Simulation results for a single UMTS frequency

We modified our simulation parameter. For static mode with single ORs link, we consider 64 UMTS licensed PU terminals in each cell (R= 2000 m). OR Rx gets interference from PUs in central and in 6 adjacent cells. The ORs are within an ad-hoc network service area (R= 100 m). OR Rx is 10 m away from OR Tx and OR Tx power constraint by UMTS BS as shown in Figure 5.19.

Based on Shannon formula, link capacity that can be achieved between two OR nodes is given by equation (5.20)

$$C_{Mbps} = Blog_2\left(1 + \frac{L_2 P_{OR\_Tx}}{Nth + I_{UMTS}}\right) \quad, \quad B = 5\ MHz \tag{5.20}$$

$$Nth = -107.01\ dBm$$

On the other hand, the interference at the UMTS BS cannot be higher than equation (5.21)

$$10log\left(10^{\frac{Nth+\mu}{10}} - 10^{\frac{Nth}{10}}\right) = -116.15\ dBm \quad \mu = 0.5\ dB \tag{5.21}$$

Then, if the non-interference rule is satisfied the required service is guaranteed for the single OR TxRx pair: as shown in equation (5.22)

$$10log\left(10^{\frac{P_{OR_{TARGET}} + G_{OR} + G_{BS} - \hat{L}p(k)}{10}}\right) < -116.15\ dBm \tag{5.22}$$
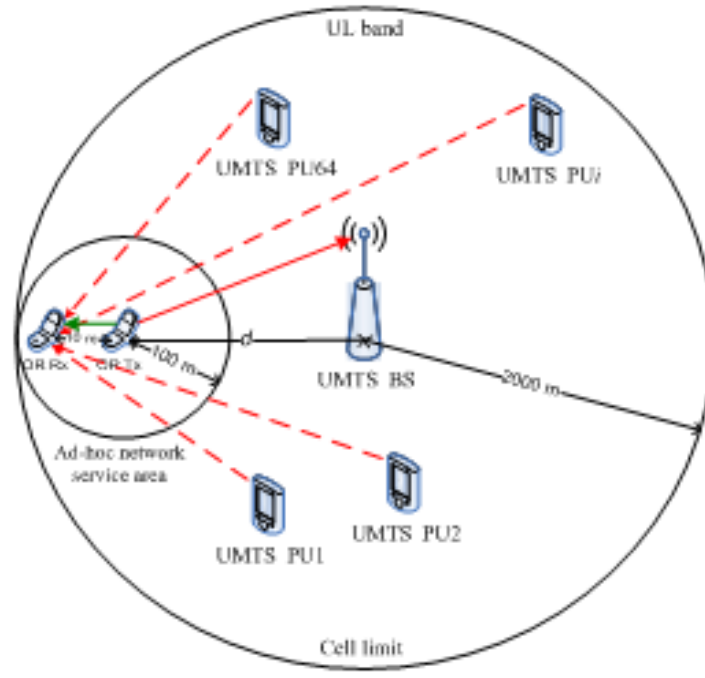
Figure 5-19: Ad-hoc Network Single Link Scenario

The Figure 5.20 shows the CDF of the interference computed at the UMTS BS due the OR network activity. The results show that an 8 Mbps OR's link capacity is guaranteed for approximately 98% of the time without exceeding the UMTS BS interference limit (-116 dBm). However, this percentage decreases to 60% when an OR link with 32 Mbps is established.
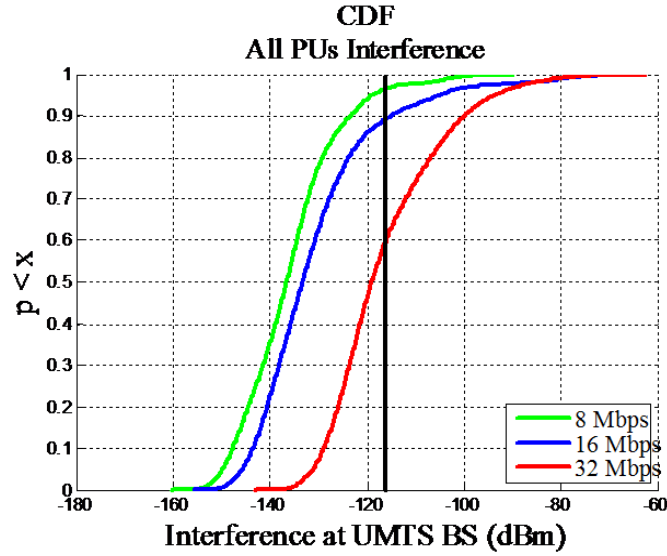
Figure 5-20 : Interference at UMTS BS

5.5.3 Simulation results for a UMTS spectrum pool

To analyse the impact of the available licensed frequencies on a OR duration call, simulations have been performed considering differents P_o OR duration call equal to 90 seconds and PU and OR terminals velocity equal to 30 km/h. Figure 5.21 shows the influence of the available licensed frequencies on a 90 seconds OR call. We can observe that the OR communication increases with the number of available frequencies as expected. To achieve a communication time more than 70 seconds, in a scenario with P_o equal to 0 dBm, at least 3 licensed frequencies must be available. On the other hand for P_o=5 dBm scenario, the same communication time is only reached with 5 available frequencies.
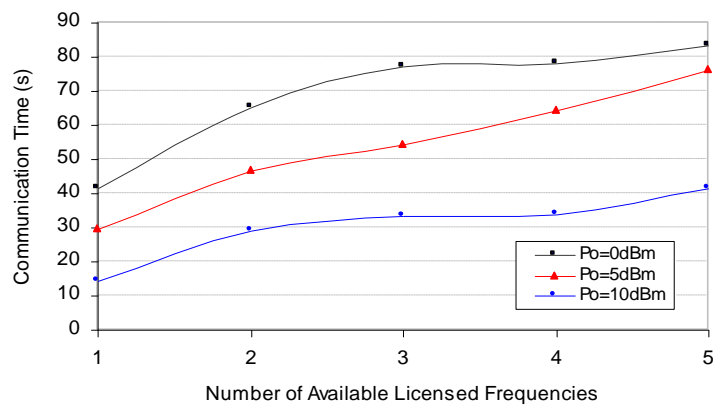
Figure 5-21: Influence of available frequencies

Figure 5.22 shows the impact of the power transmitted by an opportunistic radio on the probability of existing an opportunity higher than 20 seconds, considering 1, 2 and 3 available frequencies in

the spectrum pool. As we can see, to achieve an opportunity duration above 20 seconds, the opportunistic radio must transmit at most, -36 dBm with 1 available frequency, -28 dBm with 2 available frequencies and -26 dBm with 3 available frequencies. Thus, an opportunistic radio that operates in a given available frequency can obtain the same opportunity duration transmitting more 8 dBm of power, if there is one more available frequency and more 10 dBm of power if there are two more available frequencies. Based on these results, we can say that the higher the number of available frequencies the more transmission power is allowed. Notice that to obtain an opportunity equal to the duration call (any interruption), the opportunistic radio must transmit at most -40, -30 and -26 dBm for 1, 2 and 3 available frequencies, respectively.
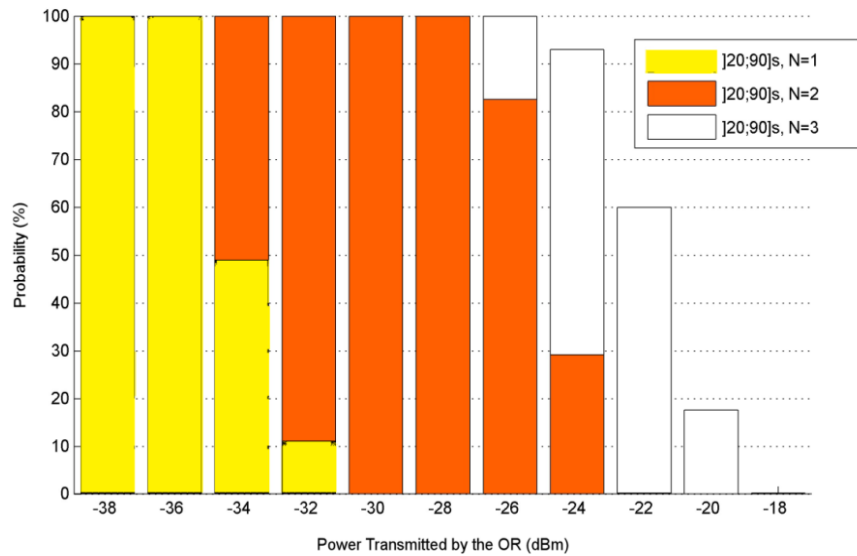
Figure 5-22: Impact of Power Transmitted by ORs

Now we show the effectiveness of different routing mechanisms previously proposed. A uniform geographic distribution of traffic is assumed to be generated by the ORs nodes simultaneously in the UMTS system. This is a worst case scenario, where the system is running at its maximum overall capacity and most severe interference level is experienced from the UMTS base station. First, we look at the scenario without capacity constraint in the system. In other words, there is no shortage of bandwidth for each transmission in the system and all the originating traffic of each node will be delivered without constraint. We look at a case where the sink (landmark node) of the system is placed in the corner as shown in the Figure 5-23 using hop-based routing. We can see that the closer a node is to the sink, the more traffic going through the node and most severe interference level is experienced with victim device, i.e. the UMTS base station, as indicated by higher requested capacity of the node. It shows the requested capacity of each node is dependent on the topological location of the OR node in respect of the topological location of the sink.

Instead of using conventional hop-based routing mechanism, we implement the capacity-based routing in the previous scenario. Figure 5-24 shows the requested capacity performance for the case in which the sink is in the corner. We can see that this time the routing pattern has been dramatically changed compared with the one with hop-based routing strategy. The route discovery process in this scenario tends to choose the routes along the edge of the system, in order to reduce severe interference from UMTS base station. It shows that by adjusting the cost function based on interference, the capacity-based routing strategy can effectively reduce the overall interference level in the system by shifting traffic to the edge of the network.
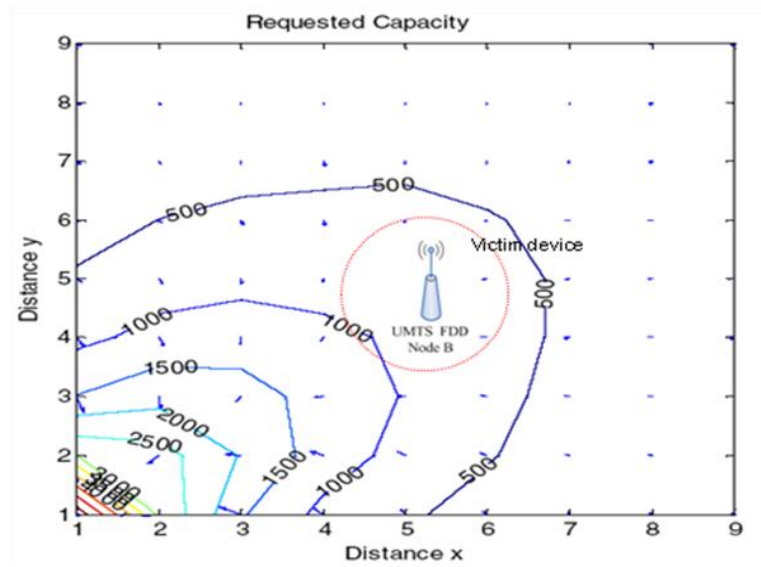
Figure 5-23: Contour plot with gradient arrows to show the requested capacity performance using hop-based routing strategy in a system

.
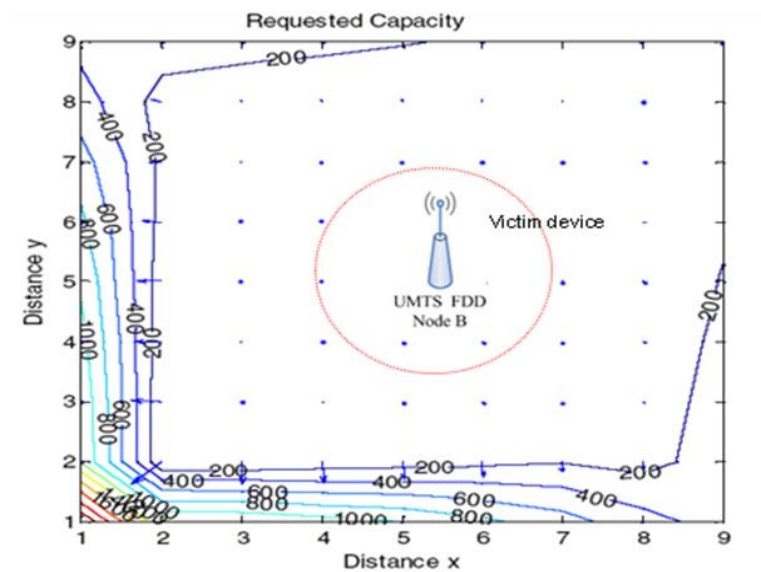
Figure 5-24: Contour plot with gradient arrows to show the requested capacity performance using capacity-based routing strategy in a system.

5.6   Conclusions

This chapter covered the simulation scenarios, which are implemented in order to observe the system performance as, OR terminals are introduced in the system level simulator. Therefore, we enhanced System level simulator that allows us to assess the coexistence of OR as a secondary system together with primary (victim device) UMTS FDD cellular systems. We explained how we use frequency band in UMTS by using OR, we enhanced our SLS by adding different classes to achieve our goals and we have also considered an ad hoc behaviour in the opportunists radio (ORs) and suggested that by implementing ad hoc features in the ORs will improve the overall performance of the system. The OR network has Cognitive radio Ad Hoc Networks (CRAHNs) topology and exploits opportunities from a spectrum pool of several UMTS frequencies. Different routing strategies for this CRAHNs network are investigated and compared.

We implemented routing feature of an ad hoc network in ORs, which dramatically reduce the server interference from the UMTS BS. Routing strategies based on hop count and capacity information were investigated and compared. The capacity-based routing strategy can reduce the overall interference level in the system by shifting traffic to the edge of the network.

.

5.7   References

[1]   Diepenbeek C. Flexibility in frequency Management. 11th CEPT Conference, October 2004.

[2]   Shared Spectrum Company. New York City Spectrum Occupancy Measurements September 2004.

[3]   M. Wellens, ☐Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio☐, Crowncom conference, August 2007.

[4]   R. W. Thomas, D. H. Friend, L. A. DaSilva, and A. B. MacKenzie, "Cognitive Networks Adaptation and Learning to Achieve End-to- End Performance Objectives," in IEEE Commun. Mag., vol. 44, Dec.2006, pp. 51-57.

[5]   J. Mitola and G. Q. Maguire, Jr., "Cognitive radio: making software radios more personal," IEEE Personal Communications, vol. 6, pp.13-18, Aug 1999.

[6]   Improving Capacity for Wireless Ad hoc Communication Using Cognitive Routing ICS, Singapore , 2007

[7]   C. Cormio, K.R. Chowdhury, A survey on MAC protocols for cognitive radio networks, Ad Hoc Networks 7 (7) (2009) 1315–1329,doi:10.1016/j.adhoc.2009.01.002.

[8]   S. Haykin, J.H. Reed, G.Y. Li, M. Shafi, Scanning the issue, Proceedingsof the IEEE 97 (5) (2009) 784–786, doi:10.1109/JPROC.2009.2015701

[9]   K. Chowdhury, I. Akyildiz, Cognitive wireless mesh networks with dynamic spectrum access, IEEE Journal on Selected Areas in Communications 26 (1) (2008) 168–181, doi:10.1109/JSAC.2008.080115.

[10]   I.F. Akyildiz, W.-Y. Lee, K.R. Chowdhury, CRAHNs: cognitive radio ad hoc networks, Ad Hoc Networks 7 (5) (2009) 810–836, doi:10.1016/j.adhoc.2009.01.001

[11]   E. M. Royer and C.-K. Toh, "A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks," *IEEE Personal Communications*, vol. 6, pp. 46-55,   Apr. 1999.

[12]   Satapathay D & Peha J (1998) Etiquette modification for unlicensed spectrum: approach and impact. Proc. IEEE Vehicular Technology Conference, 1: 272-276

[13]   Satapathay D & Peha J (1997) Performance of unlicensed devices with a spectrum etiquette. Proc. IEEE Global Telecomm. Conf., 1: 414-418

[14]   Satapathay D & Peha J (1996) Spectrum sharing without licenses: opportunities and dangers. TPRC.

[15]   Akyildiz IF, Lee WY & Chowdhury KR (2009) CRAHNs: cognitive radio ad hoc networks. Ad Hoc Networks, 7: 810-836

[16]   Weiss TA & Jondral FK (2004) Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency. IEEE in Communications Magazine 42(3): 8-14.

[17]   Haddad M, Hayar A & Debbah M (2008) Spectral efficiency of spectrum pooling systems.IET Communications, 2(6): 733-741

[18]   Pereirasamy M, Luo J, Dillinger JM & Hartmann C (2004) An approach for inter-operator spectrum sharing for 3G systems and beyond. Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 3: 1952-1956

[19]   IEEE 802.22 Working Group http://www.ieee802.org/22.R.

[20]   FCC, ET Docket No 03-237 Notice of inquiry and notice of proposed Rulemaking, ET

Docket No. 03- 237, 2003.

[21]   T. Clancy , "Spectrum Shaping for Interference Management in Cognitive Radio Networks", Software Defined Radio Forum Conference 2006.

[22]   T. Clancy "Formalizing the Interference Temperature Model," Wiley Journal on Wireless Communications and Mobile Computing, vol. 7, (9),pp. 1077-1086, November 2007.

[23]   Paul J. Kolodzy, "Interference temperature: a metric for dynamic spectrum utilization", INTERNATIONAL JOURNAL OF NETWORK MANAGEMENT , January 2006

[24]   http://www.ist-oracle.org/,D2.4, ORACLE Deliverable.

[25]   C. Huang and A. Polydoros, Likelihood methods for MPSK modulation classification , IEEE Transaction on Communications, vol 43, 1995.

[26]   "IEEE recommended practice for local and metropolitan area networks coexistence of fixed broadband wireless access systems", IEEE 802.16.2TM-2004 working group, March 2004.

[27]   I.F. Akyildiz, W.-Y. Lee, M.C. Vuran, M. Shantidev, NeXt generation/dynamic spectrum access/cognitive radio wireless networks: asurvey, Computer Networks Journal (Elsevier) 50 (2006) 2127–2159.

[28]   J. P.Monks, J. P. Ebert, A. Wolisz, and W. W. Hwu, "A study of the energy saving and capacity improvement potential of power control in multi-hop wireless networks," Proc. of the 26th Annual IEEE Conf. on Local Computer Networks, November 2001, pp. 550-559.

[29]   R. Brodersen et al, "CORVUS: A Cognitive Radio Approach for Use of Virtual Unlicensed Spectrum", University of Berkeley

[30]   Weiss TA & Jondral FK (2004) Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency. IEEE in Communications Magazine 42(3): 8-14

[31]   C. Huang and A. Polydoros, Likelihood methods for MPSK modulation classification , IEEE Transaction on Communications, vol 43, 1995.

[32]   K. Chowdhury, I. Akyildiz, Cognitive wireless mesh networks with dynamic spectrum access, IEEE Journal on Selected Areas in Communications 26 (1) (2008) 168–181, doi:10.1109/JSAC.2008.080115.

[33]   E. M. Royer and C.-K. Toh, "A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks," *IEEE Personal Communications*, vol. 6, pp. 46-55,   Apr. 1999.

Chapter 6: Enhanced System Level Simulator with Relay Nodes

Future wireless communication systems are expected to provide more stable and higher data rate transmissions in the whole network. It covers the high dense urban areas and provides access over large geographic regions in remote rural areas. However, in both scenarios it is challenging to cover the entire service area. In both cases, relays are foreseen to extend the range of the BS and to increase the system throughput allowing for a cost-efficient deployment and service. Simulations tend to be the preferred way of assessing the performance of relays. In this chapter, we explain, how the SLS design of the previous chapter has to be modified to allow the performance evaluation of relay enhanced cellular systems. The sls was then used to evaluate the performance and the results of relay enhanced cellular system which shown the significant improvement in the Average throughput, Outage and Fairness.

6.1 Introduction

In chapter 4 we focused on the development of system level simulator for conventional cellular systems and respective performance evaluation. However, conventional deployments are known to have problems in providing fairness (users closer to the BS are more benefited relatively to the cell edge users) and in covering some zones affected by shadowing. The use of relays has been proposed as a solution for these problems [1]. However, as happens with conventional cellular, the environment is still very complex (in fact more complex) and performance evaluation has to resort to simulations. In this chapter we extend the work already done at the simulation level for conventional systems by enhancing the use of relays in the simulator.

6.1.1 Motivation & Objectives

In a conventional cellular network, as shown in Figure 6.1, a base station (BS) controls a number of mobile stations (MS) within its own coverage area and all the terminals communicate directly with the BS. The current conventional deployment of cellular systems exhibits certain inherent problems such as low signal-to-noise-ratio (SNR) at the cell edge, lack of fairness, coverage holes that exist due to shadowing and non-line-of-sight (NLOS) connections. As the area up to a distance *"d"* increases quadratically then assuming uniform distribution, the density of the users' increases linearly with *"d"* as shown in Figure 6.2. The available capacity decreases with the distance *"d"* and if continuous MCSs are available, then the decrease will follow the behavior of PathLoss. If all users request a capacity *c'*, then all users at a distance *d>d'*, will not be satisfied. In the conventional system, all users, even those far away from their BS have to communicate directly with the BS, and if the same capacity has to be provided at all points in the cell, then high levels of radiated power are required, causing considerable interference to the neighboring cells.
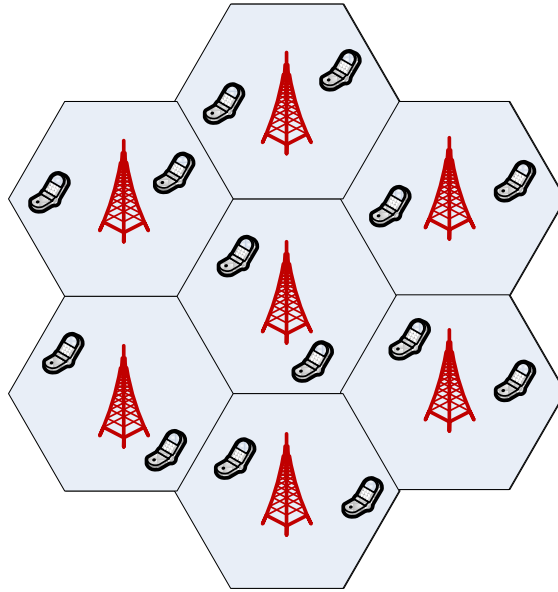
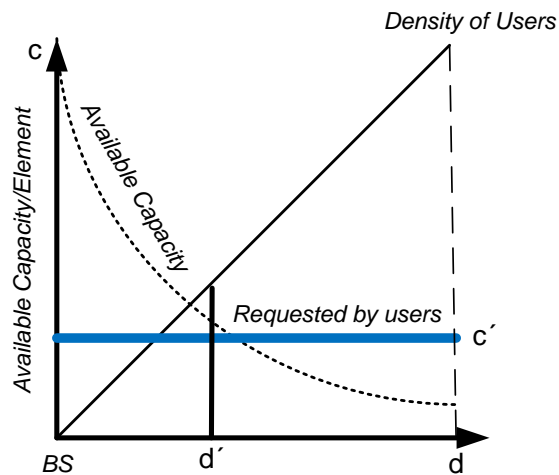**Figure 6-1: Conventional Cellular Deployment**

Figure 6-2: Available Capacity Vs Requested Capacity

In practice, a finite number of MCSs are available, and in fact, the capacity available at distance *"d"* is a staircase function, as shown in Figure 6.3.The users close to the BS can communicate with the most efficient MCS while the users at the edges use lower capacity MCS. This means some lack of fairness in the access to the network resources, as the users closer to the BS are better served than the users closer to the cell edges.
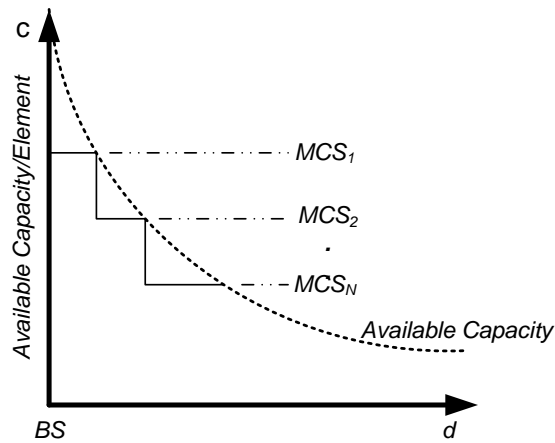
Figure 6-3: MCS According to Cell Distance

Shadowing is another problem in conventional system which may be caused due to a large obstruction such as a hill or large building which obscures the main signal path, preventing the BS signals to reach some of shadowing zone as shown in Figure 6.4.
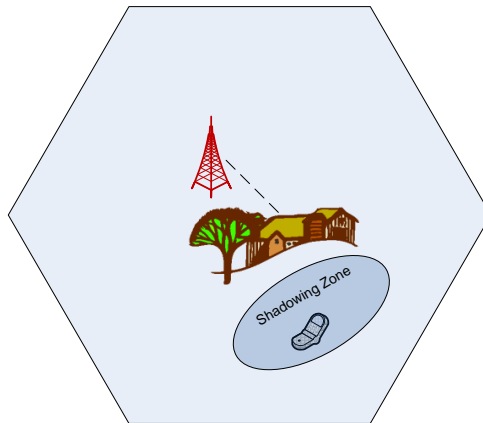
Figure 6-4: Shadowing Zone

The demand for ubiquitous high capacity and fairness requires solutions which mitigate the above mentioned problems.One brute-force solution to the problems discussed in the previous section (shown in Figure 6.2) is to place more BSs per area, but this comes with much higher deployment costs because every BS needs its own access to the fiber backbone network. Pico-cell or micro-cell infrastructures are examples of denser BS deployment. Another option is to increase the transmitted power which causes the impact on mobile battery life and more energy consumption in the BSs. However, none of these are ideal solutions due to the high deployment costs and energy consumption.

Multi-hop or relaying technology [2] are shown in Figure 6.5, where the Base station is deployed in the center and relays are deployed around the BS.So the users closer to are BS communicate directly with BS and the users which are on the edges communicate via Relays. All nodes including the BS, RSs and MSs work in the half-duplex mode thus they cannot transmit and receive simultaneously(i.e in first time slot BS send data to MS and RS and in second time slot data is transfer from RS to MS). We do not consider full-duplex mode since they are hard to implement due to the dynamic range of incoming and outgoing signals and the bulk of ferroelectric components like circulators [3].
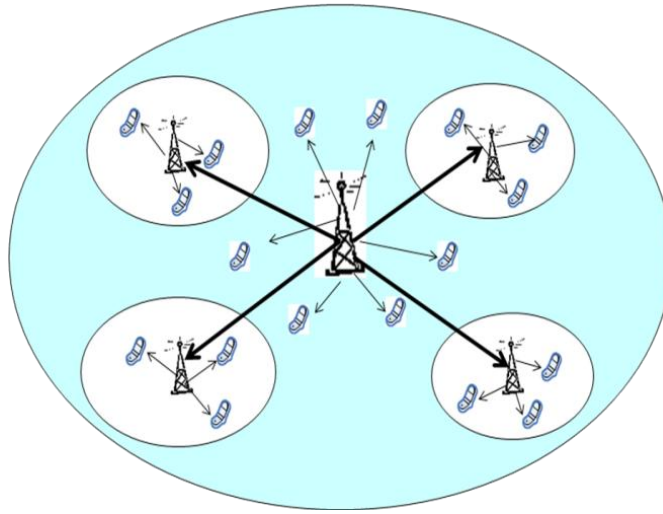
Figure 6-5 : Relay Cellular System

Relays are one of the methods which can solve the above-mentioned problems in conventional cellular system (low signal-to-noise-ratio (SNR) at the cell edge, fairness, and coverage holes that exist due to shadowing and non-line-of-sight (NLOS) connections). With relaying capabilities in the cellular network, the users are able to send their data directly to the BS or use some relay stations (dedicated or mobile) to relay their data to the BS , as link is broken down into shorter paths so they require less power and hence creating less interference to the neighboring cell.

Relays are placed at distance d_R as shown in Figure 6.6. If only BS is used then the requested capacity c' at distance d would not be satisfied, by putting relays per distance d_R the requested users capacity will be satisfied and overall capacity will be maximize.Relays greatly aid the signal transmission between the BS and UE and guarantee stronger and more stable receiving signals, especially for the UEs near the edge of the cell, thereby improving the overall system throughput.
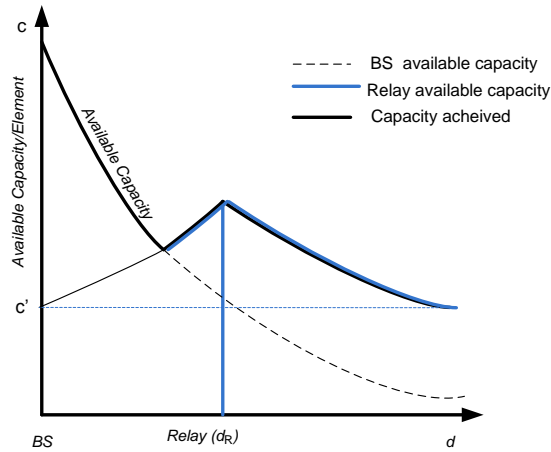
Figure 6-6: Relays Available Capacity Vs Requested Capacity

Relays also solve the problem of some covering zone due to NLOS connection. Relay provides an alternative path for communication between BS to MS, as shown in the Figure 6.7. Relays are the cost effective approach since its functionality is much simpler than BSs.
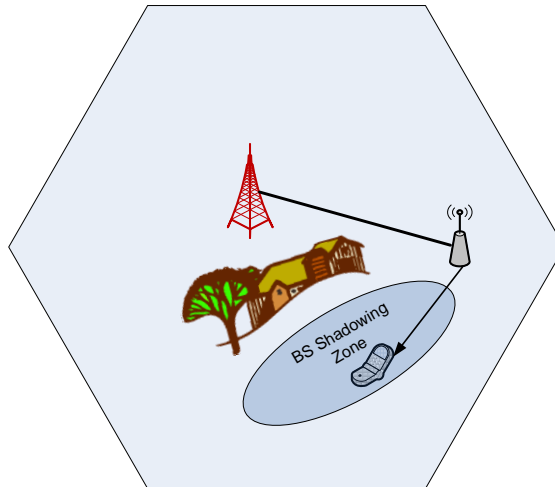

Figure 6-7: Relays: Shadowing Zone

As the use of relays is likely to have the impact in the future deployment thus, the objective of this chapter is to enhance a system level simulation tool, which replicates in an efficient manner all the processes of real networks with relaying capabilities. The tool will also allow the collection of relevant performance metrics that will shed light on how relay based schemes may improve the operation of current and future wireless networks.

6.1.2 Standarization and Preferable Deployment Scenario

In the emerging OFDMA-based standards such as 3rd Generation Partnership Project (3GPP) Long Term Evolution (LTE) [4] [5] and IEEE 802.16j [6][7], the multi-hop relay concept has been introduced to provide ubiquitous high-data-rate coverage. IEEE 802.16j was approved and published by IEEE in 2009 as an amendment to IEEE Std 802.16-2009 [8]. The purpose of IEEE 802.16j is not to standardize a new cellular network that includes multi-hop capability, but instead to expend previous single-hop 802.16 standards to include multi-hop capability [9].The two standards supporting multi-hop relaying, LTE-Advanced and IEEE 802.16j, are amendments of LTE and IEEE 802.16-2009, respectively. Therefore, they must have backward compatibility.

Multi-hop relays not only can be used in fixed infrastructures, but also can provide in-building coverage, coverage on mobile vehicle, and temporary coverage for emergency and disaster recover. Four typical usage models [8] for multi-hop relays are shown in Figure.6.8 are:

1. *Fixed Infrastructure*
2. *In-building Coverage*
3. *Coverage on Mobile vehicle*
4. *Temporary Coverage*

In the *fixed infrastructure* usage model of multi-hop relays, relay stations (RSs) are deployed in the cellular infrastructure to improve system capacity and coverage by dividing one long path into several shorter links and by offering alternative paths to users located in shadow areas.This kind of model are consider for our simulation.

In *In-building Coverage Model*, Relays are deployed to provide better coverage and higher throughput in a building, tunnel or underground such as on a subway platform.

Coverage on Mobile Vehicle refers to the type of deployment scenarios in which mobile is typically attached or mounted to transport vehicles such as commuters, long-haul trains, buses, ships, private cars, etc., to provide efficient cellular access coverage for users on board of those transport vehicles.

In *Temporary Coverage model*, Relays are deployed temporarily to provide additional coverage or capacity in an area where the Mobile relay-BS and fixrd Relay do not provide sufficient coverage or capacity.

The deployment of multi-hop relay can decrease the deployment costs because the conventional cellular system requires a very higher density of Base Station (BS) to provide sufficient coverage, and the deployment cost of a BS is higher than that of a RS. Moreover, the flexibility in relay positioning allows a faster network construction. By introducing multi-hop relaying to OFDMA cellular networks, larger capacity and coverage can be expected; however, there are still lots of

challenges. The new standard 802.16j must not only be compatible with old devices such as 802.16e devices, but also satisfy cooperative relaying functionality.
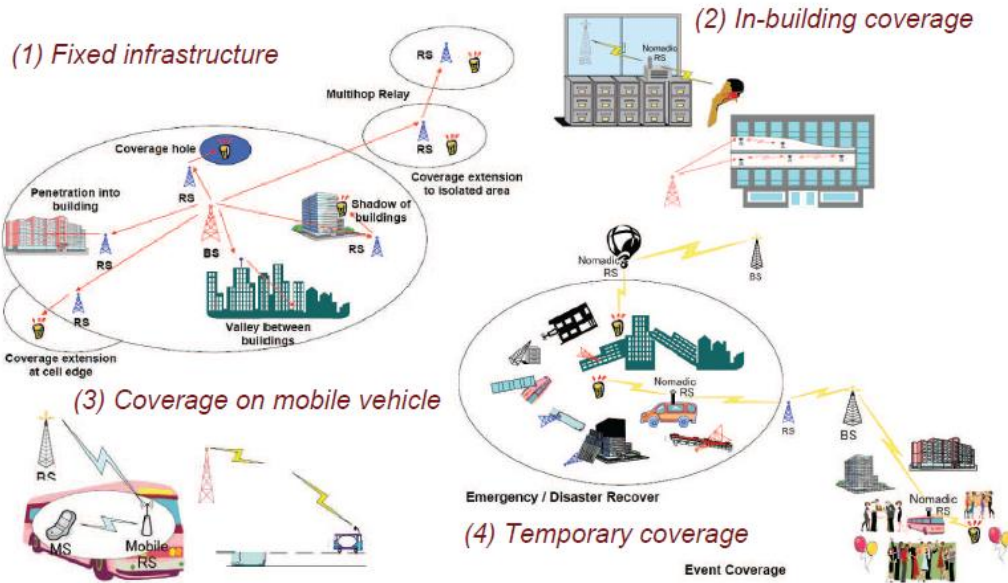

Figure 6-8: Typical Usage models for IEEE 802.16j systems [8]

From the physical layer perspective, backward compatibility requires every RS should be able to support all the modulation and coding schemes in the old standard. Moreover, since every Mobile Station (MS) may receive from the BS and a RS in the same frame, this raises more strict requirements regarding channel estimation, synchronization and frequency offset. From the Media Access Control (MAC) layer perspective, an entirely new set of messages specific to relaying must be created in 802.16j without overlapping with the existing set of MAC messages in IEEE 802.16-2009 [11]. The new MAC not only is responsible for ensuring a required Quality of Service (QoS) over multi-hops and allowing handovers among BS and RSs, but also should maintain Hybrid automatic repeat request (HARQ) over multiple hops. Technical issues such as frequency reuse, relay placement, resource allocation and scheduling are very difficult, yet extremely important, problems that IEEE 802.16j has left to manufacturers and providers to solve [9].

In the ISO–OSI reference model terminology, the physical layer, or layer 1, is responsible for the transport of bit streams over the channel, modulation and channel coding, equalization, synchronization, orthogonal frequency division multiplexing (OFDM) symbol creation, and decoding. This layer is common in all station types—BS, RN, or UT. They may only differ in points like transmit power level, RF (radio frequency) filter quality, and antenna gains. Thus, for a link between Communication peers, the link between BS and RN is just like that from BS to UT, as is the link between RN and UT. It is the higher layer that gives the link a meaning. Whether it is a broadcast or uncast transmission or a data or control channel, uplink or downlink does the medium

access control (MAC) functionality of the data link layer (DLL), or layer 2 decide something. This is where the radio system becomes a character and distinguishes LTE from WiMAX [13,14] (worldwide interoperability for microwave access), Wi-Fi, or other OFDM-based systems. A wireless relay is simply a device that receives incoming bits up to layer 2 and forwards them down again on layer 2 of the sending side, so this decode-and-forward relay acts like a repeater, bridge, or router known in the classical wired world.

Figure 6.9 shows this layer model. In these systems, received protocol data units (PDUs) are fully error corrected, automatically repeated (ARQ), stored and scheduled for transmission, and, if necessary, even segmented and reassembled prior to their next-hop transmission.
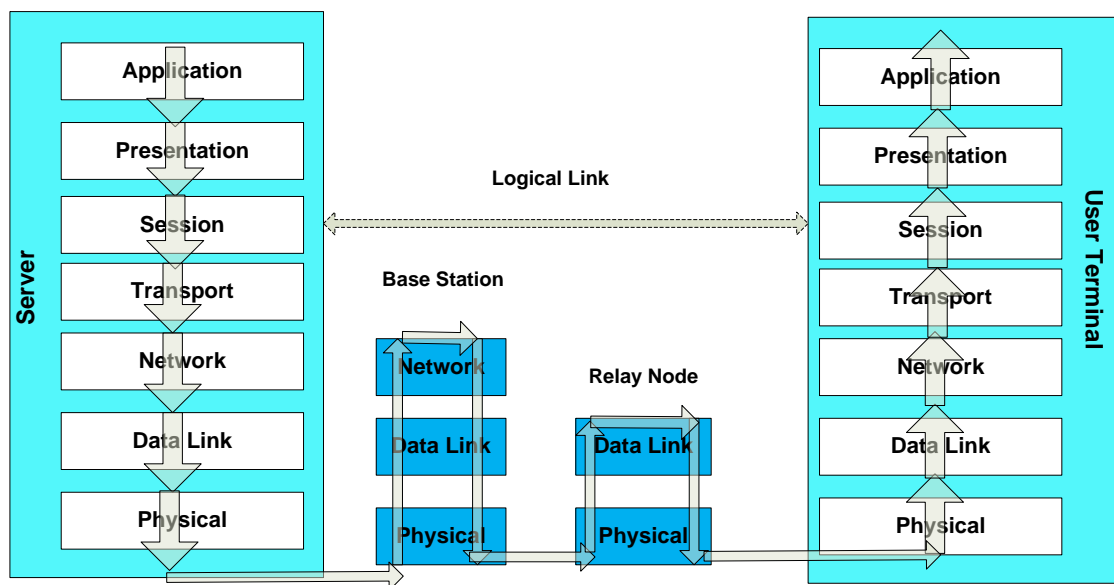
Figure 6-9: The ISO–OSI view on relays.

Conventional cellular systems use one-hop bidirectional communication between BS and UTs. UTs are immediately associated (linked) to the BS, which coordinates resources needed for the downlink (DL) and uplink (UL) transmission according to the traffic demand in both directions. Simply stated, for every certain rate to be supported, a proportional radio resource has to be allocated. Second-generation systems can only handle constant-rate traffic, while 3G systems have more flexibility in resource allocation. However, they still suffer from assumptions that came from the circuit-switched, voice-only view on the customer. Only the next-generation systems will make no difference between data and voice communication. They even sacrifice the advantages of classical digital voice transmission for a more contemporary voice over Internet protocol (VoIP) convergence. Multi-hop systems are transparent for the UT. The RN acts like a BS toward the UT on the second (last) hop. Additional resources are required only for the first hop, where the RN acts

like a UT toward the BS. This resource coordination was not foreseen in older systems—one reason why the deployments of smart relay concepts did not come at that time.

6.1.3   Classification of Relays

According to [14,1] relays can be classified into two types: *Centralized controlled Relays*: The BS has full control over the entire relay enhanced cell. The BS directly controls all UEs and all relay stations that are associated to the BS. The relay just forwards packets from hop to hop. *Decentralized controlled Relays*: Relays has full control over the UEs that are associated to it. The entire functionality that is required for the multi-hop operation is encapsulated in the relay. The BS is not affected. For the BS, a relay appears like an ordinary UE. For UEs, the relay appears to be a regular BS. Relaying can be possible through either Mobile or Fixed Relays.

Relaying through Mobile Station

In a cellular network with mobile relay station, the non-active MS (i.e. in idle state) are potential candidate to relay the traffic of the active MS to the BS. However, due to the mobility and the density of the MS, relaying with MRS brings several advantages and disadvantages (listed below) that should be considered in a realization of such a network.

1) Benefits

- The main advantage of a relaying system using MRS would be the low deployment/maintenance cost, since other user's terminals can potentially act as relays.

- The MS are able to organize themselves in order to cover some unknown dead spots, which are difficult to predict with the operators planning tool [15,16]. They can help unpredictable events like accidents or infrequently occurring such as demonstrations or sport events. They can also be used where it is not cost effective to install FRS such as mountainous environment or subway train platform.

- If there are a large number of idles MS, there are more choice to select a MRS that can optimize the system performance. For example, a relayed user can choose a MRS with whom it experiences a LOS link.

- If the source terminal and the destination terminal are close to each other's, they can communicate directly without sending their traffic through the BS as it is in conventional cellular networks.

2) Disadvantages

However, adding the communication capability between the users in cellular networks brings several disadvantages:

- It is obvious that the relaying opportunity depends strongly on the user's density.

- Relaying through other users terminals can considerably decrease the battery life of the mobile relay station.

- The relaying system performance is highly dependent of the RS selection scheme [17].

- The signaling (new channel MS-RS, inter-relay handover, RS selection) required for the relaying system might increase considerably.

- The cost of the terminal to support relaying (hardware and software) will increase. In addition, the implementations at the hardware become even more difficult if a MRS needs to relay more than one MS's traffic at the same time.

- Due to the mobility of the relayed MS and the MRS, frequent inter-relay handoffs might occurs, which might increase the signaling and affect the relaying system performance.

- Some others issues such as fast fading, power control or security need more investigations in order to make possible the communication between MSs.

Relaying through Fixed Station

1) Benefits

- The FRS are part of the network infrastructure, therefore their deployment will be an integral part of the network planning, design and deployment process. Hence, the operators will have a better control of the coverage and capacity expected in a specific area where the FRS is employed.

- The FRS can be deployed at strategic locations in order to maintain a LOS with the BS. They can also use directional antenna to improve the propagation link with the BS.

- The FRSs are less constrained by energy consumption. They may potentially be equipped with more advanced hardware, which enable them to operate in any frequency band as well as allow them to relay several MS at a time.

- The use of FRS eases the problem of RS selection since fewer number of FRS are deployed compared to MRS. Therefore the signaling overhead required for the RS selection will not be a major issue when relaying with FRS.

- The inter-relay handoffs will occur only when the relayed MS will move from one FRS to another FRS.

- Due to their sophisticated hardware, it is more secure to relay through a FRS than MRS. The data are always transferred through a known (fixed) RS.

2) Disadvantages

- The main disadvantage of using FRS is the infrastructure's cost. The dimensioning, planning, optimization and maintenance of the FRS can be expensive and cost inefficient. Furthermore, it might be cost ineffective to install FRS if there are many sparse coverage holes, in which only a few mobile terminals are located [17].

Since the number of FRS is limited and their locations are fixed, some MS might not be able to reach the FRS. Although, this depends on the number of FRS deployed in the system. There is a trade-off between the numbers of FRS vs. the infrastructure cost.

6.2   Network Architecture for the simulations

We consider an OFDMA enhanced cellular network with the BS (Base Station), six Relays (RS) and Multiple MS (Mobile Station). Similar to current cellular networks; the BS is located in the centre of the hexagonal cell. Six fixed relays are placed in each cell as new network elements enhancing the cellular infrastructure. Each relay is located on the line that connects the centre of the cell to one of the six cell vertices, and it is 2/3 away from the centre (BS) as shown in Figure 6.10, 6.11 and Figure 6.12.
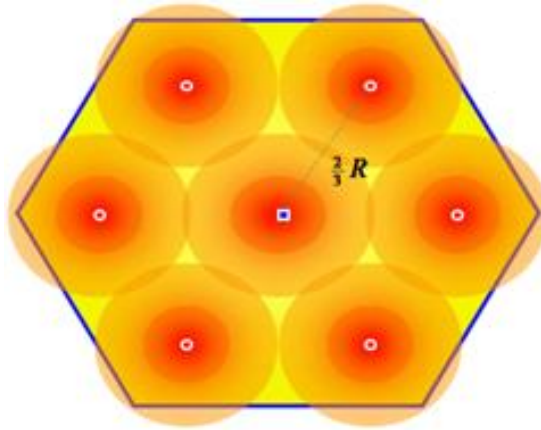
Figure 6-10: Example of hexagonal single cell deployment with Fixed Relay Nodes
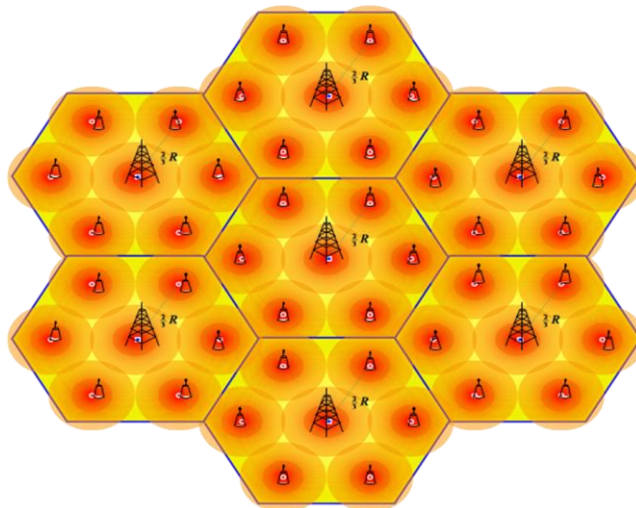
Figure 6-11: Example of hexagonal multi-cell deployment with Fixed Relay Nodes

In the downlink direction, users can receive data directly from the BS or via a RS. We call a user communicating directly with a BS a *single-hop user*, and a user that alternatively receives data via a RS a *two-hop user*. Two-hop relaying has been proven to give the highest system throughput, and when the number of hops is larger than three, the system overhead for exchanging control messages uses a great amount of resources [15] [16].
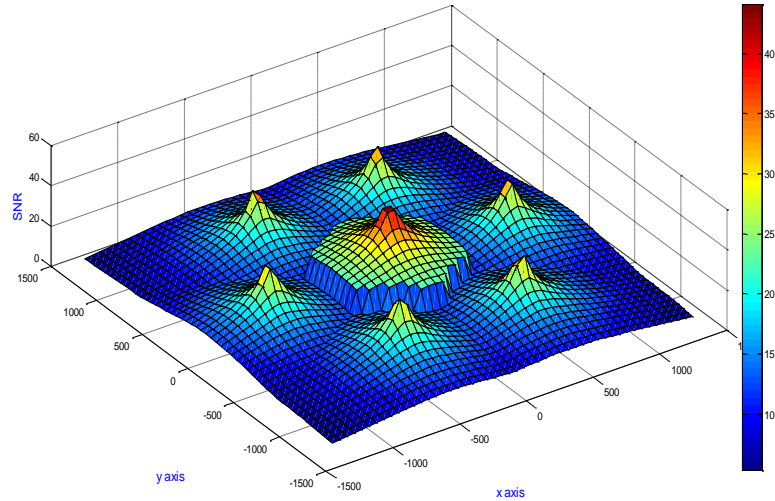
Figure 6-12: Cellular layout for Relay nodes

Table 6.1 summarizes the user-entry parameters for the configuration of hexagonal.

Parameters for the hexagonal deployment with fixed RNs
Number of RNs
Distance to cell centre
Node Tx power
Node height
Max number of users per node
Reuse Frequency

Table 6-1: Parameters for hexagonal cell deployment with Relays

6.2.1 Enhanced Relay-based System Level Simulator

In this section we show the enhancement done in the conventional SLS, which replicates in an efficient manner all the processes of real networks with Relays capabilities. The tool will also allow the collection of relevant performance metrics that will shed light on how relay based schemes may

improve the operation of current and future wireless networks. To accomplish the relays capabilities in the system level simulator some modifications (Relay node Parameter) are done as show in Figure 6.13. New classes are added to cope with relay functionalities in the SLS.
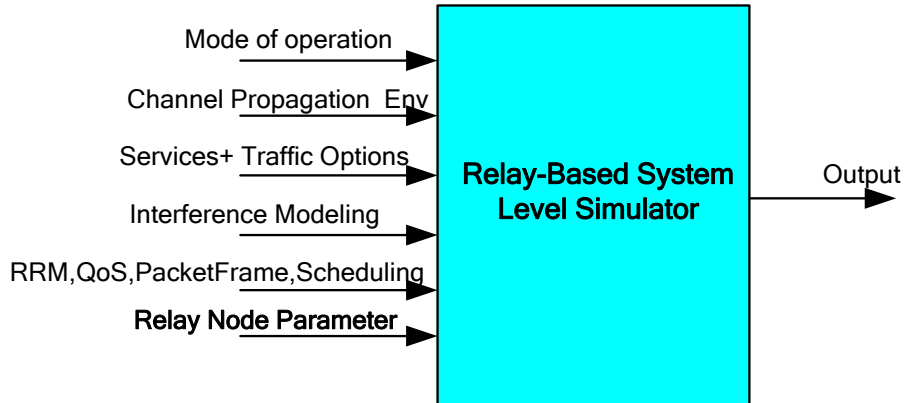
Figure 6-13: Enhance Relay based System level Simulator

Following extensions have been made in the system level simulation platform, namely in the ReadFile object, ChannelModel class and in the InterferenceModel class and added two new classes namely RelayNodeManager and RelayNode class. The TaskManager object has been also extended. The principle extensions were as follows:

6.2.2   Read/configuration file

The configuration file (./config/Wimax/sim_params.txt), contains new parameters related to the relay feature. These parameters are presented in Table 6.2. Any choice outside the scope of allowed values may interrupt (segmentation incident) the simulator or give erroneous results. Notice that in both operation modes (static and dynamic), the cell type should be omnidirectional/trisector.

Parameter	Value + comments
RelayNodeOption	(=1 Relay-based simulator is active); (=0 otherwise);
RelayNodeNumber	Number of Relay node per cell. e.g 2,3,.. etc
RelayNodeRadius	Radius of Relay Node far from base station(m) < cell Radius
RelayNodeMobile Option	(=1 Mobile Relay Node is active); (=0 otherwise);

Table 6-2: New configuration parameters

6.2.3   Channel model

The ChannelModel class is responsible for the Channel model implementation and channel related calculations for relay node

Channel Model
+calculateRelayChannel() : map<int,map<int,double>> +computePathlossNode() : map<int, double> +computeShadowinglossNode() : map<int, double> +computeShadowinglossMobileNode() : map<int, double> +fastfading_node() : map<int, map<int, double>>

Table 6-3: Relay channel Model

We are considering urban ITU Pedestrian B channel model with 6 taps [17]. Table 6.4 describes the complete set of channels between BS and RS.FRS refers to fixed relay station, MRS refers to Mobile relay station and UT refers to User terminal/Mobile.

	BS	FRS	MRS	UT
BS	NA	ITU Pedestrian B	ITU Pedestrian B	ITU Pedestrian B
FRS	Perfect	NA	ITU Pedestrian B	ITU Pedestrian B
MRS	ITU Pedestrian B	ITU Pedestrian B	NA	ITU Pedestrian B
UT	ITU Pedestrian B	ITU Pedestrian B	ITU Pedestrian B	NA

Table 6-4: Channel between BS and FRS/MRS

6.2.4 Interference model

The InterferenceModel class is responsible for the Interference model implementation and interference related calculations for relay node.

Interference Model
.
.+downlinkORintra() : double +dlInterference_node() : double +dlSIR_node() : double

Table 6-5: Interference model

6.3   Deployment scenario

6.3.1   Relay Selection Algorithms

In this architecture, a UE has two choices to receive signals: either from the BS or from one of the six relays. A specific algorithm needs to be established to determine from which node (BS or one of the relays) a UE will receive its signal. A SINR-based relay selection algorithm and pathloss-based relay selection algorithms are explained and compare through simulation. In our relay selection algorithms, we assume all six fixed relays are placed in strategic locations with good receiving signals from the BS.

6.3.2 SNR-Based selection Algorithms

The Base Station and relay link is assumed good enough to support this mode of operation (due to strategic relay locations and high-gain antennas). SINR based algorithm has two steps:

1) Out of the six relays (R1,...R6) in the cell in which the UE resides, select two that are the closest to the UE (R1, R2).

2) Compute the SINR between the two closest relays and the UE, and between the BS and the UE.

The node (relay or BS) with the maximum SINR will be responsible for transmitting signals for the UE as shown in equation (6.1)

$$R^{(k)} \equiv \arg\max_{i=0:N_R}\left(SNR_i^{(k)}\right)$$

$SNR_i^{(k)} \equiv SNR$ *measured between relay i and user k*

index 0 *refers to the Base Station*

N_R : *total number of relays* (*fixed* + *Mobile*)

(6.1)

Mobile relays are deployed in co-existence with fixed relay as shown in Figure 6.14. Mobiles are randomly deployed in hexagonal cells around the BS and fixed relay. To select which mobile becomes Mobile relay station (MRS), we follow the following approach as shown in equation (6.2)
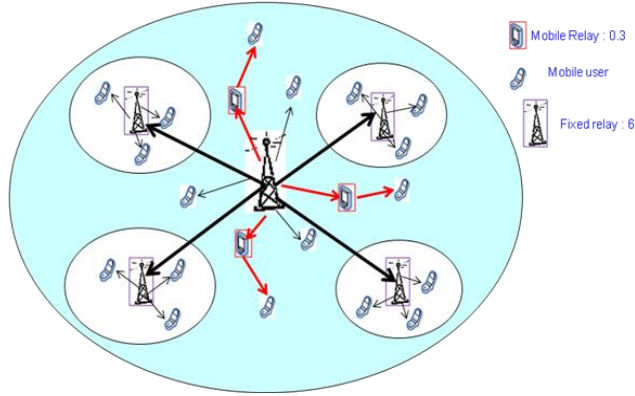
Figure 6.14: Mobile relays co-existence with fixed relay

$$N_{MR} = \sum_{i=1}^{N_T} I_i$$

N_T :*total number of mobiles*

I_i : Random binary variable that takes value 1 if Mobile i is used as relay and value 0 if used as user terminal

$$I_i = \begin{cases} 1 & with\ prob\ p \\ 0 & with\ prob\ 1-p \end{cases}$$

p parametrizable

(6.2)

Nr is the total number of mobile deployed in the cell and then we generate a random variable that takes the value 0 or 1. If mobile gets 1, then it acts as mobile relay otherwise as user terminal. (*p*=0.5)

6.3.3   PathLoss-based relay selection algorithm

The PathLoss algorithm has two steps:

> 1) Out of the six relays in the cell in which the UE resides, select two that are the closest to the user (R1, R2).
>
> 2) Compute the pathloss between the two closest relays and the UE, and between the BS and the UE.

Base station and Relay node use different Power level. The node (relay or BS) with the least path loss will be responsible for transmitting signals for that UE as shown in equation (6.3)

$$R^{(k)} \equiv \arg \min_{i=0:N_R} \left(PL_i^{(k)} \right) - \psi$$

$$PL_i^{(k)} \equiv PathLoss \; measured \; between \; relay \; i \; and \; user \; k$$

$$index \; 0 \; refers \; to \; the \; Base \; Station$$

$$N_R : total \; number \; of \; relays \; (fixed + mobile)$$

$$\psi = 10 \log_{10} \left(\frac{PL_{BS}}{PL_{relay}} \right)$$

(6.3)

Where ψ represent different power level between BS and relay.

6.4   Numerical results

Besides simulation parameters mentioned in Table 6.6, we also consider the hexagonal deployment and the collection of cells corresponding to several tires around the central cell. The numbers of tires is parametric and selected at the beginning of simulation. Simulation runs in combined snapshot mode in which mobiles are randomly deployed at each TTI, simulation runs and results are collected for each TTI, and average is performed. Fixed and Mobile relays coexist within the cell. Total mobiles are 100 and mobile deploy randomly in each cells. Number of mobiles is constant over a geographical area and deployment is random. We deployed six fixed relay in each cell and they are located 2/3 of the cell radius which are parametric. Power of BS is 12 dBW and power of fixed relay is 10% of BS (2dBW). Mobile relay can be user terminal that can be configured to act as relay. User terminal n is selected to be relay with probability p which is parametric, then the expected number of mobile relays are $N_r = N_u * p$ (N_u is the number of users terminal and N_r is the mobile relays) then the expected number of end terminals are $N_e = N_u * (1 - p)$. N_e is the end terminal.

Parameter Name	Value
Time transmission interval (T_{ti})	10 ms
Cell type	Omni or TriSector
Cell radius	100,200,…1000 m
Radio Resource Management	Best select CQI
Nominal bandwidth (W)	10 MHz
Base Station Power	10 W
Relay Node Power	1W,0.1W or 0.3W
Simulation type	Downlink
SINR method	EESM(Exp Effective SINR Mapping)
Channel Model	Urban
Carrier frequency	5 GHz
Channel model	ITU vehicular A
Mobile terminals velocity	30 km/h
Number of Users	40 per cell
Modulation and Coding	AMC
Traffic model	web3GPP No full Queue Option Cell Arrival Rate:10 user/cell/Sec
Propagation model	Multipath channel model Path loss, shadowing and fast fading
Antenna gain	14 dBm
Duration call	90 s

Table 6.6: Simualtion Parameters

We are using 1024 FFT size, 5 OFDM symbol per frame, 720 data subcarrier with PUSC scheme with 128 cyclic prefix lengths. The cluster size of the cellular network is taken to be equal to 4. Seven clusters are deployed, but the data is collected for the innermost cluster; i.e., the first tier. Both the BSs and UEs are assumed to use omni-directional antennas with a gain of 10 dB. Power control is not considered in this simulation, because when AMC is used, power control does not contribute much towards the throughput increase. For the fixed and mobile relay transmit power, we consider a variety of values to compare the system throughput and performance: 0.1 W, 0.3 W

or 1 W. Only the downlink scenario is considered. The UEs are placed randomly across the cluster with 40 UEs per cell.

We consider the following output metrics, which are calculated as follows.

Average spectral efficiency equals to total *number of bit transfer/total time/bandwidth* (bit/sec/Hz)

We also consider average thought and average spectral efficiency as output metrics, which is calculate as follows.

Service Throughput S(i,j) and OTA (over-the-air) Throughput OTA(i,j):

At the end of a given iteration *i* of the simulator, the OTA throughput and Service Throughput for sector *j* are computed. OTA (i,j) is defined as the ratio of the number of bits transmitted in sector *j* over the time required to transmit them.

Average OTA and Average Service Throughput (per cell):

Obtained by averaging OTA (i,j) and S(i,j) over all sectors and all runs.

User Service Throughput:

This is defined as the number of correctly received bits per user divided by the simulation duration.

Average spectral efficiency:

It equals to total number of bit transfer/total time/bandwidth (bit/sec/Hz).

As a performance metric of the network coverage property, the **outage probability** can be expressed by

$$P_o = P_r\left[\gamma < \gamma_o\right] = P_r\left[MCS = 0\right]$$

Where, γ_o denotes the minimum SINR requirement for the receiver to obtain service, which is set to 0 dB. The outage threshold is set to 10%, because full traffic load is adopted in our simulation and the resulting outage probability tends to be overestimated.

In this simulation, we compare the with-relaying case (fixed and mobile relay) with the without-relaying one. For the without-relaying case, all UEs will receive signals from the BS. In this case the interference comes from the other six BSs. For the with-relaying case, as described above, six relays are placed in each cell. In this case, the received signal can be either from the BS or from a relay depending on which node the UE is communicating with Figure 6.14 and Figure 6.15 compare the average throughput for the two relay selection algorithms, pathloss and SINR, when *R* = 1000 m and without /with considering the interference from other cells(central/multi cell approach).

The throughput of the SINR-based algorithm is slightly better than that of the pathloss-based one, so from now we just use SNR-based algorithm. The relays are placed on the segment between the centre and one corner of that cell. Here the fixed relay position is indicated by a parameter *m*

showing how far away the relay is from the BS. $m = \dfrac{a}{R}$, where a is the distance from the BS to

the relay, R is the distance from the BS to one corner of the cell. In the simulation, the following m values are examined: 0, 0.1, 0.2, … , In addition, the results corresponding to the m values that are 2/3, 3/4 and 7/8 are also given out in Figure 6.16. The results for the $m = 0$ case corresponds to the no relay case. The simulation results show that the best range for the relay position is 2/3 to 3/4 away from the BS, when the transmit power ratio of the BS and the fixed relay is changeable.
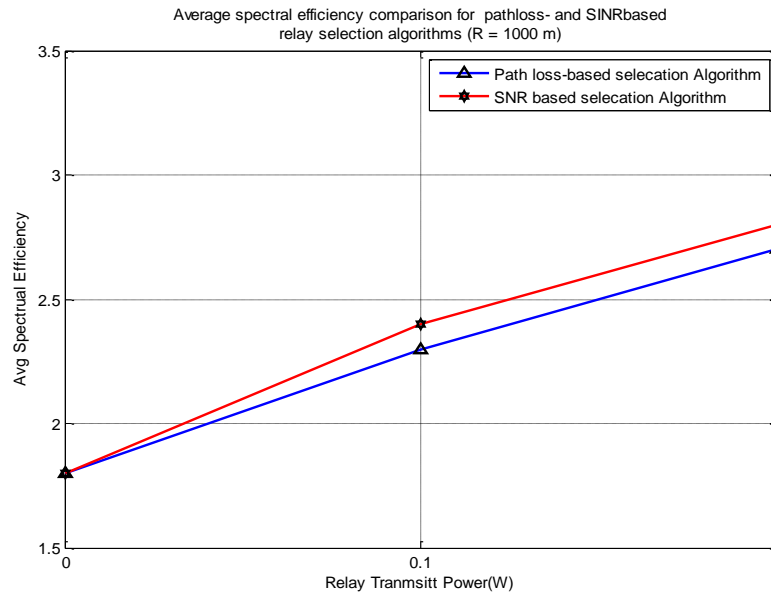
Figure 6-14: PathLoss Vs SINR based algorithm (without Interference)

Fixed relays can also help extend the system coverage. Figure 6.17 (central cell), an average spectral efficiency of 3.7 bits/sec/Hz can only be achieved in a small cell (500 m) without relays, while the same amount of average spectral efficiency can be achieved in a medium cell (1000 m) with the help of relays. In other words, the system coverage with relays is four times as much as the coverage without relays, because the coverage area is proportional to the square of the radius.
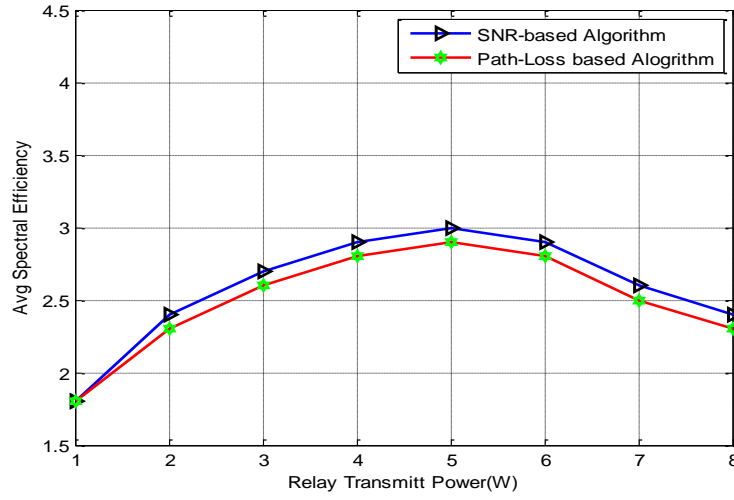
Figure 6-15: PathLoss Vs SINR based algorithm (with interference)
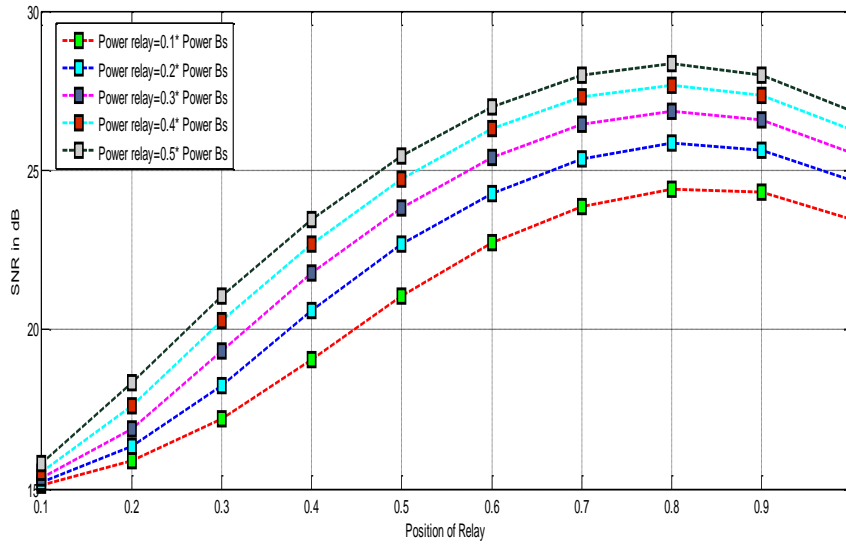
Figure 6-16: Relay Position with respect to BS

Figure 6.17 shows the average spectral efficiency for different cell radius. Figure 6.18 shows that when deploying fixed relay overall system spectral efficiency will increase and outage will decrease as shown in Figure 6.19.

We also consider the case where mobile and fixed relay co-exist together. The % of un-satisfied users, by deploying fixed and mobile relay decrease as compared to without relaying case as shown in Figure 6.20. Number of un-satisfied users is calculated according to equation (6.4).

$$\% \ of \ \ Un \text{-} \ satisfied \ user = \frac{n : drop \ users}{m : total \ number \ of \ users}$$

$$n = \left\{ \begin{array}{c} Packet \ cannot \ tranfer \ 'k' \ time \\ OR \\ SINR < \text{-} \ 10.15 \ db \end{array} \right\}$$

(6.4)

'*k*' is total number of time packet transfer (k = 5) If k is more than '5', then we consider packet is dropped.

.
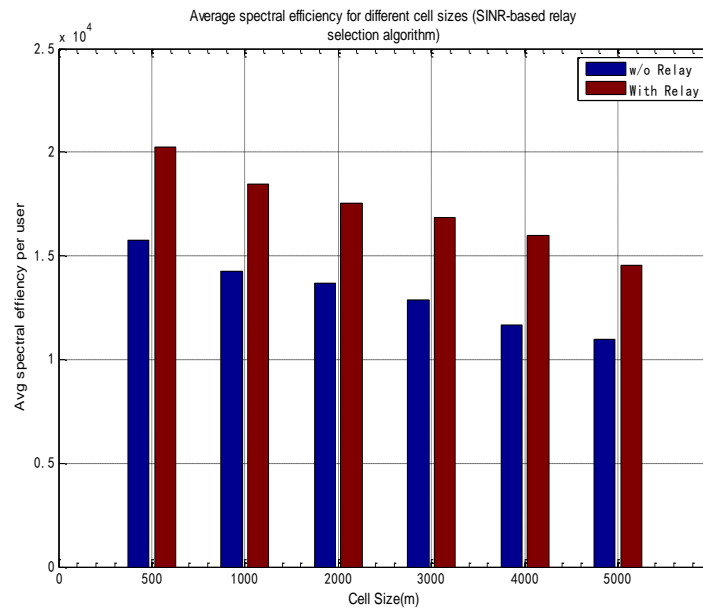
Figure 6-17 : Average Spectral efficiency for different cell size
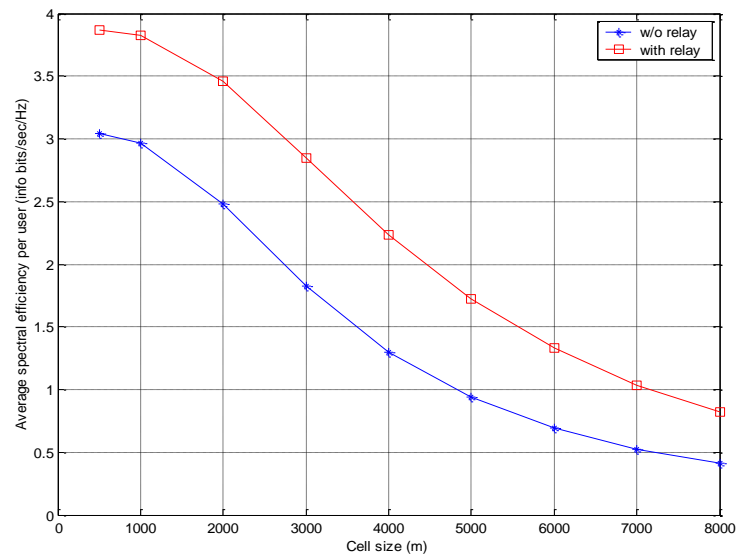
Figure 6-18: Overall system spectral efficiency
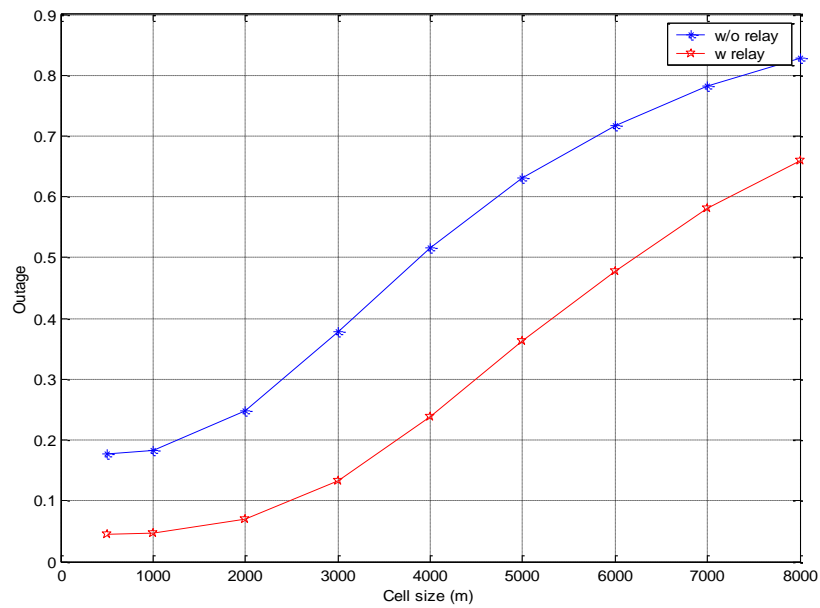
Figure 6-19: Outage probability of system
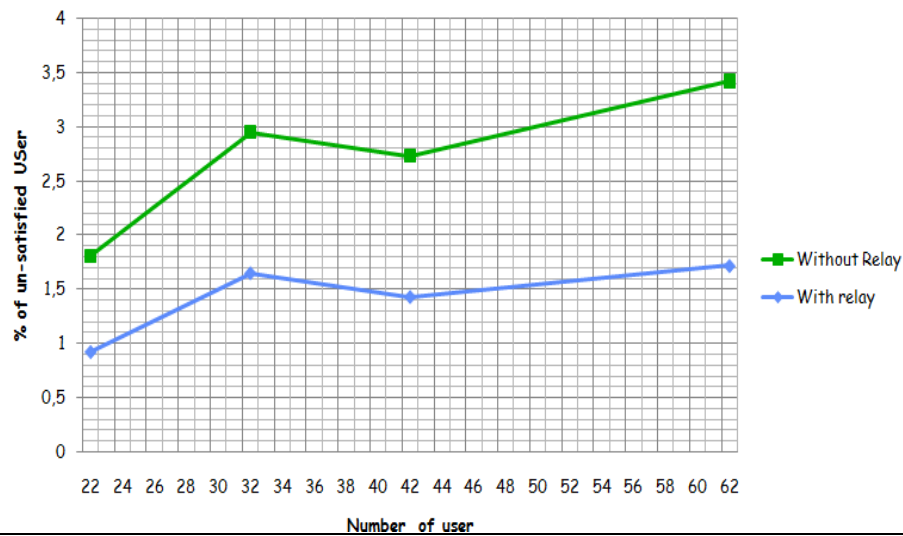
Figure 6-20: Mobile + Relay Node Vs Number of un-satisfied User

Figure 6.21.shows the results for half duplex mode.In first TTI BS send data to relay node and destination node and in second TTI, data is transfer from Relay to destination.
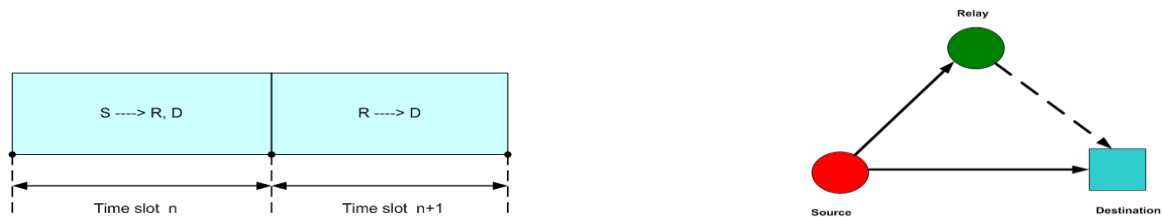
Figure 6-21: Half duplex

In the simulation result as show in figure 6.22.we observe direct link is better when we have a higher SNR values and when our cell radius start increasing, then half duplex perform better (low SNR).
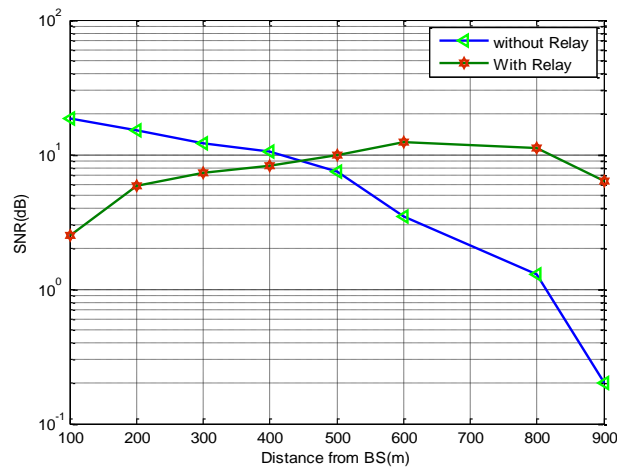
Figure 6-22: Half duplex Vs direct link

Figure 6.23 shows the performance at the edge of the cells. With no relays, we see that 20% of users are below the threshold (SNR=0db, not communicating with base station), As we increase the number of relays, the number of users who are not communicating also decreases. (with 3 relays only 8% of users are below the threshold).
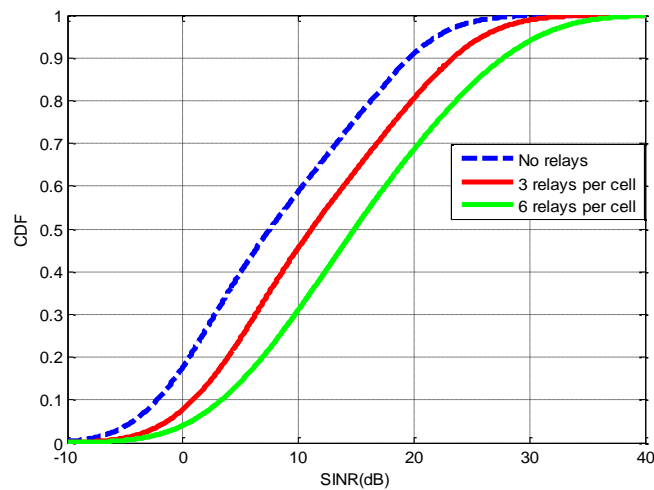
Figure 6-23: SINR at the edge of cells

Figures 6.24 to 6.25 show the result for various combinations of coding and modulation schemes used 1000m cell radius. In Figures 6.24 (with relaying case) the combinations of 64-QAM with different code rates are used more often than without relaying case; this is due to the help provided by relays.
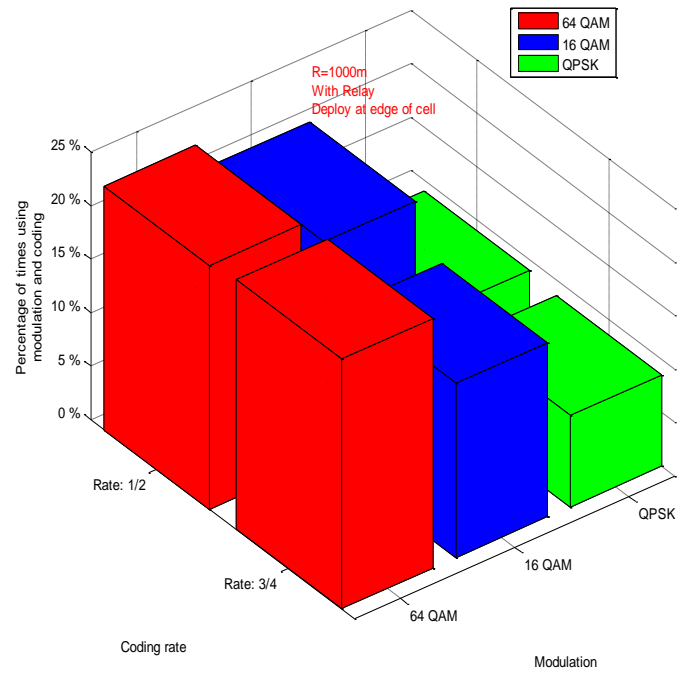
Figure 6-24: Percentage of time using a combination of modulation and coding (with relaying, SINR-based relay selection algorithm).
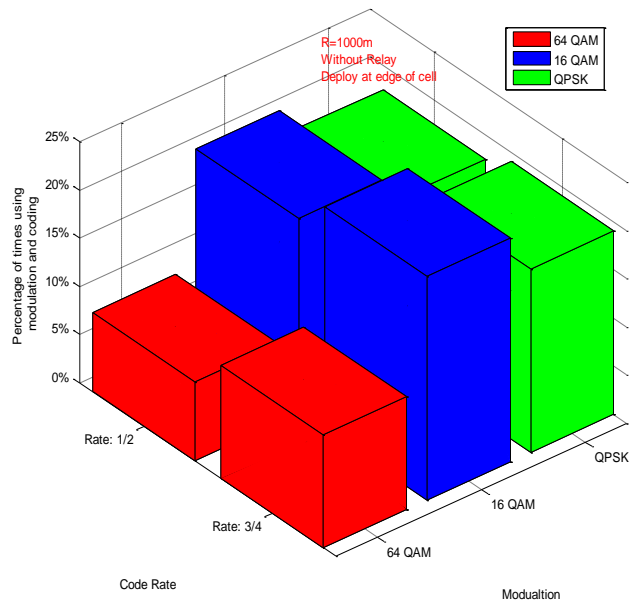

Figure 6-25: Percentage of time using a combination of modulation and coding (without relaying, SINR-based relay selection algorithm).

6.4.1 Fairness using Gini-Coefficient

The Gini coefficient was developed by the Italian Statistician Corrado Gini (1912) as a summary measure of income inequality in society. It is usually associated with the plot of wealth concentration introduced a few years earlier by Max Lorenz (1905). Since these measures were introduced, they have been applied to topics other than income and wealth, but mostly within Economics (Cowell, 1995, 2000; Jenkins, 1991; Sen, 1973).

G is a measure of inequality, defined as the mean of absolute differences between all pairs of individuals for some measure. The minimum value is 0 when all measurements are equal and the theoretical maximum is 1 for an infinitely large set of observations where all measurements but one has a value of 0, which is the ultimate inequality (Stuart and Ord, 1994). When G is based on the Lorenz curve of income distribution, it can be interpreted as the expected income gap between two individuals randomly selected from the population (Sen, 1973). The Lorenz curve (shown in Figure 6.26) is plotted as the cumulative proportion of the variable against the cumulative proportion of the sample (i.e. for a sample of 30 observations the cumulative proportion of the sample for the 15th observation is simply 15/30). To get the cumulative proportion of the variable, first sort the observations in ascending order and sum the observations, then each kth cumulative proportion is the sum of all x_i/xsum from i=1 to k.

The classical definition of G appears in the notation of the theory of relative mean difference as shown in equation (6.5):

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|}{2n^2 \bar{x}}$$

(6.5)

- where x is an observed value, n is the number of values observed and x bar is the mean value.
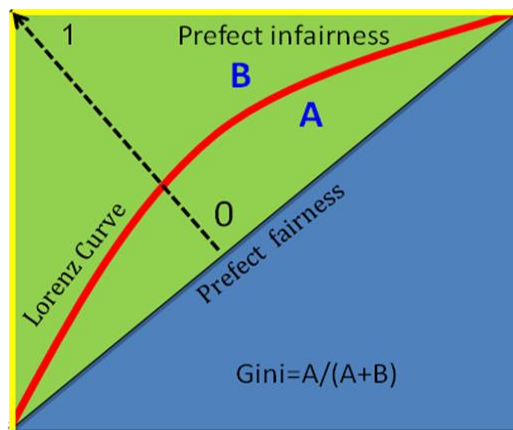

Figure 6-26 : Lorenz Curve

According to our scenario we use Gini as a fairness metrics to calculate the Gini coefficient with and without relay. We deployed 3000 user in the cell and measured the Gini coefficient with and without relay nodes using SNR as the measuring criteria of every users. Gini Index value lies between 0 and 1 if we go toward line of prefect fairness (closer to 0) our fairness increase.0 mean prefect Fairness if we go away from line of prefect fairness(closer to 1) our fairness decreases. Prefect in-fairness Lorenz Curve tells us how much data is deviated from the Line of Prefect fairness. The Gini coefficient is defined graphically as a ratio of two surfaces involving the summation of all vertical deviations between the Lorenz curve and the perfect Fairness line (A) divided by the difference between the perfect Fairness and perfect unfairness lines (A+B).

6.4.2 SNR Based Fairness using Gini Formula

SNR based fairness formula is shown in equation (6.6)

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |SNR_i - SNR_j|}{2n^2 \overline{SNR}}$$

where
'SNR' observed SNR value.
'n' is the number of SNR oberved.
$\overline{SNR}$ *is the mean value*

(6.6)

Cell Radius (m)	Gini Index (Without Relay)	Gini Index (With Relay)
500	0.12	0.10
1000	0.16	0.12
1500	0.22	0.17
2000	0.26	0.20
2500	0.32	0.29
3000	0.39	0.32

Table 6.7: Gini Index with and without Relay

These gini index are calculated using 3000 users deployed in the cell using SNR as measuring criteria for each users (with and without relay).From this table we observe that, more is the cell size ,more is unfairness between the users and with relay nodes fairness is better than without relay nodes. Figure 6.27 also shows the same behaviour.
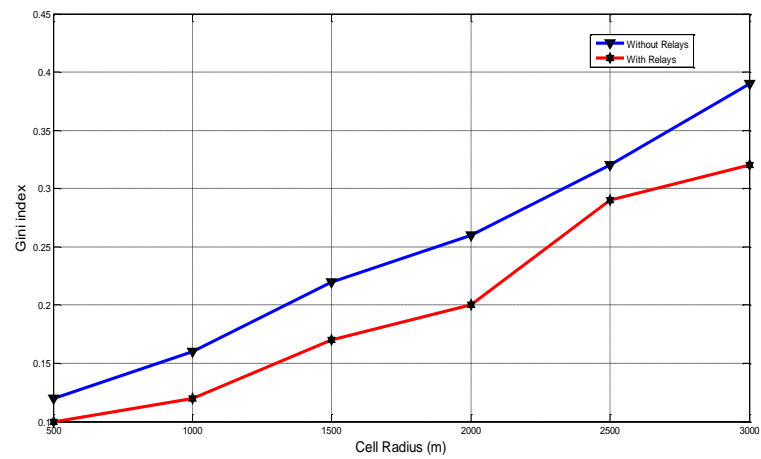
Figure 6-27: Gini Index with and without Relay

6.5   Summary and Conclusions

This chapter explained that the current conventional deployment of cellular systems exhibits certain inherent problems such as low signal-to-noise-ratio (SNR) at the cell edge, lack of fairness, coverage holes that exist due to shadowing and non-line-of-sight (NLOS) connections.

Relays are introduced to slove above mention problems in coventional cellular system.Relays are foreseen to extend the range of the BS and to increase the system throughput allowing for a cost-efficient deployment and service. Relays have been proven to substantially extend the radio coverage of a BS, especially in highly obstructed service areas. In addition to that, introducing relays to cellular networks can reduce transmission power for mobile users and offer additional diversity for improving the end-to-end radio links To evaualte the performance evaluation of relay enhanced cellular systems we modified our SLS with relay capabilites. The SLS was then used to evaluate the performance and the results of relay enhanced cellular system which shown the significant improvement in the Average throughput, Outage and Fairness.

6.6   References

[1]   IEEE 802.16j-2007, IEEE Standard for Local and metropolitan area networks. Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access systems Multihop Relay Specification

[2]   Opportunity driven multiple access," 3GPP Technical Specification Group Radio Access Network (RAN) 3GPP TR 25.924 v1.0.0, Dec. 1999.

[3]   A. Host-Madsen and J. Zhang, *Capacity bounds and power allocation for wireless relay channels*, IEEE Transaction on Information Theory, 51 (June 2005), pp. 2010– 2040.

[4]   *3GPP LTE*, http://www.3gpp.org/article/lte.

[5]   S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: FromTheory to Practice*, John Wiley & Sons, 2009. 17

[6]   *IEEE 802.16's Relay Task Group*, http://www.ieee802.org/16/relay/.

[7]   IEEE Std 802.16j, *Amendment to IEEE Standard for Local and Metropolitan AreaNetworks - Part 16: Air Interface for Broadband Wireless Access Systems – Multi-hopRelay Specification*, Jun. 2009. 15

[8]   I. B. W. A. W. Group, *Harmonized contribution on 802.16j (mobile multihop relay)usage models*, Sep. 2006. IEEE 802.16j-06/015

[9]   S. W. Peters and R. W. Heath Jr, *The future of WiMAX: multihop relaying with ieee802.16j*, IEEE Communications Magazing, 47 (Jan. 2009), pp. 104–111.

[10]   IEEE Std. 802.16-2009, *IEEE Standard for Local and Metropolitan Area Networks -Part 16: Air Interface for Broadband Wireless Access Systems*, 2009

[11]   R. Pabst, B. Walke, D. C. Schultz, and et al, "Relay-based deployment concepts for wireless and mobile broadband radio," IEEE Communications Magazine, pp.80– 89, Sep 2004.

[12]   J. Vidal et al., "Multihop networks for capacity and coverage enhancement in TDD/UTRAN", MedHocNet 2002, Sardegna, Italy, 5-7 September 2002.

[13]   H. Nourizadeh, R. Tafazolli "Capacity improvement of WCDMA cellular systems through different relaying strategies", MWCN 2005, Marrakech, Morocco, 19 September.

[14]   V. Sreng, Capacity Enhancement through Two-Hop Relaying in Cellular Radio Systems, Master's thesis, Carleton University, January 2002.

[15]   IEEE S802.16j-08/050, *Maximum number of hops for centralized scheduling mode*, Jan. 2008.

[16]   J. Cho and Z. Haas, *On the throughput enhancement of the downstream channel in cellular radio networks through multihop relaying*, IEEE J. Sel. Areas Comm., 22 (Sept. 2004), pp. 1206–1219.

[17]   3GPP R1-030224, Nortel Networks, "Update of OFDM SI simulation methodology".

Chapter 7: Conclusions &Future Work

7.1 Conclusions

Following are the main conclusion from each chapter:

Chapter 2 presented an overview of Wireless Technologies especially IEEE802.16e, IEEE802.16j and IEEE802.22 and set the stage for more detailed exploration in subsequent chapters. WiMAX is based on a very flexible and robust air interface defined by the IEEE 802.16. IEEE 802.16 has formed a task group to extend the IEEE 802.16e-2005 standard to include multi-hop communication, indicating that the field has reached a significant level of maturity. This amendment is called IEEE 802.16j. The evidence of the change and evolution in the approaches of spectrum management can already be seen in the development of the IEEE 802.22 cognitive radio based standard for fixed, point-to-multipoint, wireless regional area networks. The WiMAX physical layer is based on OFDM, which is an elegant and effective technique for overcoming multipath distortion. The physical layer supports several advanced techniques for increasing the reliability of the link layer. These techniques include powerful error correction coding, including turbo coding and LDPC, hybrid-ARQ, and antenna arrays.

WiMAX supports a number of advanced signal-processing techniques to improve overall system capacity. These techniques include adaptive modulation and coding, spatial multiplexing, and multiuser diversity. WiMAX has a very flexible MAC layer that can accommodate a variety of traffic types, including voice, video, and multimedia, and provide strong QoS. WiMAX has several features to enhance mobility-related functions such as seamless handover and low power consumption for portable devices. WiMAX offers very high spectral efficiency, particularly when using higher-order MIMO solutions.

Chapter 3 explained all the necessary information for modeling SLS. Simulations in wireless and cellular networks can be classified in three categories (Link level, Network level, and System level) with respect to the targeted layers in the protocol stack and with respect to the considered level of abstraction and we also explained that the implementation of a system level tool requires complex modeling of several issues. The minimum modeling involves three aspects: the users' behaviors; the traffics' behavior for the application considered and the radio aspects involved in the transmission of the signals from users to destination.To evaluate the performance of SLS a large number of statistics are collected for the computation of the metrics. These performance statistics

are generated as outputs from the system level simulations and are used in the performance evaluation of the used scenarios and proposed algorithms.

Chapter 4 presents in a detailed way the design, modeling and implementation of the Link level and System Level simulators. Traffic models and channel models for SLS were also discussed in this chapter. The cellular layout architecture for SLS was fully described. Of particular interest in the realization of system level simulations is the definition of the proper link to system level interface and the definition of the proper set of look up tables used in the mapping of physical layer performance. The procedure followed into the derivation of the Signal to Interference plus Noise Ratio (SINR) and the mapping function used to map the vector of SINRs into a single scalar to be inputted into the look-up tables is detailed. This chapter also describes in detail all steps followed in each frame period by the communication protocol, between the base and the mobile stations, for the transmission of the information associated to the scheduled users.

Chapter 5 covered the simulation scenarios, which are implemented in order to observe the system performance as, OR terminals are introduced in the system level simulator. Therefore, we enhanced System level simulator that allows us to assess the coexistence of OR as a secondary system together with primary (victim device) UMTS FDD cellular systems. We explained how we use frequency band in UMTS by using OR, we enhanced our SLS by adding different classes to achieve our goals and we have also considered an ad hoc behaviour in the opportunists radio (ORs) and suggested that by implementing ad hoc features in the ORs will improve the overall performance of the system. The OR network has Cognitive radio Ad Hoc Networks (CRAHNs) topology and exploits opportunities from a spectrum pool of several UMTS frequencies. Different routing strategies for this CRAHNs network are investigated and compared.We implemented routing feature of an ad hoc network in ORs, which dramatically reduce the server interference from the UMTS BS. Routing strategies based on hop count and capacity information were investigated and compared. The capacity-based routing strategy can reduce the overall interference level in the system by shifting traffic to the edge of the network.

Chapter 6 explained that the current conventional deployment of cellular systems exhibits certain inherent problems such as low signal-to-noise-ratio (SNR) at the cell edge, lack of fairness, coverage holes that exist due to shadowing and non-line-of-sight (NLOS) connections. Relays are introduced to slove above mention problems in covential cellular system.Relays are foreseen to extend the range of the BS and to increase the system throughput allowing for a cost-efficient deployment and service. Relays have been proven to substantially extend the radio coverage of a BS, especially in highly obstructed service areas. In addition to that, introducing relays to cellular networks can reduce transmission power for mobile users and offer additional diversity for

improving the end-to-end radio links To evaualte the performance evaluation of relay enhanced cellular systems we modified our SLS with relay capabilites. The SLS was then used to evaluate the performance and the results of relay enhanced cellular system which shown the significant improvement in the Average throughput, Outage and Fairness

7.2   Future Work

There exist multiple lines of investigation to continue the research carrier out in this work. For examples

1. Optimization of SLS

2. Validate SLS for other 4G tecnlogies i.e LTE

3. Commom Radio Resource Management (CRRM) for different 4G wireless technologies (i.e LTE,WiMax and HSPDA) with efficient load balancing algorithm

4. Energy Efficient RRM (Green RRM)

Optimization describes the process of finding, for a particular problem, the best solution with respect to a particular goal from a set of possible alternatives. Optimization is an inherent part of many real-life problems. For example, the production of most industrial goods requires a sophisticated, geographically distributed supply chain. An optimal planning of such a supply chain could for example be done with respect to minimizing the overall cost of transportation and temporary storage. This problem is known as the warehouse problem. A large number of optimization problems also exist in the area of mobile communications. Offline problems occur during the network planning process. Examples are the assignment of frequency blocks to cells such that a certain minimum SINR is achieved at any location, or the optimization of the size of location areas. Online problems occur during the resource assignment process, for example when using channels borrowing schemes that allow a cell to use radio channels originally assigned to neighboring cells. At the frame level, the scheduling of radio blocks constitutes a highly dynamic optimization problem, often aiming at maximizing the system throughput while taking into account several constraints, such as QoS requirements.

SLS is validated for the LTE during the course of this dissertation and work on CRRM is going now with efficient Load balancing algorithm. The heterogeneous network concept facilitates the utilizationof a common manager of the radio resources in each radio access network (RAN). Following the 3GPP approach, CRRM strategies are considered to co-ordinately manage the radio resources belonging to multiple radio access technologies (RATs) in an efficient way. CRRM is then a general concept, applicable to any combination of RATs, although the specific implementation and the degree of coordination highly depend on the coupling between the specific

RANs.Research on the design of green radio solutions is still in its infancy, significant results already exist on energy efficient designs to conserve battery power for mobile terminals and wireless sensors. Because of the focus on mobile terminals, prior work has primarily addressed uplink communications. However, from an overall perspective, it is necessary to consider the downlink, as base stations are the primary energy consumers of cellular networks. Energy efficiency from the perspective of optimized downlink communications, with a direct objective of reducing energy consumption at the base station is an area of future research.

Demonstration of SLS

This section presents the demonstration of SLS in pictorial form.

1. Load necessary parameter in Configuration file, one wants to simulate as show in Figure 8.1.
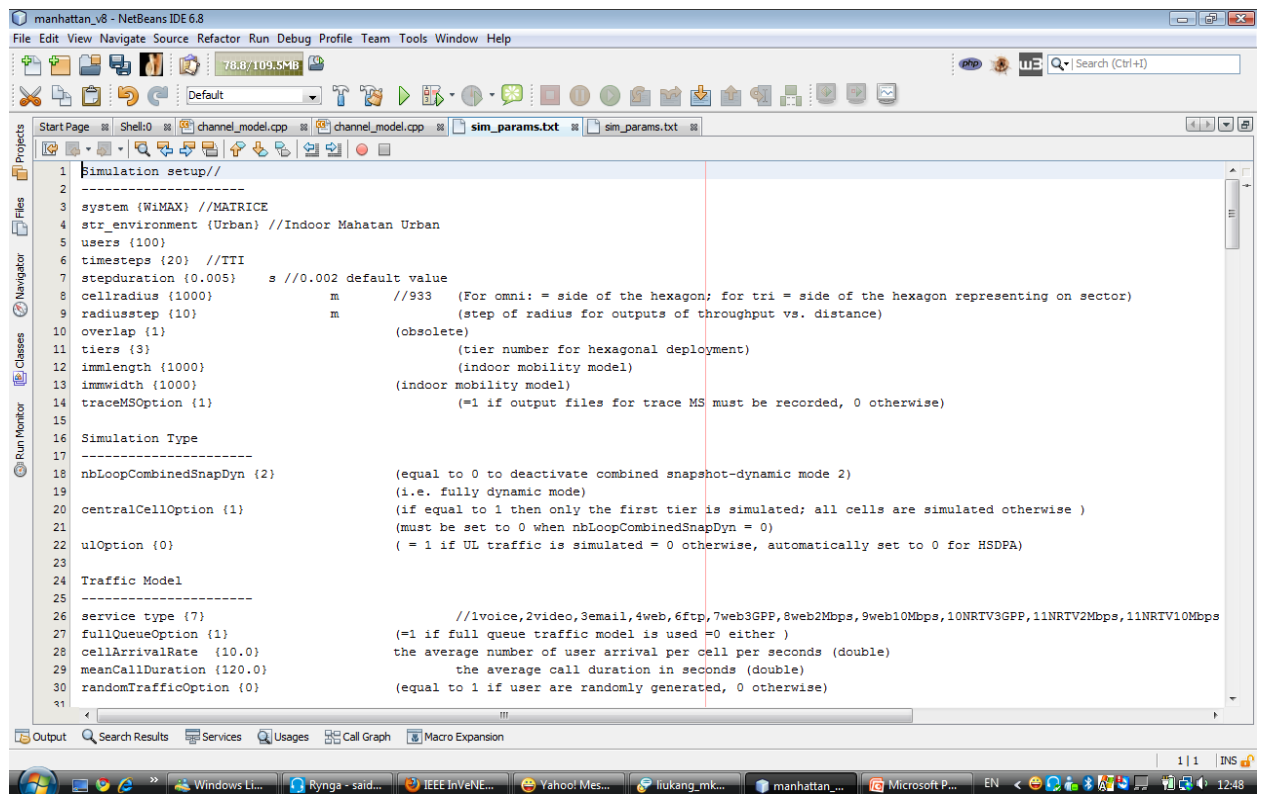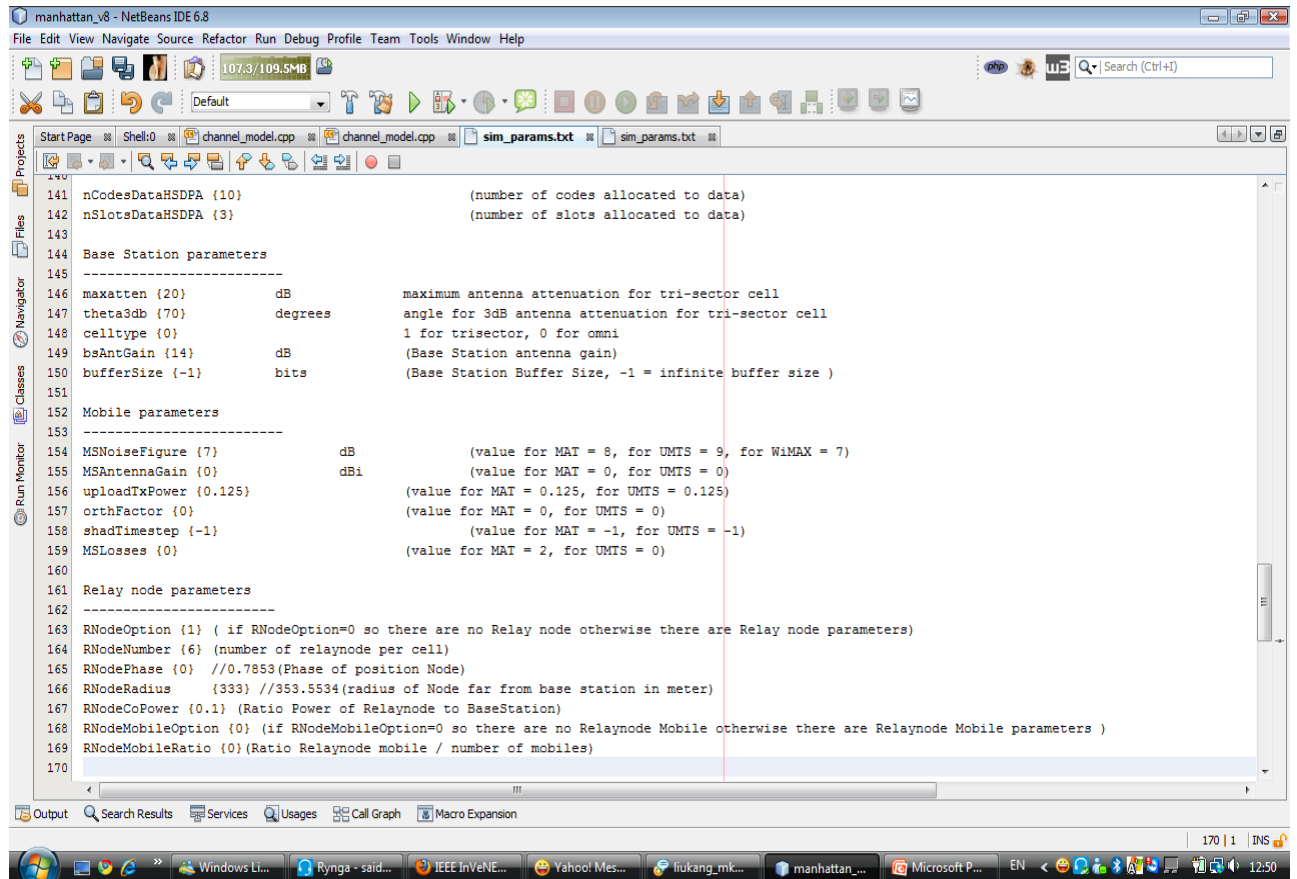
Figure 8-1: Demostration on System Level Simulator

2. Choose Relay assisted Mode/OR Modes of SLS and configure Relay parameter as show in Figure 8.2.

Figure 8-2

3.Run the Simualtion and collect desired output mertics in output file as shown in Figure 8.3.

Figure 8-3

ANNEX: A   (OFDMA Sub-Channelisation)

As a matter of fact, the OFDM PHY includes some OFDMA access. Subchannelisation was included in 802.16-2004 for the uplink and also for the downlink in amendment 802.16e. The principle is the following. The 192 useful data OFDM subcarriers of OFDM PHY are distributed in 16 subchannels made of 12 subcarriers each. Each subchannel is made of four groups of three adjacent subchannels each . A subchannelised transmission is a transmission on only part of the OFDM subcarrier space. The subchannelised transmission can take place on 1, 2, 4, 8 or 16 subchannels. A five-bit indexation shown in Table 1. Indicates the number of subchannels and the subcarrier indices used for each subchannel index for the uplink. As shown in this table, one or more pilot subcarrier(s) (there are eight in total) are allocated only if two or more subchannels are allocated. The subcarriers other than the ones used for subchannelised transmission are nonactive (for the transmitter). The five-bit subchannel index is used in the uplink allocation message UL-MAP.

Subchannell index			Pilot frequency index	Subchannel index (continued)	Subcarrier frequency indices
	0b0010	0b00010	−38	0b00001 0b00011	100: −98; −37: −35; 1: 3; 64: 66
				0b00100	−97: −95, −34; −32, 4; 6, 67: 69
		0b000110	13		−94: −92, −31: −29,7: 9, 70: 72
		0b01000	0b00111		−91: −89, −28: −26, 10: 12, 73: 75
	0b01000	0b01010	−88	0b01001	−87: −85, −50: −48, 14: 16,51: 53
				0b1011	−84, −82, −47: −45, 17: 19, 54: 56
		0b01100		0b01101	−81: −79, −44: −42, 20: 22, 57: 59
		0b01110	63	0b01111	−78: −76, −41: −39, 23: 25, 60: 62

Subchannell index			Pilot frequency index	Subchannel index (continued)	Subcarrier frequency indices
0b10000 (no subchannelisation					−75: −73, −12: −10, 26:28,89:91
		0b10010	−13	0b10011	−72: −70, −9: −7, 29: 31. 92: 94
		0b10100		0b10101	−69: −67, −6: −4, 32: 34, 95: 97
		0b10110	38	0b10111	−66: −64, −3: −I. 35: 37, 98: 100
		0b11000		0b11001	−62: −60. −25: −23, 39: 41, 76: 78
		0b11010	−63	0b11011	−59: −57, −22: −20, 42: 44, 79: 81
	0b11100			0b11101	−56: −54, −19: −17, 45: 47, 82: 84
		0b11110	88	0b11111	−53: −51, −16; −14, 48: 50, 85: 87

Table 1: Number of Subchannel and Subcarrier indices used for each(five bits) subchannel index

Subchannelised transmission in the uplink is an option for an SS. It can be used only if the BS signals its capability to decode such transmissions. The BS must not assign to any given SS two or more overlapping subchannelised allocations in the same time.The standard indicates that when subchannelisation is employed, the SS maintains the same transmitted power density unless the maximum power level is reached. Consequently, when the number of active subchannels allocated to an uplink user is reduced, the transmitted power is reduced proportionally, without additional power control messages. When the number of subchannels is increased the total transmitted power is also increased proportionally. The transmitted power level must not exceed the maximum levels dictated by signal integrity considerations and regulatory requirements. The subchannelisation can then represent transmitted power decreases and, equivalently, capacity gains.The 802.16e

amendment defined an optional downlink subchannelisation zone in the OFDM PHY downlink subframe. Uplink subchannels are partly reused.

The Main Permutation Modes in OFDMA

Subtracting the guard subcarriers and the DC subcarrier from N_{FFT} gives the set of 'used' subcarriers $N_{used.}$ For both the uplink and downlink, these subcarriers are allocated as pilot subcarriers and data subcarriers according to one or another of the defined OFDMA permutation modes.

Two families of distribution modes can be distinguished:

- Diversity (or distributed) permutations. The subcarriers are distributed pseudo-randomly. This family includes: FUSC (Full Usage of the SubChannels) and PUSC (Partial Usage of the SubChannels), OPUSC (Optional PUSC), OFUSC (Optional FUSC) and TUSC (Tile Usage of SubChannels). The main advantages of distributed permutations are frequency diversity and intercell interference averaging. Diversity permutations minimise the probability of using the same subcarrier in adjacent sectors or cells. On the other hand, channel estimation is not easy as the subcarriers are distributed over the available bandwidth.
- Contiguous (or adjacent) permutations. These consider a group of adjacent subcarriers. This family includes the AMC (Adaptive Modulation and Coding) mode. This type of permutation leaves the door open for the choice of the best-conditions part of the bandwidth. Channel estimation is easier as the subcarriers are adjacent.

Mandatory permutation modes of the presently defined mobile WiMAX profiles are:

- for the downlink: PUSC, FUSC and AMC;
- for the uplink: PUSC and AMC.

A subchannel is the minimum transmission unit in an OFDMA symbol. Each of the permutation modes of OFDMA has its definition for a subchannel. There is also a difference between allocation of the data and pilot subcarriers in the subchannels between the different possible permutation modes:

- For (downlink) FUSC and downlink PUSC, the pilot tones are allocated first. What remains are data subcarriers, which are divided into subchannels that are used exclusively for data.
- For uplink PUSC, the set of used subcarriers is first partitioned into subchannels and then the pilot subcarriers are allocated from within each subchannel.

Thus, in the FUSC mode, there is one set of common pilot subcarriers, while in the uplink PUSC mode, each subchannel contains its own set of pilot subcarriers. For the downlink PUSC mode, there is one set of common pilot subcarriers for each major group including a set of subchannels .

Slot and Burst (Data Region)

A slot in the OFDMA PHY has both a time and subchannel dimension. A slot is the minimum possible data allocation unit in the 802.16 standard. The definition of an OFDMA slot depends on the OFDMA symbol structure, which varies for uplink and downlink. for FUSE and PUSC, and for the distributed subcarrier permutations and the adjacent subcarrier permutation. See table 2 for the different possibilities.

Permutation mode and communication way	Slot definition
Downlink FUSC; downlink OFUSC	1 subchannel × 1 OFDMA symbol
Downlink PUSC	1 subchannel × 2 OFDMA symbol
Uplink PUSC, uplink additional PUSC, downlink TUSCI and TUSCI	1 subchannel × 3 OFDMA symbol
AMC (uplink and downlink)	1 subchannel × (1, 2 or 3) OFDMA symbol

Table 2:Slot defintion

In OFDMA, a data region (or burst) is a two-dimensional allocation of a group of slots. i.e. a group contiguous subchannels, in a group of contiguous OFDMA symbols as shown in figure 1 and the end of the PUSC section below for an example).
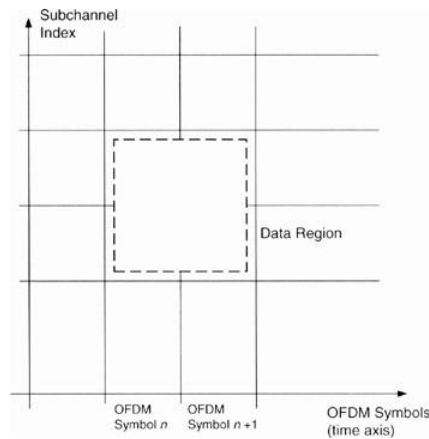
Figure 1 : Example of the data region that defines the OFDMA burst allocation

A segment is a subdivision of the set of available subchannels, used for deploying one instance of the MAC. A permutation zone is a number of contiguous OFDMA symbols, in the downlink frame or the uplink frame, that use the same permutation mode. A downlink frame or an uplink frame may contain more than one permutation zones as shown in Figure 2, providing great malleability for designers.
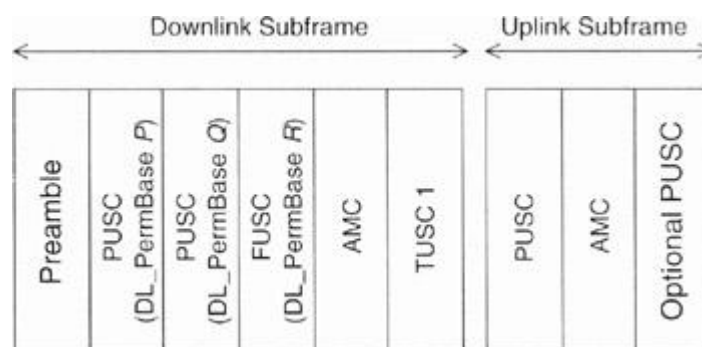
Figure 2 : Example of different permutation zones in uplink and downlink frames

PUSC Permutation Mode

The global principle of PUSC (Partial Usage of SubChannels) is the following. The symbol is first divided into subsets called clusters (downlink) or tiles (uplink). Pilots and data carriers are allocated within each subset. This allows partial frequency diversity. Some main MAC messages and some PHY subframe fields are transmitted in the PUSC mode: FCH, DL-MAP and UL-MAP . Downlink PUSC subchannel allocation will now be detailed which we used in our simulation, which is illustrated by an example.

The global principle of downlink PUSC cluster and subcarrier allocation is illustrated in Figure 3. Considering, for example, a 1024-FPT OFDMA Symbol, the number of guard subcarriers + DC carrier is (in the case of 1024 FFT) $92 + 91 + 1 = 184$. Therefore, the number of pilot and data carriers to be distributed is $1024 - 184 = 840$. The parameters of this numerical example are given in Table 3.

Parameter	FFT Size	BW	G	N	Pilot + data subcarriers	N_{FFT}	f_s	Δf
Value	1024	10MHz	1/8	28/25	840	1024	11.2MHz	10.9375 kHz

Table 3: Numerial Parameter of downlink PUSC example
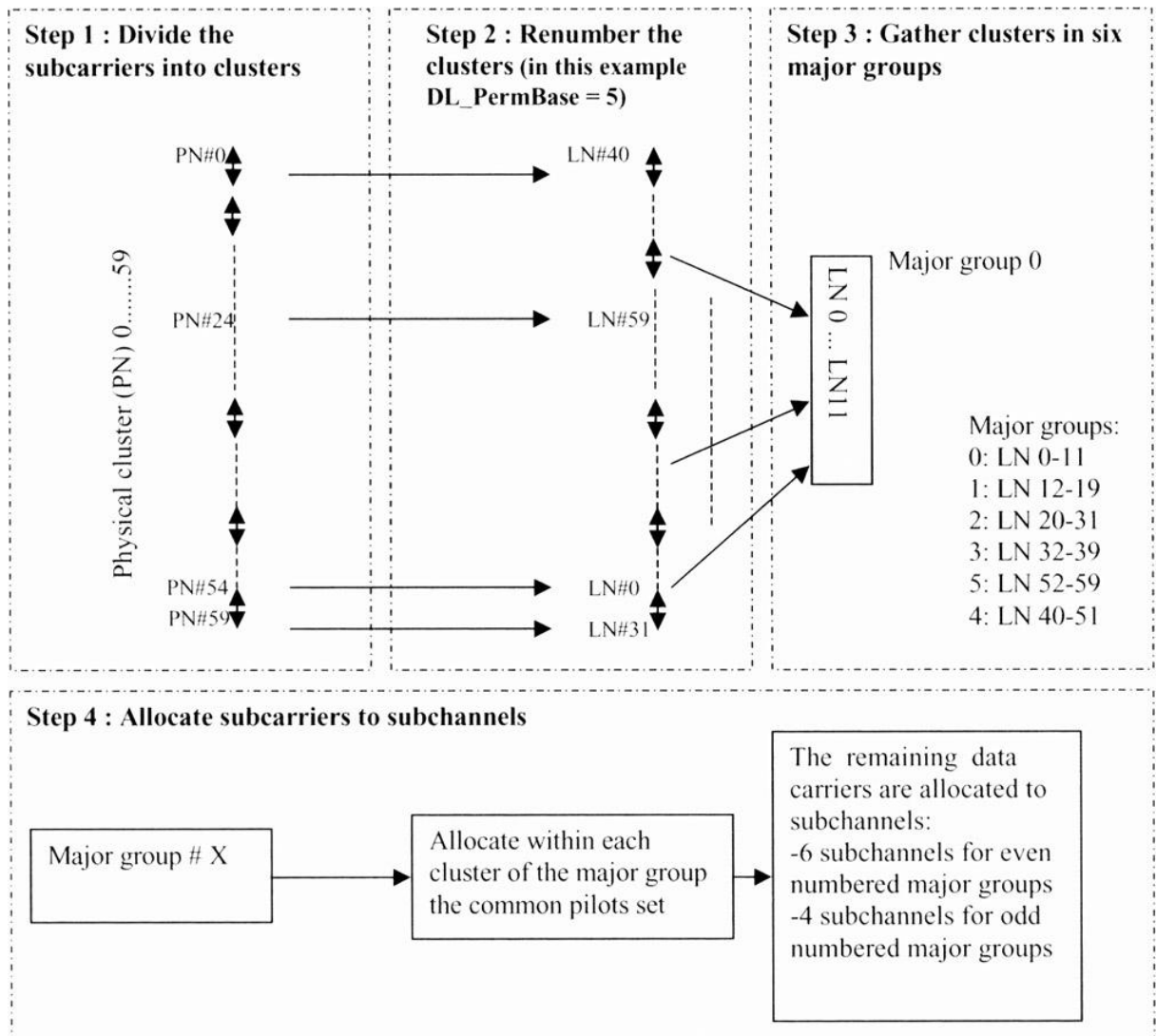
Figure 3: Illustration of the downlink PUSC Cluster and subcarrier allocation

Allocation Steps

Step 1. Divide the Subcarriers into Clusters

After removing the guard and DC subcarriers, the 840 (pilot and data) subcarriers are divided into 60 clusters of 14 adjacent sub carriers each ($14 \times 60 = 840$) as shown in Figure 4. We here mention that a PUSC cluster has nothing to see with a cluster of cells. The Physical Cluster number is between 0 and 59. Pilot subcarriers are placed within each cluster depending on the parity of the OFDMA symbols, as shown in Figure 5.
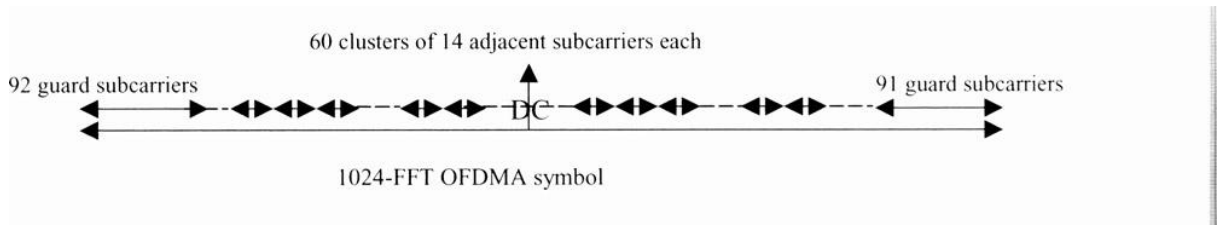
Figure 4 :Cluster allocation
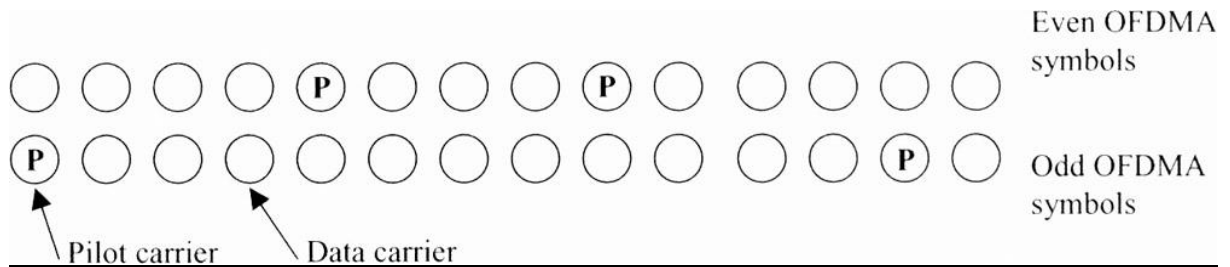
Figure 5: Cluster structure

Step 2. Renumber the Clusters

The clusters are renumbered with Logical Numbers (LNs). The cluster LN is also between 0 and 59. In order to renumber the clusters, the DL_*PermBase* parameter is used.*DL_PermBase* is an integer ranging from 0 to 31, which can be indicated by DL_MAP for PUSC zones. The clusters are renumbered to LN clusters using the following formula (denoted Formula (0) in the following):In the case of the first downlink zone (containing the FCH and DL-MAP).

$$Cluster\ Logical\ Number = Renumbering\ sequence\ (Cluster\ Physical\ Number) \qquad (0)$$

else: *Cluster Logical Number*
$$= Renumbering\ sequence(((Cluster\ Physical\ Number)+13*DL\_PermBase)\ mod$$
$$Nclusters))$$

where the *Renumbering sequence (j)* is the *jth* entry of the following vector:

[6, 48, 37, 21, 31, 40, 42, 56, 32, 47, 30, 33, 54, 18, 10, 15, 50, 51, 58, 46, 23, 45, 16, 57, 39, 35, 7, 55, 25, 59, 53, 11, 22, 38, 28, 19, 17, 3, 27, 12, 29, 26, 5, 41, 49, 44, 9, 8, 1, 13, 36, 14, 43, 2, 20, 24, 52, 4, 34, 0]

It should be remembered that, for 1024-FFT, *Nclusters* = **60**, so the above vector has 60 elements.

Step 3: Gather Clusters in Six Major Groups

The renumbered clusters are then gathered in six major groups, using the LN, as shown in Table 4.

Group	Cluster index
o	LN 0–11
1	LN 12–19
2	LN 20–31
3	LN 32–39
4	LN 40–51
5	LN 52–59

Table 4.Downlink PUSC cluster major groups(FFT 1024)

Step 4. Allocate Subcarriers to Subchannels

In the downlink PUSC the number of subchannels per OFDMA symbol is 30, numbered from 0 to 29. A subchannel is made of 24 data subcarriers, which represents the data subcarriers of two clusters. It can be verified that: $30 \times 24 = 720$ data subcarriers (720 data subcarriers + 30×4 pilot subcarriers = 840 subcarriers). For the downlink PUSC, each major group is used separately in order to have a number of subchannels; i.e. one subchannel does not have subcarriers in more than one major group. In addition, all the subcarriers of one subchannel belong to the same OFDMA symbol. The pilot and data subcarrier allocations to subchannels are done as follows. The pilot subcarriers are allocated first within each cluster, placed as shown in Figure 5. In the downlink puse, there is one set of common pilot subcarriers in each major group. The remaining data

subcarriers are first renumbered from 0 to 143 or 95 depending on the parity of the major group. Then the subcarriers are allocated within each subchannel using the following formula.

$$subcarrier(k,s) = N_{subchannels} * n_k + \{p_s[\ n_k \bmod N_{subchannels}] + DL\_PermBase\} \bmod N_{subchannels}$$

where $N_{Nsuhchannels}$ is the number of subchannels in the partitioned major group, equal to 4 or 6. depending on the parity of the major group; $subcarrier(k,s)$ is the subcarrier index of subcarrier k, varying between 0 and 23, in subchannel s, whose value ranges between 0 and 143 or 95 depending on the parity of the major group; s is the subchannel index varying between 0 and 29, and so

$$n_k = (k + 13s) \bmod N_{subcarriers}$$

where $N_{subcarriers}$ is the number of data subcarriers allocated to a subchannel in each OFDMA symbol (= 24 in this case); $Ps[j]$ is the series obtained by rotating the basic permutation sequence cyclically to the left s times, which is given in the following: in the case of an odd numbered major group the basic permutation is *PermutationBase6* (3,2,0,4,5,1), while for an even numbered major group it is *PermutationBase4* (3,0,2,1).

For even numbered major groups, the 12 clusters contain the data subcarriers of 6 subchannels:

- $6 \times 24 = 144$ data subcarriers;
- $144 + 6 \times 4 = 168$ (data and pilot) subcarriers.

For odd numbered major groups, the 8 clusters contain the data subcarriers of 4 subchannels:

- $4 \times 24 = 96$ data subcarriers;
- $96 + 4 \times 4 = 112$ (data and pilot) subcarriers.

The correspondence between subchannels and major groups is given in Table 5 (for 1024-FFT OFDMA). A numerical example of the downlink PUSC allocation is proposed below.

Major group (subchannel group)	Subchannel range
0	0–5
1	6–9
2	10–15

Major group (subchannel group)	Subchannel range
3	16–19
4	20–25
5	26–29

Table 5: Correspondence between subchannel and major group

Numerical Example

Based on comprehension of the IEEE 802.16 standard, a numerical example is proposed. A start is made with step 4, the previous steps having fixed values. The aim is to find the 24 physical (data) subcarriers of subchannel 16 of the downlink PUSC. It is assumed that *DL_PermBase* = 5 (indicated in the DL-MAP MAC Management Message) and that the OFDMA symbol considered is odd numbered. Subchannel 16 is in major group 3, as shown in Table 5. Therefore, *basic permutation sequence* = (3,0,2,1), $N_{subcarriers}$ = 24 (this is the case for all subchannels) and $N_{subchannels}$ = 4 (odd numbered major group). In major group 3, the correspondence between the Logical Number (LN) and the original Physical Number (PN) is obtained by applying the equation of step 2, using the LN and its position in the renumbering sequence. Thus the correspondence is as shown in Table 6.

Cluster LN	Logical subcarrier index	Cluster PN formula (0)	Cluster physical subcarrier index
32	0–13	3	42–55
33	14–27	6	84–97
34	28–41	53	742–755
35	42–55	20	280–293
36	56–69	45	630–643
37	70–83	57	798–811
38	84–97	28	392–405
39	98–111	19	266–279

Table 6: Original cluster numbering (major group 3)

Table 7 depicts n_k and the physical subcarrier index corresponding to the each subcarrier k in subchannel s (= 16). For each subcarrier, the LN cluster of this subcarrier major group is used in order to find the physical subcarrier index (Table 6 is also used). For the pilot set for major group

3, using Table 6 values and the principle of Figure 4 gives the physical indices of each cluster pilot subcarriers. These indices are proposed in Table 8.

Logical subcarrier index (k) in the considered subchannel ($s = 16$)	n_k (formula)	Logical subcarrier index in the major group and corresponding cluster LN (using Table 6)		Physical subcarrier index (using Table 6 and Figure 4)
		Subcarrier	Cluster LN	
0	16	67	37	805
1	17	69	37	807
2	18	76	38	396
3	19	78	38	398
4	20	83	38	403
5	21	85	39	267
6	22	92	39	274
7	23	94	39	276
8	0	3	32	45
9	1	5	32	47
10	2	12	32	55
11	3	14	33	86
12	4	19	33	91
13	5	21	33	93
14	6	28	34	746
15	7	30	34	749
16	8	35	34	753
17	9	37	35	281
18	10	44	35	288
19	11	46	35	290
20	12	51	36	633
21	13	53	36	635
22	14	60	36	643
23	15	62	37	800

Table 7: Subcarrier Allocation

Cluster PN	Pilot subcarrier physical index
3	42 and 54
6	84 and 96
53	742 and 754

Cluster PN	Pilot subcarrier physical index
20	280 and 292
45	630 and 642
57	798 and 810
28	392 and 404
19	266 and 278

Table 8: Pilot Subcarrier physical index

The data rate corresponding to one slot (equal to one subchannel over two slots) will now be computed. One OFDMA symbol duration is 102.9 μs (see the numerical example parameters in Table 3. A slot contains 48 subcarriers and then 48 modulation symbols. The instantaneous data rate of one subchannel can then be computed. The obtained values (no repetition) are given in Table 8.

	BPSK 1/2	QPSK 1/2	QPSK 3/4	16-QAM 1/2	16-QAM 3/4	64-QAM 2/3	64-QAM 3/4
Instantaneous data rate	116.6	233.3	349.8	466.5	699.75	932.9	1049

Table 8: Instantaneous data rate of one subchannel(kb/s)

A lot is, by definition, in the downlink PUSC, made of 48 subcarriers:

$$1 \text{ subchannel} \times 2 \text{ OFDMA symbols} = 24 \text{ subcarriers/OFDMA symbol} \times 2 \text{ OFDMA symbols}$$
$$= 48 \text{ subcarriers (or 48 modulation symbols)}$$

The data region information indicated in a DL-MAP MAC Management Message for an (OFDMA) downlink user is:

- OFDMA symbol offset. The offset of the OFDMA symbol in which the burst starts, measured in OFDMA symbols from the beginning of the downlink frame in which the DL-MAP is transmitted.
- Subchannel offset. The lowest index OFDMA subchannel used for carrying the burst, starting from subchannel 0.

- Number of OFDMA symbols. The number of OFDMA symbols that are used (fully or partially) to carry the downlink PHY burst. The value of the field is a multiple of the slot length in symbols.
- Number of subchannels. The number of subchannels with subsequent indexes used to carry the burst.

ANNEX: B (Methods for Effective SINR Mapping)

Introduction

In the following sub-sections some proposed methods for SINR mapping in the literature are presented.

Mean Instantaneous Capacity Mapping Method (MIC)

In the Mean Instantaneous Capacity Mapping Method (MIC) compression method the mapping is performed by computation of the AWFGN channel capacity for each one of the N SINR values corresponding to the data sub-carrier set into which the data block is mapped into. Then an average over the set of capacity values is performed in order to find the effective SINR. The compressed SINR value is given by equations (1) and (2).

$$C(SINR_{eff}) = \frac{1}{N}\left(\sum_{k=1}^{N} C(SINR_k) \right)$$

(1)

$$C(SINR_k) = \log_2(1 + QSINR_k)$$

(2)

The parameter Q is a channel-specific correction factor that accounts for the variations in SINR between transmissions.

Effective SINR Mapping Based on Mutual Information (MI-ESM)

It is a one-dimensioning effective SINR mapping, very similar to EESM. It employs Bit Interleaved Coded Modulation (BICM) as the compression function. The MI-ESM is computed according to equation (3).

$$SINR_{eff} = I_{m_{ref}}^{-1}\left(\frac{1}{P} \sum_{p=1}^{P} I_{m_p}(SINR_p) \right)$$

(3)

The term $I_{m_p}(x)$ is the mutual information function for a given modulation and coding scheme and $I_{m_{ref}}(x)$ is the mutual information function for the modulation and coding scheme used as reference for the whole block. The value of m_{ref} can be set to the average number of transmitted bits per resource element.

The mutual information function for BICM is computed according to equation (4).

$$I_{m_p}(x) = m_p - E_Y \left\{ \frac{1}{2^{m_p}} \sum_{i=1}^{m_p} \sum_{b=0}^{1} \sum_{z \in X_b^i} \log \frac{\sum_{\hat{x} \in X} \exp\left(-\left|y - \sqrt{\frac{x}{\beta}}(\hat{x} - z)^2\right|\right)}{\sum_{\tilde{x} \in X_b^i} \exp\left(-\left|y - \sqrt{\frac{x}{\beta}}(\tilde{x} - z)^2\right|\right)} \right\}$$

(4)

Where:

· I_{m_p} is the mutual information of the applied modulation alphabet of size 2^{m_p} data symbols at the pth data symbol.

· m_p is the number of bits per symbol transmitted in the elementary resource.

· X is the constellation set of 2^{m_p} data symbols.

· X_b^i is the set of symbols for which bit i equals bit b.

· Y is a zero-mean unit variance Gaussian random variable.

· β is a parameter whose optimal value is found such that it results in the minimization of the gap between the predicted and measured BLER.

· E_Y is the average of the mutual information. The averaging is performed along the set of reserved resource elements.

Due to its complexity, calculation of the mutual information at real time is not efficient. A better approach is to compute it offline by means of Monte-Carlo simulations and storing the results in proper look-up tables.

In the performance of the MI-ESM method was conducted where the authors suggest that the performance of both methods EESM and MI-ESM are similar. However, it is important to mention that a point in favor for the use of this method is that, differently from the EESM mapping method, the MI-ESM includes the possibility of performing adaptive modulation and coding inside the coded block while it is being transmitted.

Two-Dimensional Data Compression and Mapping
Proposals for the SIR mapping also include the two-dimensioning data compression and mapping of the vector of received SINRs. In this case one more degree of freedom is available for the optimization and a better approximation, in terms of BLER, can result. As an example, it is proposed to use the mean, as given by equation (6) and the normalized variance (by the mean value), as given by equation (7) of the vector of SINR values of the frequency selective fading channel in the computation of the compressed effective SINR.

$$SINR_{\text{var},\beta} = \frac{1}{P} \sum_{p=1}^{P} SINR_p$$

(6)

$$SINR_{\text{var},\beta} = \frac{1}{P} \sum_{p=1}^{P} \left(\sqrt{\frac{SINR_p}{SINR_{avg}}} - E\left(\sqrt{\frac{SINR_p}{SINR_{avg}}} \right) \right)^{\beta}$$

(7)

ANNEX: C (Summary of SLS)

Table 1 shows the summary of Complete List of Parameter Implemented in SLS.

Features	Description	IEEE802.16e,j and NGMN (Baseline)	SLS
Topology	Network topology for cellular deployment	Network of 7 clusters with wraparound, each cluster 19 cells, each cell of three sectors	Implemented
Topology(Fixed Relay)	Network topology for cellular deployment	6 Fixed relay deployed 2/3 radius from BS(Parameterized)	Implemented
Site-to-site distance	Distance from one BS to another	1000m(Parameterized)	Implemented
Operating Bandwidth	bandwidth	10 MHz for TDD deployment	Implemented
Carrier frequency	2 GHz	2 GHz	Implemented
Distance-dependent pathloss	Pathloss from BS to MS	$PL = A + 10\gamma \log_{10}(\dfrac{d}{d_o}) + X_f + X_h$	Implemented
Lognormal shadowing	shadowing	WINNER model for log normal shadowing with standard deviation 8 db	Implemented
UE speeds of interest	speed	3 km/h(Parameterized)	Implemented
Total BS Tx power	Power of BS	46 dBm(Parameterized)	Implemented
UE Tx power	Power of MS	24 dBm(Parameterized)	Implemented
Power of Relay	Relay Power	0.3 of BS(Parameterized)	Implemented
Relay Selection Algorithm	How relay selected	Best SINR	Implemented
Resource Allocation	RRM	One user per TTI	Implemented
MS dropping	Dropping MSs in the service area	Following a uniform distribution throughout the service area	Implemented
Channel Mix	MSs channel models assignment	Fixed for all MS or mixed scenario random channel models according to channel mode mix or Random channel realizations of single type of channel model.	Implemented
Shadowing correlation	Inter-site correlation and shadowing correlation distance	Inter-site correlation equal to 0.5 and shadowing correlation distance equal to 50 m	Implemented
Traffic mix	MSs traffic profiles assignment	Specified traffic mix	Implemented

Simulation time	Necessary time to end simulation and output overall performance results	Simulation time for each drop should ensure convergence in the user performance metrics. A sufficient number of drops are simulated to ensure convergence in the system performance metrics	Implemented
Statistics collection	Collection of statistics from simulated MSs	Collection from all MSs	Implemented
Channel model	Ray-based (summing over multiple rays: 3GPP and SCME (SISO,MIMO)	Correlation-based with spatial correlation specific to each user but same for all taps	Implemented
Interference model	Co-channel interference model, fading model for interferers, number of major interferers, threshold receiver, interference awareness	From each MS into each cell and from each cell to each MS for each simulation interval - Same fast fading channel model but different realizations to each link between an MS and all BSs in the network -Generate fast fading channel realizations for strong interferers	Implemented
Channel estimation Model	Model to account for the impact of channel estimation errors on the performance	Effective noise variance equal to noise variance plus MSE of channel estimation errors	Implemented
PHY model	PHY abstraction method: EESM, MMIB, RBIR, etc.	RBIR	Implemented
Traffic model	MS traffic arrival processes and service profiles	Full queue	Implemented
Mobility model	Single or multiple moving MS models for MS trajectories and velocities	Location of each MS remains unchanged during a drop. The speed of an MS is only used to determine the Doppler effect of fast fading	Implemented
Link direction	DL, UL	DL, UL	Implemented
Duplexing scheme	TDD, FDD	TDD	Implemented
Multiple access	Multiple access scheme	OFDMA	Implemented
Subchannelization	Sub-carrier permutation	FUSC ,PUSC and AMC	Implemented
Resource allocation granularity	Smallest unit of resource allocation	30 sub-carriers x 30 OFDM symbols	Implemented
Basic modulation	Modulation schemes for data and control	QPSK,16QAM, 64QAM	Implemented
Data channel coding	Channel coding schemes	Convolutional coding (CC) , Turbo Coding (CTC)	Implemented
Multi-antenna Transmission Format	Multi-antenna configuration and transmission scheme	MIMO (2x2)	Implemented
Pilot structure	Pilot structure, density etc.	Specific to the subchannelization scheme and MIMO mode	Implemented

Receiver structure	MMSE/ML/MRC/Interference cancellation	MMSE (Matrix B data zone), MRC (Matrix A data zone)	Implemented
AMC	Adaptive Modulation and Coding	QPSK (1/2) 1/2/4/6, QPSK (3/4) 16QAM (1/2, 3/4) 64QAM (1/2, 2/3, 3/4, 5/6)	Implemented
H-ARQ	Chase combining, incremental redundancy, synchronous/asynchonous, adaptive/non-adaptive, ACK/NACK delay, Maximum number of retransmissions, retrasmission delay	Chase combining asynchronous, non adaptive, 1 frame ACK/NACK delay, ACK/NACK error, maximum 4 HARQ retransmissions, minimum retransmission delay 2 frames	Implemented
Channel Quality Feedback	Channel quality indicators, feedback delay, feedback rate, feedback error	CQI = Effective SINR per resource unit, no feedback delay, no prediction errors	Implemented
Power loading	Sub-carrier power allocation	Equal power	Implemented
Scheduling	The scheduler allocates system resources for different packet transmissions according to a set of scheduling metrics. The scheduling metric is traffic specific and accounts for the reported metrics (CQI and other metrics)	Max C/I for full buffer data	Implemented
Power control	Open loop / Closed loop	Maximum power	Implemented
Frequency reuse	Frequency reuse pattern	Frequency reuse of 1/3	Implemented
Radio Resource Management Algorithm	Radio Resource Management Algorithm	Relay Assignment Algorithm Packet Scheduling	Implemented