Cooperative control of relay based cellular networks

Jiang, Peng

Cooperative Control of Relay Based Cellular Networks

Peng Jiang

Submitted for the degree of Doctor of Philosophy

School of Electronic Engineering and Computer Science
Queen Mary, University of London

2009

To my beloved family

Abstract

The increasing popularity of wireless communications and the higher data requirements of new types of service lead to higher demands on wireless networks. Relay based cellular networks have been seen as an effective way to meet users' increased data rate requirements while still retaining the benefits of a cellular structure. However, maximizing the probability of providing service and spectrum efficiency are still major challenges for network operators and engineers because of the heterogeneous traffic demands, hard-to-predict user movements and complex traffic models.

In a mobile network, load balancing is recognised as an efficient way to increase the utilization of limited frequency spectrum at reasonable costs. Cooperative control based on geographic load balancing is employed to provide flexibility for relay based cellular networks and to respond to changes in the environment. According to the potential capability of existing antenna systems, adaptive radio frequency domain control in the physical layer is explored to provide coverage at the right place at the right time.

This thesis proposes several effective and efficient approaches to improve spectrum efficiency using network wide optimization to coordinate the coverage offered by different network components according to the antenna models and relay station capability. The approaches include tilting of antenna sectors, changing the power of omni-directional antennas, and changing the assignment of relay stations to different base stations. Experiments show that the proposed approaches offer significant improvements and robustness in heterogeneous traffic scenarios and when the propagation environment changes. The issue of predicting the consequence of cooperative decisions regarding antenna configurations when applied in a realistic environment is described, and a coverage prediction model is proposed. The consequences of applying changes to the antenna configuration on handovers are analysed in detail. The performance evaluations are based on a system level simulator in the context of Mobile WiMAX technology, but the concepts apply more generally.

Acknowledgements

I would like to express my deep and sincere gratitude to my supervisor, Dr. John Bigham for his helpful advice and direction throughout the entire work. His wide knowledge and logical way of thinking have been of great value for me. His understanding, encouraging and personal guidance have provided a good basis for the present thesis.

I would also like to thank all my friends at QMUL, with whom I have spent four pleasurable years.

Finally, I would like to dedicate my thesis to my family for their continuous support and love.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

16QAM	16-state quadrature amplitude modulation
3GPP	Third generation partnership project
AMC	Adaptive Modulation and Coding
ASA	Authentication and Service Authorization
BER	Bit Error Rate
BS	Base Station
CAC	Call Admission Control
CBR	Call Blocking Rate
CCI	Co-Channel Interference
CDMA	Code Division Multiple Access
CSN	Connectivity Service Network
DCPC	Distributed Constrained Power Control
DoA	Direction of Arrival
DSE	Dynamic Service Establishment
DL	Downlink
FBSS	Fast base station switching
EDT	Electrical down-tilt
FDD	Frequency Division Multiplex
FRS	Fixed Relay Station
FUSC	Fully Used Subchannelization
GLB	Geographic Load Balancing
GPC	Grant per Connection mode
GPSS	Grant per Subscriber Station mode
GSM	Global System for Mobile Communications
HARQ	Hybrid Automatic Repeat Request
LOS	Line of sight
MAC	Media Access Control
MAN	Metropolitan Area Network
MCN	Multi-hop Cellular Network
MDT	Mechanical down-tilt
MDHO	Macro Diversity handover
MIMO	Multiple input multiple output
MS	Mobile Station
NLOS	Non-Line-Of-Sight
NRS	Nomadic Relay Station

OFDM	Orthogonal Frequency Division Multiplex
OFDMA	Orthogonal Frequency Division Multiple Access
PC	Power Control
PMP	Point-To-Multipoint
PUSC	Partially Used Subchannelization
QPSK	Quadrature phase shift keying
RBCN	Relay based cellular network
RF	Radio Frequency
RNC	Radio Network Controller
RNG	Random Number Generator
RRM	Radio Resource Management
RS	Relay Station
RSSI	Received Signal Strength Indicator
SDMA	Spatial Division Multiple Access
SINR	Signal to Interference and Noise Ratio
SNR	Signal to Noise Ratio
S-OFDMA	Scalable Orthogonal Frequency Division Multiple Access
SS	Subscriber Station
TDD	Time Division Duplex
UE	User Equipment
UP	Uplink
UMTS	Universal Mobile Telecommunication System
WCDMA	Wideband Code Division Multiple Access
WiMAX	Worldwide Interoperability for Microwave Access

Chapter 1 INTRODUCTION

With frequency reuse and seamless handover features, mobile cellular networks are the most common wireless communication systems. In recent years, mobile operators not only provide voice calls, web surfing services, but also enable several real-time multimedia services, such as video calls and Mobile TV. The need for high-speed data services in mobile networks has increased remarkably. However, mobile cellular systems were originally deployed to provide good coverage for voice services, providing the required minimal data rate everywhere. Although Mobile WiMAX [1] and future 4G mobile networks [2] promise to provide higher data rates than previously, the cost of backhaul equipment and their installation and lower speed of services around the cell boundary make it expensive to provide high speed coverage over the whole network.

The integration of multi-hop capability into conventional cellular networks has been seen as an effective way to develop high speed wide-area wireless networks [3-10], where many advantages are expected in terms of coverage, throughput and QoS provisioning. For example, a Relay Station (RS) situated between a mobile station (MS) and a base station (BS) performs a relay of the transmission. Such relay based cellular networks (RBCNs) can provide high data rate services and expansion of coverage in an economic and feasible manner as the RS does not have a wired connection to the backhaul. The basic hexagonal cell structure is still assumed for the RBCNs and frequency reuse strategies are still used in this topology.

Generally, each cell in a cellular network has a predefined traffic capacity, limited by the available number of bandwidth units or channels, which could be frequencies, time slots or codes depending on the radio access technique used. If a larger number of MSs happen to attach to a single cell or if some of the attached MSs generate excessively large traffic demand[1], the attached cell can experience

[1] Hotspot area is used to represent the potential traffic congestion area in this thesis.

congestive overload, resulting in denial of service or decrease of connection throughput which could result in customers' dissatisfaction. Maximizing the service providing probability and spectrum efficiency are always major challenges for wireless network operators and engineers because of heterogeneous traffic demands, hard-to-predict user movements and complex traffic models. Although introducing RSs into a cellular network can enhance system capacity and coverage quality, servicing heterogeneous demand and demand that changes both in quantity and in geographic location over time is still a challenge for network operators. Even if the channel capacity increases using broadband communication technologies, it may not cope with heavy traffic congestion sufficiently since the wireless network environment is more complicated and the communication load is not always uniform distributed. These all have created the requirement to increase adaptability of the wireless network through more efficient ways of using the limited radio resources.

Previous research on wireless networks has led to many schemes to optimize the radio resource and increase the system capacity [11-16]. Traffic balancing is thought of as one way to handle the unexpected and uneven traffic demand and many methods have been proposed to address this problem. Previous studies on real time control of radio coverage patterns for geographic load balancing with semi-smart antennas [17, 18] have shown that the system performance can be improved by balancing traffic among different cells. In the research in this thesis, adaptive control is investigated in more detail and extended. Cooperative control is defined as a control mechanism with network wide optimization features to adaptively control the Radio Frequency (RF) domain of the physical (PHY) layer of the wireless network (such as the transmit power, tilting angle, radiation pattern) according to the network situation (traffic demand changes, propagation environment changes).

Normally, RF domain control or optimization is mainly considered in the network planning stage using radio link budget calculations to guarantee the ability of the network to ensure the availability of the services over the entire service area [19]. High spectrum efficiency will require good engineering of the cells by proper choices for site locations, antenna beam width, tilt angles, orientation, etc. However, sometimes it is impossible to acquire the best choice and the static control limits the ability of the network to cope with the variable wireless environment. Due to the large QoS requirement differences in the resulting radio link budgets, uniform coverage and capacity designs as estimated for the conventional voice-only radio networks, become more and more limited in their ability to provide an efficient cellular network. The cooperative control for geographic load balancing borrows some ideas from both dynamic radio resource management (RRM) and smart antennas. It is an approach for traffic load balancing that provides flexible radio coverage according to the current geographic traffic conditions. In the network planning stage, the capacity and coverage are coupled together for the local area. The radio resource allocation is optimized by traffic load demand redistribution for each cell with the whole network radio coverage controlled by cooperative changes to the RF domain of the PHY layer in order to achieve cell-level spatial multiplexing. Cooperative control tries to make full use of the total radio resources of the whole network by dynamic reshaping the radio coverage to reallocate the traffic demand between cells. This process needs to be performed cooperatively, as the local cells have very limited capabilities to resolve traffic hot spots independently.

In this work, the cooperative control idea has been extended into a more general cooperative format and applied in RBCNs to obtain system improvement and adaptability. Firstly, the cooperative control process applies not only for the BSs, but the RSs will also be working in close collaboration with each other and the BSs to achieve more efficient radio resource usage. Secondly, the cooperative optimisation not only responds to the traffic load changes but also responds to environment and network changes. The radio coverage pattern will adapt if a BS (or RS) fails, a BS (or RS) moves, MSs move, or simply demand changes. Thirdly, in previous research in this area [17, 18], the semi-smart antenna system was mainly considered as the vehicle to provide the cooperative control and flexible coverage. In this work, simple antenna systems, such as the omni-directional antenna and sector antenna, which are already mainly used in existing cellular networks, are considered.

The objective of this work is to improve the capacity, performance and adaptability of RBCNs by using the variable RF domain control of the PHY layer to mitigate problems of load balancing and to increase the flexibility of coverage so that coverage is provided at the right place at the right time. Dynamically changing the RF configurations also mitigates the problem of static network planning in the context of a variable wireless environment. The aim of the cooperative control process is to make full use of the potential capability of RBCNs. There is no modification required to the wireless standards or protocols when this form of cooperative control is applied. According to different situations (upgrade the existing cellular network with dumb RS or fast deployment with smart RS), several approaches are proposed with the adaptive concept. For practical usage, the radio coverage and handover issues are also analysed to assess the potential risks of the cooperative control when applying changes in a realistic environment. Although large network cooperative optimization is the main objective of this work, local cell optimization is also investigated and the latter focuses on the heterogeneous demand in a single MS. This work investigates the above issues in the content of Mobile WiMAX network standards.

Contributions:

The main objective of this work is to introduce an adaptability concept into RBCNs and investigate the potential adaptability of RBCNs using cooperative control. The contributions of this thesis are summarised as follows:

- An investigation into upgrading existing cellular networks with dumb RS. According to the BS antenna model and adaptability mechanism, two kinds of simple antenna system are considered, namely omni-directional antennas with variable pilot signal power and sector antennas with simple tilting capability. Two system wide cooperative algorithms are proposed to perform geographical load balancing between network components. Perceived traffic load is used to eliminate the requirement to exchange traffic condition information and reduce communication overheads and system complexity. The cooperative pilot power approach extends the "cell breathing" idea to a more cooperative way and a large scale network optimization. Despite its simplicity, limited improvement was shown for the case of cooperatively adjusted variable power using an omni-directional antenna. The cooperatively tilting approach shows a marked improvement in the system capacity, especially in hot-spot areas.

- A novel fast deployable RS approach is proposed with phased array antenna and passive direction of arrival (DoA). The fast deployment and adaptive routing ideas are taken from the ad hoc network. With the adaptive antenna and cooperative algorithms, RSs could be fast deployed with plug and play convenience and RBCNs can be more efficient, more tolerant of environment changes and discover its local area propagation environment. The hope is that this can reduce the cost and even the need for detailed network planning. Such an approach could be incrementally deployed into existing cellular networks. It is inexpensive as the RS does not need a wired backhaul and there is no expensive antenna system requirement. It is scalable because it retains a cellular control structure and has a simple routing path to the BS. The objective is not only to take advantage of geographic load balancing to maximize the efficiency of radio resource usage, but also to provide adaptability and hence resilience to faults and attacks. This approach is very suitable for scenarios that need quick deployment and have high data rate requirements.

- To support the flexible radio coverage feature, a radio coverage prediction model is proposed. Linear and logistic regression methods are used to learn coverage prediction models using survey data or data acquired during network operation with cooperative control helps. A key point is that the model can be learned through a bootstrapping process. Such a prediction model is important for intelligent radio resource management as it allows more accurate hypothetical reasoning and hence the discovery of optimal solutions.

- From the network level aspect, a detailed analysis of the effects on handovers when control actions are applied. To mitigate the burst handovers that would occur if handovers were applied when the antenna configuration changes, a scheduling scheme is proposed. Based on the Mobile WiMAX standards, three different handover algorithms are implemented and analysed in the context of cooperative control in the RBCNs.

- An adaptive radio resource allocation with serving node selection is proposed to mitigate the different QoS requirements and traffic demands for single users in a local cell area. The scalable solution illustrates an effective and efficient approach for QoS provision in RBCNs. Important concerns are: (a) the selection of the node to service the mobile user with a QoS that guarantees the time delay requirement for multimedia services and eliminates bottleneck links; and (b) the choice of the optimal slot allocation scheme that optimizes the radio resource allocation to maximize user satisfaction levels and balances the QoS requirements.

The author's publications are listed in Appendix A.

The thesis is organised as follows:

Chapter 2 presents the background of the thesis. It reviews the literature in the relay based cellular networks domain and gives a brief overview of radio network planning, radio resource management and Mobile WiMAX networks. More details are given in later chapters where appropriate.

A detailed description of cooperative control is presented in Chapter 3. A comparison is made with other adaptive methods and its distinctive features in this thesis are discussed. To evaluate cooperative control strategies, a comprehensive Mobile WiMAX network simulator is built to credibly validate the proposed approaches.

In chapter 4, the approach for upgrading existing cellular networks with dumb RS is investigated. Two cooperative algorithms are introduced, which are based on an omni-directional antenna with variable pilot signal power and a sector antenna with real-time tilting ability. The performance of the algorithms is compared with a "conventional" network, i.e. one with no optimization applied.

In chapter 5, a rapidly self-deployable and reconfigurable approach using a phased array antenna system is introduced for the relay based cellular network. Passive direction finding is used in this approach. The re-configurability of the BS assignment makes the network more flexible than one where a RS is statically assigned to one BS. Simulations are also performed to evaluate this approach, and results are presented.

In chapter 6, the radio coverage and handover issues are analysed with respect to the potential risks of using cooperative control in a real environment. A statistical estimation approach for radio coverage prediction and a scheduling scheme for managing handover demand mitigation are proposed. Different handover algorithms are also investigated for their behaviour in the context of cooperative control.

In chapter 7, an adaptive radio resource algorithm is evaluated for relay based cellular networks that address MS assignment concerns. It uses a global optimization method to allocate the radio resource to balance the users' QoS requirements and maximize the radio resource efficiency.

Finally, the last chapter concludes the thesis and some suggestions are made as to how the work could be extended.

Chapter 2 BACKGROUND

2.1 Introduction

In this chapter, background on relevant research fields is introduced. Topics covered include relay based cellular networks, radio network planning, radio resource management and Mobile WiMAX networks. Only brief reviews are given here. Details of each technology particularly related to the techniques developed are described in the relevant chapters. Section 2.2 gives an introduction to the basic concepts of cellular networks and ad hoc networks and reviews literature in the relay based cellular networks domain. In section 2.3 and section 2.4, some essential principles of radio network planning and resource management are reviewed. Since cooperative control research is demonstrated in the context of a Mobile WiMAX network, section 2.5 gives a brief introduction of Mobile WiMAX technology. Finally, section 2.6 summarizes this chapter.

2.2 Relay Based Cellular Networks

2.2.1 Cellular Networks and Ad Hoc Networks

Cellular networks were first introduced in the 1980s. The motivation came from the growing demand for wireless communications over a wide area, together with limited frequency availability [20]. The main aim was to provide an efficient use of spectrum and at the same time expand the coverage area by allowing spectrum to be reused. They were also designed to handle mobility, so that users were able to roam from cell to cell and connections were handed over to maintain the communication.

The key cellular concept is radio resource reuse where the same set of channels can be reused in different geographical locations as long as they are sufficiently far apart from each other so that interference is within tolerable limits. So, advanced network planning is very important for the performance of a cellular network.

An ad hoc network is defined as the category of wireless networks that utilize multi-hop radio relaying and are capable of operating without the support of any fixed infrastructure [21]. An ad hoc network is formed to provide an alternative communication infrastructure for mobile or fixed nodes and users, without the spectrum reuse constraints and requirements of network planning of cellular networks. The topology of ad hoc networks provides many alternative paths for a data transfer session between a source and destination, resulting in quick reconfiguration of the path when the existing path fails due to node failures. The absence of any central coordinator or BS makes the routing a complex issue compared to cellular networks.

The difference between cellular networks and ad hoc wireless networks are listed below:

- Cellular networks use a fixed infrastructure. There is one hop link (BS to MS) in the communication. It is designed for voice traffic, and it can guarantee bandwidth and has seamless connectivity. There is a high cost of deployment and network maintenance. The major goals of routing (here call admission) are to maximize the call acceptance ratio and minimize the call blocking ratio. It is easy to achieve time synchronization.
- Ad hoc networks do not use any fixed infrastructure. An ad hoc network can have multiple hops in the wireless links. It is based on packed-switching and shared radio channels, more suitable for best-effort data traffic. Self-organization and maintenance properties result in quick and cost-effective deployment. The main aim of routing is to find paths with minimum over head [22] and also quick reconfiguration of broken paths [23]. Time synchronization is difficult and consumes bandwidth [21].

2.2.2 Hybrid Networks

Multi-hop wireless networks have traditionally been studied in the context of ad hoc and peer-to-peer networks. In recent years, there has been an upsurge of interest in the integration of multi-hop capability into conventional cellular networks—Hybrid wireless networks.

In hybrid wireless networks, several architectures such as multi-hop cellular networks (MCNs) [7, 10, 24] and integrated cellular ad hoc relay (iCAR) networks [5, 6] have been proposed. The use of multi-hop connections, in which a station situated between MSs and BSs perform hopping relay of the transmission, can expand the coverage area and reduce the transmit power required per hop. The capacity of a cellular network can be increased if the network incorporates multi-hop relaying. The MCNs or iCAR networks can provide high capacity, resulting in lowering the cost of communication to less than that of single hop cellular networks [5, 7]. The base station maintains the information about the topology of the network for efficient routing. In MCNs and iCAR, the base station may not be involved in a multi-hop path if the communicated MSs are near to each other. As the mobile users perform the relay function in these architectures, the main problem arises if there are few MSs. In this case an efficient routing path may not be found and there can be a low signal to interference plus noise ratio (SINR) around the boundary area.

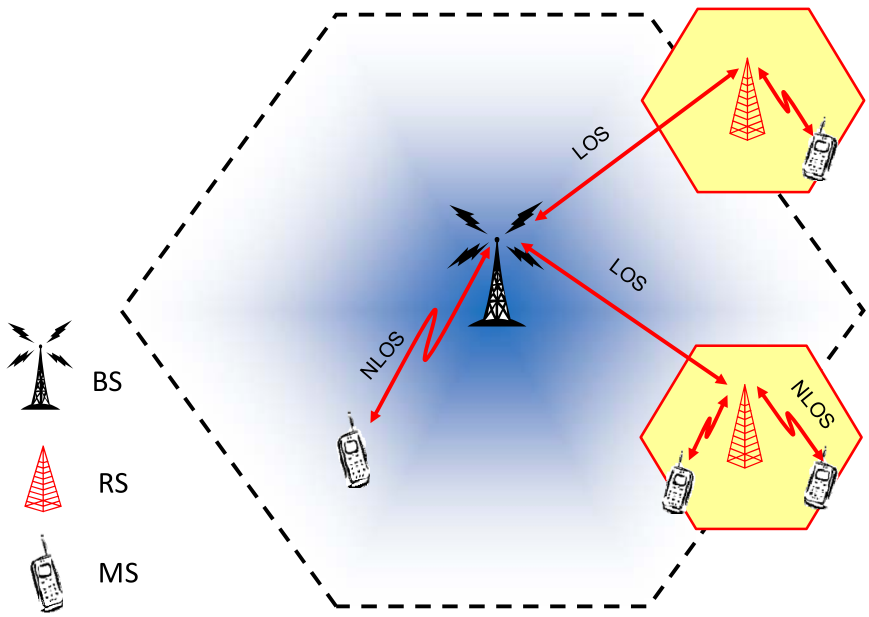


Figure 2.1 The structure of a relay based cellular network

In this thesis, the relay based cellular network architecture puts the fixed relay stations near the cell boundary and limits the number of hops in a connection to two, as shown in Figure 2.1. So, from the transmitter to receiver, there is no more than one relay station. In every communication, a BS must be involved. Hence two types of connection is established in such a relay based cellular network and these are called single-hop and two-hop connections. Using a single hop connection, a MS is directly connected to a BS. A two hop connection is defined as a connection of a MS to a BS via a RS. Especially in shadowed areas, a two hop connection may yield a higher throughput between a BS and a MS than a single hop connection. A MS at the cell border can use a two hop connection with a bandwidth efficient modulation and coding scheme.

The fixed relay station (FRS) is chosen in the proposed relay based cellular network as it has some advantages. For example, a fixed relay can handle a large number of transmissions, and does not need to consider the increased energy consumption (it is assumed that a FRS always uses a constant power supply). They can be deployed at strategic locations that have good links between themselves and the BS. This will reduce the propagation losses between the relaying station and user terminal and thus reduce the required number of expensive cell sites. Consequently, the use of FRSs can achieve higher link data rates and potentially increase the high data rate coverage in a larger cell [8, 25].

2.2.3 Relay Station in Relay Based Cellular Network

With the help of OFDM technology and adaptive modulation and coding, the relay station has been seen as an effective way to improve the system capacity of the wireless network. The better the signal quality the higher the modulation level that can be used for the communication channel and this can result in maximizing system capacity. To clearly distinguish the relay station concept from those of a repeater, Pico-Cell and FemtoCell, a comparison is given below.

RS Versus Repeater

The relay station in this thesis is based on the concept defined in IEEE 802.16j [26] where the main functions are 1) to relay user data and possibly control information between BSs and MSs (or other RSs), and 2) to execute processes that indirectly support mobile multi-hop relays. For the main functions, the RS is very similar to a repeater. They both relay the signal from the BS to the MS or vice versa. However, a RS is more complex than the repeater and has the additional functions listed below [27]:

- Flexible radio resource assignment;
- Scheduling;
- Bandwidth request and re-allocation;
- QoS support;
- MS handover support.

A repeater is essentially a low noise amplifier. By placing the repeater in a coverage hole, the overall signal in both directions will become magnified, allowing better quality service. However, anytime you add a repeater device in your network, you risk adding interferers [28]. These repeaters are “dumb” in the sense that they blindly retransmit all the signal received from the base station. When you use a repeater, it increases not only your desired signals, but the noise floor as well. Also, because repeaters simply expand the coverage area of an already existing sector, they do not improve the system capacity.

RS Versus Pico-Cell

To overcome a small coverage hole or shadowing in a network, many manufacturers and operators now have Pico-cells [29]. The Pico-cell BS is similar to a small BS, which has nearly all the capability of the normal BS but a smaller coverage. The RS is similar to the Pico-cell with small coverage and limited capability, communication directly to the BS with a wireless link to the BS (wireless backbone).

The main advantage over Pico-cellular BSs is that the RSs do not need a wired network connection, which is the determining cost factor. The other advantages such as the lower site acquisition costs and less costly antenna structures are very attractive in their own right but they also mean that a RS can be deployed faster.

RS Versus FemtoCell

In recent years, the FemtoCell has attracted a lot interest from research and industry as a means to provide sufficient performance in indoor environments. A FemtoCell is defined as a low power wireless access point that operates in licensed spectrum to connect standard mobile devices to a mobile operator’s network using residential DSL or cable broadband connections [30, 31]. The FemtoCell uses an existing internet connection, and this can obviously decrease the cost of installation and deployment. Because of the required wired backhaul connection, the FemtoCell is limited to the home or office environment and competes with WiFi networks.

2.3 Radio Network Planning

The radio network is part of the wireless network, which includes the BSs, MSs and the interface between them. The result of radio network planning will affect the performance of wireless networks. A basic network planning process is shown in Figure 2.2.

The main objective of radio network planning is to provide a cost-effective solution for the radio network in terms of coverage, capacity and quality [32]. Because of the complexity of the real environment and different wireless technologies used, the network planning process and design criteria vary from region to region depending upon the dominating factors and network operators' requirements.

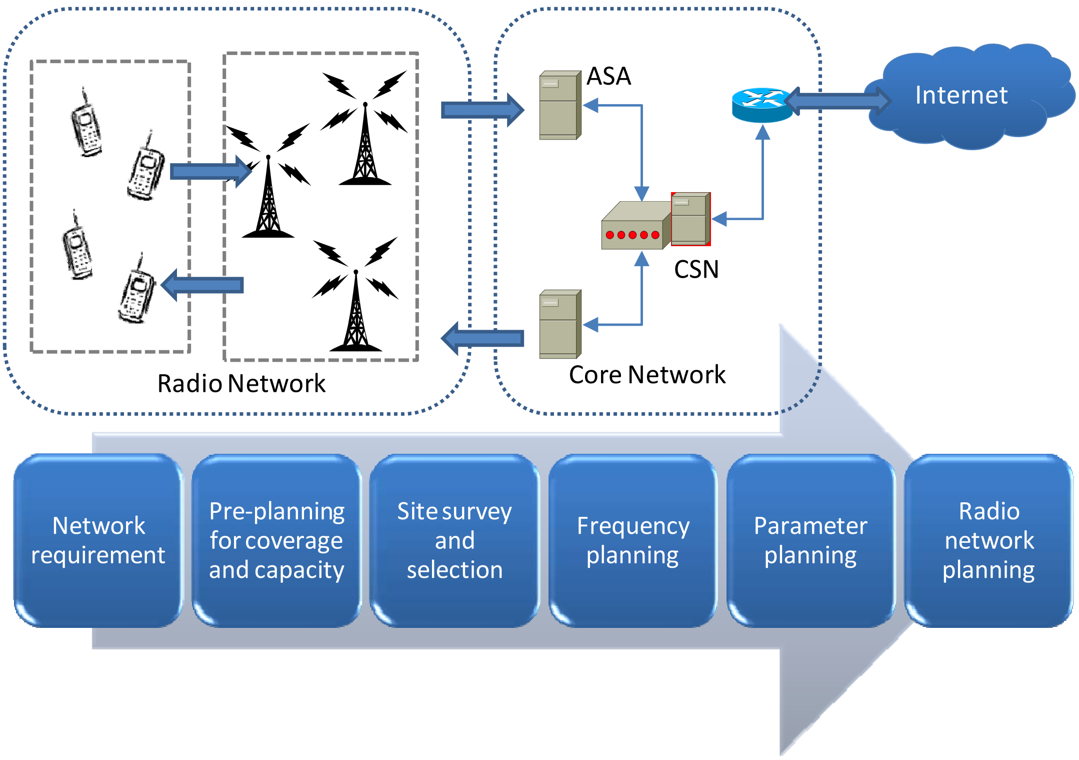


Figure 2.2 The radio network and network planning processes

The network requirement process in Figure 2.2 makes an assessment over the planning area of potential traffic demands, service type provision, etc. The potential traffic demand is dependent on the user communication rate and user movement in the planning area. After the pre-planning for coverage and capacity according to the network requirement, specific areas are identified for prospective sites. The parameter planning pre-defines and optimizes the parameters to provide the seamless communication and minimize interfere. These parameters include those related to signalling, radio resource management, handover, etc.

One objective of the research reported in this thesis is to adapt the radio coverage to achieve traffic load balancing, so the radio coverage and capacity planning part are given more focus here. Radio coverage planning includes defining the coverage areas, service probability and related signal strength [19]. The antenna system used by the cell and local propagation environment are also considered at this stage in radio coverage planning. Radio coverage will be discussed in further in chapter 6 where a radio coverage prediction approach is described.

A definition of capacity must include the subscriber and traffic profile in the region and whole area, the availability of the frequency bands, and frequency planning methods [33]. Capacity planning should not aim only to meet current demands and state, but the solution should also aim to comply with future requirements by providing an acceptable development path. So, a good network plan will consider and optimize network capacity and growth, architect network resilience and survivability [34].

The provision of high data rate services and multiple services in 3G and future wireless networks means that network planning and optimization become more important and difficult than ever. The optimisation will be a capacity–quality trade-off instead of a plain quality improvement process [34].

2.4 Radio Resource Management

The role of RRM is to provide QoS guarantees to mobile users according to their bandwidth requirements while maintaining the high utilization of network resources [35]. The objective of a RRM scheme is to assign a cell, a channel and a pair of transmitter powers to as many communication links as possible at any moment of time subject to the constraint that the quality of all links should be above a certain threshold. The RRM process for an MS starts when the MS is trying to make a phone call and ends when the call is finished or dropped. This is shown in Figure 2.3. RRM is important from the perspective of air-interface resource utilisation, offering ways to achieve optimum coverage and capacity for a quality of service guarantee.

The RRM processes in a wireless network could be divided into two levels [36], the macro-level and micro-level. The macro-level is to control the connection access and QoS requirement of the MSs, which involves connection admission control (CAC), radio resource allocation (RRA), etc. The micro-level is to guarantee the QoS of MSs when connected, and includes power control and package scheduling. This thesis focuses on the global optimization for radio resource optimization with RF domain of PHY layer control, so the macro-level resource management is more relevant and so a detailed description is given below.

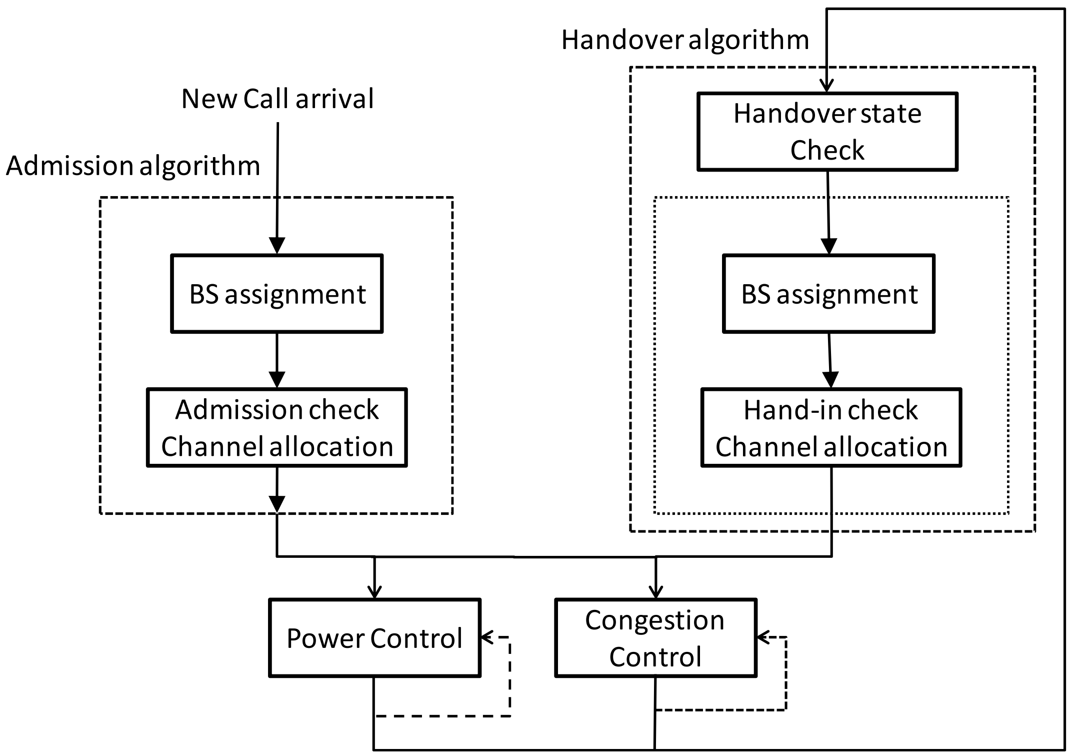


Figure 2.3 The flowchart and processes in the radio resource management

CAC is to provide QoS guarantees for the MSs that request admission to access the network while efficiently utilizing radio resource of the local cell [37, 38]. Based on the QoS requirement of MSs and available radio resources, CAC decides on acceptance or rejection of the new connection and handover requests. Rejecting a new call request leads to call blocking at service initiation and rejecting a handover call request leads to call dropping in the middle of service. If the connection request is accepted, radio resource allocation will allocate the radio resource according to the QoS required. However, if more MSs request access or handover for the local cell, this local cell starts to become congested, and congestion control action needs to be taken to mitigate the effects. Some existing methods for congestion control have been given in [14, 39]. The research described in this thesis could be thought of as a form of congestion control, but using an entirely different approach: adaptively changing the antenna configuration of the antenna system to optimise the radio coverage (and hence received power) in order to minimise the traffic congestion.

2.5 Mobile WiMAX Network

Broadband wireless access (BWA) has become the best way to meet escalating business demand for rapid internet connection and to integrate data, voice and video services. The Institute of Electrical and Electronics Engineers Standards Association (IEEE-SA) published the first version of its IEEE Standard 802.16 in December 2001 [40].

WiMAX (Worldwide Interoperability for Microwave Access), is based on the IEEE 802.16 standards. It is a wireless technology and designed from the ground up to provide wireless last-mile broadband access in a Metropolitan Area Network (MAN) [1]. The Mobile WiMAX network is based on the IEEE802.16e standard, which was published in February 2006 [41]. The Mobile WiMAX air interface adopts Orthogonal Frequency Division Multiple Access (OFDMA) for improved multi-path performance in non-line-of-sight (NLOS) environments. Adaptive modulation and coding (AMC), Hybrid Automatic Repeat Request (HARQ), Fast Channel Feedback, and a full range of smart antenna support were introduced with Mobile WiMAX to enhance coverage and capacity for Mobile WiMAX networks.

For enhancement of coverage, throughput and system capacity in Mobile WiMAX networks, the Relay Task Group (802.16j) approved in March 2006 [26] the development of a Multihop Relay Specification. Since a simulation network is used to validate this thesis is based on a Mobile WiMAX network, more details about the Mobile WiMAX will be described in chapter 3. More information for the Mobile WiMAX and IEEE 802.16e standard can be found in [40-43].

2.6 Summary

In this chapter, some essential principles and advantages of RBCNs are briefly explained. To distinguish the concepts of RS, repeater, pico-cell BS and FemtoCell, a brief comparison is also given. Following this, the basic goals of network planning and radio resource management are described. Finally, as the simulation network used in this thesis was based on the Mobile WiMAX technology, Mobile WiMAX network is introduced.

Chapter 3 MOBILE WIMAX SYSTEM LEVEL SIMULATION FOR COOPERATIVE CONTROL

3.1 Introduction

In the previous chapter, the essential principles of radio network planning and radio resource management were briefly described. Cooperative control is seen as a flexible and economical solution for geographic load balancing to make full use of the limited spectrum resource and mitigate the problem of static network planning. In the first part of this chapter, a detailed explanation of cooperative control and discussion of other adaptive methods are given. Following this, the distinct features of cooperative control in this thesis are explained. A key point of the cooperative control research is to develop optimisation algorithms for cooperatively adjusting the physical antenna coverage patterns. As described in chapter two, Mobile WiMAX is a broadband wireless solution and introduces relay stations to enhance the flexibility of the network architecture. In order to test the performance of the optimization algorithms in the context of Mobile WiMAX networks for non-uniform traffic distributions, a comprehensive WiMAX network simulator has been developed by the author. Its features include connection admission control, power control, and handover management. In the last two sections of the chapter, there is a detailed explanation of the architecture of the system level simulation tool and how it works. To validate the simulator, the simulation result is compared with data published by the WiMAX forum [1] using the same network parameters.

3.2 Cooperative Control

Larger demands of heterogeneous wireless services drive current and future radio technologies to increase the cellular network capacity. The traditional way to increase network capacity is to increase the number of carrier frequencies per cell. While this is the most straightforward method it is clearly limited by the scarce spectrum resource. Another way is to use the *cell splitting* technique [44], which decreases the size of the cells by splitting and inserting new transmitting sites in the network. This is quite costly, and when we consider the backhaul and the cost of hardware and maintenance, it usually does not make sense unless the network operator can be sure that this area will continue to have a high level of demand. The challenge involves meeting an evolving and generally increasing demand that may vary dramatically between different geographic areas of the network.

In recent research, dynamic radio resource management is seen as a cost effective and flexible way to optimized spectrum utilization and improve the system capacity. Dynamic radio resource management mitigates the problem by employing adaptive approaches in channel assignment, re-allocation and sharing aspects of wireless communication, and typically takes one of two approaches [45]:

- Capacity Adaptation: An overloaded cell can try to increase its own capacity by borrowing capacity from neighbouring cells. Examples are channel borrowing [11, 46], channel sharing [12, 47], dynamic channel allocation [13, 14, 48-50].
- Load Adaptation: An overloaded cell can try to reduce its own load by forcing or directing some or all of its associated wireless devices to switch to alternative neighbouring cells. Examples are cell breathing [51] and soft handover schemes [16, 52, 53].

Most of the literature in load-balancing schemes has focused on the capacity adaptation solution: an overloaded cell borrows excess capacity from neighbouring underutilized cells. When overloaded, a cell borrows idle communication channels (essentially additional capacity) from neighbouring cells. The cell coverage, however, remains unchanged and the co-channel interference may be increased and leave a hotspot area in a worse situation. Such schemes also work only when an individual cell has expensive hardware and specialized software to support multiple simultaneous channels.

Recent studies also show the potential opportunity for application of smart antennas [54] to achieve load balancing, by employing spatial division multiple access (SDMA) schemes. They have greater compatibility with network evolution as the use of RF to access mobile networks is unlikely to change. However, most work on smart antennas only considers the radio propagation channels within one cell and these severely limit their efficiency. With the multi-hop feature, in [55], [56] and [57], the authors present load balancing algorithms for efficient routing in multi-hop wireless access networks. However, the route balancing is limited by the traffic distribution and if there are few MSs, it may not find a routing path. The battery consumption and complicated mobile relay communication are also constraints for the routing balancing schemes.

Cooperative control for geographic load balancing is a kind of load adaption solution, which borrows some ideas from both dynamic radio resource management and smart antennas. It is a control mechanism to apply real time adaptive control at the RF domain of PHY layer (such as the radiation pattern, transmit power or tilting angle) according to the traffic demand and adjusted to support load balancing over the network. A central feature of the mechanism is the development of algorithms to coordinate the radiation patterns of an antenna that can learn from the environment during operation, so that the radiation patterns generated provide coverage where and when it is needed. This is could be thought as a novel way for network optimization that adapts the RF domain parameters of the PHY layer with self-configuration features for each local cell. The cooperative optimization is with respect to the whole network rather than optimization based on the local on-site cell. (However, the cooperative optimization algorithm can be distributed so the cells can make their decisions autonomously. A distribution approach based on colour labelling is proposed in [58]).

The main objective of cooperative control is to adaptively control the RF domain to achieve radio network performance optimization, and is divided into two parts: one for spectrum utilization optimization with geographic load balancing; and the other for real time RF domain optimization of the radio network.

Geographic load balancing

To achieve the geographic load balancing from the network level point of view, cells (sectors) with an excess of demand will let MSs that are near the boundary handover from the higher traffic demand cell to a neighbouring fewer traffic demand cell (sector). To support the collaboration feature, the antenna configuration responds to the handover processes to support the link setup. Several algorithms, such as [53] and [12], also propose that the BS directs the MSs to handover. They make use of the inevitable overlap area that exists among the cells, and direct calls from one cell to some of the neighbouring cells. The retry [53] and load sharing scheme [12] balances the traffic, but the optimization is limited because the traffic in the overlap area and in a high traffic load cell still may have a high blocking rate as the cell coverage is unchanged. The traffic balancing is also only limited to the neighbouring cells rather than providing a network level optimization. The difference between the above schemes and the cooperative control described in this thesis is that the cooperative control in this thesis uses the RF domain to optimize the radio coverage and the handovers of MSs are the network level consequence of the change in radio coverage. Such cooperative control could improve the signal quality of the whole network whatever the traffic distribution. Cooperative control does not need to modify the handover algorithms. This means the cooperative control is handover algorithm independent, although different handover algorithms will have different performance, which will be investigated in the chapter 6.

In the cooperative process, adjacent cells will change their radio coverage and partially focus their radiated power onto the hotspot, e.g. one might cover more and another back off. More capacity is indirectly given to the hotspot by the congested cell through reducing its coverage. Meanwhile other neighbouring base stations expand their coverage to fill any gap left by the changes thus lending some capacity to the hotspot. The dynamic coverage approach delivers optimized network capacity through traffic load balancing. This is the basic benefit from adaptability of the network.

Real time RF domain optimization

In the network planning stage, a network covering a big region will require analysis for each part of the region separately, as propagation conditions (apart from topography) will change from one region to another. This could lead to a long period of analysis and the analysis may only consider the static situation. Although BS site optimization takes a lot of work to ensure signal reception, holes in radio coverage are still inevitable. An on-site engineer making changes to the RF is the normal way to achieve RF domain optimization. By changing antenna angles, adjusting power levels, and changing antenna heights, with the aim of increasing signal quality, an engineer can limit the amount of dropped and poor quality calls [32]. However, this method has no real time capability. It is typically carried out in off-peak hours and only when poor performance has already happened. Flexible radio coverage could mitigate the cost and time to deployment of a network, by changing cell shape to adapt to perceived traffic distribution and propagation environment fluctuations. According to the perceived performance of a network, cooperative control could optimize the RF domain in "near" real time.

3.3 System Level Simulation for Cooperative Control

3.3.1 The Architecture of the Simulation Tool

To evaluate performance capabilities of optimization algorithms for Mobile WiMAX technology in the presence of non-uniform traffic distributions, a comprehensive WiMAX network simulator has been developed by the author. The simulator is mainly based on the OFDMA–TDD (time division duplex) system according to the IEEE802.16e [41] and Mobile WiMAX [1, 59]. In this section, the basic simulator architecture is described. Additional information related with the Mobile WiMAX standards will be described in section 3.4.2.

There are several reasons for the system level simulation used in this thesis:

- Firstly, system level simulation is used for radio network optimization purposes, i.e. changing certain system parameters in order to reach an optimal configuration for the given network. This is very suitable for cooperative control research with variable antenna configuration settings and considering the impact of different policies over the whole network.

- Secondly, system level simulation can be used to evaluate the performance of a specific network by collecting QoS measures of all MSs in the network. In this thesis, the perceived QoS parameters for end users are used for the QoS measures, such as: call blocking rate, call dropping rate and system throughput. It offers much insight into the overall system-level performance of a wireless network as a whole.
- Multiple cells and a large number of mobile users were included in the system level simulations in contrast to link level simulations, which commonly only evaluate the performance of signal transmissions. A combination of the link level and system level would be a feasible option, but the complexity of such simulator (including everything from the waveforms to a network with many cells and thousands of mobile users) was considered too high. Link level or package level simulation was considered over-complex to obtain simulation results for the heterogeneous traffic demand, where the traffic demand changes over time and geographic location.

This system level simulation uses Monte Carlo techniques [60]. The Monte Carlo method is commonly used to sample from different probability distributions in order to simulate a desired system and makes extensive use of random number generators to generate the samples. Stochastic models are useful for obtaining numerical solutions to problems that are too complicated to solve analytically [61] such as heterogeneous traffic distributions. The traffic scenario models used will be explained in section 3.3.3.

This simulation system needs to be capable of simulating Mobile WiMAX networks with unevenly distributed traffic, with different antenna systems. To evaluate the variable traffic distribution situation, the discrete event approach is used. Within the context of discrete-event simulation, an event is defined as an incident that causes the system to change its state in some way, e.g. a call request, a handoff request, a movement of a MS, and the chronological sequence of events represent the dynamics of the system [62]. The discrete events in this simulation are defined by the traffic distribution changes. A traffic snapshot is used to represent the traffic distribution instance for a traffic scenario in this thesis. The sequence of traffic snapshots simulate a dynamic traffic scenario, where the traffic distributions vary from uniform to heterogeneous. Typically, in a sequence of snapshots the traffic is initially spread uniformly over the geographical area, and as time progresses clusters of high demand traffic are formed by the active moving and active stationary MS.

The system level simulator contains three main components: the network initialiser, the network optimiser and the network simulator. The simulation is performed iteratively to calculate both uplink and downlink system capacity for the traffic configuration in each snapshot, as illustrated in Figure 3.1.

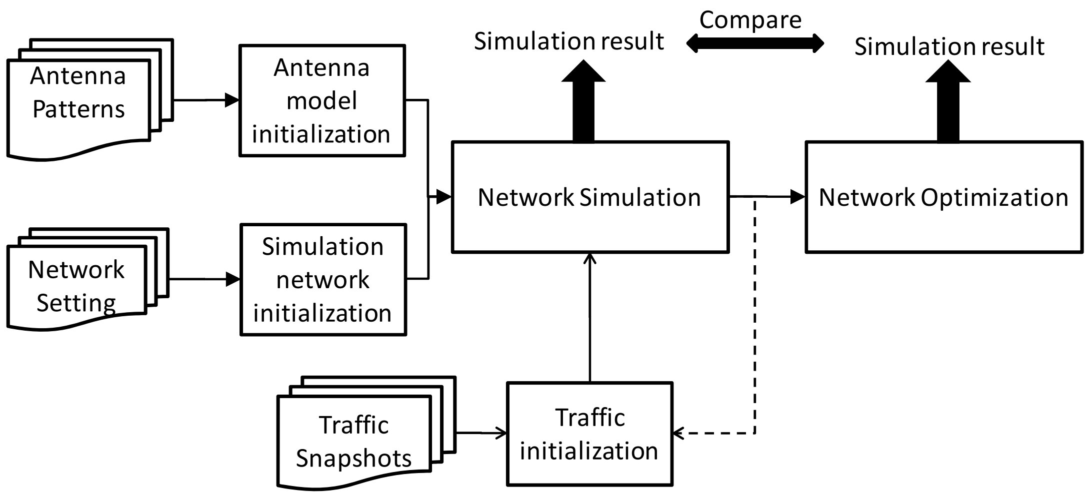


Figure 3.1 The simulator architecture

At first, the simulation network needs to be initialised with a network topology (the locations of BSs and RSs) and antenna model (the horizontal and vertical radiation pattern). According to the traffic scenario setting, such as hotspot locations, number of hotspots, population in hotspot area, etc, a set of traffic snapshots will be built and stored in text files, one snapshot per file.

After the network initialization, the Mobile WiMAX simulation will run based on the snapshots created for the traffic scenario. For each traffic snapshot, the network simulation evaluates the QoS performance of the simulation network. Experiments using the chosen optimization algorithm are then performed for the network under the same traffic conditions as for the conventional network. The conventional network simulation is normally a simulation with conventional usage of antennas and relays. All the other parameters are kept the same for the optimizing network simulation and the conventional network simulation so that the improvements in system capacity and other statistics can be fairly addressed.

3.3.2 The Processes of the Network Simulator

Mobile WiMAX simulations are first performed for a conventional network simulation just after the network initialisation. This simulation performs a system level simulation for the Mobile WiMAX network. The Mobile WiMAX simulation includes two parts: uplink and downlink. In the uplink and downlink, the following processes are done: connection admission control, handover, power control and adaptive modulation and coding. The flowchart of the simulation processes is illustrated in Figure 3.2 and the detail of implementation is explained in section 3.4.2.

Compared to the processes of radio network management (described in the Chapter 2), the simulator evaluates the probability of providing service in the simulation network. For each active MS, the simulation processes illustrated in the figure are applied and a decision to accept or block (or drop) this MS is made. The simulator performs a sequence of distinct discrete event simulations centred around the configuration provided by snapshot files. The simulations are used to determine the performance of the system for the network state indicated by the snapshot. In each simulation, as time passes, the traffic distributions change continuously from the beginning of each traffic snapshot. Each traffic snapshot is the basis of a period of discrete event simulation, starting at the time of the snapshot. The duration of the window is chosen so as to be able to collect a reliable estimate of the network performance for the demand at the snapshot in question. There are 100 replications of the discrete event simulation process at each traffic snapshot. After the discrete event simulations the simulator advances to the next snapshot and the discrete event simulations at the new snapshot is started. For each discrete event, the processes of simulation described above are performed for each active MS. For each snapshot, multiple discrete event simulations are run for the active MSs in the simulation network. The duration of each event simulation is 600 million seconds (this could be changed). One hundred replications of the discrete event simulations are made and the statistics on call lost rate and errors are computed from the results. 100 replications are chosen as this ensures a small standard error of the statistics computed.

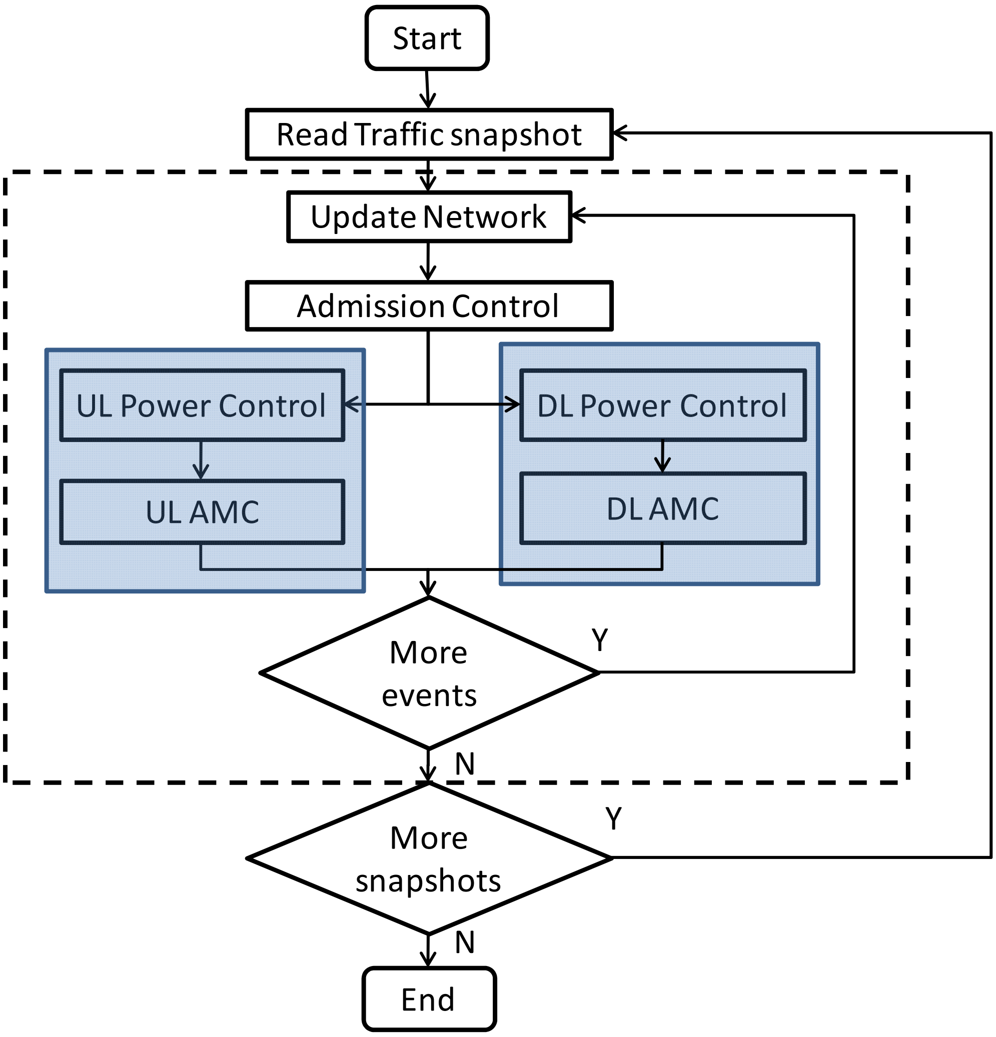


Figure 3.2 The flowchart of simulation processes

3.3.3 Traffic Model

The voice call is by far the most popular and most commonly used service in mobile network. In the multi-cell simulations, real time voice traffic is assumed for all the mobile users. It has fixed bandwidth demand for the uplink and downlink and its continuous characteristic can well represent the state of channel when cooperative control is applied. The proposed optimization algorithms for cooperative control did not depend on the traffic types, and the different services have been investigated in the Chapter 7. The common model for a single voice source uses an ON-OFF process [63], which also applied in this thesis. The traffic of each user is randomly generated using an ON/OFF model with the interval between connection requests following a negative exponential distribution. The formula used to generate samples from the exponential distribution is shown in equation (3.1), which is obtained by inverting the cumulative probability distribution.

$$Y = \frac{-1}{\lambda} \cdot \ln(X) \quad \lambda > 0 \; X \in [0,1.0] \tag{3.1}$$

where X is generated by sampling from the uniform distribution using a pseudo-random number generator. λ is the rate of call requests. Y is the time to the next connection request.

The random number generator (RNG) algorithm used in the simulation was developed by Jacob S. Rosenberg from the University of California, Berkeley. More detail about this algorithm could be found in the book [64] and GUN project website [65]. The RNG used is the random function from BSD UNIX systems [66]. This BSD RNG is a combined multiple recursive generator, which is used for the cryptography of the BSD system [67]. The period if this BSD RNG is $16^{(232-1)}$, which is enough for the simulations, and the statistical features of it are also good for Monte Carlo simulations [60].

The Mobile WiMAX simulator implements the RRM functions and captures the dynamic end-to-end behaviour of the overall simulation network. The RRM mechanisms are responsible for ensuring that a certain planned coverage for each service and QoS of connection will be provided and keep the blocking as low as specified. The traffic scenario is used to represent the traffic demand changing over geographic location and time. The simulator creates a specific scenario for evaluating capacity and coverage, and the cooperative control algorithms optimise the radio resources usage for this specific scenario.

Different traffic scenarios can be used in the simulations. Each scenario is like a story. Typically, in a scenario initially the traffic units are uniformly distributed all over the network and the MSs start to move. Then hotspots are gradually formed by selecting coordinates and making predefined number of users move to approach these, also potentially moving, hot spot coordinates. According to the user movement research in [68, 69], the locations of MSs in the hotspot area are in a

Poisson distribution relative the centre of the hotspots. The MSs not in the hotspot area still uniformly distributed geographically. The formed hotspots usually move in a linear pattern through the network. The scale of each hotspot is adjustable. Generally 100 traffic snapshot files are generated for each scenario. The snapshot files store the coordinates, the traffic type and the demand of each mobile user. Each file is created by the scenario builder and describes the geographic distribution of traffic units and hotspots of a scenario. Although the number of traffic snapshots for each scenario could be more (or less), this only modifies the length of simulation story. The interval between each scenario is usually notionally 60 seconds, so in total one scenario and hence simulation corresponds a 100-minute long scene. However, a snapshot could be generated for e.g. every second.

The traffic scenario files are built using a tool developed by the author, called the scenario builder. This was developed using the C++ language. A traffic scenario is defined by the number of hotspot areas and the number of MSs in each hotspot area. The location of a hotspot can be defined randomly or manually according to the simulation requirements. When the random option is chosen, the (x,y) coordinates are each sampled from the uniform distribution. The size of the network can be specified either in terms of the total number of MSs or the number of cells. Based on the input parameters of a scenario, the coordinates of all the MSs in the simulation network are generated by the scenario builder. The speed of the MSs is 30km/hour and the directions of movement are based on the location of the hotspot area. All MSs are all moving and movement is such that MSs converge onto the nearest hotspot area.

The typical behaviour of hotspots in the experiment to be described is shown in Figure 3.3 (the red points represent the congested MSs): The sketch does not give the exact activity times for hotspots. It is presented simply to give a better understanding of the evolution of the simulation scenarios used. For example, at the start traffic units are all uniformly distributed and they can form hotspots, then the hotspots may move through the network area. Usually in the models built, hotspots attract each other and intensify the traffic demand density of some cell clusters. Traffic units not belonging to any hotspots will also move around in the cells. Traffic units congregate in the hotspot areas while the traffic units in other areas are still uniformly distributed.

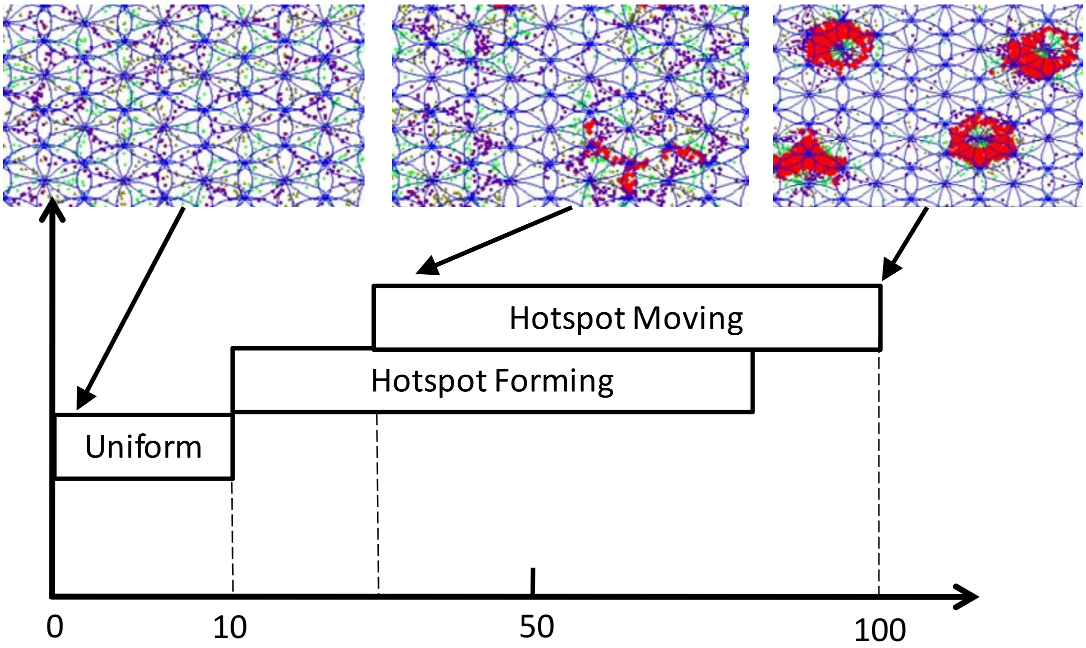


Figure 3.3 Typical hotspots activity

3.4 Mobile WiMAX Simulation and Simulation validation

3.4.1 Mobile WiMAX Simulation

In the simulator, the simulation network deployment is specified in terms of BS and RS placement. The effects of channel variations such as path loss and shadow fading are incorporated into the analysis. The traditional hexagonal cellular model is used as illustrated in Figure 3.4. In each cell, there is a BS in the centre and some fixed relay stations near the cell boundary. BSs are directly connected to their backbone network and fixed RSs are located within BS coverage.

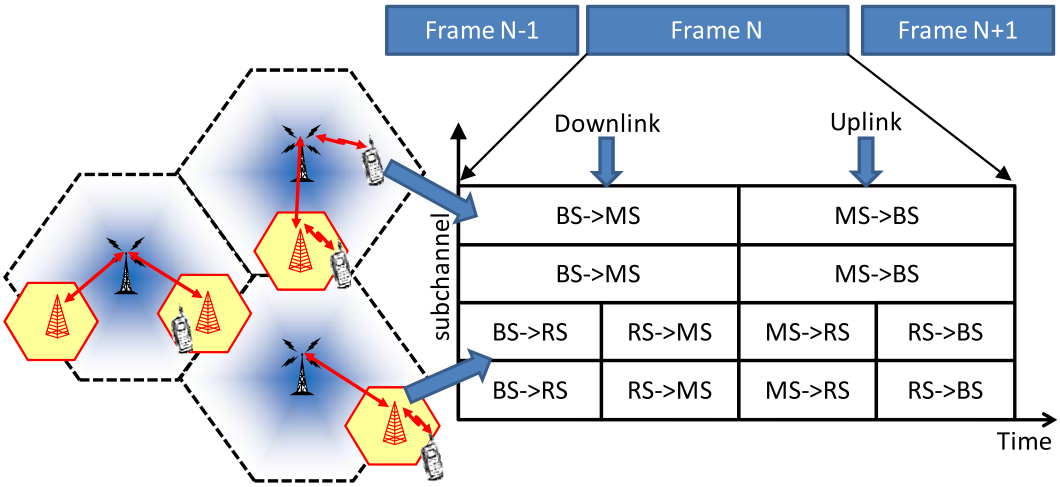


Figure 3.4 The network component topology and channel model of the simulator

Uplink and Downlink

As the initial release of Mobile WiMAX certification profiles only support the TDD operation, the TDD duplex method is used in the simulated Mobile WiMAX. In the case of TDD, the 802.16e standard divides transmission time into super frames and each super frame can be divided into a downlink sub-frame and an uplink sub-frame [41]. The uplink and downlink transmissions share the same frequency but are separated in time. A downlink sub-frame starts with a preamble, which helps MSs perform synchronization and channel estimation. The downlink scheduling is simple because the only sender is a BS. The data packets are broadcast to all MSs and a MS only picks up the packets destined for it. For the uplink, the BS determines the number of time slots that each MS will be allowed to transmit in an uplink sub-frame

However, when RSs are introduced, the normal frame structure could not be used. There are several proposed frame structures in the multi-hop group [26], a simple but sensible one is used as illustrated in Figure 3.4. If the MSs directly link to the BS, the sub-channel and time slot usage are the same as for the normal cellular network. When the MSs indirectly link to the BS by the RS relay, each time slot is divided into two parts. One is for when the BS sends the data to the RS. Within a given frequency channel and antenna sector, all RSs and MSs (directly connected with BSs) receive the same transmission. The other time slot is for when the RS relays the data to the MSs. For the uplink, the MSs and RSs share the uplink to the BS on a demand basis.

Abstract PHY and Channel Model

One of the biggest challenges in system level simulation is to develop a reliable PHY abstraction method that can support intensive system level simulation. The objective of the PHY abstraction method is to accurately predict link layer performance with a computationally simple model, which is critical for system level simulation. The PHY of Mobile WiMAX is based on the OFDMA or scalable OFDMA (S-OFDMA) technology [1]. In the simulator, there is no implementation of the OFDMA, only the essential principals and features of OFDMA are used in the simulator. Following the key OFDMA orthogonality feature, perfect orthogonal sub-channels are assumed and intra-cell interference is ignored. The inter-cell interference is the main interference for the OFDMA system, which is calculated based on [70].

Accurate description of wireless channels is also essential for the evaluation of system performance. In the simulator, perfect channel modelling is assumed. A perfect channel here means that if a mobile user's connection request is accepted, the small-scale behaviour of the system that is affected by instantaneous variations in the channel is not considered. There will be no package or frame loss, no delay will occur except the delay in the RS, and the signal attenuation is only caused by the propagation effects. Only when the transmit power of a MS reaches its limit (through uplink power control) or handover fails (as a result of target cell congestion), will it lose its connection. Because of the simple assumption for the PHY layer, the real time performance of the OFDM signals cannot be included in the simulation model. Hence, this is all about the radio resource usage rather than low level link management.

3.4.2 Simulator Implementation and Parameter Settings

This section presents the details of the Mobile WiMAX simulator and the parameter settings used. Most of the simulator implementation is according to the Mobile WiMAX standard [41], Mobile WiMAX white papers [1, 59] and the documents contributed to IEEE 802.16's Relay Task Group [26] to make sure of an accurate representation of Mobile WiMAX system.

Connection admission control

Mobile WiMAX is a connection-oriented system, and the CAC mechanism deals with the management of a new connection requests. In the proposed relay based cellular network, the admission control module at BSs accepts or rejects the new connection. The RSs only relays the connection request messages to the BS. Because there are two types of links in the communication, the processes of requesting connection and granting bandwidth are implemented separately.

When a MS wants to request a connection, it has two choices: one is to link directly with BS, the other one is to link indirectly with the BS via the RS. The selection is achieved by comparing the pilot signal strength. As the direct link with the BS is like a traditional cellular network, the processing detail will not be described here (more detail can be found in [41]). The process for the indirect link with BS via RS is illustrated in the Figure 3.5.

If the RS has the bigger pilot signal strength, the MS will synchronize with the RS and send the connection request message. The RS will not make any decision on the request and simply relay the message to the BS that is serving it. The BS will make the decision on the request according to the available radio resource and the state of the MS. If the admission control module accepts the new connection, it will notify the uplink packet scheduling component at the BS and provide the token bucket parameters to the traffic policing module at the MS. The BS then broadcasts the UL-MAP to all MSs and RSs in the downlink subframe. As radio resource usage and connection is the focus of the system level simulation, perfect OFDMA scheduling is assumed for the uplink and downlink to guarantee the QoS (more information about the OFDMA scheduling can be found in [71] and [72]).

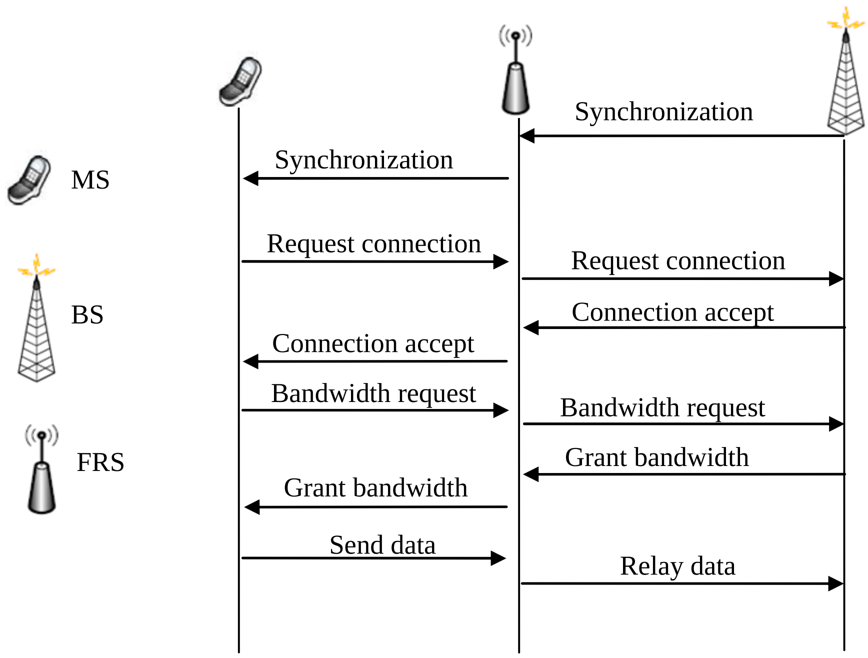


Figure 3.5 The message signalling for the two hop link

When a MS wants to establish a connection with a BS or RS, the CAC module located at the BS determines whether the system accepts the new connection or not. Before the decision, CAC will collect the state information of the system and confirm that the new connection does not degrade the QoS of current connections (this naturally depends on the business policy) and the system can provide the QoS requirements of the new call. So, a connection is admitted if:

- There is enough bandwidth to accommodate the new connection while also keeping enough reserved bandwidth for potential handover users.
- There is enough SNR to support the new connection setup
- The QoS of existing connections is maintained

Suppose there are j^i MSs using the i^{th} classes of service in the coverage of cell m (including both BS and RSs served by the BS), the current radio resource usage for Cell m could be calculated by:

$$C_{total}^{m} = \sum_{i=1}^{I} \sum_{j=1}^{j^i} r(i,j) \tag{3.2}$$

where $\mathrm{r(i, j)}$ is the used bandwidth of the $\mathrm{j^{th}}$ MS for the $\mathrm{i^{th}}$ class of service, i is the number of service classes. When a new connection comes, the following principle should be held:

$$\begin{cases} C_{total,old} + \Delta C < C_{total,Max} \\ SNR_{DL}^{j} > Min_{SNR_{DL}} \forall j \\ SNR_{UL}^{j} > Min\_SNR_{UL} \forall j \end{cases} \quad (3.3)$$

where $C_{total,old}$ is the current total of used bandwidth, ΔC is the bandwidth of the new connection request and $C_{total,Max}$ is the maximum bandwidth that the Cell m can supply. SNR_{DL}^{j} and SNR_{UL}^{j} are the downlink and uplink SNR for the MS j, Min_SNR_DL and Min_SNR_DL are the minimum required SNR for DL and UL (using the lowest modulation and coding scheme).

Bandwidth Granting Scheme

In the IEEE 802.16 standards, regarding the granting of bandwidth requested, there are two modes of operation: Grant per Connection mode (GPC) and Grant per Subscriber station mode (GPSS) [42]. In the first mode, the BS grants bandwidth explicitly to each connection. Only voice service is assumed for each MS in most of the simulations, hence there is only one connection for each MS in this simulation. So, it is feasible for the GPC granting mode to be used for MSs with a direct link to the BS. In the second mode the bandwidth is granted to all the connections belonging to the subscriber station. This mode is suitable when there are many connections per terminal and allows more sophisticated reaction to QoS needs. GPSS is used for RSs as this allows an RS to re-distribute bandwidth among its connections, maintaining QoS and service-level agreements. A RS requests the BS for bandwidth on behalf of all the MSs attached to it, and the BS grants it a bandwidth to be divided between all the connections belonging to that RS. Then the RS redistributes this sum-total granted among its users according to the service class of the user's connection and its QoS requirements. This granting mode is used for the RS to give the RS more freedom to control the bandwidth usage and guarantee the QoS.

Adaptive Modulation and coding

AMC is used to dynamically adapt the modulation and coding scheme according to the channel condition so as to achieve the highest spectral efficiency at all times. AMC changes the modulation level or coding scheme depending on the channel state, choosing it in such a way that it squeezes the most out of what the channel can transmit. In OFDMA, modulation and coding can be chosen differently for each sub-carrier and can change with time. At the system level, connection is the main unit, and the implementation of the AMC is based on each connection and the coherent modulation scheme selection is dependent on the state of the connection (SNR of the channel).

When the MS creates a direct connection to the BS or an indirect one by the RS, the AMC controller will choose a modulation level according to the SNR of the connection. For every iteration of the simulation, the AMC controller will check the connection (The MS gives feedback to the AMC modulation), and a new modulation level will be chosen for the connection. The AMC schemes used in the simulation and the corresponding maximum data rate and required SNR are listed in Table 3-1. In the uplink, only QPSK and 16 QAM are supported in Mobile WiMAX.

Table 3-1 AMC schemes used in the simulations

Modulation	Code rate	5MHz		10MHz		SNR(dB)
		Downlink	Uplink	Downlink	Uplink	
QPSK	½	3.17	2.28	6.34	4.70	9.4
	¾	4.75	3.43	9.5	7.06	11.2
16 QAM	½	6.34	4.57	12.07	9.41	16.4
	¾	9.50	6.85	19.01	14.11	18.2
64 QAM	½	9.50		19.01		22.7
	¾	14.26		28.51		24.4

Power Control

In the 802.16-2004 standards, standard defines the requirement for uplink closed-loop power control [42]. In the 802.16e standard, it adds open loop power control for uplink, but it is optional. The exact power control algorithm is not defined in the 802.16 standards, and the algorithm implementation is vendor-specific. In this simulator, perfect uplink power control (closed-loop) is assumed and the distributed constrained power control (DCPC) scheme is used as a reference algorithm to make sure the MS transmit power does not exceed the maximum one [73].

In each discrete event of the simulation, the location of MS is constant for the served BS or RS. The uplink power control algorithm will choose the modulation level according the distance from the service station (each target modulation level corresponds to a target SNR value). In the DCPC, each of the MS adjusts their transmitter powers synchronously at discrete time instants. The power adjustment made by the *i* th MS at the *n* th time instant is given by

$$\mathrm{P}_{\mathrm{i}}^{(\mathrm{n})} = \gamma_{\mathrm{t}} \frac{\mathrm{P}_{\mathrm{i}}^{(\mathrm{n}-1)}}{\gamma_{\mathrm{i}}^{(\mathrm{n}-1)}} \quad 1 \leq \mathrm{i} \leq \mathrm{M}\ \mathrm{n} \geq 1 \tag{3.4}$$

where $\mathrm{P}_{\mathrm{i}}^{(\mathrm{n}-1)}$ is its last times instant transmit power according to the feedback of the MS, $\gamma_{\mathrm{i}}^{(\mathrm{n}-1)}$ is its resulting SNR, and γ_{t} is the target SNR. In the power adjustment process, the following principle should hold:

$$\mathrm{P}_{\mathrm{i}}^{(\mathrm{n})} < \mathrm{P}_{\max} \tag{3.5}$$

$\mathrm{P}_{\max}$ is the maximum transmitter power of MS. If the transmitter power $\mathrm{P}_{\mathrm{i}}^{(\mathrm{n})}$ is bigger than the maximum transmitter power, the MS will keep the lower modulation level or be dropped (the minimum requirement is not met).

Although the downlink power control is useful for multi-cell interference reduction and [74, 75] has analysed and proposed some schemes, the complex environment and movable users may limit the system performance. To maximize the throughput of the system, a maximum power is set for the BS and RS. There is not an exact implementation for the downlink power control in the simulation, but a power constraint is still imposed on each BS and RS. The sum of data transmit power and pilot power should not exceed the total maximum transmit power.

$$P_{max}^{DL} = P_{data} + P_{pilot} \tag{3.6}$$

Handover Management

In the simulation, only the handover to a different cell is considered. The handover in the same cell is assumed to happen automatically and there is no connection drop in the handover in the same cell. For the handover to a different cell, a simple hard handover scheme is implemented. Before explaining the process of handover for the proposed network, some terms are defined here to make the explanation clearer.

- Performing channel measurements for an MS: checking the QoS of the connection, the SNR here, and the pilot signal strength of this MS. If one of them is below the prescribed threshold, the handover process is applied.
- Neighbour station searching: searching BSs and RSs around for those that can provide service to this MS.
- Active station: a BS or RS that can potentially serve this MS
- Active set: for each MS a list of active stations the MS can be connected to. The active set is saved.

A handover begins with a decision for an MS to handover from a serving BS to a target BS. For each discrete event of the simulation, the active MS will perform channel measurement to check if handover is needed. If the channel cannot guarantee the QoS, the MS will need a handover to link to a stronger BS or RS. In this case, the MS will perform neighbour station searching to find the active stations nearby and save the information on each active station in the active list. The active list is sorted according to the received pilot power value. The MS chooses the highest pilot power active station from the list and sends a request for a connection to the active station. If the active station cannot accept the MS, the MS will choose a new active station from the list, and then request a connection. The process can be repeated until there is no active station left in the list. If no one can accept the MS, the MS will drop. If the active station can accept the MS, then it updates the subscription status of this MS.

Table 3-2 Parameter settings for the simulations

System Setting		
Cell radius	1500(BS) 500(RS)	m
Distance between BSs	2800	m
DL:UL	1:1	
Operating Frequency	3.5	GHz
Bandwidth	5/10	MHz
Minimum MS to BS Distance	20	m
Frequency-reuse factor	1 or 1/3	
Propagation Environment		
Thermal Noise diversity	-174	dBm/Hz
Pilot Power	30	dBm
Pilot threshold	-80	dBm
Shadowing correlation	0.5	
Load Control		
Load factor	75%	
Required Minimum SINR	7	dB
Uplink		
Maximum Transmit power	23	dBm
Antenna gain	-1	dBi
Noise figure	8	dB
Antenna Height	1.5	m
Downlink		
Maximum Transmit power	45(BS) 41(RS)	dBm
Antenna gain	17(BS) 7(RS)	dBi
Noise figure	4(BS) 6(RS)	dB
Antenna Height	30(BS) 15(RS)	m
Handover		
Pilot power threshold	-74	dBm
Number of Active station	3	

3.4.3 Verification and Validation of Simulation Model

For simulator development, one of the most important aspects is the credibility of the simulation model and its implementation. To ascertain the simulation accuracy, verification and validation of the simulation model are essential.

Simulation verification is to determine whether the simulator performs as intended [76]. During the simulator development and cooperative control research, several optimization algorithms were implemented and tested in the simulator. At the first traffic snapshot of each scenario, the total traffic demand is carefully selected and uniformly distributed geographically over the simulated network to ensure with high probability that all the mobile users can be served. This uniform allocation can provide an initial verification as the behaviour can be predicted. For each optimization algorithm, the simulation results from a particular state were printed out and compared with predicted calculations to make sure that the simulation is operating as intended.

There are several methods for validating the simulation model. In this case, comparing with results from similar simulation models has been utilized to determine whether the implemented model is an accurate representation of the real system [76] . There is little information available on system level simulations for Mobile WiMAX, especially for the evaluation of end users' experience with variable traffic distributions. The developed simulator can be thought of as the combination of a network planning tool and system wide simulator. This thesis investigates the relationship between RF domain configuration and network performance. Network planning capability is also included, so that a RBCN could be set-up based on the locations of BSs and RSs. To simulate the variable traffic distribution, snapshot samples of the traffic distribution over the network are chosen at different times.

To validate the simulation model, a Mobile WiMAX cellular network simulation, without RS, is tested in order to validate the results against the results presented from [1]. To do this validation, the system environment and the assumptions for the main parameters must be set to the same values as those given in [1]. These are listed in Table 3-3.

Table 3-3 Simulation parameters[1, 59]

Parameters	Value
Operating Frequency	2500 MHz
Channel Bandwidth	10 MHz
DL/UL	1:1
Num of Null Sub-Carriers	184(DL) 184 (UL)
Num of Pilot Sub-Carriers	120(DL) 280(UL)
Num of Data Sub-Carriers	720(DL) 560(UL)
BS-to-BS Distance	2.8 km
Minimum MS to BS Distance	36 m
BS Height	32 m
MS Height	1.5 m
Number of Sector	3
BS Antenna Gain	15 dBi
MS Antenna Gain	-1 dBi
BS Maximum Power Amplifier Power	43 dBm
MS Maximum Power Amplifier Power	23 dBm
Log-Normal Shadowing	8 dB
BS Shadowing Correlation	0.5
Thermal Noise	-174 dBm/Hz

In [1, 43] there is no exact value given regarding the realistic service the system could support. Only the peak rate for the uplink and downlink at different situations are presented. To compare with the reference value and to test the traffic scenario model, a simulation was built by the author of this thesis to evaluate the average cell throughput[2] (kbps/cell) when the traffic demand is increased. . So,

[2] The average cell throughput R is measured as

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

Where b is the total number of data bits received (DL) or transmitted (UL) by all MSs in the simulated system over the whole simulated time. k is the number of cells and T is the simulated time.

this experiment validates the Mobile WiMAX model in a more realistic environment. For simplicity, the active MSs with a constant bandwidth demand are designed according to the different number of sub-carriers for DL and UL. The traffic demand of each MS is a bandwidth of 19 sub-carriers (around 185 KHz) for DL and 14 sub-carriers (around 137 KHz) for UL, so the maximum number of supported MSs is 40. There are a total 19 cells in the simulation network and the uniform geographical traffic distribution is used for the whole simulation. The simulation results are plotted in Figure 3.6**.**

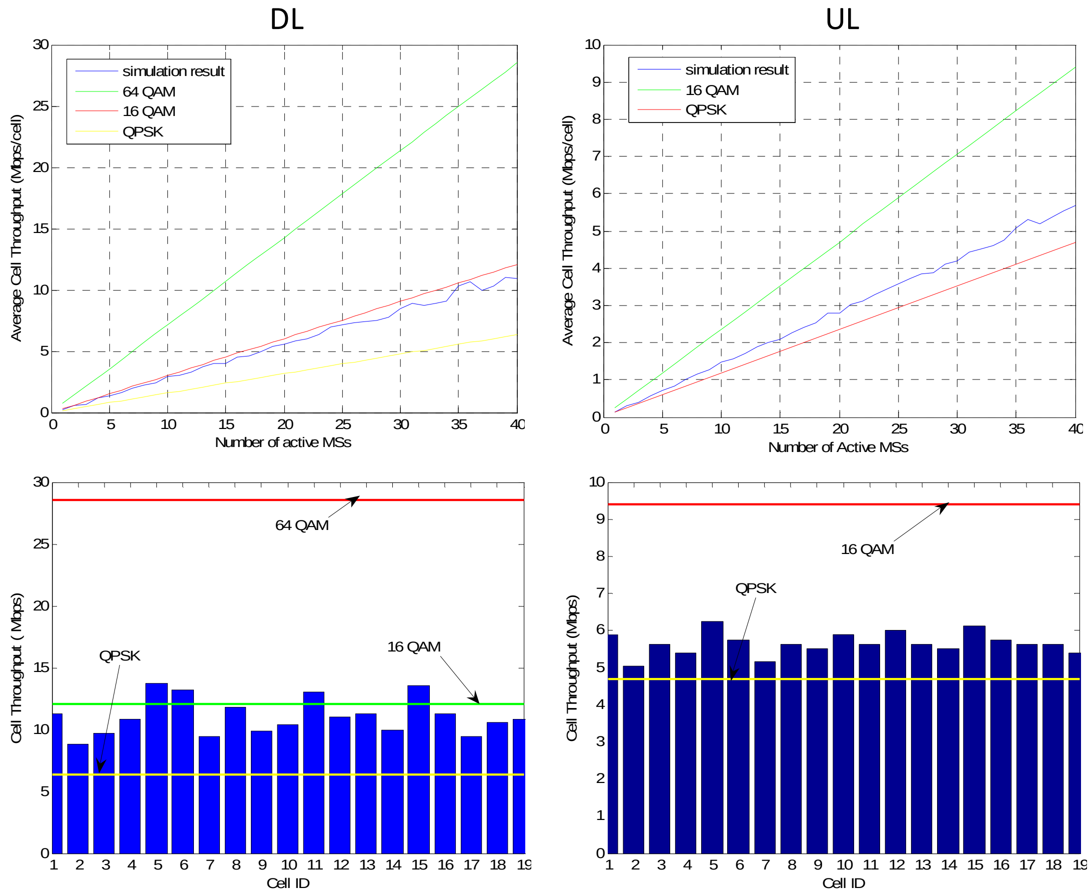


Figure 3.6 Average cell throughputs with different traffic demand

Seen from the top figures, with the number of active MSs increased, the cell throughputs of DL and UL both increased. When the number of active MSs is 40 and there is no available radio resource left, the average cell throughput of reaches its maximum point. In the bottom figures, the maximum cell throughputs for each cell compared to peak rates of different modulations are plotted. Compared to the reference values from [1] the values output from the simulation are similar (same trend and within a reasonable range) although they are not exactly the same because of differences in some details: To achieve spectrum efficiency, different modulation levels are selected based on the channel state with the help of AMC. In the Mobile WiMAX, the highest modulation level used for DL is 64 QAM (Quadrature Amplitude Modulation) and UL is 16 QAM. The peak rate of each modulation level is calculated based on all the channels using the same modulation level. However, this seldom happened in the physical network and used as benchmark values for the validation. Seen from the above figures, the average cell throughput is within the reference value range of 64QAM and QPSK (Quadrature Phase Shift Keying) as most of the MSs could not select the highest modulation level. It is also clear that results from the simulation model show the same trend as the corresponding results from [1] under a similar system environment. The similar trends and reasonable range validate the simulator and give credibility to its results.

In [77], the white paper for Mobile WiMAX products presents a cell throughput (15.75 Mbps) based on a similar system environment, which is nearly equal to the simulation result obtained here (average cell throughput is around 15.6 Mbps, including both uplink and downlink).

3.5 Summary

In this chapter, the distinct features of cooperative control as used in this thesis are given. As the optimization algorithms for different antenna models are the key point of this thesis, a system level simulator is built to evaluate the performance of the optimization algorithms. The architecture of the simulator, the Mobile WiMAX simulator and traffic models have been described. The last section provided simulation model verification and validation relevant to the research.

Chapter 4 BS CONTROL OPTION

4.1 Introduction

Introducing RSs into a cellular network has the potential to improve the system capacity and signal quality in boundary areas. However, as the location of the traffic demand can change with time, simply upgrading an existing cellular network and configuring the RBCNs statically for access, only partially solves the problem. This chapter focuses on exploiting the potential flexibility of existing antenna systems already used in cellular networks and upgrading with dumb (and so cheaper) RSs. There are many complexities to realizing the potential of a network in a dynamic context. In this chapter, antenna system control for BSs is investigated with a view to obtain increased performance by using

- system wide cooperative algorithms to perform geographical load balancing in large scale network;
- integration of the cooperative use of RS within a hybrid cellular structure;
- simple and cheap forms of antenna shaping that do not require knowledge of the exact location of MSs.

As the research will be approached in stages linked to the capabilities of the antennas, in this chapter only the most immediate possible deployments are considered, viz. through omni-directional antennas with variable pilot signal power and sector antennas with simple electric tilting capability. The communication between BSs and RSs use the same antenna, as that used to communicate with MSs. So, the dumb RS could be thought of as a bigger MS from BSs' point of view.

Tilting antennas have already been deployed in sections of 3G networks in the UK, albeit without automatic real time cooperative control. Power control of the pilot signal to keep a proper number of active MSs in the local cell are described in [78] and [79]. Improving the system capacity in RBCNs equipped with such existing antenna systems by cooperative control is the main research topic of this chapter. The proposed approach does not require any modification either at the user side or of the standards. Existing cellular networks can be upgraded through incremental deployment of the antenna control as both the antennas and the cooperative algorithms can easily work with non adaptive BSs.

The key purpose of this chapter is to demonstrate the gains obtained by using new cooperative optimization algorithms for adjusting the physical antenna coverage patterns on simple antenna systems in networks with BSs and RSs. The physical layer is assumed to be controlled either by omni-directional antennas with variable pilot signal power or by sector antennas with simple tilting capability. Two cooperation algorithms for cooperatively adjusting the physical antenna coverage patterns in real time for relay network structures are presented. Two experiments are used to demonstrate the approach:

- Both BS and RS have simple omni-directional antennas. Here there is the possibility of cooperative real time load balancing by adjusting the pilot signal powers.
- Each BS has sector antennas with simple tilting capability and RSs have simple omni-directional antennas.

The results of both experiments are compared with the performance of a conventional network, where adaptive cooperation is not applied and the reduction in blocking and dropping rates are quantified.

The rest of the chapter is organized as follows: in Section 4.2, the cooperatively variable pilot signal power approach is described. In Section 4.3, the cooperatively adaptive tilting approach is described. Finally, Section 4.4 compares the two optimization approaches and concludes the chapter.

4.2 Cooperative Pilot Signal Power Control

4.2.1 Pilot Signal Power Control

In cellular communication, the pilot signal is used by mobile terminals for channel quality estimation, cell selection, and handover [80]. A mobile terminal measures and compares the pilot signals that it can detect, and typically attaches itself to the cell with the best quality pilot signal.

Factors that determine the level of the received pilot signal include the transmission power used for the pilot signal, signal attenuation between the BS and MS, and the effect of thermal noise [80]. In a simple propagation scenario, where signal attenuation is essentially determined by distance, the mobile terminal will be attached to a cell that belongs to the 'closest' BS (as perceived by signal strength), if all the cells use the same level of pilot power. With uniformly distributed traffic and equally-spread BSs, the coverage of the cells will be roughly the same. Increasing or decreasing the pilot power makes the cell larger or smaller, so allowing control of cell coverage.

The variable pilot signal approach uses an idea similar to cell breathing, which is already used in normal cellular networks. Cell breathing is a mechanism that attempts to keep the forward and reverse link handover boundaries balanced by changing the forward link coverage according to the changes in the reverse link interference level [51]. With the help of cell breathing, a heavily loaded cell shrinks by reducing its power and subscriber traffic is then redirected to a neighbouring cell that is more lightly loaded. In [81, 82], adjusting pilot power for load balancing is introduced for WCDMA networks. It is based on the target values for coverage and traffic load and attempts to minimize the deviation from the target values by adjusting the levels of pilot power. It is similar to the problem studied in this work. However, their solutions are mainly focused on a pure cellular network and local cell performance optimization.

In previous pilot signal studies, [81, 83] are also interested in the problem of choosing the optimal levels of pilot power, which involves the trade-off between full coverage on one hand, and the power consumption on the other hand. The more power allocated for pilot signals, the better the coverage obtained. On the other hand, a higher value of the pilot power level in a cell means higher pilot pollution in the network and less power available to serve user traffic in the cell.

In the work of this thesis, the key issue is how to achieve geographic load balancing using variable pilot signal power for a large scale network. So, the traffic in the local cell and neighbour cells both need to be considered. When a heavily loaded BS decreases its pilot signal power, the neighbouring cells can cooperatively increase the pilot power to avoid coverage holes in the network. The proposed scheme does not require any modification at the user side nor of the Mobile WiMAX standard. It only requires the ability to dynamically change the transmission pilot signal power. The MS will automatically request connection or handover to the neighbouring cells according to the pilot signal strength. Unlike existing cell breathing methods, which utilize local optimization heuristics, a centrally controlled algorithm finds the optimal pilot power settings that balance the load of the heavy traffic loaded cells. So, this is a global optimization problem for the wide area network optimization.

Before the description of the proposed cooperative variable pilot algorithm, the potential of variable pilot signal strength in Mobile WiMAX is investigated below. In Mobile WiMAX networks, there are three types of sub-carrier in the S-OFDMA PHY layer [84]. The pilot sub-carriers are used for various estimation and synchronization purposes. The pilot allocation is performed differently in different subcarrier allocation modes [41]. For DL Fully Used Subchannelization (FUSC), the pilot sub-carriers are allocated first and then the remaining subcarriers are divided into data sub-channels. There is one set of common pilot subcarriers in FUSC. For DL Partially Used Subchannelization (PUSC) and all UL modes, the pilot subcarriers are allocated from within each sub-channel and each sub-channel contains its own set of pilot subcarriers [41]. In the *AWE* network planning tools for Mobile WiMAX [85], the transmit power for pilot subcarriers is also separately defined and transmit power for data and pilot subcarriers can be set individually. So, providing variable pilot signal strength in Mobile WiMAX is feasible from the perspective of standards and practical usage. In the work below, the FUSC model is used in the DL and it is assumed that the transmit power of the data sub-carriers will be constant when the pilot sub-carrier transmit powers are changed.

4.2.2 Cooperative Pilot Signal Power Control Algorithm

In the proposed RBCN, the selection criterion of BS and RS is mainly based on the received pilot signal strength perceived by the MS. The MS compares the path gains between BS and MS or RS and MS, and chooses the station with the strongest pilot signal strength.

Since the main objective is to balance the geographic load, the traffic conditions in local cell and adjacent cells must be considered. For each MS, the two closest cells (directly or indirectly) that could potentially provide the service are used. They need to cooperate with each other to provide the service for a traffic unit without generating any coverage holes. When one of them is heavily loaded it has less potential to serve a new traffic unit. Thus the other cell should serve such traffic. The RSs have more potential ability to cooperate with the neighbouring cell(s) as they are near the boundary of a cell and have more overlapping coverage with the neighbouring cells.

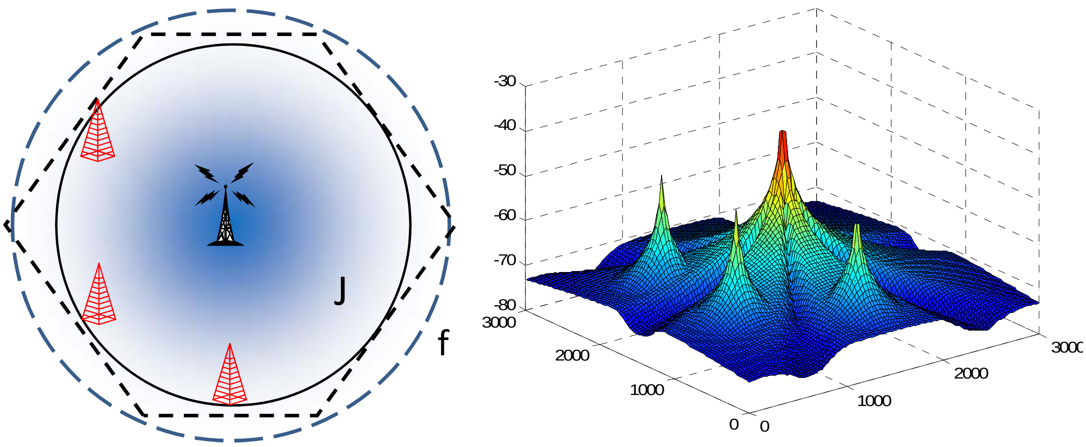


Figure 4.1 The cell architecture and corresponding pilot signal strength distribution

The perceived pilot signal strength of each MS is assumed to be known by each BS, which is based on the feedback of received signal strength indicator (RSSI) of the strongest pilot signal. So, the number and rough distribution of potential MSs can be known with the feedback of each MS, at least if the MS is within a range where it is conceivable that the BS can cover it. The frontier for a BS is defined as the boundary beyond which it would always be impossible for the BS to cover the MS in practice. According to the variable pilot signal power range, the feasible cell coverage, i.e. up to the frontier, is divided into bands as shown in Figure 4.1. Using j as the index of the bands, j ranges for 1 to f, where f is the index of the band at the frontier. The band that each MS belongs to is determined by the perceived pilot signal strength. Note that in a real physical network the bands will not be circular, and the method here does not depend on a circular shape.

From the mathematical aspect, the geographic load balancing problem could be defined as below:

$$\mathrm{Min}\sum_{\mathrm{n}=1}^{\mathrm{N}}\left|\mathrm{d}_{\mathrm{n}}^{\mathrm{total}}-\bar{\mathrm{d}}\right| = \mathrm{Min}\sum_{\mathrm{n}=1}^{\mathrm{N}}\left|\mathrm{d}_{\mathrm{n}}^{\mathrm{BS}}+\mathrm{d}_{\mathrm{n}}^{\mathrm{RS1}}+\mathrm{d}_{\mathrm{n}}^{\mathrm{RS2}}+\mathrm{d}_{\mathrm{n}}^{\mathrm{RS3}}-\bar{\mathrm{d}}\right| \tag{4.1}$$

where $\mathrm{d}_{\mathrm{n}}^{\mathrm{total}}$ is the total served traffic for the cell n, $\mathrm{d}_{\mathrm{n}}^{\mathrm{BS}}$ is the number of MS directly connect to the BS in cell n, $\mathrm{d}_{\mathrm{n}}^{\mathrm{RS}}$ is the number of MSs directly connected to the RS and $\bar{\mathrm{d}}$ is the average cell traffic demand where the average is taken over all the cells of the network, N is the number of cells in the network. For each cell, there are four parameters that could be controlled: the pilot transmit power of the BS and the pilot transmit power of the three connected RSs. To achieve geographic load balancing, for each cell, an optimal setting will be chosen to minimize the difference between the served traffic and average cell traffic demand.

To represent the traffic condition between the local and adjacent cells, the cooperative algorithm uses a cost function. Instead of using the real value of traffic load from the neighbouring BSs, the BS's perceived traffic load is used. This is reasonable since there are lots of overlapping areas between each BS & RS up to the frontiers. This potentially eliminates the requirement to exchange traffic condition information between adjacent cells, so reducing communication overheads and system complexity, i.e. I have cooperation without communication. At any one time a RS is served by only one BS. In this comparison the assignment of a RS to a BS is fixed. When tilting of sector antennas is allowed this requirement is relaxed as this allows the flexibility to be exploited better.

The cost of a single BS k only serving all bands up to and including band j is defined as C_j^k. C_j^k is a weighted sum of

- the demand (the bit rate requirement) of MSs up to and including band j (i.e. those that are going to be served) and

- the demand of all MS beyond band j up to the frontier (i.e. those that are not served.), as expressed in the equation below:

$$C_j^k = d_j + w \cdot d_j^f \tag{4.2}$$

Each unit of demand beyond band j is given greater weight (*w* in the equation (4.2)) as this represents a deficiency in the coverage – at least from the perspective of the BS in question. The RSs have bands as well, and a similar cost function is used for RSs. So, the total cost for a cell *i* to serve the MSs up to band (j for BS, m for RS1, n for RS2, t for RS3) could be calculated as below:

$$C_i^{total} = C_{BS}^{j} + C_{RS1}^{m} + C_{RS2}^{n} + C_{RS3}^{t} \tag{4.3}$$

Therefore, the cell with the higher cost value should give some help to the neighbours as the overlap areas have some more un-served traffic demand.

Initially, all the BSs start with coverage at some historically sensible band and the costs of this coverage is computed. The BS in the cell with the highest C_i^{total} tries to increase its coverage by one band to potentially serve more MSs. If the capacity of the cell is reached then the coverage cannot be increased further. If a BS has reached its frontier band then the coverage of a RS with highest cost value in this cell is increased. The above processes will cycle through the cells one by one.

This greedy search will always try to increase the coverage band to serve more traffic until the frontier band is reached. A cooperative control scheme is applied to negotiate with neighbouring cells to make sure the coverage band does not always increase to the frontier and to guarantee all the desired areas are covered. The size of overlap area between the adjacent cells is assumed constant even after boundary changing. So, if the neighbouring coverage is increased, the local cell will decrease the band to mitigate the overlap area increase. Note that this could also keep the size of handover area constant, which mitigates the handover rate increase. The proposed algorithm is shown in procedural form below.

Algorithm: Variable Pilot Signal Strength
1: Set the maximum iteration number: Iter_Max, n←0;
2: Repeat
3: Calculate the cost value C_i^{total} for each cell;
4: Sort the cells in descending order according to the cost value and save in the list L;
5: Repeat
6: Get a cell *m* from list L;
7: IF (BS currently covers up to band j & does not reach frontier band)
8: increase cover by a band; j←j+1
9: Else
10: Sort the RSs in cell *m* according to descending cost value;
11: Get a RS C_{RS}^{t} from list
12: IF (RS in this cell has not reached its frontier band)
13: increase cover by a band of the RS; t←t+1
14: Until no cell is in the list
15: Adjust the bands to avoid the maintain the size of the handover area
16: n←n+1
17: Until n reaches Iter_Max or there are no blocking MSs in the whole network.

4.2.3 Simulation and Results

Simulation Configuration

To evaluate the system performance, a 6×6 hexagonal Mobile WiMAX network was used in the simulation (as shown in Figure 4.2). A traditional hexagonal cellular model is used. In each cell, there is a BS in the centre and 6 RSs around, but only 3 belong to the BS and the others belong to the neighbouring cells. The important simulation configuration parameters are given in Table 4-1.

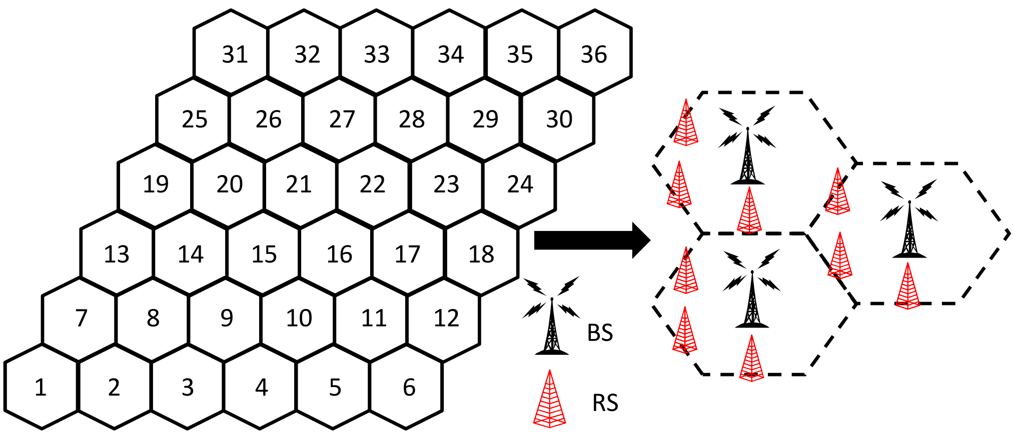


Figure 4.2 The architecture of simulation network

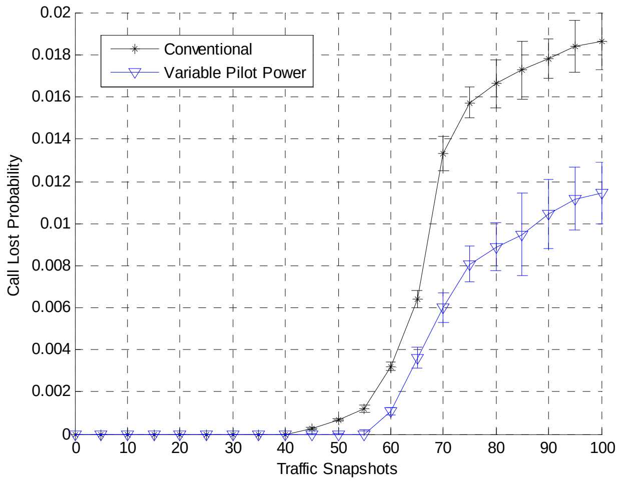
Table 4-1 Configuration parameters used in the simulation

Parameters	Values
Cell radius	1.5km(BS) 500m(RS)
Sector Number	Omni (BS) Omni (RS)
Distance between BS to BS	2.8 km
Distance between BS to RS	1 km
Operating Frequency	3.5 GHz
Channel Bandwidth	5 MHz
Noise Spectral Density	-174 dBm/Hz
Antenna Gain	17dBi(BS) 7dBi(RS) 0dBi(MS)
Antenna Altitude	30m(BS) 15m(RS) 1.5m(MS)
Pilot Power	32 dBm(BS) 27 dBm(RS)
Minimum Mobile-to-BS Distance	20m
Pilot threshold	-80dBm

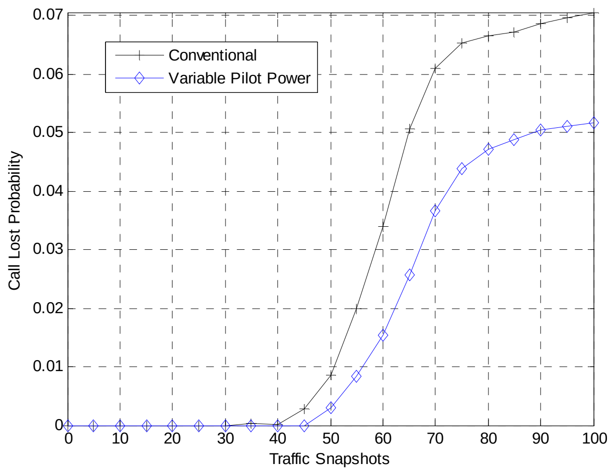
Simulation Results

In this experiment, BS and RS both use omni-directional antennas. The results of using the variable pilot power algorithm in a Mobile WiMAX network are compared with the results from a conventional network, where the pilot signal power is constant.

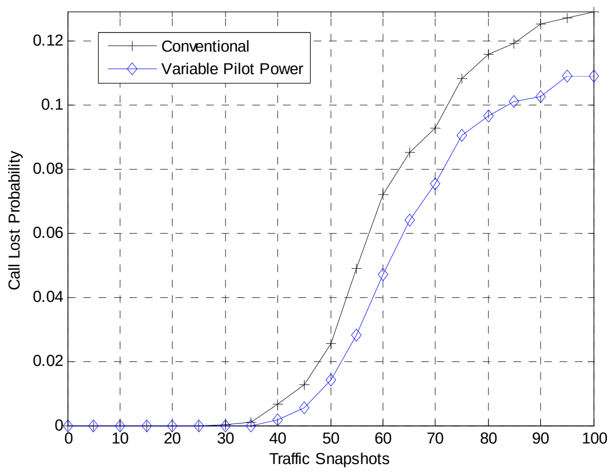
The call loss probability results for two simulation networks are presented in Figure 4.3 for the traffic scenarios where each hotspot area has a population size of 500, 750 and 1000 MSs. It shows a simulation of 100 snapshots, each 60 seconds apart. The horizontal axis is time, measured in snapshots. It demonstrates benefits that are consistent with the system capacity results. At the beginning of the simulation, when the traffic is nearly uniformly distributed geographically, the blocking rate of the balanced network is the same as the conventional one, just as one would expect. As the hotspots form, the blocking of the RBCNs increases slowly. When there is more traffic demand in each hotspot, i.e. as the traffic becomes more heterogeneous, the performance improvements are more significant. In the Figure 4.3 (a), the confidence interval with 95% confidence level is also plotted to indicate the reliability of the simulation results.



(a) Traffic hotspot with a population size of 500 MSs



(b) Traffic hotspot with a population size of 750 MSs

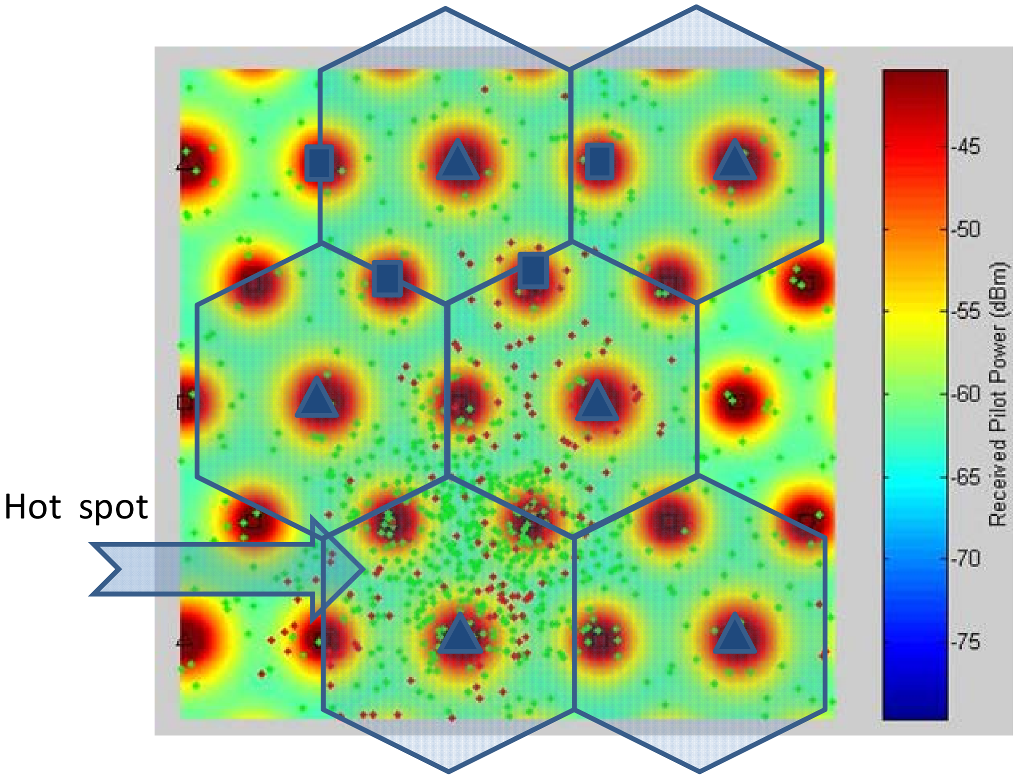


(c) Traffic hotspot with a population size of 1000 MSs

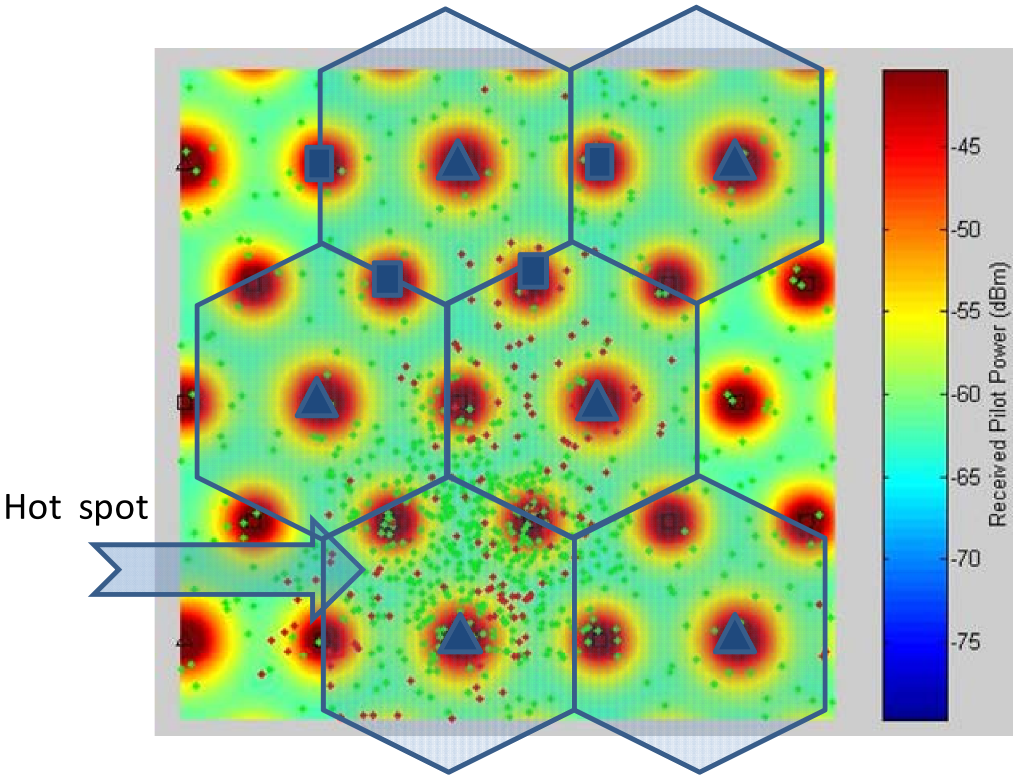
Figure 4.3 Simulation results when using variable pilot signal power

In a conventional network, all the BSs or RSs use constant pilot signal power. The MSs will choose the nearest BS or RS to request the connection. If the BS is overloaded, the MSs will be blocked. The BSs and RSs in the cooperative network can change pilot signal power according to the current traffic condition. The BSs or RSs in the hot spot area will decrease the pilot signal power, as there are too many MSs in their coverage and there is not enough radio resource for the new or handover connections. The neighbouring BSs with light traffic loads will increase the pilot signal power. If the MSs are in the centre of overloaded BSs, they cannot get help from the neighbouring RS or BS, as they are too far from the MSs and cannot support the link setup.

The pilot signal power distribution for one hotspot area in a conventional network and an optimized network are shown in Figure 4.4. The big triangles represent BSs, the squares represent RSs. The red and green dots represent the blocked and served MSs respectively. In the conventional network, if the cell is overloaded, the MSs will be blocked (red dots shown in Figure 4.4 (a)). The BSs and RSs in the optimized network can change pilot signal power according to the current traffic distribution. The BSs or RSs in the hot spot area will decrease the pilot signal power and the neighbouring cells will increase the pilot signal power in coordination. When the pilot signal power decreases, the MSs in the boundary of BS will request connection or handover to the neighbour BS or RS (as shown in Figure 4.4 (b)).



(a) Conventional network



(b) Network with variable pilot power

Figure 4.4 The received pilot power distribution

4.3 Cooperative Electric Tilting Control

4.3.1 Electric Tilting

Base station antenna down-tilt is a common practice in cellular networks. By changing the outer frontier of a sector through tilting up or down the antenna's vertical radiation angle, a network operator has the capability to alter power distribution and cell coverage. This is guided by measurements and network planning, which is an essential support in cellular networks. As shown in Figure 4.5, the principal mechanism on down-tilt is to change the 'look direction' of the main beam, which is the direction where the antenna's directional gain is at its maximum.

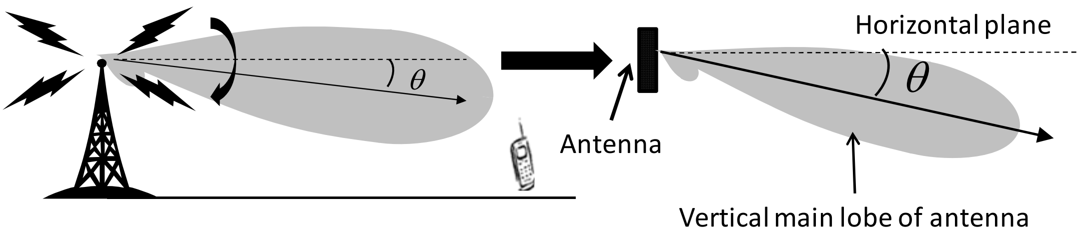
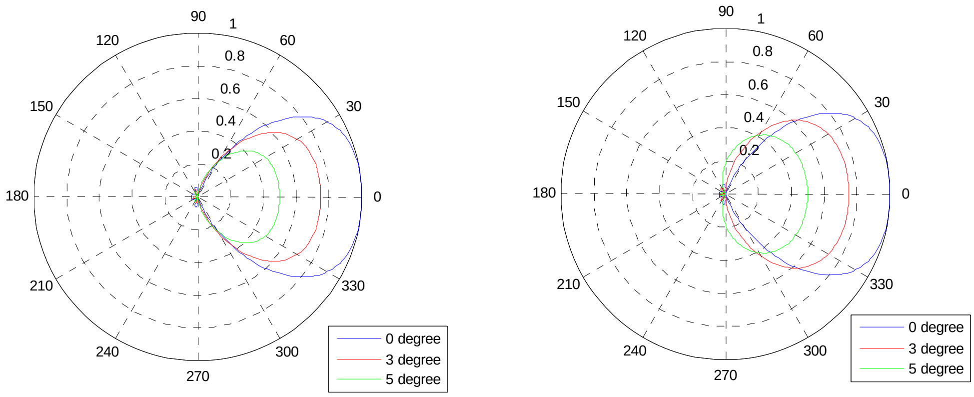


Figure 4.5 Illustration of the antenna tilting concept

Antenna down-tilt can be implemented mechanically—mechanical down-tilt (MDT) or electrically—electrical down-tilt (EDT) [86]. Mechanical down-tilt is usually accomplished by physically tilting the antenna using brackets or shims. When the antenna has down-tilt, the antenna main lobe is more precisely directed towards the intended area. However, the mechanical down-tilt is exclusively in the main-lobe direction and the antenna radiation pattern is not down tilted at all in the side-lobe direction [87]. The antenna mechanical down-tilt can only be conducted by an on-site visit of engineers. Most commonly BS antennas' tilt angles are selected in advance during network planning, and go through an adjustment process only when an antenna is being installed. Therefore, electrically down-tilted antennas might become an attractive choice for antenna selection. Antenna electrical down-tilt is carried out by adjusting the relative phases of the antenna elements of an antenna array in such a way that the radiation pattern can be down-tilted uniformly in all horizontal directions [88].



(a) Electrical down-tilt approach (b) Mechanical down-tilt approach

Figure 4.6 The impact of antenna tilting on the horizontal radiation pattern

A 60 degree sector antenna is assumed to be used by the BSs in this approach. The coverage areas of MDL and EDL under different tilting angles are shown in Figure 4.6. The measurement is taken in an assumed free-space environment and only line-of-sight (LOS) propagation happens. The sector antenna is assumed to have the ability of electric down tilt, rather than mechanical, for the following reasons. First of all, it can be controlled by a simple cooperative algorithm in real time. This gives us more flexibility to service heterogeneous demand and demand that changes over time. Furthermore, the electric approach can be more precise in the down-tilt angle and the changed coverage and hence help ensure there are no holes in the network.

4.3.2 Cooperative Electric Tilting Control Algorithms

In a mobile network, the outer frontier of a sector (this now refers to a real BS sector antenna) can be efficiently modified by tilting the antenna radiation angle up or down, so distributing the energy either closer or further from the BS. This has an impact on different mobile communication systems. In the WCDMA network, the capacity and coverage are strongly based on the level of interference. The purpose of adjusting the antenna tilt is to optimize power distribution and hence reduce the other-to-own cell interference ratio [89, 90].

In the proposed RBCN, each RS can potentially be served by two BSs. The antenna down-tilt can be used to reassign the RSs by tilting at different angles. There are two states of BS antenna tilting for RS assignment. The first state is tilting for assignment, i.e. antenna down tilt θ_B degrees used for covering and serving the RS (see the BS on the right in Figure 4.7). The other state is down tilt θ_A degree to avoid covering of the RS (see the BS on the left in Figure 4.7). The θ_A and θ_B tilt angles are selected in advance during network planning, and only go through an adjustment process when an BS antenna or RS is being installed.

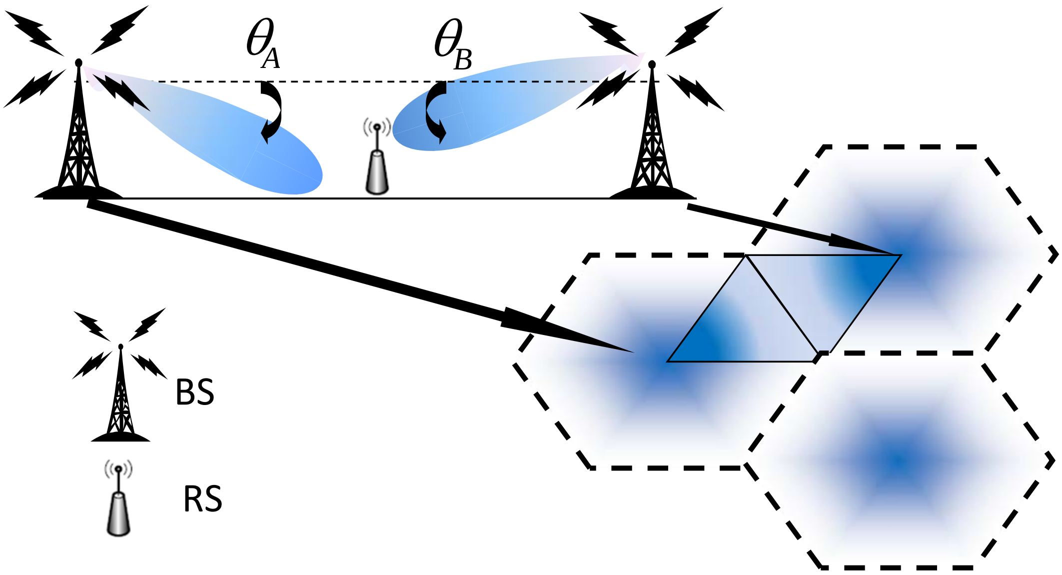


Figure 4.7 Antenna tilting for RS reassignment

Initially each BS serves three RSs as illustrated in Figure 4.1. When the moving traffic forms into hotspots, the BSs will reassign the RS according to the demand from the MSs. The main objective is to increase system capacity by splitting any hot-spots, getting neighbouring BSs to serve the RSs in hot-spot areas, so improving the load balancing in the network. As described above, the cooperative titling control approach aims to gain the following improvements:

1. Split the hotspot traffic area and so achieve load balancing
2. To focus the radiation power to the high traffic load area to improve the system capacity by making the sector antenna tilt down,
3. Ensure no hole in the network coverage through cooperative control

In this approach, it is not necessary to know the exact location of mobile users. It is sufficient to know just the demand of all the mobile users in each BS and served RSs. According to the demand of each BS, the RS of the higher traffic load BS is reassigned to the neighbour base station. An improvement in network capacity results because the main-beam of the sector antenna is angled to direct energy towards the relay station. The antenna of the neighbouring base station tilts down to avoid covering the RS and focuses the radiation power to the nearby area to gain the improvement.

When the RS is reassigned to the target BS, all the MSs attached to the RS also need to handover to the target BS. An efficient group handover procedure as described in [91] is used for RS handover in the simulation. When a RS hands over from the source BS to the target BS, it hands over on behalf of all the attached MSs with the source BS initiating the handover. The BS moves all of the MSs together with one set of messages with the RS, instead of individual handover messages with each MS. The group handover is faster and more efficient than the normal handover process, as only one station (RS) is involved instead of multiple MSs.

A simple cooperative algorithm that has been developed to demonstrate the potential is described. For simplicity only the number of MSs is used to make decisions about tilting angle and which BS a RS is assigned to. The approach generalizes readily to bit rate requirements rather than simple counts. D_i is defined as the number of MSs within the frontier of BS i, and all its served RSs. It includes both the served and un-served MSs.

$$D_i = \sum_{k=1}^{6} S_k + \sum_{j} R_{ij} \tag{4.4}$$

where S_k is the number of MSs within the frontier of sector k but not assigned to any RS. R_{ij} is the number of MSs within the coverage frontier of RS_j and RS_j is currently served by BS i.

When the MSs move and form hot-spots, some MSs are blocked by the BS as the radio capacity is reached. At this time, or possibly at some time before, triggered by a capacity threshold, the cooperative tilting algorithm is applied to the network. Equation (4.4) is used to calculate the demand for each BS. According to the demand, the BSs are sorted in descending order. If the BS capacity is exceeded, the BS iterates though each of the RSs it currently serves and finds the one with the smallest demand and assigns it to the neighbouring BS that can serve the RS. After all the BSs that have their capacity exceeded have been considered in this way, the demand for each BS is calculated again depending on the reassignment results. If BS capacity is still exceeded, then the process of reassigning the smallest RSs is repeated. The iteration will stop after all the MSs are served or the maximum iteration number has been reached. The adaptive tilting algorithm is shown in procedural form below

Algorithm: Adaptive Tilting for reassignment

```
Set the maximum iterative number: Iter_Max, n←0;
Repeat
    Calculate the demand D_i for each cell;
    Sort the cells in descending order according to the demand and put in the list L;
    Repeat
        Retrieve a cell m from list L;
        IF (demand D_i of cell m extends the capacity)
            Sort the current connected RSs into descending order
            Reassign the RS with lowest traffic
    Until no cells remain in the list
n←n+1
until n reaches Iter_Max or there are no blocking MSs in whole network
```

Note that this algorithm does not use the locations of the MSs. Also the algorithm is the same for each sector. As described this is a centralized algorithm, e.g. running at an authentication and service authorization (ASA) server.

In this approach, the antenna tilting control is simple. As there are only two states for the antenna tilting control, this is very easy to implement. Even in a real network, with appropriate preliminary network planning, the operator can readily decide the tilting angle. This tilting angle will be varied according to the location and altitude of the antenna system and there is not a constant angle for the reassignment. When an operator decides the initial tilting angle, it needs to make sure it leaves no coverage hole in the network. The cooperative tilting control will make sure the opposing sectors work in cooperation, when one is tilted up for the RS assignment and the opposite one must tilt down. As the coverage of RSs will be constant and between the opposite sectors, this could potentially mitigate the coverage hole effect when cooperative tilting is applied.

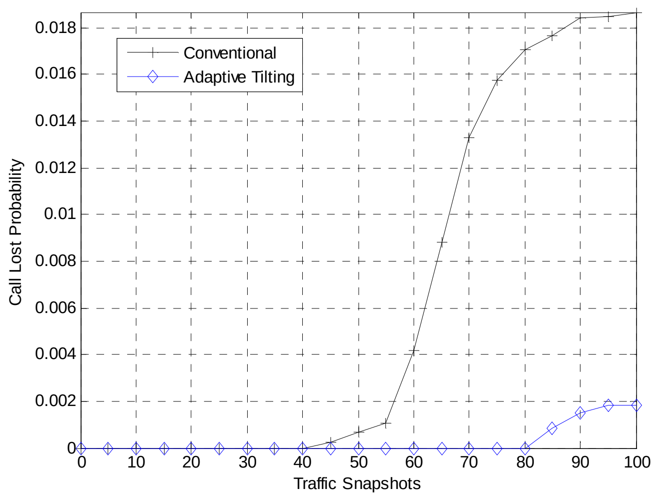
4.3.3 Simulation and Results

Simulation Assumptions

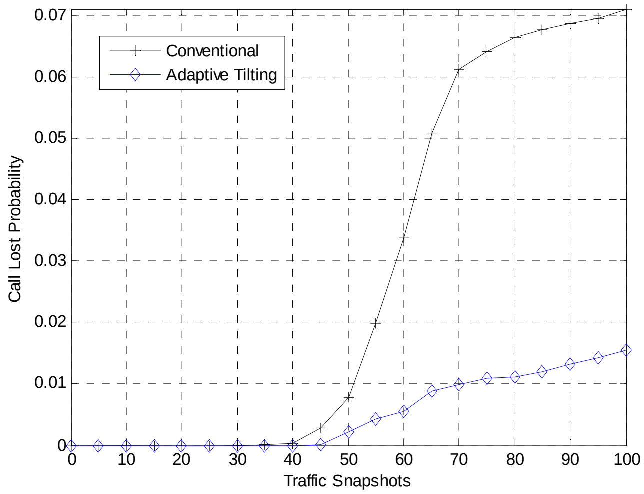
In this simulation, each BS has six sectors and each sector has a section antenna with tilting ability. RSs still use omni-directional antennas but with constant pilot signal power. Initially three RSs are assigned to each BS. The locations of the RS are such that there can be zero to six RSs assigned to any BS. A RS can only be assigned to one BS at a time. The other experiment parameters are the same as the cooperative pilot signal experiment described previously.

Adaptive Tilting Experiments

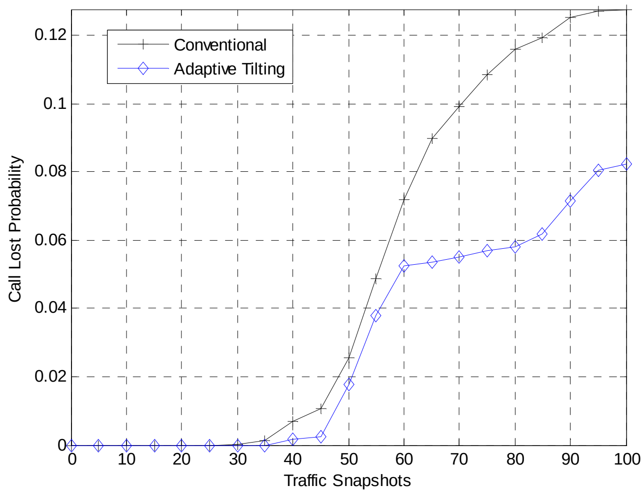
As shown in Figure 4.8, the cooperative tilting algorithm approach achieves a near optimal result. Similar results are found for different variations of the simulation parameters. As the hot-spots form, the blocking rate of the network increases. However, the blocking rate of the scheme with the cooperative tilting algorithm is always better than the conventional one, especially when hot-spots occur. When the hot-spots are relatively light, the cooperative tilting approach can almost avoid blocking users, while the conventional network has lots of blocked or dropped users. When the hot-spots are very heavy, the cooperative tilting approach reaches its optimization limit, and its blocking rate also increases. However, it still performs much better than the conventional one. Note that the first term of the cost function (4.4) helps reduce unnecessary handovers.



(a) Traffic hotspot with a population size of 500 MSs



(b) Traffic hotspot with a population size of 750 MSs



(c) Traffic hotspot with a population size of 1000 MSs

Figure 4.8 The simulation result with adaptive tilting

The antenna coverage of BSs and RSs for a hotspot area is illustrated in Figure 4.9. The big triangles represent BSs, the diamonds represent RSs. The larger red dots represent blocked MS and the smaller dots served MS. Registered users not making a call are not shown. The left of Figure 4.9 shows the coverage for a snapshot without cooperative tilting and there are many blocked MSs in the hotspot area due to the limited radio resource. When the adaptive tilting algorithm is applied in the network to the same scenario, the BSs rearrange the RS according to demand. The BS with a light traffic load tilts up to serve more RSs, and the BS with a heavy traffic load tilts down to serve the MS around the BS (right of Figure 4.9). This splits the hotspot, so balancing the traffic in the hotspot area and decreasing the blocking rate.

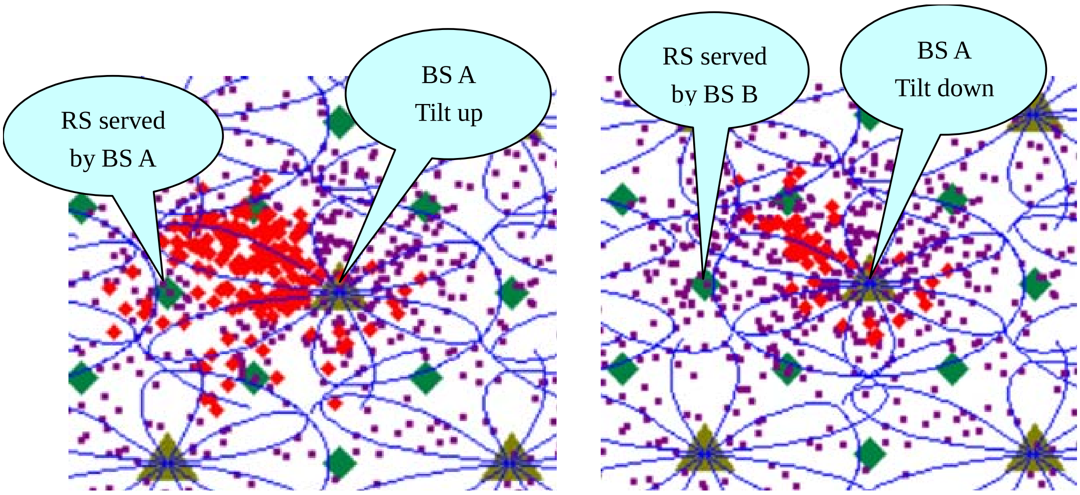


Figure 4.9 The antenna coverage in conventional network and network with cooperative tilting

4.4 Summary

A potentially cheap approach to improving QoS that exploits the flexibility of controlling the antenna patterns to find near optimal cooperative coverage patterns has been illustrated. It is a very efficient way to enhance system wide resource utilization and enhance the possibility to fulfil QoS guarantees in Mobile WiMAX. According to the capability of two kinds of antenna, two cooperation algorithms for RBCNs have been presented.

Despite its simplicity, improvement was shown for the cooperatively adjusted variable power algorithm using an omni-directional antenna. The simple algorithm performed well in the (rather static) simulated environment, but there are some concerns for the proposed approach when applied in the realistic network. Firstly, the load balancing with variable pilot signal strength cannot itself ensure that enough resources released as the real propagation environment is complex and may not be adequately modelled by the simple propagation model used. So, the proposed algorithms may have a poorer performance in a physical network than the simulation result. Secondly, the dynamic behaviour of the cell also makes the users at the edge of the cell more vulnerable to being dropped or terminated even when the cooperative control scheme is applied. Once the MS at the edge of a cell experiences difficulty in maintaining the connection, it tries to increase the transmit power of the hand set to its maximum value. So, this may increase the handover rate and consume the user's battery.

Compared to the variable pilot signal strength, the cooperative tilting approach shows a marked improvement in the system capacity, especially in the hot-spot areas. If hotspots have a small population or are located around the boundary of cells, the cooperative tilting algorithm can achieve a near optimal result. As the adaptive tilting only reassigns the RSs according to the traffic distribution, there is little effect on the MSs connected to the reassigned RSs. The transmit powers of the MSs remain the same. With the help of a group handover scheme, the radio resource released in high traffic cells could also be more efficient and predictable.

What is especially interesting about this kind of approach is that it can be incrementally deployed into the existing cellular networks. It is inexpensive as the RS does not need a wired backhaul and there is no expensive antenna system requirement. It is scalable because it retains a cellular control structure and has a simple routing path to the BS. It inherently provides flexibility and hence resilience to faults and attacks.

Chapter 5 RS CONTROL OPTION

5.1 Introduction

To enhance capacity and coverage quality for cellular networks, additional RSs may be built on selected sites to improve service probability and QoS. How to create a RBCN or add RSs into an existing cellular network quickly and provide high capacity and QoS is a network planning problem. Issues in RBCN planning have been investigated in recent research, and for example [4] gives a detailed description for RS planning based on the Manhattan situation (i.e. bad urban environment). However, if time is short and the demand distribution changes then the planned deployment may be sub-optimal. This chapter looks at how to put a "smart" RS into an existing cellular network more efficiently with limited network planning by using:

- simple and cheap forms of phased array antenna with passive direction of arrival estimation;
- a system wide cooperative algorithm to respond to the environment changing and perform geographical load balancing for congestion situations.

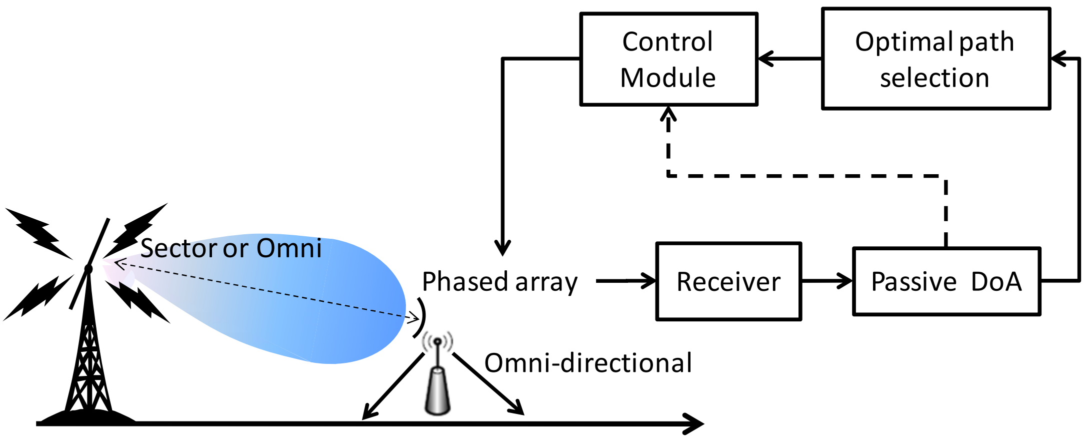


Figure 5.1 The architecture of proposed RS antenna control system

In this work, a RS has two antenna systems, where one is an omni-directional antenna system for communication with the MSs surrounding it, and the other one is a phased array antenna system with electric scan ability for communication with the BS. The antenna control system of RS is illustrated in Figure 5.1. The antenna system used by BSs could be sectorised or omni-directional antennas.

Compared to conventional cellular networks, deployment of ad hoc wireless networks offers low cost, plug-and-play convenience, and flexibility [21]. In this chapter, a self-organizing approach for RBCN is described. Although prior planning will normally always be advantageous, the aim of the approach developed is to allow efficient and quick deployment of a RBCN that is inherently more flexible than conventional deployments, at a reasonable cost. As a special case of ad hoc networks, wireless mesh networks provide flexible routing for the wireless communication to avoid transmission bottlenecks and increase the service probability [92]. The concept of optimal path selection is used in RSs to provide the flexible capability to adapt to a changing environment in real time. The phased array antenna with beam steering lets the RS continually adjust its radiation pattern so that the main beam always points at the direction of a suitable BS receiver or transmitter. The proposed approach is also suitable for emergency situations and when temporary coverage is needed and where automatic deployment and recovery from failure of elements in the system is extremely important.

Here, self-organizing refers primarily to RSs. Similar to the node in ad hoc networks, each RS has self-deployment abilities such as path discovery, connection establishment, scheduling, and topology management. The use of smart antennas and beam forming technology to improve the propagation characteristics in wireless networks has been investigated in [54, 93], but most of that research is concentrated on the smart antenna usage in BSs and entails the expensive costs of smart antenna development. In the approach taken in this thesis, a phased array antenna is assumed to be used at the RS. The potential capability of beam steering lets the main beam of a RS change direction and hence the assignment of a RS to a BS in a coordinated manner. With the help of a passive DoA algorithm, RSs can explore the network environment and find the best directions to connect to nearby

BSs. When a new RS is dropped into a desired coverage area, it will be able to integrate itself into an existing network quickly and reliably.

Wireless networks are also expected to maintain communication service even in emergencies or disaster situations, since reliable communication is necessary for speedy and effective rescue operations. However, current cellular networks may not satisfy the requirement to maintain sufficient connectivity in the aftermath of a natural disaster or emergencies due to communication congestion or damage to facilities. In the proposed approach, actions at the RSs are coordinated with other network components to optimize the radio coverage and radio resource utilization. The performance evaluations indicate that the proposed approach can be expected to be more flexible, reliable, and have better performance than the traditional cellular network.

The rest of the chapter is organized as follows: in Section 5.2, the phased array antenna and electronic scanning technology are described. In Section 5.3, RS self deployment with phased array antennas and passive DoA estimation is described. Section 5.4 proposes a cooperative control approach for congestion and emergency situations so that the service blocking and dropping rates are minimized. In Section 5.5, the performance of the cooperative approach is evaluated in the system level simulator. Section 5.6 concludes the chapter.

5.2 Phased Array Antenna with Electric Scan

5.2.1 Basic Array Theory

Before describing the phased array antenna and electric scanning, the basic theory of the antenna array is covered. An antenna array is an antenna system composed of several similar antennas (a.k.a elements) to obtain the degree of directivity or beam-width required [94]. An antenna array can be arranged in any arbitrary fashion, but the most preferred geometries are linear and circular. A linear array with uniformly spaced sensors (elements) is the most commonly used structure as it is simple to implement. The geometry of a linear array (with elements placed along a line) consisting of N identical elements spaced equidistantly with distance

ΔZ between consecutive elements, assuming that the reference and the first antenna coincide, is shown in Figure 5.2. Elements of the array are assumed to be isotropic point sources.

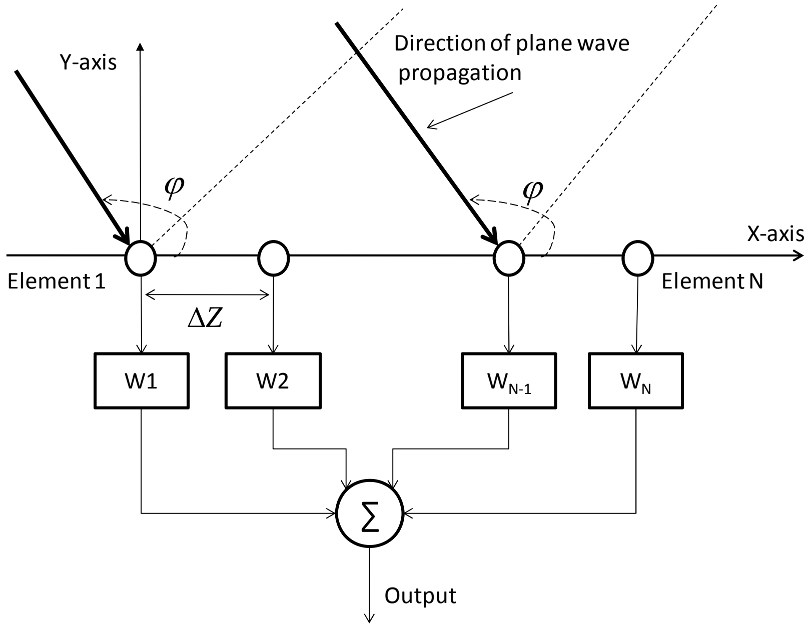


Figure 5.2 Geometry of a linear array with equidistantly spaced identical elements

The mathematical model for the antenna factor (AF) of the linear array is already given by [94]. All the elements are assumed to have identical amplitudes but each succeeding element has a β progressive phase lead for current excitation relative to the preceding one (β represents the phase by which the current in each element leads the current of the preceding element).

$$AF = 1 + e^{j(kdcos\theta+\beta)} + e^{j2(kdcos\theta+\beta)} + \cdots\cdots\cdots$$

$$+ e^{j(N-1)(kdcos\theta+\beta)} = \sum_{n=1}^{N} e^{j(n-1)(kdcos\theta+\beta)} \quad (5.1)$$

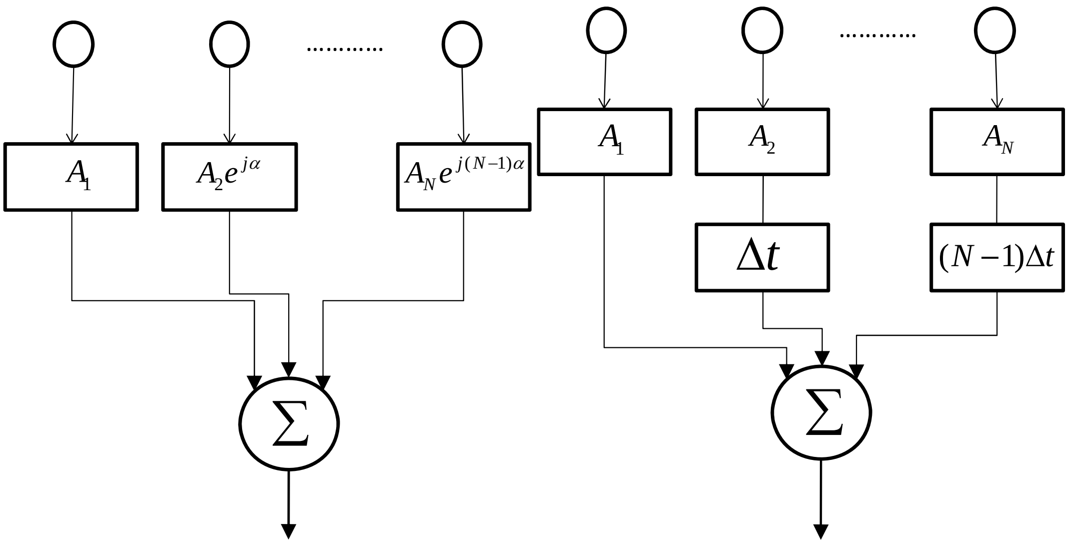
The total radiation pattern for the antenna array is the product of the individual element pattern and the array factor, which is at the phase centre of the individual source with the relative amplitude and phase of the source.

$$\mathrm{E(total)} = [\text{element pattern}] \times \mathrm{AF} \quad (5.2)$$

The element pattern reference here is the radiation pattern from a single isotropic point source. If only the phases are changed, with the amplitude weights remaining fixed as the beam is steered, the array is commonly known as a phased array [95].

5.2.2 Electronic Beam Scanning

For a given array the main beam can be pointed in different directions by mechanically moving the array. This is known as mechanical steering. An engineer can control the main lobe of the antenna system to point in any direction with mechanical steering. However, this will not provide immediate real time control and rapid response to traffic demand changing over time. Beam steering can also be accomplished by appropriately delaying the signals before combining. The process is known as electronic steering, and no mechanical movement occurs [94]. Array beams can be formed or scanned using either phase shift or time delay systems, which is shown in the Figure 5.3. For simplicity only one dimensional steering is considered in the electronic beam scanning below.



(a) Phased Scanning Array (b) Time Scanning Array

Figure 5.3 Two electric beam scanning approaches

Phased Scanning

Beam forming and beam scanning are generally accomplished by phasing the feed to each element of an array so that signals received or transmitted from all elements will be in phase in a particular direction. This is the direction of the beam maximum. This can be accomplished by changing the phases of the signals at the antenna elements. A phased array antenna uses the phase shifters to control either the phase of the excitation current or the phase of the received signals [95]. When all the signals are combined, a beam is formed in the desired direction.

The equation (5.1) can be written as

$$AF = \sum_{n=1}^{N} e^{j(n-1)\psi} \tag{5.3}$$

where $\psi = kdcos\theta + \beta$. If the maximum radiation of the array is required to be oriented at an angle $\theta_0 (0 \leq \theta_0 \leq 180^o)$, then the phase excitation β between the elements must be adjusted so that

$$\psi = kd\cos\theta + \beta|_{\theta=\theta_0} = kd\cos\theta_0 + \beta = 0$$

$$\beta = -kd\cos\theta_0 \tag{5.4}$$

Thus by controlling the progressive phase difference between the elements, the maximum radiation can be directed in any desired angle to form a scanning array.

Time Scanning

The array factor of a time scanned linear array is given by [95]

$$AF = \sum_{n=1}^{N} e^{j(n-1)\psi} \quad where\ \psi = kdcos\theta + \omega\Delta t \tag{5.5}$$

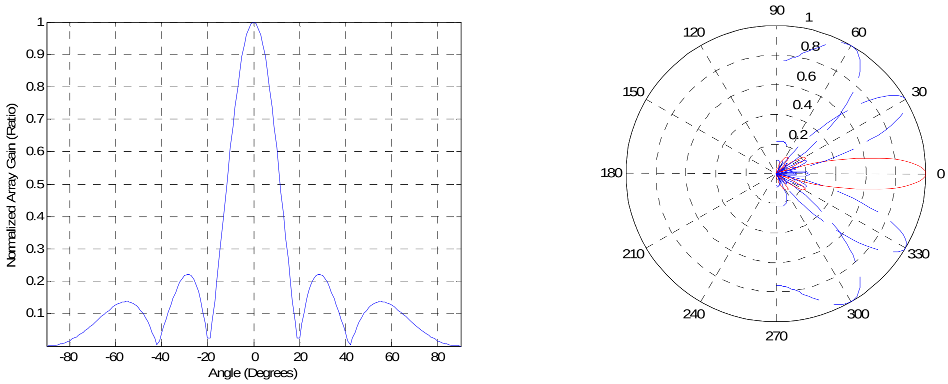
If the direction is set

$$\psi = kd\cos\theta + +\omega\Delta t|_{\theta=\theta_0} = kd\cos\theta_0 + \omega\Delta t = 0$$

$$\beta = -kd\cos\theta_0 / \omega \tag{5.6}$$

Time delays are introduced by switching in transmission lines of varying lengths. It delays the incoming signal from each array element by a certain amount of time, and then adds signals together. If we wish to adjust the angle at which the maximum energy is emitted, we need only to adjust the time delay $\beta = -kd\cos\theta_0 / \omega$ between successive antenna elements.

In principle, the phased array antenna can use both electric scan methods. However, in the research reported in this thesis, only the phased scan is considered. The smaller beam-width, the more cost of the phased array antenna. For an economic design, a 20 degree beam-width phased array antenna is used in the work below. The phase change effects on the radiation pattern for 30 degree steps are plotted in Figure 5.4.



(a) Radiation pattern of a phased array antenna (b) Phased scanning with 30 degree steps

Figure 5.4 The phased array antenna used in the simulation

5.3 Self Deployment with Phased Array Antennas

When designing a wireless network, the main aim of network planning is to provide a cost-effective solution for the radio network in terms of coverage, capacity and quality [19]. The network planning process and the design criteria vary from region to region. When pre-planning is complete, based on the coverage plans, the network operator identifies specific locations for prospective sites. As the RS does not need the wired line link, when setting up such a RS in the prospective area, the main objective is to find which path or direction provides the best link to the BS.

In the proposed approach, a phased array antenna with electric beam steering ability is assumed to be used by RSs to communicate with BSs. The phased array antenna creates an effective antenna pattern at the receiver with high gain in the direction of the desired signal and low gain in all other directions [95]. Hence, the exploitation of directional transmissions could suffice to ensure a wireless backbone with high speed and a high degree of spatial reuse. So, quickly and efficiently finding the optimal path and steering the main beam to this direction for RS setup is a key point in the fast deployment of RS.

The optimal path mainly depends upon radio propagation characteristics in the given area. The signal that is transmitted from the transmitting antenna and received by the receiving antenna travels many complex paths. This signal is exposed to a variety of buildings, passes through different types of terrain, and is affected by a combination of propagation environments [96]. When planning the network, the propagation model needs to be studied carefully. As the radio propagation varies from region to region, creating individual propagation models would be immensely time-consuming. Usage of standard models is economical, but these models have limited accuracy. Therefore, a direction of arrival algorithm is needed for the phased array antenna system to find the optimal path in the variable propagation environments by itself.

In wireless communication, DoA algorithms are usually considered in the context of smart antenna systems to find the direction of target and interference signals. Many DoA algorithms have been developed for smart antenna systems. The main principle is to calculate the direction by measuring the Time Difference of Arrival (TDoA) at individual elements of the array [93]. However, the TDoA measurement requires each antenna element to have a receiver (or sensor) to measure the difference in received phase at each element in the antenna array. A real time digital signal processing (DSP) processor is also needed to adjust the element weights towards some optimization of output signal. These requirements make the smart antenna very expensive, and the DoA algorithm used by smart antennas cannot be used in the phased array antenna systems because of the hardware limitations assumed in this thesis. To find the optimal path, a cheap and simple passive DoA estimation algorithm is used that needs neither costly equipment nor an adaptive beam forming method.

Passive DoA was first developed for military radar systems (passive radar or passive covert radar) used in the navy. It is used to detect and track objects by processing reflections from target sources [97]. It has several advantages compare to the smart antenna DoA algorithms used for the optimal path finding. Firstly, the phased array antenna system only perceives the direct signal from the transmitters

(BSs) when finding the direction of the optimal path. This is very suitable for quick RS deployment and network upgrades, as no output signal means there will be no interference to the existing network. Secondly, the equipment is considerably cheaper than smart antenna systems. It does not need adaptive beam forming schemes nor a real time adaptive DSP processor for each antenna element. Although the precision and speed of the passive DoA is not as good as the algorithms used in smart antenna systems, such as the MUSIC algorithm [98], as will be seen later it is arguably still enough for finding a good, even an optimal path.

Passive DoA is analogous to mechanically steering the main beam direction towards the target direction to maximize the received signal quality. Phase shifters control either the phase of the excitation current or the phase of the received signals. When all the signals are combined, a beam is formed in the desired direction. If the main lobe is directed towards the target site, it will maximize the transmitted or received signal.

The steering locations that result in maximum received power at the RS yield the DoA estimates. It is also possible to use SNR, but the simulations presented here use received power. Using progressive phase differences, the phased array antenna combines the received signal power from its different antenna elements. In the electric scan, the receive power for each steering angle can be calculated by:

$$P_r(\theta) = P_t + G_t(\theta) + G_r(\theta) - P_{loss}(\theta) \tag{5.7}$$

All powers are in dB. $P_r(\theta)$ is the received power at the RS from the BS at degree θ where θ is the angle with respect to a reference direction of arrival at the RS.

P_t is the transmit power of the BS.

$G_t(\theta)$ is the gain of the transmit antenna of the BS at degree θ.

$G_r(\theta)$ is the gain of the receive antenna of the RS at degree θ.

$P_{loss}(\theta)$ is the path loss between the transmitter and receiver.

In free space, if the location of BS and RS is fixed, the transmission power of the BS and the path loss will be constant, ignoring small effects of frequency changes. For each steering angle, according to the above equation, a receive power $P_r(\theta)$ can be measured for this angle (only an angle of 180 degree will be considered).

$$\theta_m = \mathrm{argmax}\big(P_r(\theta)\big) = \mathrm{argmax}\big(G_t(\theta) + G_r(\theta)\big) \tag{5.8}$$

When the phased array antenna electrically scans the network, if the main lobe of the RS points to the BS, the maximum received power can be detected and the corresponding θ_m will be chosen as the DoA.

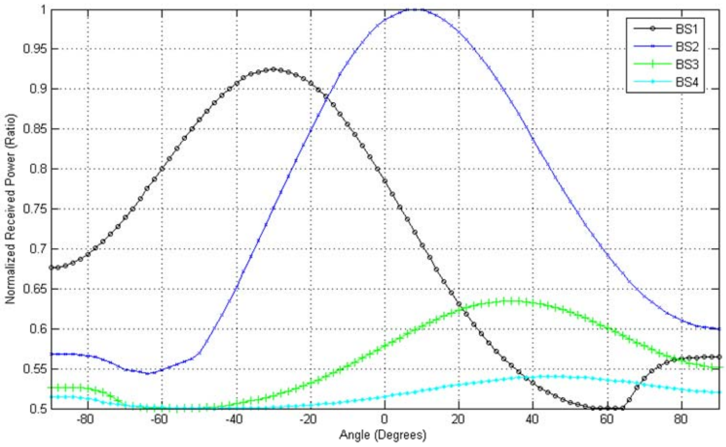
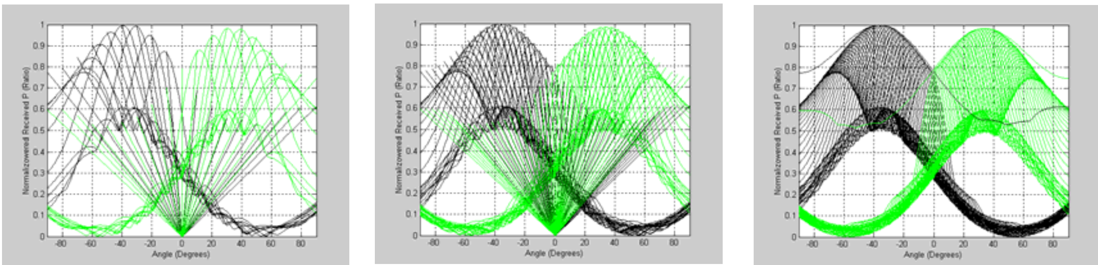


Figure 5.5 The performance of passive DoA : normalized received power from four neighbouring BSs using a 20° beam width.

The performance evaluation of passive DoA with four neighbouring BSs around the RS is illustrated in Figure 5.5, which is obtained using Matlab. The x-axis represents the scanning angle and y-axis represents the normalized received power (the power ratio of the received power to maximum received power). According to the pilot signal, the RS can distinguish between the BSs and calculate the perceived pilot power strengths. As can be seen from the results, a precise determination of angle may not be achieved by the passive DoA because of the wide beam-width. In the real world, when there are more sources of interference it is anticipated that this would makes the peaks broader limiting the accuracy of passive DoA. Obversely, the wider beam width of phased array antenna can make the approach tolerant to inaccuracy of the DoA. The phased array antenna can cover a wide set of angles and the best direction will be selected at the centre of the range of high powers.

The precision of the passive DOA estimation depends on the beam-width of the phased array antenna used by the RS and the angle between the directions where the power is measured. The scanning is performed over a range of angles and the angle between each measurement of power is called the scan step. The smaller the granularity chosen for the step size, the more accurate the determination is in principle, but the phased array antenna becomes more expensive. The performance of passive DoA with different scan step sizes is plotted in Figure 5.6.



(a) 10 degree (b) 4 degree (c) 2 degree

Figure 5.6 The performance of passive DoA with different scan step sizes

With the help of passive DoA and phased array antennas, the following RS deployment process for adding an RS to the network is proposed:

- With simple planning and design, initialize the RS at what appears, a priori, to be a sensible location (fixed).
- The RS electronically scans the network with passive DoA estimation, finds the BSs around, and saves the corresponding directions and signal powers.
- Steer the main beam to the BS with greatest signal power, scan for the downlink channel and establish synchronization with the BS
- Perform the RS initialization procedure for a point-to-multipoint model as defined in [41] to register the RS and set up connections.

The above process has been described in the context of relay networks where there is only one RS between BS and MS. There is no modification to the MAC or PHY layer required. However, by extending the functionality a little, as well as discovering the best BS, a RS has the ability to discover the nearest RS. In a simple policy, the RS would always choose a BS in preference to another RS unless the difference in power exceeds a prescribed bound. So if more than two hops are allowed between the BS and the MS a chain of RSs could be used. Since the network has a cellular structure with BSs already in place, loops of relays can be avoided. When a new RS wants to enter the network but is too far away from the BS (received signal power is below some threshold), the new RS will communicate with the nearest RS. After the new RS has steered its main beam and synchronized with the nearby RS, the RS initialization procedure as defined in the mesh model [41] can be performed.

After RS setup, a connection is created between the BS and RS. The performance of the wireless system in multipath and NLOS positioning is a key issue for self-deployment. The propagation environment for the BS to RS link will be predominately NLOS in many urban core deployments. In direction estimation, multiple incoming signals can result in an incorrect angular position being determined. The passive DoA becomes increasingly unfavourable as the angle between the wanted and interfering paths decreases. For the communication channel, multipath fading will cause amplitude and phase fluctuations, and time delay in the received signals [96]. A common method used to reduce multipath fading is the use of high-gain narrow-beam antennas [93, 97].

To test the performance of phased array antennas on the RS in a multipath environment, a multipath propagation model was implemented according to [99, 100]. The channel model in [101] combines various propagation effects including shadowing, path loss, angular and delay spread, scattering and polarization into one general framework. In this work, as we only consider a single user (RS), the dispersion characteristics of a propagation channel in delay and angle of incidence are described by the directional impulse response function, hence the channel model is given by:

$$h(t, \tau, \Omega) = \sum_{j=1}^{J} \alpha_j \delta(\tau - \tau_j)\, \delta(\Omega - \Omega_j) \tag{5.9}$$

where J represents the number of paths and each path is characterized by an excess delay τ_j and a fast fading α_j;

t is the time passing;

Ω is the direction and $G(\Omega)$ is RS phased array antenna pattern at direction Ω.

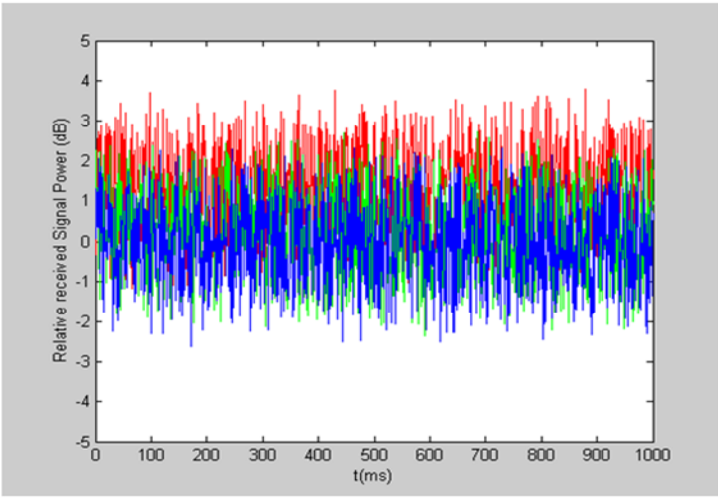


Figure 5.7 The performance of the phased array antenna in multipath environment

The above function is valid if only the multi-paths whose amplitudes are above a certain threshold (e.g., the noise level) are considered [100], so the six highest power multi-paths were used (J=6) in the following simulations. The direction of each path is assumed to be known by the RS in the experiments. Figure 5.7 shows the results of three experiments. The x-axis represents the time in milliseconds and the y-axis represents the relative perceived signal strength. A relative perceived signal strength value of 0dB represents the case when the received power of RS corresponds to a line of sight environment. The red graph displays the perceived signal power for a phased array antenna with a 10 degree beam-width and adaptively steering the beam to the direction of path which has the highest power. The green graph is for an antenna with a 20 degree beam-width, also with adaptive steering. The blue graph is also for a 20 degree beam-width, but steering the main beam to one of the six paths randomly, i.e. each path is equally likely to be chosen. (This is rather pessimistic.) When the beam-width is 20 degrees there is nearly no improvement when the direction of the main beam is dynamically changed to match the best path. This indicates that directing the main beam dynamically in the direction of incidence of the best path has little effect when the beam-width is broad. However, if the beam-width is narrow, then there is an advantage, but this is expensive as the narrow beam equipment and fast phase shifter cost more. So, when the RS is set up in urban areas, a broad beam directed in one fixed direction appears sensible, at least until the RS moves.

5.4 Self Organization in Congestion and Disaster Situations

Network planning work remains nontrivial, is hard to reverse after deployment and does not provide the kind of flexibility that a robust and adaptive network needs to provide [19]. The optimization process should be on-going, in order to sustain efficiency of the network and ensure maximal revenue generation from the network.

Once a radio network is designed and operational, its performance is monitored. The process of network performance monitoring consists of two steps: monitoring the performance of the key parameters, and assessment of the performance of these parameters with respect to capacity and coverage. In the below analysis, dropped calls and connection congestion are the metrics used for performance assessment.

In a Mobile WiMAX network, each connection (both management and data) is identified by a Connection ID (CID) [41]. There is no routing required; data is transmitted solely between the BS and the MS. When multi-hop communication is to be used in a cellular network, routing is very important for the network performance as one or more RSs may exist between a BS and an MS in the relay based cellular networks. In relay enhanced cellular network architectures, a BS has the capacity to concentrate the traffic of multiple RSs and to directly serve mobile users within its own coverage footprint. With the help of beam steering, the RS can link to a BS according to the location and propagation environment in order to optimize the communications channel and radio resource usage.

To balance the load across BSs (and RSs), the traffic conditions in each cell and its adjacent cells must be considered. Notionally the BSs need to cooperate with each other to provide the service without generating any coverage holes. (The term notionally is used as it may be via a central coordinator.) When a BS is heavily loaded, it has less potential to serve new mobile users and so call blocking increases. Overloaded BSs therefore need to seek help from neighbouring BSs or RSs. As the ASA server can know the traffic load distribution for the entire network (or large area) and monitor dropped calls and connection congestion, it is sensible to implement this controller as a centralized optimization algorithm at the

ASA to maximize the bandwidth usage over all BSs. The optimization process can be considered as the process of finding a simple routing algorithm based on QoS and load balancing. The routing algorithm is in the central control (ASA server) and the paths are formed using phased array antenna beam steering. The BS establishes the path by either informing all the RS along the path of relevant path information or embedding path information as part of connection management. The process of optimal path management is:

1. According the perceived pilot signal strength, a MS will request the connection to the BS or RS.
2. If a cell or an area has a high traffic load (or high blocking rate), the ASA server will run the optimization algorithm and reassign the RSs according to the optimization assignment scheme.
3. The ASA server sends management messages to the RS and BSs, and hands over the RS to the target BS.
4. The phased array antenna in the RS will steer the direction of its main beam towards the selected BS.

The efficient group handover scheme in [91], also used in the adaptive tilting approach (Chapter 4.3), is used when the RS is reassigned to the target BS.

Rather than finding the physical location, the user demand (bandwidth requirement) and a measure of "perceived" distance between the MS and BS or the MS and the RS are used in this approach. The perceived distance performs a mapping of traffic distribution in the real world to a virtual domain where path loss determines the 'distance' between BS or RS and the MS. It represents a normalized distance in the simulation propagation model and the pilot signal strength feedback from the MS is used to compute the radial distance from the BS or RS in the virtual world. This mapping is valid even in complex propagation environments as we are simply establishing a distance in a uniform virtual world.

The optimization algorithm is straightforward and does not necessarily need to be ultra-fast, as the need for re-allocation can be anticipated. When triggered the ASA has to

1. Sort the BS IDs according the traffic load, in ascending order.

2. Take each BS in order and for each BS use the simulated model of the environment to establish how many RSs and MSs as it can serve, so the more lightly loaded BS will typically be serving more RSs.
3. Find the blocking rate for all the BSs; save the routing path for each RS.
4. Repeat steps 1-3 until the configuration does not change or the iteration limit is reached.
5. Use the best assignment scheme.

In a real network, the optimization algorithm would not need to be executed frequently and not over the entire network. When an ASA server detects that some BSs or coverage areas have a high blocking rate, reconfiguration only needs to be performed for those areas. This local optimisation will be faster than the whole network optimization and decrease the handover frequency of RSs.

Experience from operating over a period in a specific environment, can be used to describe the traffic demand distributions along with their associated RS assignment schemes and the performance recorded. If traffic demand distribution around a relay station for example, can be predicted reasonably accurately then this can be used to make proactive changes in the topology of the network, before the blocking rate becomes high. In this way the network does not wait for unacceptable performance before making changes. Such intelligent self-adjustment can be put into each RS based on past cases experienced. This can also decrease the computer requirements for the ASA server, but the ASA server does have to monitor demand as it has overall control.

A requirement for future wireless networks is to be capable of maintaining reliable communication services even in emergencies, since quick, accurate damage assessment information is necessary for speedy and effective rescue operations. A simple scheme is put forward for RSs to let the proposed network tolerate network component failure. The proposed scheme is based on the self deployment and cooperative control mechanism described above and tries to minimize the number of users affected by network component failures.

In an emergency or disaster situation, some BSs may fail as the equipment is broken or the backhaul is severed. RSs and MSs in the coverage of the failed BSs cannot get service. If the link between a RS and a BS is broken down, the RS will run its self deployment process immediately and find another neighbouring BS that can help. The electronic scan and passive DoA will be used to find the nearby BSs. It there is a surviving BS around, the RS will steer the main beam and request service from the surviving BS. Heavy communications may cause congestion in the surviving system, so cooperative control for geographic load balancing is still needed to coordinate with the surviving cells to decrease the call blocking rate.

Although the scheme is simple and the MSs near the failed BS may still not regain coverage, the MSs in the coverage of RSs can tolerate the BS failure and obtain services from the surviving system. The temporary coverage is very important for the rescue operation and sometimes this can save lives.

Hosting the cooperative controller at the ASA server is consistent with current commercial deployments where it is expected that the intelligence will be hosted in semi-centralized systems much as they are today. However, the intelligence could be distributed to the BS and these can act cooperatively but autonomously. Distribution could be particularly important in emergency contexts. The criteria for distributing the intelligence will be based on the degree of mobility of the nodes (BS and RS), the nature of the backhaul, the nature of inter BS communication and the required resilience.

5.5 Simulations and Results

A 4×4 hexagonal model (16 cells) is used to form a Mobile WiMAX network. In each cell, there is a BS in the centre and six fixed relay stations around the BS (as illustrated in Figure 5.8). BSs are directly connected to their backbone network. Fixed relay stations are located within BS coverage. Each relay station is located within the coverage of two BS. A RS can only connect to one BS at a time. In the simulation described below, a 5 MHz bandwidth profile is used in each cell and other parameters settings are the same as in section 3.4.2.

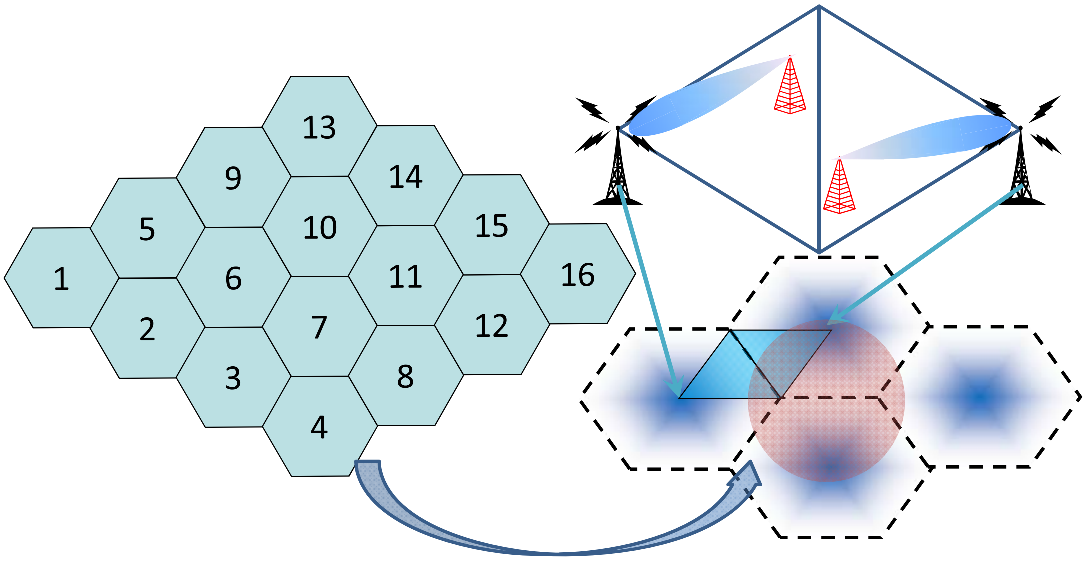


Figure 5.8 The scenario structure in the simulation network

5.5.1 Traffic Congestion Simulation

In this experiment, to emulate traffic congestion, a scenario that builds up traffic hotspots is used. Each cluster has a population size of 150 active MSs. As the algorithm results are sensitive to the location and distribution of hotspots, in the results used below, the hotspots are put near the boundary of the cells (such as cell 3 and cell 8 in Figure 5.8) and the size of a hotspot area is approximately the area of one cell. There is no torus assumed around the simulation network to avoid the boundary effect. The required region in many deployment situation will be not larger. Using the wrapped hexagonal cell, the inaccuracy of measurement can be limited even though the number of cells in the simulated network is small [102]. In the simulation, the location of a hotspot is far from the boundary of the simulation network and the adaptive configuration of the RSs will not change their own radio coverage, so that the boundary effect for the experiment is very limited. The constant number of active MSs and the MSs not in any hotspot areas are still uniformly distributed and will also mitigate the boundary effect.

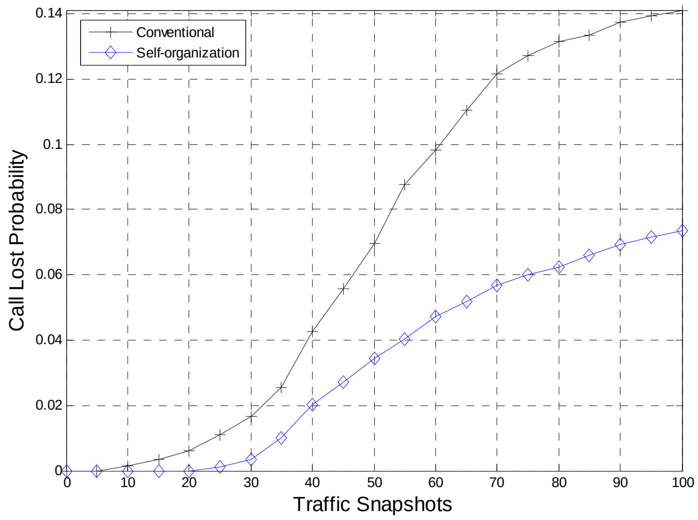


Figure 5.9 The call lost probability in a traffic congestion scenario

Simulation results assessing call lost probability are shown in Figure 5.9. There are two simulation experiments: one on a conventional cellular network and one on a self-organizing network. Both networks have the same number and distribution of BSs and RSs, but the conventional network has constant RS assignment for each BS. The RS assignment for each BS is changeable in the self-organizing network. At the beginning of the simulation, when the traffic is nearly uniformly distributed, the blocking rate of the load balancing network is the same as the conventional one, just as one would expect. As the hotspots form, the difference between a conventional network and a load balancing network becomes clearly seen. Indeed evolution from the uniform distribution is chosen as it forms a control and also indicates how the relative performance improves with respect to the degree of heterogeneity. The cooperative RS assignment using phased array antennas always outperforms the conventional network. As the traffic becomes more heterogeneous, the performance improvements are more significant.

For a further investigation for the radio resource utilization (RRU) improvement with the cooperative algorithm, the difference of RRU rate in this scenario is plotted in Figure 5.10. The difference of RRU rate for each traffic snapshot is defined as the difference between the maximum cell RRU rate and minimum cell RRU rate in the two simulation networks, expressed as below:

$$D_{RRU} = \max_i RRU_i - \min_i RRU_i \tag{5.10}$$

where the RRU_i is the cell RRU rate for cell i, calculated as the allocated bandwidth divided by the total bandwidth in the cell i. The number of active MSs in the simulation is carefully selected to fully use the available spectrum when the traffic is uniformly distributed. So, the difference in RRU rate expresses the waste of spectrum resource.

From Figure 5.10, it can be seen that when the traffic demand is still nearly uniform at the beginning of the scenario, the radio resource utilization for each cell for the two (simulated) networks is very similar. When MSs move to form hotspot areas, the difference of RRU rates increases with the traffic snapshot index. The cells in the hotspot area have higher demand, and the RRU rates are nearly hundred percent. In the cells with less demand, the RRU rate become lower and some spectrum resource is not used. With the help of geographic load balancing, the limited spectrum is used more efficiently. In the most heterogeneous situation, the spectrum efficiency improvement is as much as 60%.

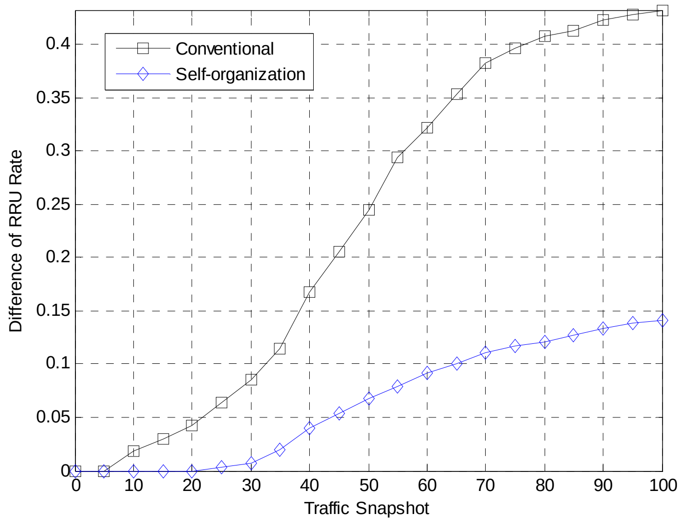


Figure 5.10 The difference of RRU rate for each traffic snapshot

This principle of this cooperative algorithm is similar to the adaptive tilting one (described in Chapter 4.3). The main idea is to split the hotspots with the help of RSs. With the help of the phased array antenna at the RS, the RS has more ability to choose the best BS to connect to. So it is not surprising that the gains are more than in the tilting experiments. In a realistic network, the locations of cells and RSs will not be the orderly distribution in the simulation. The RSs would be put in high traffic load areas to increase the system capacity and there will be more BSs around.

The cooperative approach is not based on specific assumptions on the location of BS or RS, and in fact the assignment of RSs will be more flexible as there could be more BSs to choose to connect to. The tilting approach needs the two neighbouring BS antennas that cooperate in the tilting to be at the angles required to cover the RSs and so avoid a hole in the network. The approach is this chapter allows adjustment in more general failure cases.

5.5.2 BS Failure Simulation

In a disaster situation, such as immediately after an earthquake, a wireless network in such an area may be damaged, e.g. severed fibre or tilted masts. Several BSs may not be able to link to the backhaul but the self-organizing RSs could still search for help from the neighbouring BSs.

According to the relative location of failed BSs, two experiments were designed to test the proposed approach performance on a BS failure situation. Both experiments are based on the same network configurations as the congestion experiment above but now also one to four BSs are allowed to fail. MSs in the two experiments are uniformly distributed. In one experiment it is assumed the failed BSs are not adjacent and in the other the failed BSs must be neighbours. The latter is a more stringent requirement.

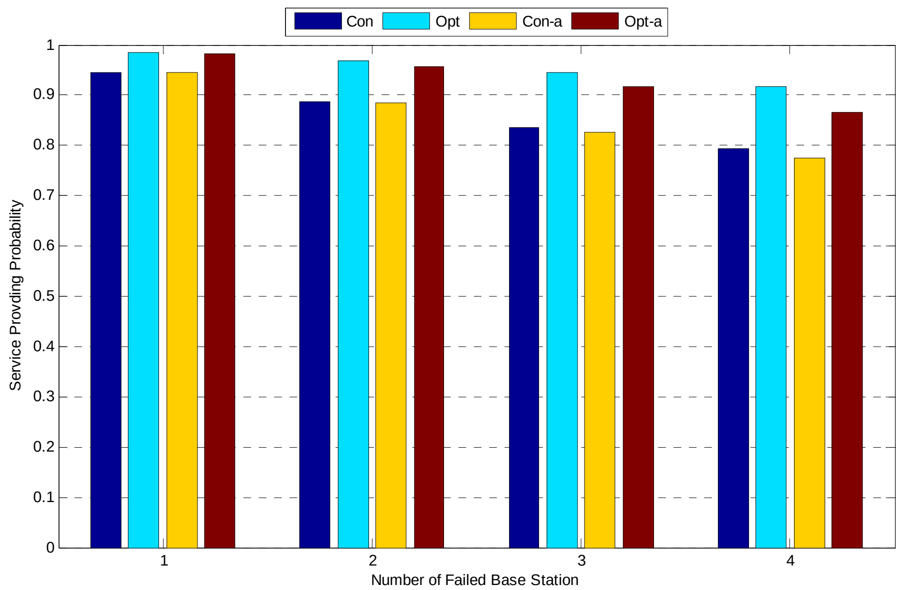


Figure 5.11 The service provision probabilities for the BS failure scenarios

The simulation results showing the service provision probabilities are plotted in Figure 5.11. The x-axis represents the number of failed BSs and the y-axis represents the probability of service availability. The “Con” label signifies the conventional network and “Opt” represents the network with self-organizing RSs. The postfix “-a” means the failed BSs are adjacent in the experiment. The results show that the proposed network is capable of maintaining service provision above 90% even when four BSs have failed. If a larger area of wireless network is broken down (where the failed BSs are adjacent), RSs in the disaster area cannot find help from the neighbours. The self-organizing approach can still improve the service providing the service provision probability of around 90% compared to 75% for the conventional network.

5.6 Summary

A potentially cheap approach to improving QoS that exploits the flexibility gained by allowing a RS to find optimal BSs has been described. Such an approach could be deployed quickly into existing cellular networks. The efficient set up of wireless relay networks in existing cellular networks using only rudimentary network planning was a main concern in this chapter. The approach described allows the network to adapt if a BS fails, a BS or RS moves, MSs move, or simply in response to demand changes. It inherently provides flexibility and hence resilience to faults and attacks. This approach is very suitable for scenarios that need quick deployment and have high data rate requirements, such as immediate emergency coverage (earthquakes, typhoons) or temporary coverage (Olympic Games, conferences).

With adaptive antenna and cooperative algorithms, a relay based cellular network can be more efficient, more tolerant of the environment changing and can discover its environment. The hope is that this can reduce the cost and even the need for network planning. This topic is addressed further in Chapter 6.

Chapter 6 RADIO COVERAGE AND HANDOVER ISSUES FOR COOPERATIVE CONTROL

6.1 Introduction

In previous chapters, owing to the need for simplifying assumptions to handle the complexity of network scenarios, the effects of signal fluctuations are not accounted for, and the consequent handover effects caused by cooperative control for the whole network are not considered. However, when cooperative control is applied in the physical network, the network operators or engineers need to consider these practical issues. Firstly, when considering the flexibility of radio coverage provided by the cooperative control techniques described, network holes are the main concern, as increasing capacity in certain locations can adversely affect the coverage in other areas. When an antenna configuration is changed, it may not be able to cover a region or the received power cannot support link set up. It is difficult for a network operator to find a coverage hole as a mobile user in a blind spot cannot be detected. Also, normally, only when a mobile user finds service is not available, can it determine that it is in an uncovered area. The collaboration between the cells also depends on the accuracy of radio coverage prediction, so a reliable radio coverage prediction model is required to achieve the collaboration and mitigate the complex propagation effects for optimization performance. When an antenna configuration is changed, the radio coverage will be changed and some boundary MSs might be involved in handover processes. The burst in handover requirement across the entire network and different handover algorithms may affect both the network performance and optimization performance. In this chapter, the radio coverage and handover issues are analysed to determine the potential risks and to consider how to make the operational decisions suggested by the cooperative control algorithms when applied in a real environment.

Cooperative control algorithms act to minimize the un-served MSs and improve the spectrum efficiency. The algorithms described in Chapter 4 and Chapter 5 did not depend on a predictive model. Only the mobile demands and rough locations were required for the optimization input, which could be easily achieved by the feedback from the MSs. However, the accuracy of radio coverage is based on the simple propagation model assumption in the simulations.

To improve the accuracy of radio coverage and avoid a network hole occurring in the realistic environment, a prediction model is proposed based on the received signal power feedback of the MSs. The proposed prediction approach exploits the difference between the received signal power in different antenna configurations to investigate the relationship between the antenna configuration and the propagation environment. This is done through a statistical approach that builds the prediction model, taking path loss and shadowing into account. The statistical model, or to be more precise the set of statistical models, is determined using data collected during a training stage. Each model predicts, for a small area, the mean received signal power and the signal outage probability. Linear regression and logistic regression are used to estimate the coefficients of a linear model where the independent terms are the proposed parameters for the propagation model. This does not mean that a linear propagation model is being assumed, only that inside each small area a generalised linear model in terms of the independent variables is adequate. The prediction model is verified with theoretical and simulated radio coverage comparisons, and the survey data for prediction accuracy is investigated in different scenarios. Ideally these comparisons would be based on real signal strengths collected over a geographical area, using different antenna configurations, but such data would be very difficult to obtain unless one has access and control of a real network. In principle, the proposed statistical prediction model is a general model for radio coverage prediction applicable to different cooperative control mechanisms, and is not just limited to the study of the kind of RBCNs described in this thesis. The capability of an antenna system affects the flexibility of the antenna configuration and hence the flexibility of the radio coverage. Specifically, a conventional 60 degree sector antenna is assumed to be used by the BSs and omni-directional antennas at the RSs. Flexibility is obtained by changing the transmit power at the BSs and RSs. For the handover issues, the burst handover is handled by a restructuring schedule scheme according to the traffic load distribution. Three handover algorithms defined in the Mobile WiMAX standard [41] are investigated according to the handover rate, utilization and overhead.

The rest of the chapter is organized as follows: Section 6.2 presents the statistical prediction model for radio coverage, the model verification and survey data analysis. Section 6.3 gives a detailed description and investigation of the handover effects when cooperative control is applied. Following this, the two approaches applied in the physical network and some standard issues related to the practical usage of cooperative control are described. Finally, the conclusions are given in Section 6.4.

6.2 Radio Coverage Prediction for Cooperative Control

In order to apply flexible radio coverage in a realistic environment, an approach for learning an efficient and accurate radio coverage prediction model is described in this section. The relationship between the radio coverage and the antenna configuration for a local area (which is assumed to be known and created based on the simulated antenna model) is discovered and then used for radio coverage prediction.

6.2.1 Statistical Prediction Model

The objective of network planning is to make sure there are no network coverage holes in the desired coverage area and to provide seamless handover [103]. From the PHY layer aspect, this means that there should be enough radiation power distributed in the planning area to support reliable transmission and achieve the minimum SNR requirements. Cooperative control could be thought of as an optimization scheme for radiation power distribution.

The concept of radio coverage is simplified here, to only the "coverage requirement" for adequate reception. This simply requires that the instantaneous received signal at the MS and the service station (BS or RS) be greater than some minimum signal level. This "simplified" radio coverage mainly depends on the propagation environment and antenna configuration. In the previous chapters, an idealized free space propagation model was used in most of the simulations. Based on the relationship between the path loss and distance, the radio coverage was very predictable. However, for most practical channels, the free space propagation model is inadequate to describe the channel and predict system performance. The mean received power and the signal variance affect the service area and prediction accuracy. The received signal is more likely to suffer from the effects of the channel in the form of short term and long term fading, also known as fast fading and shadowing [44, 99].

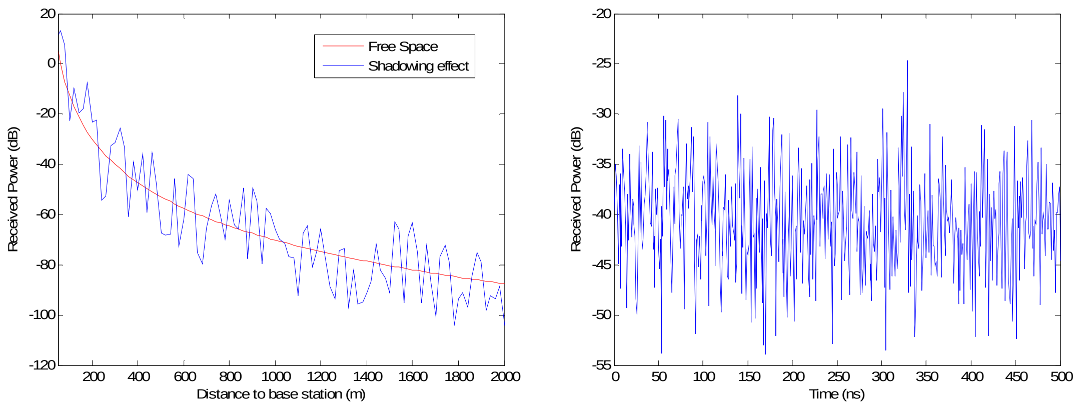


Figure 6.1 The different effects of the propagation components

The overall propagation in a realistic environment can be characterised as having three basic components: attenuation, shadowing (slow fading) and fast fading [104], which are illustrated in Figure 6.1 and obtained using Matlab. The sum of these components describes the resultant overall path loss between transmitter and receiver. The popular stochastic path loss models assume that the magnitude of a signal that has passed through the channel model will vary randomly according to some stochastic distribution, such as the *Rayleigh* or *Rician* distribution. The accuracy of the model calculation depends strongly on the accuracy of the topographical data and predictions between different models are often not consistent. These models are often used for general types of propagation environment simulation, such as an urban or rural environment. Complex terrains, different objects around the transmitter and receiver, and weather all affect the real world propagation environment making it difficult to predict. A radio coverage survey is commonly used in network planning [19, 57] for radio coverage analysis. This takes a long time for sample collection and only reflects the variability at the time of the survey. The radio coverage survey is also mostly based on a constant antenna configuration such as fixed transmission power and radiation pattern.

For shadowing effects, the received signal power change imposed by the channel can be considered roughly constant over the period of use at a constant location. For fast fading, the amplitude and phase change imposed by the channel varies considerably over the period of use. Therefore, the distance between the transmitter and the receiver is used for the long-term fading description and the fast fading changes with time.

$$\mathrm{P_{loss}} = \mathrm{P_{attenuation}} + \mathrm{P_{shadowing}} + \mathrm{P_{fastfading}} = \mathrm{P_{long}(d)} + \mathrm{P_{short}(t)} \qquad (6.1)$$

All powers are in dB. The received signal power depends directly on the transmitted power level in addition to the orientation and polarization properties at the transmitter and the receiver, as described by the antenna radiation patterns, and the signal attenuation (including path loss, shadowing and multipath effects) predicted by the propagation model [105]. In theory, for a specific radio link when the number of electromagnetic paths is *N*, the received power at time *t* can be calculated as:

$$P(t)_r = \sum_{n=0}^{N-1} (P_t \cdot V_n + G_t + G_r - P_{loss}^n(t)) \qquad (6.2)$$

where $P(t)_r$is the received power of the mobile user, P_t is the transmit power , G_t and G_r are the transmit and receive antenna gain values, $V_n$is the radiation pattern value of antenna in transmission path n , $P_{loss}^n(t)$ is the sum of the signal attenuation in transmission path n at time t.

If the distance between the transmitter and received is also included in equation (6.2) and transmit power and radiation pattern are assumed constant, substituting equation (6.1) into equation (6.2), gives

$$P(d,t)_r = \sum_{n=0}^{N-1} \left(P_t \cdot V_n + G_t + G_r - \mathrm{P_{long}(d)}\right) - \sum_{n=0}^{N-1} \mathrm{P_{short}(t)} \qquad (6.3)$$

Therefore, the received signal power can be divided into two parts. One includes the longer term effects that depend on the distance and location with respect to the BS or RS, which can be assumed to be constant during the received transmission. The other one is the short term effect, which changes during the transmission time.

To accommodate the different effects of long term and short term propagation, two parameters, mean received signal power and signal outage probability, are used by the radio coverage prediction in the proposed model. Firstly, the mean received power gives a prediction of received signal level over the area. Here the effects of shadowing and fast fading disappear through averaging. The variability of signal strength, which is determined by the slow fading and fast fading, is expressed by a signal outage probability. The signal outage occurs if the signal drops below some specific power threshold or some noise level, expressed as the SNR, which is required for reliable communication. The received power used in the radio coverage analysis can be easily achieved by user feedback in response to the antenna configuration.

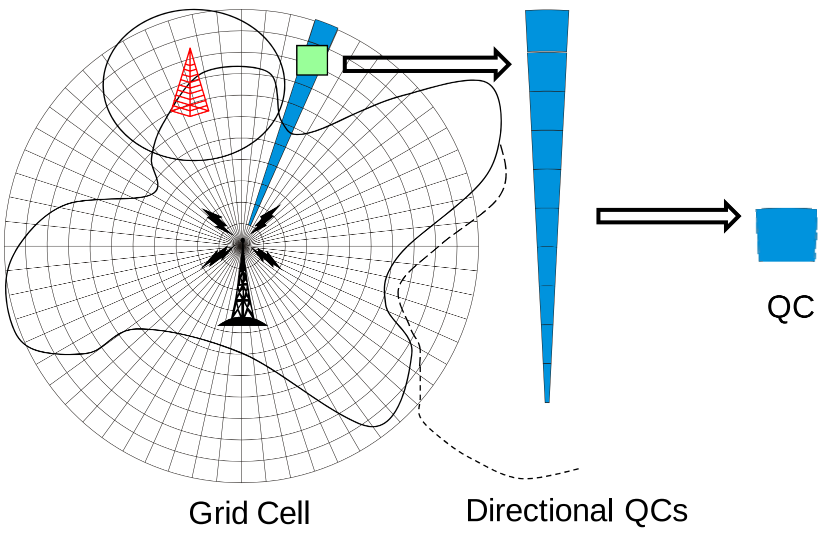


Figure 6.2 The wireless network coordinate system with QC concept

Due to the limited capability of the antenna system and complicated propagation environment, the radio coverage will be modelled in terms of a small area rather than an exact location. The cell coverage is divided into a grid of small areas and each cell of the grid is called a quantization cell (QC), which is illustrated in Figure 6.2.

Each model is determined using data collected in a training stage over a small area and each model predicts for that small area. So, the proposed model is a set of statistical models for each QC in a cell. Each model predicts, for a small grid area, the mean received power of radio coverage and the probability of the signal outage. The exact values for a specific location are not predicted as the value is a variable changing with time and it is the aggregate values that are useful for the network optimization. It is assumed that the set of received power measurements collected over the training period is a representative sample. Based on the capability of the simulated antenna model and the accuracy of network based location estimation algorithm, in most of the simulations below (except the survey data analysis), the sample sizes are selected and listed below:

- Number of samples collected in each QC=30
- Number of QCs in each cell=600
- Number of regression antenna configurations used=10

For each QC, the mean received power denotes the average signal quality level in this area and the signal outage probability describes the variability of signals in this area. This variability is due to both fast and slow fading, but it primarily reflects slow fading. With sufficient training data, signal outage probabilities are typically more accurate than their predicted received power counterparts because they directly handle the uncertainty of received power measurements. The number of QCs needed in practice is a function of the terrain. Highly variable propagation environments would benefit from more QCs. Observation of performance in the real world is needed to determine the best number of QCs to use. A mechanism for merging QCs with low prediction error and splitting a QC with a high prediction error could be constructed if real data were available. Similarly, the number of antenna configurations, depends on the observed predictive accuracy that can be obtained by interpolation, and this could be determined using the errors found if a real system were used.

According to the above analysis, the statistical prediction models have been designed with the following features.

1. Propagation environment models for local areas are used to predict mean signal strength, learned using sample data collected during the survey or training stage;
2. Signal outage probability models based on the survey data, training and feedback, are used to estimate the variability induced by shadowing and fast fading;
3. Fast response to environment changes using feedback from mobile users and data of network performance.

The regression idea used in the proposed prediction model is similar to the adaptive temporal radio maps scheme [106, 107] and cell boundary determination method [108], that also used regression methods for the survey data analysis and radio coverage prediction. However, they only give a prediction for the large-scale path loss and do not take into account the actual multipath fading or shadowing variation. Only received signal strength is used in the above literature and the samples are based on reference points.

The potential signal quality prediction in individual QCs according to the current antenna configuration is the main objective of the prediction models. It allows cooperative control algorithms to predict signal coverage based on hypothetical settings, determine achievable data rates, compare performance of different radiation pattern schemes and hence make the best choice of action. The prediction approach will take time to collect and analyse the signal feedback to build the models. Nevertheless, if the prediction model is built and working in the operation stage, this should lay a good foundation for improving performance of network optimization and allow fast response to the propagation changes in the environment. For the downlink, the optimal antenna radiation pattern can increase the transmission power efficiency, where the radiation power distribution is based on the traffic load distribution. For the uplink, as the radiation pattern can be more sensitive to the propagation environment, this may increase received power and decrease the multipath effect for the channel model.

6.2.2 General Regression Methods for Prediction

A typical lognormal path loss model has been demonstrated in (6.4), which uses a parameter *n* to denote the power relationship between distance and received power [105]. As a function of distant d, path loss (in decibels) is expressed as

$$\mathrm{PL(d)} = \mathrm{PL(d_0)} + 10 \cdot \mathrm{n} \cdot \log\left(\frac{\mathrm{d}}{\mathrm{d_0}}\right) + \mathrm{X_\sigma} \tag{6.4}$$

where $\mathrm{PL(d_0)}$ is the power loss at a reference distant $\mathrm{d_0}$and $\mathrm{X_\sigma}$ denotes a zero mean Gaussian random variable that reflects the variation in average received power.

Seen from the equation (6.3), the receive power of mobile user is not only related to the physical environment, but also the antenna setting in the transmit antenna (the antenna gains of receiver and transmitter are constant). The mean received signal power will be a function of antenna configuration and distance between the receiver and transmitter, expressed as:

$$\mathrm{P_r} = \mathrm{F(P_t \cdot V_r, d)} = \mathrm{P_t \cdot V_r + G_t + G_r - PL(d_0)} - 10 \cdot \mathrm{n} \cdot \log\left(\frac{\mathrm{d}}{\mathrm{d_0}}\right) \tag{6.5}$$

The random variable $\mathrm{X_\sigma}$ in (6.4) is averaged to disappear. Because of the small area restriction and lognormal feature of propagation [109], a linear relationship is assumed between the mean signal strength and antenna configuration in each grid area (the distance could be assumed to be constant as the location of a QC relative to a BS or a RS is fixed). For the investigation of relationships between variables, regression analysis is a common statistical tool. General linear regression and logistic regression are used to estimate a linear model in terms of the estimated parameters for the propagation model. This does not mean that a linear propagation model is being assumed. Each model is *only* assumed linear in its coefficients. It can have non-linear terms.

For parsimony, two simple linear predictors for QC *i* mean received power are evaluated below.

$$\mathrm{P_r^i} = \beta_0 + \beta_1 \cdot \mathrm{P_{BS}} + \beta_2 \cdot \mathrm{P_{RS}} \tag{6.6}$$

$$\mathrm{P_r^i} = \beta_0 + \beta_1 \cdot \mathrm{P_{BS}} + \beta_2 \cdot \mathrm{P_{RS}} + \beta_3 \cdot \mathrm{P_{BS}^2} + \beta_4 \cdot \mathrm{P_{BS}} \cdot \mathrm{P_{RS}} + \beta_5 \cdot \mathrm{P_{RS}^2} \tag{6.7}$$

where P_r^i is the expected received signal power of QC i, P_{BS} and P_{RS} are the transmit power of BS and RS. For each QS i, the tuple (P_r, P_{BS}, P_{RS}) is called an observation. Given n independent observations $(P_r^n, P_{BS}^n, P_{RS}^n)$ of the antenna configuration and the response received power (here the superscript indicates observation numbers rather than powers), the linear regression model becomes an $n \times p$ system of equations and the β coefficients can be estimated using linear regression methods.

Multivariate linear regression is useful for the cooperative control as it:

- Gives which terms β_j (transmit power of BS or RS) have greatest effect on the response
- Finds the direction of the effects (signs of the β_j)
- Predicts unobserved values of the response (the received power for a new antenna configuration)

Network based location estimation methods utilize the radio signals transmitted between a mobile user and a set of BSs with known locations. With network based location estimation (such as the triangle method), the approximate location can be calculated for each mobile user with signal feedback. As described above, the location information is mainly used to find which QC a MS actually belongs to. Therefore, the analysis could feasibly work during the operational stage for continuous analysis and fast response to changes in the propagation environment.

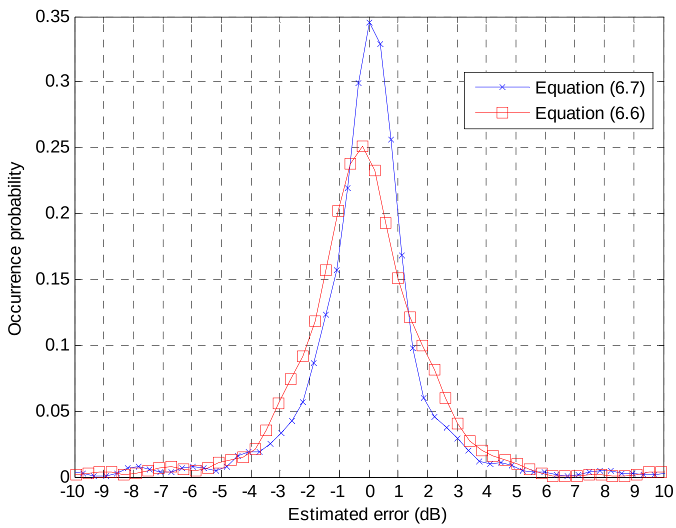


Figure 6.3 The error distribution probabilities of two linear models

In Figure 6.3, the error distribution, viz. predicted value minus the observed value, is plotted using Matlab. The result is based on error calculations for each QC of the entire cell. The x-axis is the value of the estimated error and y-axis is the error occurrence probability based on this error estimation test, the second linear function (6.7) can achieve more accurate predictions than the first one (6.6) with a smaller bias and a smaller standard deviation of the error. Hence, the analysis and simulations below are all based on the linear function (6.7).

The mean received power prediction represents the long term propagation characteristic of each QC, which includes the signal attenuation and shadowing effects. The cell radius could be estimated based on the average value of the received signal power, but an important question is how to determine the "reliability" of the radio coverage over the area. The signal outage probability is used as a measure of the probability of reliable communication in each QC. The signal outage probability is predicted using logistic regression with an error that is assumed to be binomially distributed. According to the signal outage probability concept [110], the signal outage probability for QC i is expressed as:

$$P_{SO}^{i} = E\left(\frac{y}{C}\right) \tag{6.8}$$

where C is the number of samples collected in QC i, and y is the number of MSs whose received power is below the threshold, i.e. $P_{received} < P_{threshold}$. Based on the previous proposed equation (6.7), the logistic model is expressed is:

$$\ln\left(\frac{P_{SO}^{i}}{1 - P_{SO}^{i}}\right) = \beta_0 + \beta_1 \cdot \mathrm{P_{BS}} + \beta_2 \cdot \mathrm{P_{RS}} + \beta_3 \cdot \mathrm{P_{BS}^2} + \beta_4 \cdot \mathrm{P_{BS}} \cdot \mathrm{P_{RS}} + \beta_5 \cdot \mathrm{P_{RS}^2} \tag{6.9}$$

where $\mathrm{P_{so}^i}$is the signal outage probability for QC i. Expressed in terms of P_{SO}^{i} this is a non-linear relationship and requires an iterative solution. Therefore, a general linear model can be built between coverage odds and antenna configuration ($\mathrm{P_{BS}}$, $\mathrm{P_{RS}}$). The predicted signal outage probability could be calculated based on the estimated β^* as in the equation below:

$$\mathrm{P_{SO}^i} = \frac{1}{1 + \exp(\beta_0 + \beta_1 \cdot \mathrm{P_{BS}} + \beta_2 \cdot \mathrm{P_{RS}} + \beta_3 \cdot \mathrm{P_{BS}^2} + \beta_4 \cdot \mathrm{P_{BS}} \cdot \mathrm{P_{RS}} + \beta_5 \cdot \mathrm{P_{RS}^2})} \tag{6.10}$$

These models are fitted to the whole area of cell coverage (including the BS coverage and RSs coverage). As the received signal power gives a prediction for the cell radius and the signal outage probability represent the reliability of coverage, a combination could take advantage of the two prediction mechanisms. In practical usage, the two prediction models will be considered independently with their own requirements.

At first, the predicted mean signal power for each QC is used to give a rough prediction of the boundary of radio coverage that could be achieved, which could be used as the benchmark for coverage prediction. Based on the required reliability of the coverage (e.g. 95%, so the signal outage probability will be 5%), the variability around the boundary could be estimated. The cell radio coverage probability could be estimated by the cell edge coverage probability. This is based on the fact that usually the signal strength degrades and varies the most at the cell border. So, if there is reliable coverage at the cell edge then it is expected that there will normally be enough coverage in the whole cell.

The proposed prediction approach does not depend on such assumptions. This is useful as in practical usage, the deep fading areas will be not just around the cell edge, and the signal outage probability computed for each QC could be used to estimate these areas.

The accuracy of combined prediction and mean received signal power prediction are compared. The cell coverage and the error distribution are plotted in Figure 6.4. The estimated error is the simulated coverage distance minus the predicted coverage distance. The mean error for the combined prediction is around 55m, and the mean signal power prediction is around 95m (these results are reasonable as they are correspond to the data in [111]. though that work only used mean signal power regression for the radio coverage prediction).

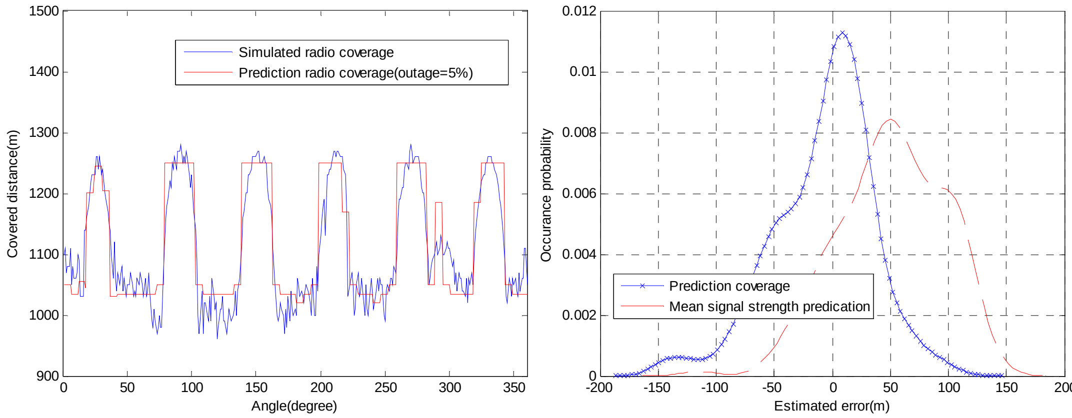


Figure 6.4 The comparison of combined model and only mean signal strength prediction

The quality of signal can vary substantially due to the characteristics of the environment in which it is measured according to the discussion in [112]. For the complex propagation environment in a real wireless network, the prediction model will be initialized by the training data at first, which is extracted either from real-world data (as done in conventional network planning or during an experimental training period) or from simulation models. A key feature of the approach is that the accuracy of the prediction model can be improved by an iterative learning scheme based on the user feedback in response to the antenna configurations, and this can be more robust as time evolves and the environments changes. As mentioned in the previous section, the prediction model is motivated by the need to give coverage prediction under variations of received power according to potentially novel antenna configurations. Most attractive is that the prediction model is simple and no accurate location requirements means the prediction model could be used if the adaptive antenna power system were deployed and further optimizations could be made during real operation. In the operational stage, the new received power samples are compared against the received power distributions over these areas, and the coordinates of the best matches are averaged to improve the QC estimation. So, the proposed prediction models are each local and can respond to the variables of the situation (such as season changes, day or night changes). An iterative learning scheme could build upon the previous models rather than creating a completely new model.

6.2.3 Prediction Model Verification

The proposed prediction model was verified by comparing with theoretical coverage using a hybrid propagation model for the radio coverage prediction, which combines lognormal shadowing and Rayleigh fading. This verification method is also used and recommended in [108]. Rayleigh fading assumes that the magnitude of a signal in a communication channel varies randomly according to the *Rayleigh* distribution [113]. The probability density of the amplitude ρ is described by the equation below based on the *Rayleigh* distribution:

$$f(\rho|\sigma) = \frac{\rho \cdot \exp\left(\frac{-\rho^2}{2\sigma^2}\right)}{\sigma^2} \tag{6.11}$$

Averaged over one RF-cycle, the instantaneous power $p = \frac{1}{2}\rho^2$, and the probability density of the power will be

$$f_p(P) = f_p\left(\sqrt{2p}\right)\left|\frac{dp}{d\rho}\right| = \frac{\exp\left(-{}^{p}/_{\sigma^2}\right)}{\sigma^2} \tag{6.12}$$

where σ^2 the is the local-mean power, which can be calculated according to the lognormal shadowing environment as in the equation below:

$$\sigma^2 = P_t - \mathrm{PL}(\mathrm{d_0}) - 10 \cdot \mathrm{n} \cdot \log\left(\frac{\mathrm{d}}{\mathrm{d_0}}\right) \tag{6.13}$$

where d is the distance between the MS and serving station (BS or RS). So, the coverage probability can be fairly simple to compute based on the equation below:

$$P_c = \int_{\mathrm{p_{th}}}^{\infty} f_P(P|\sigma)dp = \int_{\mathrm{p_{th}}}^{\infty} \frac{\exp\left(-{}^{p}/_{\sigma^2}\right)}{\sigma^2}\mathrm{dp} = \frac{\exp\left(-\frac{\mathrm{p_{th}}}{\sigma^2}\right)}{\sigma^2} \tag{6.14}$$

where p_{th} is the threshold of received power. The radio coverage probability here is the probability that the received power is above the required threshold. Based on lognormal shadowing, the local mean received power can be calculated for each location, and the theoretical radio coverage based on lognormal shadowing can be calculated using above equation.

There are two models that need to be verified. The first one is the hybrid propagation model implementation, the other is the prediction model based on the hybrid propagation environment. The hybrid propagation data is generated using the COST 259 model and displayed using Matlab. The verification results are shown in Figure 6.5. In the left figure, the comparison of simulated coverage (red line), predicted coverage (green line) and theoretical coverage (blue line) are plotted. The x-axis represents the cell angle from the BS's view, and the y-axis represents the radio coverage distance. The signal outage probability requirement chosen is 5 percent. In the figure on the right, the error probability distribution of the prediction error and theoretical error are plotted. The theoretical error is the simulated radio coverage minus the theoretically computed radio coverage using the model. The prediction error is the simulated radio coverage minus the predicted radio coverage.

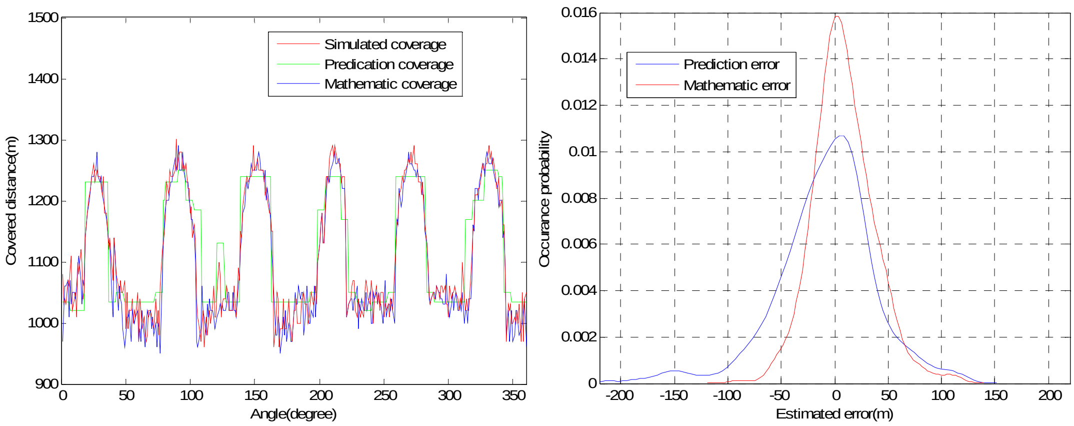


Figure 6.5 Verification results in the hybrid propagation model

As seen from the results, the theoretical coverage is nearly the same as the simulated coverage in the left figure and the estimated error distribution fits the normal distribution well (the mean value is 0). This demonstrates that the hybrid model implementation is satisfactory. The variance is larger as the simulated radio coverage is estimated based on the observed samples and it includes the fast fading effect which varies over time. Comparing the simulated coverage and theoretical coverage, indicates that the predicted coverage is reasonable and the estimated error distribution shows an acceptable standard deviation of the error. Therefore, the proposed prediction model is verified based on the above two comparisons.

6.2.4 Survey Data Analysis

The proposed prediction model mainly depends on the predicted relationship between the antenna configuration and the corresponding received power and signal outage probability, which is analysed based on survey data and regression methods. So, the quality and quantity of the survey sample are very important for the prediction model. In the following experiments, the quantity of the survey sample is considered with respect to prediction accuracy at different parameter settings.

In the previous simulations, the survey data was collected for the test antenna configurations used by the regression analysis. Although the commonly used antenna configurations can be always tested explicitly in the survey stage, a capability of the proposed model is to predict the performance of an unknown antenna setting. This is tested in the simulation below.

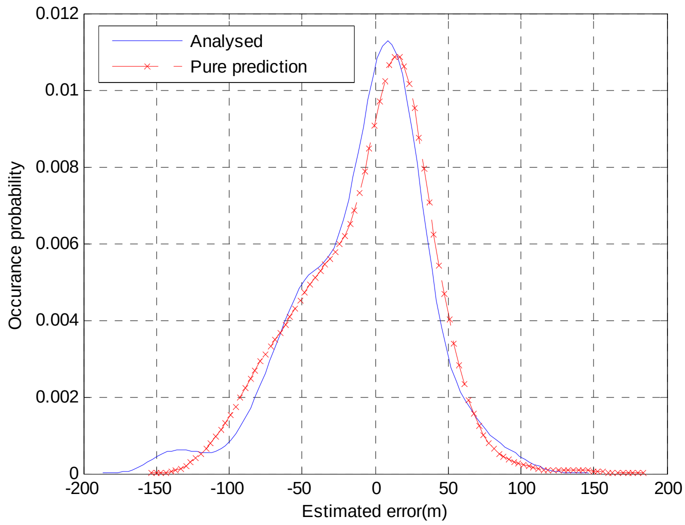


Figure 6.6 The performance of pure prediction simulation

In this simulation, a new antenna configuration is applied in the prediction model. This is pure prediction as there is no survey sample collected for this antenna configuration. The predicted coverage is totally based on the predicted relationship between the antenna configuration and radio coverage. To show a detailed comparison of the performance for known (blue line) and for unknown antenna configurations (red line), the estimated error distributions of two antenna configurations are plotted in Figure 6.6. The error distributions of the two antenna configurations are nearly the same, although the unknown one has a little wider standard deviation and bigger mean value.

The data collected in each QC is assumed to be representative of the whole QC. The number of survey data points in a QC affects the accuracy and is considered in the tests below. The number of survey samples in each cell is determined by three parameters, the number of locations where data is collected in each QC, the number of QCs in each cell and the number of the survey antenna configurations for each location. The number of the survey data points for each cell can be calculated by the number of locations used in each QC x the number of QCs x the number of antenna configurations tested in each location

Table 6-1 The parameter settings in the survey data analysis simulation

Test parameter	Test sampling parameter setting		Number of samples in each cell
Number of locations in each QC	5		30000
	15		90000
	30		180000
	60		360000
Number of the survey antenna configurations	3		54000
	5		90000
	10		180000
	20		360000
The number of QCs in each cell	Angle segment (degree)	Radius segment (m)	
180	10	300	5400
360	10	150	108000
600	6	150	180000
1200	3	150	360000

In the simulations in 6.2.2 , the number of survey data samples collected for the whole cell was 180,000. This is a large number for data collection in the survey stage. Although lots of data could be collected in the operational stage, keeping the survey data at a reasonable number is still important for prediction model initialization. To find the possibly different effects of the sampling parameters on the accuracy of the prediction model, three simulations were designed according to the parameter settings in Table 6-1. The simulation setting for each parameter is varied by changing the test parameter and fixing the other two. For example, the green line in the figure 6.7 shows the error between the real coverage and predicted coverage at different setting of number of samples collected in each QC where the antenna configuration and number of QC in each cell are fixed.

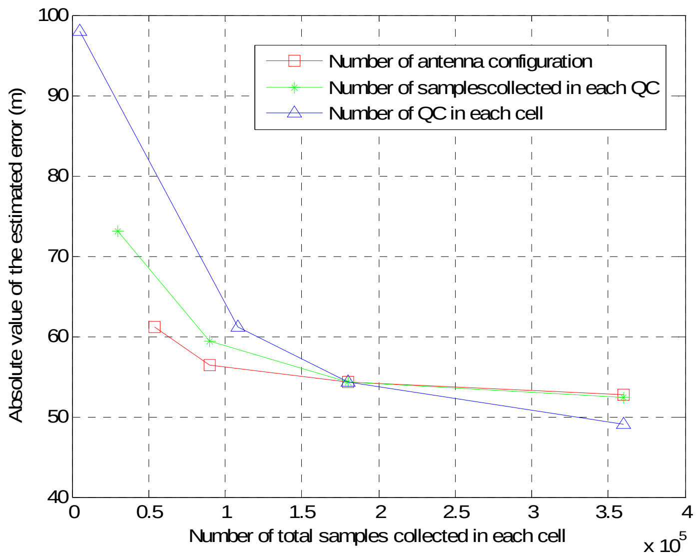


Figure 6.7 The error comparisons at different parameter settings

The simulation results are shown in Figure 6.7. The x-axis is the total number of samples collected in each cell, composed in different ways using the three parameters, and the y-axis represents the average absolute value of the error between the coverage as determined by the propagation model and coverage predicted by the regression model. In all the simulations, the signal outage probability is constant at 5%. The effect of the three parameters on the prediction accuracy shows that the QC size is the most important parameter for prediction accuracy. Therefore, increasing the number of QC in each cell can most significantly improve the prediction accuracy when the total number of samples collected in each cell is limited. When the propagation environment is more complicated, such as in a bad urban area, decreasing the QC size will more efficiently mitigate the variability of the QC and improve the accuracy. The number of antenna configurations used in the survey stage affects the prediction accuracy, which is restricted in this simulation to changing the power of a whole sector of a BS antenna and changing the power for each omni-directional RS antenna.

6.3 Handover Issues for Cooperative Control

In this section, the handover issues related with cooperative control are described. A restructuring scheduling scheme is proposed for burst handover mitigation. The different handover algorithms for cooperative control are also analysed according to the handover rate, utilization for cooperative control and overhead.

6.3.1 Burst Handover Effect

With the help of cooperative control, the radio coverage will be adaptively changed according to the traffic load distribution. The MSs near the boundary will hand over from the higher traffic demand cell to the neighbouring cells, which have less traffic demands. So, the call request success rate in the cell with the previously higher traffic load will be increased as the spare radio resources are released as a consequence of being handed over MSs. To provide seamless communication, handover sessions are often given higher priority over new call sessions from an admission control standpoint. However, too many handovers may cause signal overload and degrade service quality, particularly when there is a significantly heterogeneous traffic distribution. The majority of the handovers occur in places of poor coverage, where calls frequently become unreliable when their channel has interference or is in a fading situation.

In previous chapters, the optimal antenna configurations for each cell were considered to be applied together to avoid coverage holes occurring. However, in practice this would result in a burst in handover processing and this may limit the optimization performance and affect core network performance. To decrease the potential call dropping rate caused by applying cooperative control, a restructuring scheduling scheme is integrated into the network optimisation to mitigate bursts in handover processing. The schedule scheme will divide the handover requirements into small steps and apply the suggested changes step by step based on the traffic demand distribution.

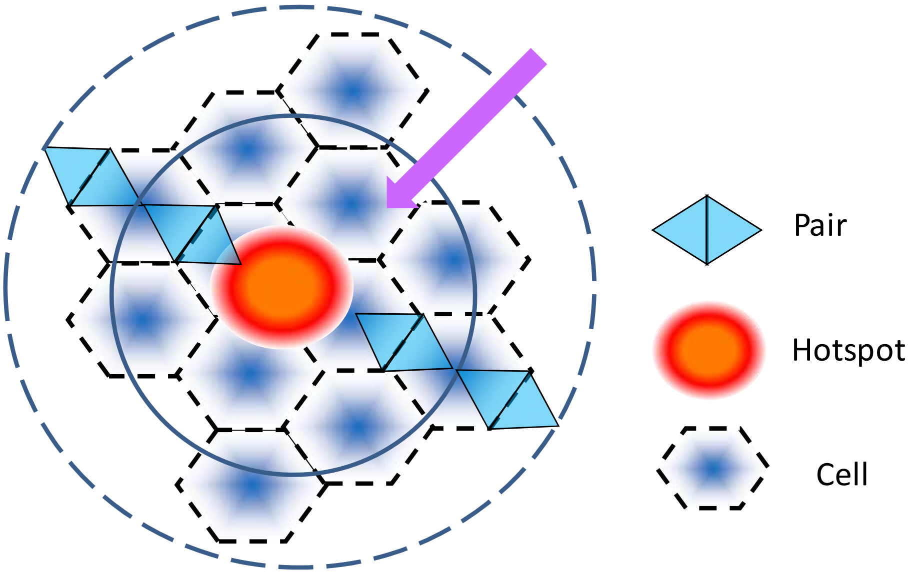


Figure 6.8 Illustration of the restructuring scheduling scheme

With the collaboration feature, adjacent cells need to respond to neighbours' actions when cooperative control is applied. The restructuring scheduling scheme first clusters the cells based on the traffic load distribution, where the hotspots are put in the centre of the cluster. As the clusters are far from each other and deal with their own hotspot, the optimal control scheme could be applied by each cluster without collaboration loss. In each cluster, the opposite sectors are made a pair and the optimal antenna configurations for each pair need to be applied together. The network hole will occur immediately between the two sectors if the opposite sectors are not in close collaboration. The restructuring scheduling scheme for each cluster area is illustrated in Figure 6.8. Based on the relative distance to the central cluster, the corresponding pairs are built in a ring. As there are fewer traffic demands in outside ring sectors, they have more capabilities to help those farther inside. The optimal antenna configuration will be applied from the outside ring of the cluster to the inside one. Therefore, the optimal antenna configurations could be applied by cluster and ring.

The cell clustering depends on the nature of the hotspot area. If the hotspot is spatially bigger there needs to be a bigger cluster. The cluster of cells contains the potential helpers for the hotspot area (maybe not just in one cell). The hotspot does not need to be exactly at the centre of the cluster (sometimes it is impossible). As the scheduling scheme applies the optimal antenna configuration ring by ring, from the outside to the centre, the hotspot area only needs to be in the centre of the cluster.

No matter whether the scheduling scheme is integrated with cooperative control or not, time synchronization is always required by the collaboration. The optimal antenna configuration needs to be collaboratively applied to avoid a hole occurring, so precise time synchronization is still required even when centralised control is used. Clustering the cells not only can mitigate the burst in handover processing, but also can decrease the requirement for precise time synchronization over the whole network. With the restructuring scheduling scheme, only the sectors in the same cluster need to be precisely synchronized. Every pair and ring needs to be synchronized to apply the optimal antenna configurations together. For the synchronization between each ring, an acceptable time interval can be inserted into the two adjacent ring action times to mitigate the burst handover processes. According to the above description, the simulations below were made to test the performance of the scheduling scheme. The simulation parameters are listed in Table 6-2.

Table 6-2 The simulation parameters for the restructuring scheduling scheme

Parameters	values
Number of Cells	19
Number of Sectors in each cell	6
Number of RSs in each cell	6
Path-loss coefficient	3.5
Standard deviation of shadow Fading	4 dB
Handover threshold (SNR)	7dB
Minimal signal level (SNR)	8 dB
Maximum speed of MS	50 km/h
Total number of MSs	6000
Cooperative control timer	400 second

The handover rate quantifies how often handover actions are made. The handover rate is expressed as below in the simulation and is updated every second using network measurements:

$$R_{HO} = \frac{N_{ho}}{N_{nho} + N_{ho}} \tag{6.15}$$

where the R_{HO} is the handover rate, N_{ho} is the number of MSs in the handover stage and N_{nho} is the number of MSs not in the handover stage. Here, the MSs in the handover stage occur during the time interval in which the MS is between the handover threshold and the minimal signal level. The handover threshold is set at the point at which the power received from the neighbouring cell has started to exceed the power received from the current cell by a certain amount. The minimal signal level is the point at which the power received from the current cell is at the minimum acceptable level.

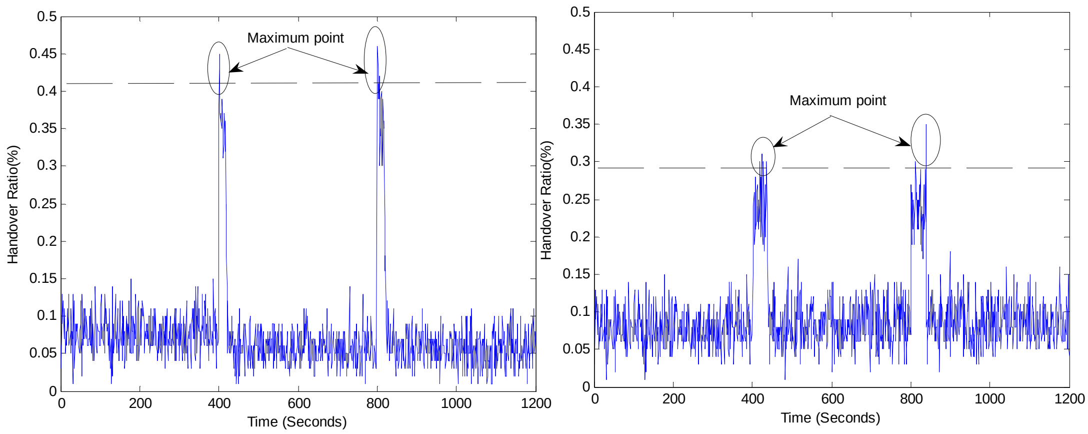


Figure 6.9 The performance of the scheduling scheme

In this simulation, 20% of all the MSs will move into the hotspot area and the MSs not in the hotspot area are still uniformly distributed. The maximum points in the Figure 6.9 represent the burst handover requirement of the radio resource release. Seen from the Figure 6.9, the most obvious change is that the burst handover rate has decreased using the scheduling scheme. The burst handover process requirements are mitigated by step scheduling. The maximum requirement of handovers has decreased around 40 percent even though the time scale of handover caused by the cooperative control has increased. The time when the maximum ratio occurs is also different when the schedule is applied. Without the scheduling scheme, the handover requirement will be immediate, just at the time that the optimal antenna configuration is applied, and so this is the time when the maximum handover requirement occurs. When the scheduling scheme is used, the optimal antenna configurations will be applied ring by ring. The outside ring will be applied first and fewer MSs involved in the handover process compared to the inside ring sectors. When the sectors in the inner ring are nearer or in the hotspot area, the handover process will increase and reach the maximum rate. Based on the propagation model used in the simulation and the radius of the hotspot area, the potential maximum handover rates are calculated and plotted in Figure 6.9 (the horizontal line in the simulation result). Because of the stochastic nature of the models and the shadowing effect, the simulated results are not exactly same as calculated ones but still near the calculated maximum handover rates.

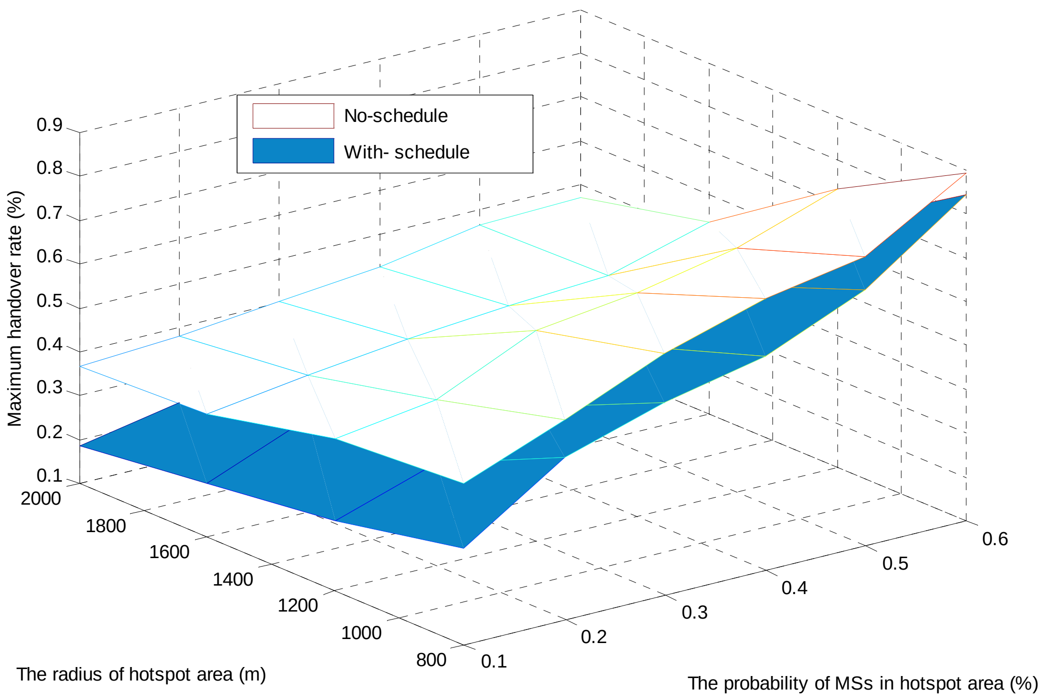


Figure 6.10 Handover performances for different traffic demand distributions

The performance of the scheduling scheme for different traffic distributions is illustrated in Figure 6.10. The maximum handover rate is the main interest here. The x-axis represents the probability of MSs being in the hotspot area. The y-axis represents the radius of the hotspot area (the hotspot area is assumed to be circular). The z-axis represents the maximum handover rate. The higher the probability and smaller the radius of the hotspot area, the more heterogeneous the traffic situation will be. As the traffic becomes more heterogeneous, more handovers will be required for the cooperative control actions. The maximum handover rate with the scheduling scheme is compared with non-scheduling one. As can be seen from the results, no matter how heterogeneous the situation is, the restructuring scheduling scheme always achieves a lower handover rate. If the hotspot area is bigger or has a low population, the scheduling scheme could schedule the optimal antenna configuration application to achieve a 50% decrease in the rate hence mitigating the burst in handovers. The burst in handover processes is smoothed by each ring, but there is no schedule for the sectors in each ring in the collaboration action. So, the optimization performance will be constrained by the ring with the highest traffic demand (i.e. the innermost ring), which is shown by the smaller gap between the results for the different methods.

6.3.2 Analysis for Different Handover Algorithms

The handover can be divided into two types: hard handover and soft handover (the intra-cell (softer) handover is not taken into consideration as it takes place local to the cell and does not require extra transmission resources from the neighbouring cells). A mandatory hard handover and two optional soft handoff procedures: macro diversity handover (MDHO) and fast base station switching (FBSS) are defined in the Mobile WiMAX standard [41].

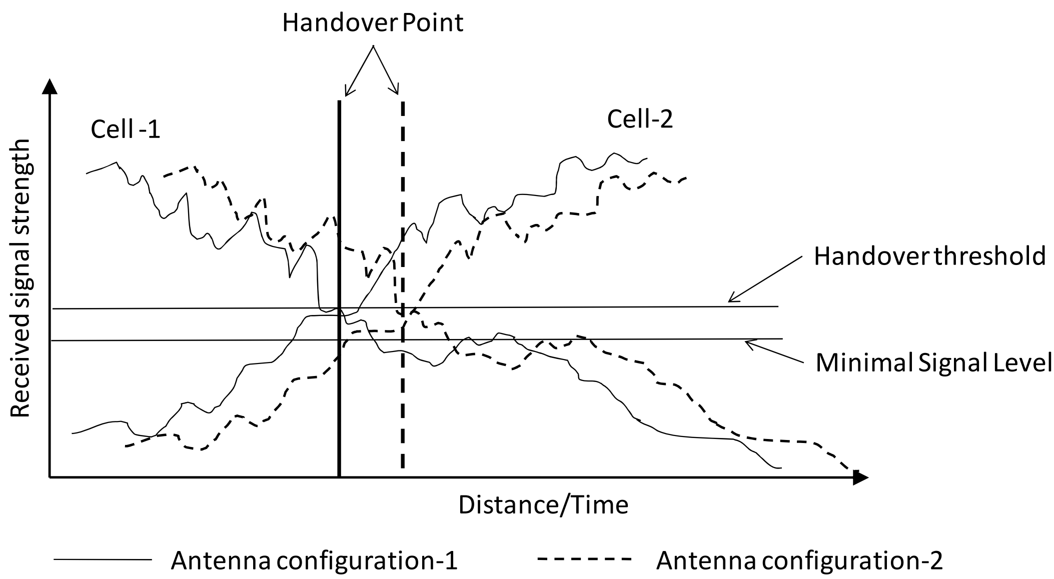


Figure 6.11 The illustration of hard handover process when cooperative control applied

Hard handover uses the "break-before-make" principle [33], which means at any moment one call uses only one channel and achieves more efficient radio resource usage. However, once the handover process fails, the call may be temporarily disrupted or even terminated abnormally. A MS can re-establish the connection to the serving cell if the connection to the target cell cannot be made. The process of hard handover and cooperative action are illustrated in the Figure 6.11. The figure shows a simple case involving two cells and an MS moving away from cell-1 (serving cell) towards cell-2 (target cell). The x-axis represents the distance and time passed between the two cells, and the y-axis represents the received signal strength (this includes the signal from a base station or relay station, depending on which is serving the MS). The softer handovers between the base station and relay station are not represented in this figure. The MSs perceive that the received signal strength from the serving cell is below the handover threshold for a time period (dropping time) and the handover procedures are initiated. The MS will scan the network and try to find another cell to handover to. When another cell that could support the transmission is found (the target cell), the MS will break the connection with the serving cell and connect to the target cell, otherwise it will be dropped. When the MS does the hard handover, a handover point could be specified when the MS does handover process (the connection breaking point). If only the long term propagation effects are considered, the handover point is roughly a constant area if there is a constant antenna configuration. With the cooperative control, the handover point will be changed based on the traffic demand distribution when the optimal antenna configurations are applied. Normally, the handover point will move towards the high traffic load cell to direct some of the MSs in the high traffic load cell to handover to the neighbouring cells. So, the cooperative control effect for the hard handover is obvious and predictable.

If a MS is moving around the handover point area, the MS may frequently handover between the two cells, and this is called the “ping pong effect” [114], which occurs mostly with the hard handover process. If more traffic load demands occur in the handover point area, the handover rate will be increased and may lead to more connections being dropped. To deal with this problem, one way is to use soft handover, another is to pay more attention when the optimal antenna configuration is applied.

The advantage of the soft handover is that the connection to the serving cell is broken only when a reliable connection to the target cell has been established, also known as “make-before-break” [115]. A similar scenario to the hard handover is illustrated for the soft handover process in Figure 6.12. If the perceived signal strength from the monitored cell is window_add(dB) below the strongest cell for a period of Δt, the cell is added to the diversity set (named as the active set in WCDMA). When one of the cells in the diversity set is window_drop(dB) below the strongest cell for a period of Δt, the cell is removed from the diversity set. In the soft handover stage, a MS will ask for radio resources from all the cells in the diversity set. Even for the Fast Base Station Switching situation, the MS has established one or more connection IDs and conducts periodic ranging with these cells in the diversity set. The connection could only fail if all of the channels are interfering or fade at the same time. Therefore the probability that the call will be terminated abnormally due to a failed handover is lower.

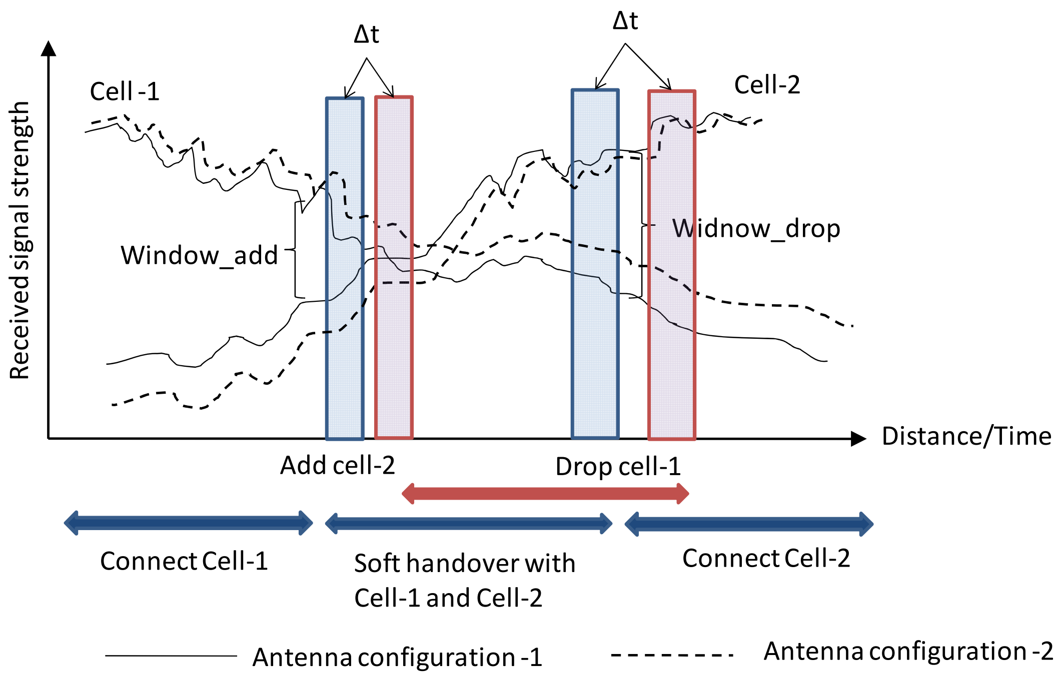


Figure 6.12 An illustration of the soft handover process when cooperative control is applied

An example of window_add , window_drop and Δt period is expressed below:

- Window_add=3dB where the MS needs to identify the add cell when it is 3 dB below the strongest cell.
- Window_drop=5dB where the MS needs to identify the drop cell when it is 5 dB below the strongest cell.

- The Δt will be different for adding and dropping the cell. Typically, the add time is Δt =320 ms and the drop time is Δt =640 ms.

With cooperative control, the time interval of adding and dropping will be changed as seen from Figure 6.12. The whole handover processing range will also move towards the high traffic demand cell. First, the add event occurs later than in the last configuration. This make the MSs in a lower traffic demand cell be in a soft handover stage later. In the same way, the dropping event in the traffic cell with the lesser of the two demands will occur earlier than normal. These could mitigate the radio resource requirement of soft handover from the high traffic cell. However, the adding and dropping event will also occur earlier in the low traffic load cell. The time duration of the MS in the soft handover process may not decrease as both events have moved. So, the efficiency of soft handover for the cooperative control is complicated and investigated in the simulations below.

According to above analysis, while soft handover could support flexible radio coverage more robustly, it may not be as efficient as hard handover. The handover is caused by the RF radio coverage changing. The MSs handed over from the high traffic load cells are already in or near the handover areas. When the radio coverage changes, these MSs are expected to transfer back to the neighbours and return the radio resource they occupied. However, the communication with more than one cell, which is a key feature of soft handover, consumes additional radio resources and leads to a decreased spectral efficiency and cooperative control performance. Because of the expensive equipment required at the BSs, RSs and MSs to provide the diversity, the diversity gain achieved by the soft handover is not considered in this analysis. Therefore, the different handover mechanisms will have a different effect on the system optimization.

The simulations below were used to test the performance of the different handover algorithms in a cooperative context. As the detail of handover implementation is not specified in the Mobile WiMAX standards, some parameters used and the algorithm implementation are based on the WCDMA and TDD-CDMA recommendations. Some typical soft handover parameters used in the simulations are listed in Table 6-3, which are based on the recommended parameters in [116, 117].

Table 6-3 The parameters used in the soft handover algorithms

	MDHO	FBSS
Window_add	3 dB	2 dB
Window_drop	5 dB	4 dB
Add Timer	300ms	300ms
Drop Timer	500ms	500ms

In the simulations, the diversity set size is limited to three. The implementation of FBSS is similar to the MDHO. One advantage of FBSS is to eliminate the various steps involved in a typical handover with their associated message exchange and offer a mechanism that is significantly faster than the conventional handover mechanism. The FBSS allows the MS to switch quickly from cell to cell, but the cells in the diversity set are required to have a better quality signal than in the MDHO. Therefore, the handover window of FBSS is more constrained than for MDHO.

To quantify the efficiency of the handover when used with cooperative control, the handover utilization U_{HO} is defined by the expression:

$$U_{HO} = \frac{N_u}{N_{ho}} \tag{6.16}$$

where N_u is the number of useful handovers for cooperative control, and N_{ho} is the total number of handovers. Here, "useful handover used for cooperative control" means the number of MSs transferred to the neighbouring cell by cooperative control. The total number of handovers includes the useful handovers and the normal handover requirements. The handover utilization measures the ratio of useful handovers to the total handovers, when the optimal antenna configuration is applied.

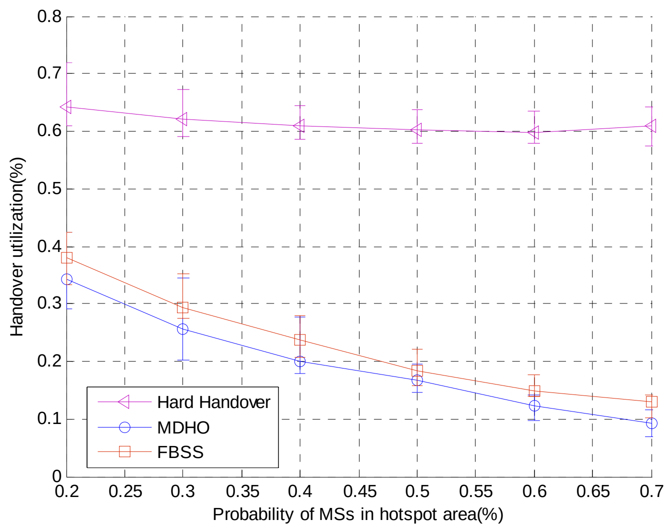


Figure 6.13 Handover utilization comparison

The handover utilizations of the three handover algorithms are plotted in Figure 6.13. The x-axis represents the probability that a MS is in a hotspot area (the ratio of traffic heterogeneity), and the y-axis represents the handover utilization. The hard handover yields the highest handover utilization. The "distinct handover point" feature lets the hard handover achieve load balancing more efficiently. When the traffic distribution becomes more heterogeneous, more MSs will be involved in the handover stage through the action of cooperative control. The hard handover shows a nearly constant performance. The handover utilization of the two soft handover methods decreases when the traffic becomes more heterogeneous and sometimes only 10% of handovers are useful for load balancing. This does not mean worse performance of cooperative control with soft handover, but more handover processes will be required for the soft handover when an optimal antenna configuration is applied. Note that the soft handover requires more radio resources than the hard handover, so cooperative control with soft handover needs to give more consideration to the management of the handoff process to mitigate unnecessary handovers.

In the case of MDHO, the MS is allowed to simultaneously communicate using the air interface with all the cells in the diversity set to achieve macro diversity. So, every cell in the diversity set will allocate the radio resource and transmission power resource to the MSs. Unlike MDHO, the MSs in FBSS only communicate with one cell (referred to as the anchor cell) in both uplink and downlink at a time and BSs involved in FBSS share all information, such as connection ID, encryption, and authentication keys. The cells involved in FBSS must be on the same carrier frequency and the serving BS (anchor cell) could be switching seamlessly [41]. The involved cells also need to reserve the sub-channel for the MSs in the diversity set.

When the MSs are in the soft handover stage, they will create a connection with each cell in the diversity set. Each connection between a MS and cell requires logical baseband resources, reservation of transmission capacity, etc. The handover rate only measures how often handover actions take place, and the radio resource usage in the handover is not determined. The soft handover overhead is the relative proportion of MSs in soft handover, and is often used to quantify the soft handover activity in a network. It is also regarded as a measure of the additional hardware/transmission resources required for implementation of soft handover. Accordingly, the soft handover overhead (β) is defined as

$$\beta = \sum_{n=1}^{N} n \cdot P_n - 1 \tag{6.17}$$

where N is the diversity set size and P_n is the average probability of a MS being in n-way soft handover. In this context one-way soft handover refers to a situation where the MS is connected to one cell, while two-way soft handover means that the MS is connected to two cells. The results of handover overhead for the two soft handover algorithms based on different heterogeneous traffic distributions are plotted in Figure 6.14. “Before” represents the situation when cooperative control is not applied and “after” represents the situation just after the new antenna configuration is applied.

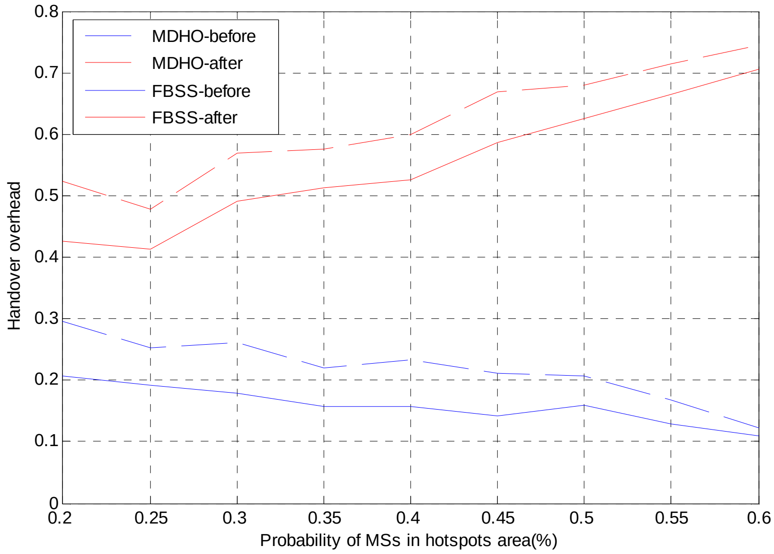


Figure 6.14 The performance of two soft handover algorithms on handover overhead

From the results, the handover overhead for the conventional network (no cooperative control is applied) is decreased when the traffic distribution became more heterogeneous. As radio resources in the hotspot area are limited, the total requirement of soft handover is decreased as some MSs will be blocked or dropped. If cooperative control is applied, in a more heterogeneous traffic distribution, more handover overhead will be required to balance the traffic demand across the network. At the beginning, the handover overhead of FBSS is nearly 10% less than MDHO. As the traffic became more heterogeneous, the overhead gap between the two algorithms becomes smaller.

After the basic insight into the handover utilization and handover overhead, it can be seen that the hard handover and FBSS are recommended for cooperative control. Hard handover is recommended for low speed scenarios, such as stationary and walking situations. These scenarios are mostly found in an urban area, which have more traffic demands and more heterogeneous traffic situations and where demand saturation may occur. The frequency of flexible radio coverage change can be fast and may achieve significant improvements with the most efficient radio resource usage and simple hardware requirements of hard handover. The "break-before-make" feature makes hard handover have a higher call dropping rate than soft handover. As there are more optional serving cells in urban area, call dropping may not be a limitation for the hard handover. Compared to the MDHO, the FBSS algorithm is chosen by the author for simulations because the MS communicates in the uplink and downlink with only one BS at a time and this could achieve more spectral efficiency with the help of the RSs. Besides, in an OFDM system, in order to demodulate the signals from the different cells that are on the same carrier frequency, the MS would require multiple antennas, with the number of antennas being equal to or greater than the number of cells in the diversity set. In either case, it is not possible to do this cost-effectively. FBSS is recommended by the author for simulations for rural area situations, such as a highway, where the movement speeds of MSs is very fast and the fast fading effect may limit the RF domain optimization. The FBSS could mitigate the call dropping rate when performing cooperative control and the frequency of antenna configuration changes will be slow.

6.3.3 Handover with Radio Coverage Prediction

According to above investigations, the handover area should be given more consideration when cooperative control is applied. In the network planning stage, the potential handover area could be carefully selected and designed. However, the handover area will be variable when the antenna configuration is changed by the cooperative control policy. With the help of the proposed statistical prediction model, the potential handover area could be estimated. One use of handover area prediction is to find deep fading areas and so pay more attention to these when the changes suggested by flexible radio coverage optimisation are applied. The other one is to diminish the handover "ping pong effect" and decrease the unnecessary handovers when applying a new antenna configuration. An illustration of the cooperative control process that addresses handover concerns and utilises coverage prediction is shown in Figure 6.15.

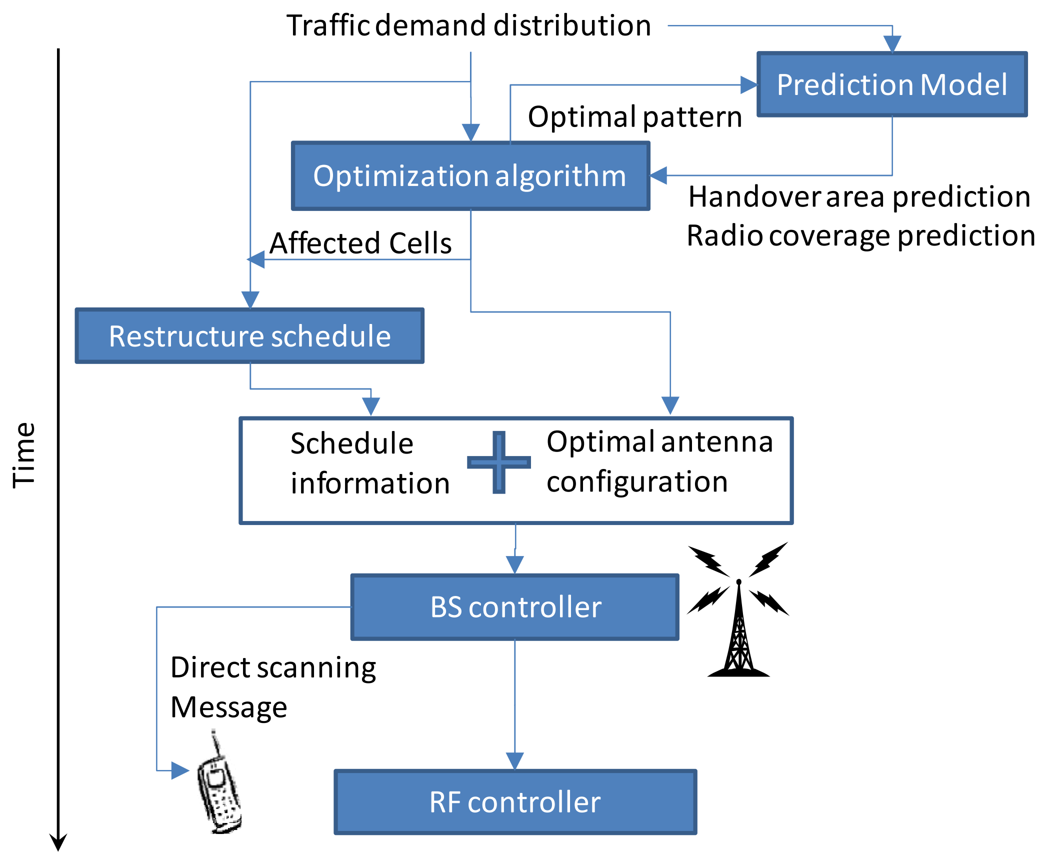


Figure 6.15 Application of cooperative control in a real environment with scheduling of handovers

Firstly, the network performance will be monitored in real time and the traffic demand distribution could be estimated based on the rough locations of demands collected using a network based location estimation method. When the network performance is below some threshold (such as the call dropping or blocking rate) or after a constant time interval, the cooperative control optimisation will be triggered. According to the traffic demand distribution, the optimization algorithm will calculate the antenna configuration for each cell. There is a loop for improvement between the optimization algorithm and radio coverage prediction model, as the radio coverage prediction model can use the feedback of the actual radio coverage and potential handover area according to the applied optimal antenna configuration and the optimization algorithm will improve the optimal antenna configuration. The optimization algorithm will utilise the handover area prediction, adjusting the handover area to avoid frequent handovers and pay more attention to the deep fading scenarios. Based on the traffic demand distribution and the affected cells the restructuring scheduling method could cluster the cells (and also the rings and pairs) and build the schedule for handovers to mitigate the burst handover impact on the core network and call dropping rate. The optimal solution includes the optimal antenna configuration and schedule information and will be sent to the BS controller. The proposed process favours cells that have acceptable signal quality and more available resources over those with the best signal quality and limited resources, and so it may be possible to mitigate the loss in spectral efficiency and optimal cooperative performance.

When the proposed schemes are applied, there is no modification required to the Mobile WiMAX standards and it is feasible for the system to continually update itself. Beside, some handover protocols defined in the standard[41] could also be used for optimization guarantees. In order to be aware of its dynamic radio frequency environment, the BS allocates time for each MS to monitor and measure the radio condition of the neighbouring BSs. This process is called *scanning,* and the time allocated to each MS is called the *scanning interval.* In order to start the scanning process, the BS issues a *MOB_SCN-REQ* message that specifies to the MS the length of each scanning interval, the length of the *interleaving interval*, and the number of scanning events the MS is required to execute. So, with the *MOB_SCN-REQ* message, the BS can direct the MSs to perform multiple scanning events to avoid a sudden handover and fading effect before the optimal antenna configuration has been applied.

6.4 Summary

In this chapter, an approach to radio coverage prediction in flexible antenna configurations is described. Essentially the model of the wireless environment at different locations can be learned through a bootstrapping process. The analytical results cover general topologies and were verified through simulations. Such models are important for intelligent radio resource management as they allow more accurate hypothetical reasoning and hence the discovery of optimal solutions. This model can be also further improved by analyzing the azimuth and elevation power distributions of the transmission antenna and introducing terrain databases.

This prediction model could also be further elaborated with research on multiple-input and multiple-output (MIMO) antenna systems (e.g. the analysis of optimal RF domain among the multiple antennas) and so improve communication performance.

The handover issues are investigated in the context of cooperative control. The optimal solutions for the whole network are carefully selected and scheduled to mitigate the burst in handover demands if the actions were to be applied instantaneously. The different handover algorithms defined in the Mobile WiMAX standard were analysed in detail with respect to handover utilization, overhead, etc. The motivation for these approaches is that they allow cooperative control to respond to situations in the realistic environment and for the cooperative control algorithms to incrementally improve performance over time.

Chapter 7 SCALABLE QOS PROVISIONING WITH SERVICE NODE SELECTION IN RBCNS

7.1 Introduction

In wireless networks, providing and guaranteeing the QoS for each mobile is very difficult, due to the limited resources in energy and bandwidth [118, 119]. In addition, the recent provision and popularity of real time multimedia services by the mobile operators exacerbates the problem. Although introducing RSs in the cellular network can expand the area of coverage and reduce the transmission power per hop [8], the consequence is a more complex network architecture and topology, and efficient QoS provisioning is more difficult.

In the previous chapters, approaches based on deploying different antenna systems to improve the total network performance for a larger scale RBCN were provided and the practical issues when cooperative control is applied in realistic networks were also analyzed. The main feature of the approaches was coordination inside cells to split the traffic hotspots to allow the RSs to support geographic load balancing. When the traffic demands are balanced according the geographic location, the QoS for each MS needs to be considered by the serving BSs and RSs.

In most research on multi-hop based cellular networks [8, 25, 120], the radio resource allocation for the multi-hop link is a simple partitioning into different time slots. The advantage of such a scheme is that it is very easy to implement and to understand. This model was also used in the previous simulations in this thesis. The demand of MSs in the previous simulations were assumed to be for a voice call service and required a constant bit rate and time delay, so the simple scheme was applicable. As there is the same traffic demand for uplink and downlink, the radio resource management can easily divide and allocate the radio resource for the BS to RS and RS to MS link. However, this is not efficient for multimedia services, where there are variable and different requirements for the uplink and downlink.

In a mobile communication network, not all users are of the same type. Browsing the web, emailing, sending/receiving video, downloading files, are all activities that might be performed simultaneously by the population of users. Each of these operations places different demands on the system. Rapid and highly variable demand induced arrivals and departures result in the total system demand varying quickly and sometimes dramatically. Some requests might require a higher data rate on download than on upload, while for others the balance is more evenly distributed. To more efficiently satisfy the QoS requirement of each user, effective radio resource allocation algorithms need to be developed that balance the QoS requirements of each application and user with the available radio resources. In other words, capacity needs to be allocated in the right proportions among users and applications at the right time.

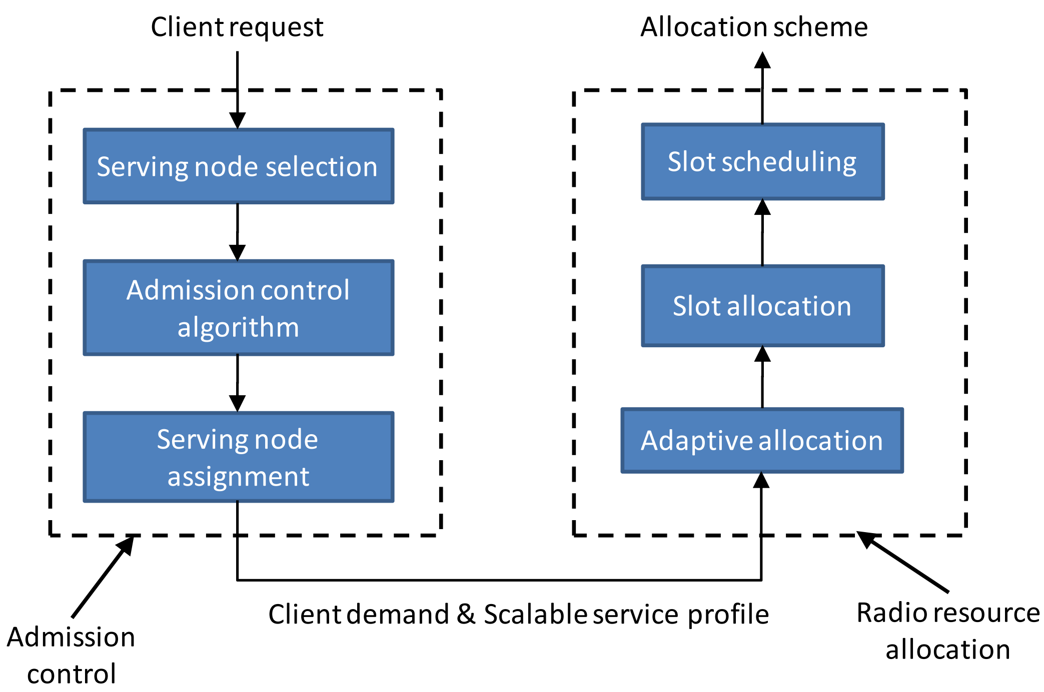


Figure 7.1 The illustration of proposed scheme

To solve the hybrid resource allocation problem in a distributed manner with less complexity and maximize the usage of the limited radio resource, a cross layer scheme for scalable QoS provision is proposed, as illustrated in Figure 7.1. At first, to guarantee the time delay requirement for real time service, a simple optimal serving node selection mechanism that addresses QoS concerns is added in the admission control process. Then an adaptive radio resource allocation is proposed to maximize the user satisfaction and guarantee the QoS requirement. As a result, the serving station that can guarantee the restricted delay requirement and use less slots is always allocated to the mobile user, which could also mitigate the bottleneck problem in RSs. The adaptive algorithm performs the functions of both maximising the number of satisfied users and improving the spectrum efficiency with linear-complexity with respect to the number of users and channel conditions. The protocol and algorithm are evaluated under different traffic scenarios with both real-time and non-real time services.

The remainder of this chapter is organized as follows. First, the related work on radio resource allocation of OFDMA and the challenge for RBCNs are reviewed in Section 7.2. Then Section 7.3 describes the proposed scheme that tries to maximise the quality satisfaction of mobile users with service node selection and the numerical results are presented in Section 7.4. Finally, Section 7.5 concludes the chapter.

7.2 Radio Resource Allocation for OFDMA in RBCN

In wireless cellular networks, radio resource management has been studied for a long time and many methods have been developed to deal with this problem. However, when RSs are added, normal cellular network radio resource management schemes cannot be used because of the more complex network architecture and topology.

Firstly, the basics of OFDMA radio resource allocation are described below. Figure 7.2 depicts a typical OFDMA-TDD system where the radio resource is partitioned in both frequency and time domains. OFDMA system lets each user occupy a subset of traffic channels and each channel is exclusively assigned to one user at any time. The signal is transmitted in an “atomic” period called a symbol. The frequency resource is divided into multiple OFDMA subcarriers and each sub-channel is a cluster of OFDMA subcarriers. The number of subcarriers in one sub-channel is constant according to the system spectrum configuration (28 subcarriers in each sub-channel in 10 MHz system bandwidth). Sub-channels may be created using either contiguous subcarriers or subcarriers pseudo-randomly distributed across the frequency spectrum [41]. For simplicity, but without loss of generality, the allocated sub-channels and OFDMA symbols are assumed to be contiguous through this thesis. The smallest resource unit through which data is transported is termed a 'slot'. A slot in OFDMA requires both a time duration and sub-channel dimension and is the minimum possible data allocation unit.

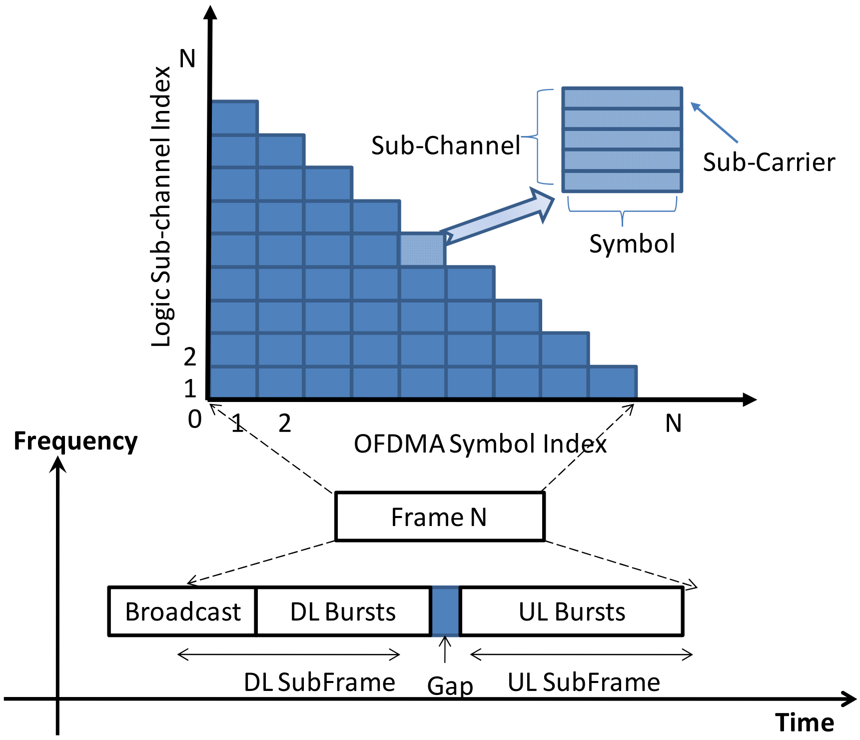


Figure 7.2 Radio resource allocation in OFDMA

The basic OFDMA radio resource allocation problem can be formulated as a slot assignment problem [121]. The objective is to find an optimal assignment from base stations to MSs in such a way that all the MSs demands are satisfied, when every slot can be assigned to at most one MS and the total interference is minimized [122]. Notice that in this problem it is not important which slot is assigned to a MS by a BS but how many slots this BS assigns and if the MS receives the number of slots corresponding to its demand. To apply the slot assignment problem in a real network such as Mobile WiMAX, this will be subject to some additional constraints to satisfy more specific system configuration requirements, the modulation level effects and variable service demand requirements.

Table 7-1 OFDMA parameters in Mobile WiMAX

Parameters	Downlink	Uplink
System bandwidth	10 MHz	
FFT size(N_{FFT})	1024	
Null Sub-Carriers	184	184
Pilot Sub-Carriers	120	280
Data Sub-Carriers	720	560
Sub-Channels	30	35
OFDM Symbols/Frame	48	
Data OFDM Symbols	40	

The above table shows the difference between OFDMA parameters in the uplink and downlink with the 10 MHz system bandwidth configuration. In the downlink, for example, 720 data sub-carriers and 120 pilot sub-carriers build 30 sub-channels. The 5 millisecond frame is divided into 48 symbols for both uplink and downlink, so each symbol lasts 102.9 microseconds. Only 40 symbols are used for data transmission. The PHY data rate is the maximum transmission data rate support by the physical layer in 10 MHz system bandwidth, which we can set as a benchmark. The sub-channel/symbol in the table represents the information bits transmitted by each sub-channel in one OFDMA symbol time.

The transmission rate for a connection (assumed that each MS only has one connection) depends on the average SNR of the link and the allocated radio resource (number of slots). Let the Boolean assignment variable Symbol_m and $\mathrm{Subchannel}_n$ represent the assignment of a MS to a sub-channel and symbol, i.e., $\mathrm{Subchannel}_n(i)=1$ indicates the sub-channel n is allocated to the MS i. Taking the downlink slots allocation problem as example, the downlink transmission rate for MS i can be calculated as the below:

$$R_i = \sum_{m=1}^{40} \mathrm{Symbol}_m(i) \cdot \sum_{n=1}^{30} \mathrm{Subchannel}_n(i) \cdot \mathrm{MCL}_i \tag{7.1}$$

where Symbol_m and $\mathrm{Subchannel}_n$ is 1 if symbol m or sub-channel n is allocated to the MS i. MCL_i is the modulation and coding level for MS i which mainly depends on the average SNR of the link. As listed in the Table 7-2, the $\mathrm{MCL_i}$ is the information bit that a sub-channel transmits in a symbol time. For example: when a MS requires a downlink bit rate of 100 kbps and the modulation and coding level of the connection is 2 (so the MCL=8.38 kbps), the required number of slots is

$$Number\ of\ Slot = \frac{100}{8.38} \approx 12$$

There are several ways to allocate the slots, such as, 4 sub-channels transmitting 3 symbols.

Table 7-2 AMC level in Moible WiMAX

Level	Modulation and coding	subchannel/symbol	SNR
1	QPSK(3/4)	6.59 kbps	11.2
2	16 QAM(1/2)	8.38 kbps	16.4
3	16 QAM(3/4)	13.2 kbps	18.2
4	64 QAM(2/3)	18.3 kbps	22.7
5	64 QAM(3/4)	19.8 kbps	24.4

Resource allocation and scheduling are central to radio resource management [118]. As this thesis is more focussed on efficient radio resource utilization, the scheduling scheme defined in [41] is used. Clearly, resource allocation plays an important role in the performance of OFDMA systems. Despite the absence of intra-cell interference, optimum resource allocation is still an NP-hard problem and difficult to tackle in practice [123]. Additionally, the heterogeneous individual users' rate requirements further complicate the problem. In the past few years, resources allocation problems in OFDM/OFDMA systems were extensively studied. Previous works dealt with a wide variety of optimization objectives, such as: subcarrier allocation [124, 125] for minimizing power (where the exact assignment is not important), and subcarrier assignment for maximizing system throughput in the presence of co-channel interference (CCI) [122, 126]. These algorithms maximize the total throughput or minimize power allocation, but do not consider heterogeneous QoS requirements and are mainly for purely cellular networks. In a fixed relay based cellular system, where the cellular spectrum is being reused for the RS to MS hop by the time slot, it is also important to ensure that the resources are assigned efficiently. Here, adaptive modulation and coding is allowed when radio resource allocation is performed to satisfy and to maintain individual QoS requirements.

However, there are several challenges in applying existing OFDMA radio resource allocation algorithms to a RBCN. First, unlike traditional cellular networks where the MS can only connect to the BS, there are several options. With the different assignments to the BS or the RS, the radio resource requirement is different, which makes the radio resource allocation more complex. Secondly, the adaptive modulation and coding utilization results in different throughputs at different SNR levels. Thirdly, the different types of service and the users' variable service requirements make it complicated to efficiently allocate the radio resource to maximize user satisfaction.

To solve the hybrid resource allocation problem in a distributed manner with less complexity, a novel algorithm for MS assignment and allocation of radio resource for RBCNs based on the OFDMA-TDD system have been devised by the author. First, a simple MS assignment scheme is formulated based on QoS concerns. Second, an optimal sub-channel allocation algorithm is formulated that tries to maximise the quality satisfaction of mobile users and balance the heterogeneity requirement.

7.3 Proposed Approach

7.3.1 Service Node Selection with QoS Considerations

In cellular communication, the pilot signal strength is used by mobile terminals for cell selection, and handover. A mobile terminal measures and compares the pilot signals that it can detect, and typically attaches itself to the cell with the best quality pilot signal. In a relay based cellular network, service node selection is performed in two situations. First is when the MS joins the network and the second is if the MS needs to switch RS or handoff to another cell. In the classical form of management of a relay system, the selection criterion for BS and RS is mainly based on the pilot signal strength level perceived by the MS user, which determines the path gain between the BS and the MS or the RS and the MS. When resource management of a relay based cellular network is focused on the impact of resource management decisions on high data rate service coverage extension, then such a scheme is an attractive solution for delay-tolerant traffic services. However, for delay-sensitive traffic services the processing and scheduling delay induced by the relay station need to be considered. Taking voice over IP service as an example, the time delay has a more important effect than the bit rate as the bandwidth required for this service is small. Also, this classical algorithm [127] does not optimize the entire system performance because it does not reflect the characteristics of multi-hop networks. Therefore, the serving node (i.e. BS or RS) selection needs to be considered carefully as part of the adaptive radio resource management for a relay based cellular network.

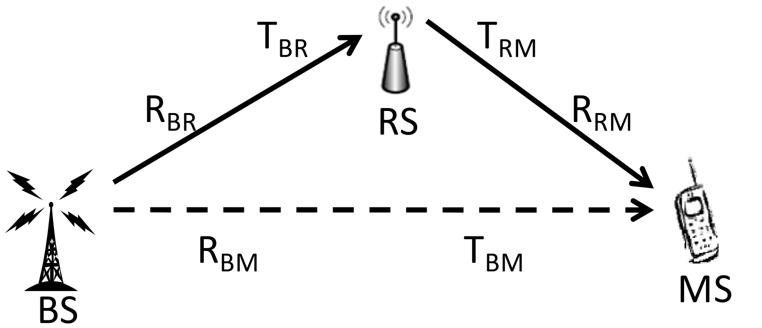


Figure 7.3 The potential paths for the MS in a single cell

In a single cell when there are relay stations, the MS has two potential paths of connection to the BS as shown in Figure 7.3. One is to directly link to the nearest BS. Another is to indirectly link to the BS through the nearest RS. If the MS wants to transmit M bits of data, the average bit rate of the indirect link can be approximately calculated as below:

$$R_{brm} = \frac{M}{T_{brm}} = \frac{M}{T_{br} + T_{rm}} = \frac{M}{\frac{M}{R_{rm}} + \frac{M}{R_{rm}}} = \frac{R_{br} \cdot R_{rm}}{R_{br} + R_{rm}} \tag{7.2}$$

where R_{brm} is the average bit rate from BS->RS->MS , T_{brm} is the time taken to transmit the M bits data. According to the Equation (7.1), the bit rate can be calculated as:

$$R_i = \text{Slots}_i \times \text{MCL}_i \tag{7.3}$$

where $\mathrm{Slots_i}$ is the number of allocated slots (the number of sub-channel*the number of OFDMA symbol) for the MS i and $\mathrm{MCL_i}$ is the bit rate per slot according the modulation and coding level of link.

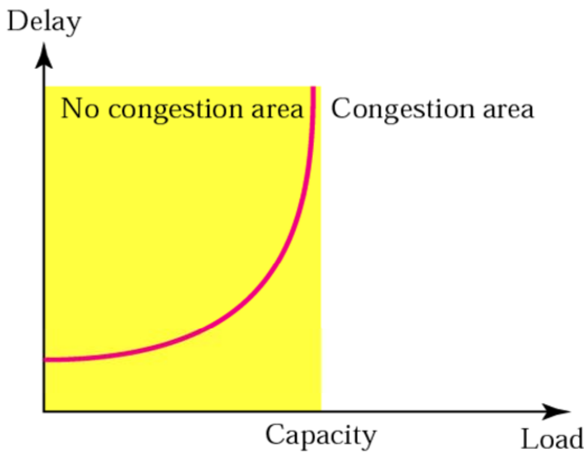


Figure 7.4 The illustration of the relationship between relay delay and traffic load

In a QoS scheduling scheme, a real time service will have a high priority for service flow scheduling and processing. However, when more real-time services are required in a RS, the time delay will increase and violate the service delay requirement constraint as illustrated in Figure 7.4. The exact time delay in RSs will depend on the capability of the RSs and the service model, such as the speed of DSP processors, and the scheduling algorithm used.

A simple scheme is taken forward for the MS assignment based on the above average bit rate and time delay description: A new connection request submits its QoS requirements to the nearest BS. The QoS requirement includes the bit rate and time delay. The optimal serving node will be chosen that can fulfil the time delay request. If two potential nodes can both fulfil the time delay request, for a constant bit rate requirement from the user, the path for which the lesser number of slots is requested, according to equation (7.3), will be selected.

The two paths selection scheme can be easily extended to several paths, as several RSs may be near the MS. In the path selection scheme, the BS enumerates all the received candidate paths $(p_1, p_2 \ldots . p_n)$, where n is the number of disjointed paths. The BS estimates the end-to-end delay and the number of slots required for the potential path, and chooses the best path $\mathrm{p_{sel}}$ i.e. the path that satisfies the required QoS metric and has the least number of slots.

$$p_{sel} = \mathrm{argmin}\{\mathrm{slot}(p_1), \mathrm{slot}(p_2), \ldots \mathrm{slot}(p_n)\} \tag{7.4}$$

where slot(p) is the number of slots used by path p. In other words, the proposed MS assignment scheme selects a path with the maximum achievable throughput and QoS satisfaction, and this should eliminate bottleneck links.

As the location of MSs will be changing with time, the service node selection should have to establish and maintain the optimal path. Hence, the service node selection scheme should not only work with admission control but also a congestion control to support the QoS guarantees and efficient radio resource usage. The proposed approach is for each end user rather than each frame, so it works for both the uplink and downlink according to the service type and requirement.

7.3.2 Radio Resource Allocation for Scalable QoS Provisioning

In the Mobile WiMAX standards, the MAC layer is connection-oriented. All data communications are in the context of a connection. Each connection is associated with a single data service. Each data service is associated with a set of QoS parameters that quantify aspects of its behaviour. Some schemes have been defined to support QoS, such as Service Flow QoS scheduling and Dynamic Service Establishment [41]. The primary purpose of the QoS specification defined in the standard is to define transmission ordering and scheduling. Four service classes are supported [1, 41]:

1. Unsolicited Grant Service (**UGS**): Limited delay and constant traffic rate (constant bit rate required)
2. Real-time Polling Service (**rtPS**): Limited delay and variable traffic rate (minimum bit rate required)
3. Non-real-time Polling Service (**nrtPS**): Tolerant to delay and variable traffic rate (minimum bit rate required)
4. Best Effort (**BE**): Tolerant to delay and variable traffic rate

To evaluate the heterogeneous traffic load situation, the traffic mixtures in Table 7-3 are used in the analysis below. Each connection is equally likely to belong to one of the four services. The bit rate required by each connection lies in the range of the maximum and minimum bit rate. For simplicity in the simulations that follow, the bit rate is sampled from the uniform distribution ranging from the lower to upper bit rate.

Table 7-3 The service traffic mixture

Service	Service class	Min bit rate	Max bit rate	Delay	Average duration(sec)
Voice	UGS	30 kbps	30 kbps	Sensitive	120
Video	rtPS	64 kbps	256 kbps	Sensitive	300
FTP	nrtPS	64 kbps	512 kbps	Tolerant	180
Email	BE	10 kbps	128 kbps	Tolerant	60

When considering a feasible solution to the slot assignment problem, a fundamental problem is to find the maximum number of MSs that can be satisfied by the given set of base stations. Since the limited radio slots that every base station owns may not guarantee full satisfaction in a high-density area, finding an assignment that maximizes the number of satisfied clients over the network can be very useful.

Consider a set L={1,…,m} of mobile users, each has a traffic demand t_i corresponding to r_i slots and suppose the slot allocation set $N(i)$ is used to satisfy the demand of mobile user i. To maximize system capacity, as many subscribers as possible will be assigned to the schemes with highest modulation and coding level. However, this solution is highly unfair because some users close to the BS or RS can monopolize all the subcarriers while others users starve, not receiving any assignment. The high bit rate of the downlink in the multimedia services and the uneven radio resource requirement for the uplink and downlink make the efficient usage of frequency resource more difficult. Thus, it is desirable that the adaptive allocation should have fairness constraints on the performance of individual users.

To operate cost efficiently under bursty QoS requests and maximize the number of serving MSs, the MSs are assumed to have a scalable QoS ability. The scalable QoS scheme used in this chapter is very similar to the guarantee and non-guarantee servers used in [128]. So, a negotiation process is involved in the determination of QoS parameters for efficient utilization of radio resource, especially for the application adaptation process. Instead of using the exact parameter settings for

QoS, and approximation of the user satisfaction of the QoS is used in the adaptive approach and this is explained below.

Normally, the ordinary users will not be interested in the exact parameter setting of QoS, they may only care about the satisfaction associated with the QoS. For video service users, they may care about the video quality, i.e., the video resolution, the time delay. The exact QoS parameter setting, such as the peak rate, the mean rate, the mean burst length, delay, jitter, and the request blocking, may be of no intrinsic interest.

From the point of view of radio resource allocation, the data rate and time delay are used to determine the QoS satisfaction. If a BS can provide the service to the MSs better than their minimum requirement, then a users' basic requirement is taken to be satisfied. To estimate the QoS satisfaction of each MSs, the QoS satisfaction factor of MS i is defined as below.

$$\mathrm{QS_i} = \begin{cases} 1, & \text{Offered QoS} > Min\ QoS \\ 0, & \text{Offered QoS} < Min\ QoS \end{cases} \tag{7.5}$$

For different service class as defined in the Table 7-3, the minimum required QoS (bit rate here) is different. For example the Voice over IP service, the minimum bit rate requirement is equal to the maximum one, and there is no possibility for the degradation of the QoS. However, for the video service, the bit rate can be decreased as long as it remains bigger than the minimum required, and the user can still be satisfied. So the allocation objective function is:

$$\mathrm{Max} \sum_{i=1}^{k} \left(\frac{\mathrm{MLC_i} \cdot \mathrm{Slot(p_i)}}{\mathrm{Require_i}} \cdot \mathrm{QS_i} \right) \tag{7.6}$$

The constraints are:

$$\mathrm{MLC_i} \cdot \mathrm{Slot_i} \geq \mathrm{Mini\_Service_i} \tag{7.7}$$

$$\sum_{i=1}^{n} \mathrm{Slot_i} \leq \mathrm{Slot_{max}} \tag{7.8}$$

where the k is the number of served users, $\mathrm{Mini\_Service_i}$ is the minimum bit rate of the service required by MS i, which are defined in Table 7-3. The $\mathrm{Slot_{max}}$ is the total number of slots in each frame.

The first set of constraints (7.7) guarantees that all service minimum requirements will be satisfied. The second set of constraints (7.8) makes sure that the sum of allocated slots will not exceed the number of available ones. Here, the indirect link to a BS needs to be paid more attention, as the link needs the slots allocated to both the BS->RS link and RS->MS link. Solving the above optimization problem will efficiently allocate the bandwidth for each connection, but also guarantee the QoS satisfaction of MSs in an individual cell.

The key steps in the algorithm to perform the allocation over a single cell are as follows:

1. Try to allocate the minimum service requirement of each MS in the entire cell.
2. If (1) can be done go to 3. If (1) cannot be satisfied for all minimum service requirements of MSs, reset the allocation and sort the MS by minimum requirement and allocate in order considering the smallest minimum requirements of MSs in the cell first.
3. If some slots are still not allocated yet, assign slots from the set of unallocated slots.
4. According to the required service types of the MSs, the MSs are divided into 4 sets. Based on the business or service requirement (in the simulations, the real time service such as voice call is given the highest priority), choose one of the sets.
5. Get the MSs from the chosen set and allocate one slot to each MS.
6. Calculate the satisfaction value of the entire network according to the optimization equation (5) and save the slots allocation scheme.
7. Repeat from (5) until all four sets have been evaluated.
8. Compare the four slot allocation schemes and find the maximum value in the optimization equation, and then allocate the slots according to the selected scheme.
9. Repeat from (3) until no slots are left or all MSs's requirements are met.

As the subcarrier spacing is fixed in S-OFDMA, the smallest unit of bandwidth allocation (based on the concept of sub-channels) is fixed and independent of bandwidth and other modes of operation. The number of sub-channels scales with bandwidth and the capacity of each individual sub-channel remains constant [129]. So the slots assignment algorithm can be easily extended in variable system bandwidth configurations.

7.4 Simulation and Results

Considering that the end-users' perception of QoS is different depending on the application requirements, it is important to define a set of common parameters that allow expression of the service requirements. From the mobile users' perspective, they want to get the radio resource fairly and as much as they require. From the service providers' aspect, they want to provide the service to the individual users as much as they can while achieving maximum throughput to ensure radio resource efficiency. To evaluate the performance of the algorithm, two simulations were designed to test the performance of the algorithm at different the QoS rates and total network throughputs.

A hexagonal cell model is used in the simulation network, and the BS is located in the centre of the cell. Six fixed relays are placed in each cell as 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.

To evaluate the performance of the radio resource allocation algorithm, three different allocation schemes are simulated.

- Conventional allocation: Allocates the slots as much as the MS requires. The sub-channels are reused for the link BS->RS and RS ->MS with different symbols.
- Minimum allocation: The scheme only allocates the slots to meet the MS service minimum requirement. This scheme can maximize the accepted number of users.
- Adaptive allocation: The scheme uses the proposed algorithm to balance the QoS requirements of each MSs and maximize the system throughput.

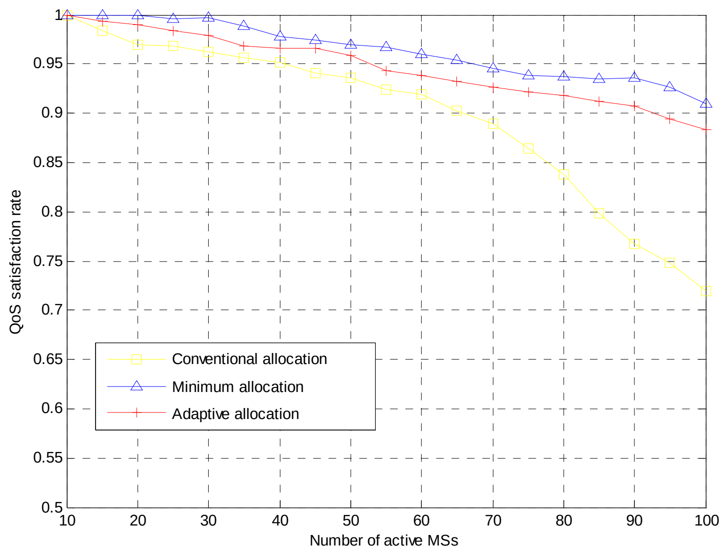


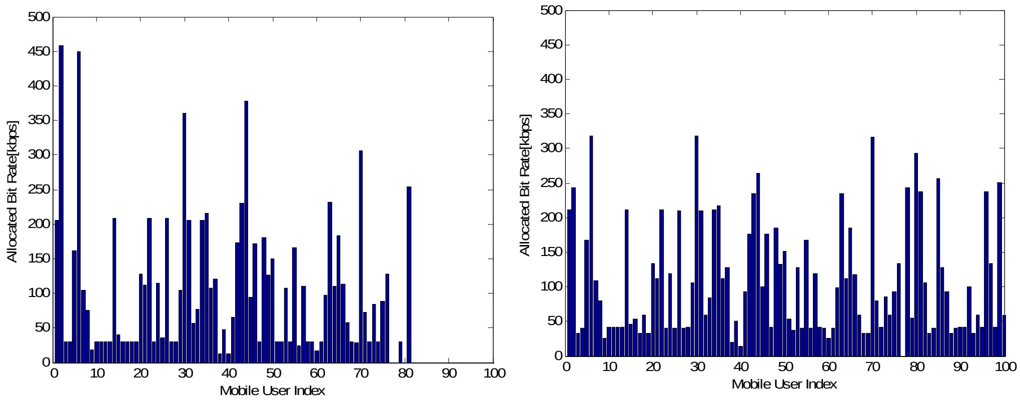
Figure 7.5 Performances of the three algorithms with respect to QoS satisfaction rate

The first simulation is used to investigate the performance of different allocation schemes on the QoS satisfaction rate with different traffic loads. In Figure 7.5, the x-axis represents the number of active MSs and the y-axis represents the QoS satisfaction rate[3]. At the beginning of the simulation when traffic demand is low, the three schemes have the same performance as there are enough slots for the MSs' requirements. With the increased traffic load, the QoS rate is decreased as the radio resource is limited. In the three schemes, the minimum allocated scheme achieves the best performance in terms of QoS rate as it only allocates the minimum slots the service required by the MS. The adaptive allocation can achieve nearly the same performance as the first priority of the scheme is to satisfy the minimum requirement of all the MSs, then the rest of the slots will be fairly allocated to balance the QoS requirements of each MS according the different services.

To evaluate the balancing of QoS requirements with respect to a fair allocation to the set of users (another aspect of the algorithm), the allocation bit rate for each MS is plotted in Figure 7.6. There are 100 active MSs in the cell. The x-axis represents the index of the mobile user and y-axis represents the allocated bit rate to each MS. For the conventional scheme (Figure 7.6 (a)), it will allocate slots, as much as the

[3] The QoS satisfaction rate is the proportion of offered users that meet their minimum requirements.

MS requires, in the order of the MS index. If there are no more slots that can be allocated, the requests of MSs will be denied. So, there are some MSs with no allocated bit rate in the conventional scheme. The adaptive scheme is plotted in the right figure (Figure 7.6 (b)). The adaptive scheme balances the QoS requirements for each MS, and all the MSs will firstly be allocated the minimum requirement.



(a) Conventional (b) Adaptive

Figure 7.6 Assessment of fairness: the spread of allocated bit rate to 100 active MSs

The system throughput under different traffic loads is shown in Figure 7.7, the x-axis represents the number of active MSs and the y-axis represents the throughput in the cell. As the time delay requirement is guaranteed by the service node selection scheme, the time delay is not considered in this simulation. Because the algorithm has been developed for the whole cell rather than single MS performance improvement, the maximum and minimum data rate for each MS are not shown in this figure. Because the minimum allocation only allocates the minimum service request to the users, the network throughput is always very low. When the radio resource still can satisfy the requirements of MSs, the adaptive allocation and conventional allocation achieve the same throughput. When the traffic load increases, there is not enough radio resource to support all the users' requirements. The cell throughput of conventional allocation scheme tends to a constant value as no more user requirements can be satisfied. However, the adaptive scheme can still increase the throughput as it balances the QoS requirement for each MS and chooses a more efficient scheme to maximize the throughput.

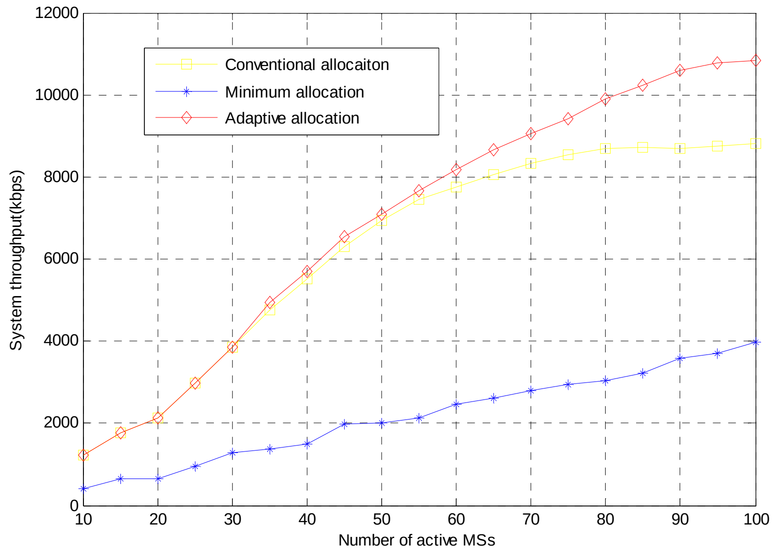


Figure 7.7 The performance of the three allocation algorithms with respect to cell throughput

7.5 Summary

This chapter has described a scalable solution to the provision of QoS in a RBCN with a radio resource allocation approach that also addresses service node selection. First, the service node selection is based on the estimation of the required slots and time delay and tries to minimize the slots required while guaranteeing the QoS requirement. The adaptive radio resource allocation dynamically changes the slot allocation to achieve the maximum throughput while also balancing mobile user satisfaction. The approach is evaluated using two hop relay communication scenarios, but the idea can be easily extended to multi-hop scenarios. The algorithm is demonstrated to have a QoS almost as good as provision of the minimum allocation and a significantly greater throughput when there are higher demand rates. The fairness of the allocation was also demonstrated.

Chapter 8 CONCLUSION AND FUTURE WORK

8.1 Conclusion

The main objective of this thesis is to explore the potential capability of RBCNs with the help of cooperative control to improve the spectrum efficiency and respond to environment changes. The proposed global optimization solutions to allocate wireless resource when there are heterogeneous traffic demands over a large scale network and within a local cell both achieve effective and efficient performance in system level simulations.

Although introducing the RSs into a cellular network could provide a higher data rate in larger scale networks and improve the system capability at reasonable cost, how to more efficiently use the limited spectrum in a non-uniform traffic situation is still a big challenge for network operators and engineers. Chapter 3 analysed the heterogeneous traffic problem in wireless networks and presented the literature on adaptive radio resource management relevant to load balancing. Comparisons with other similar ideas are given, and a detailed description for the cooperative control principle is given. Following this, the detail of the system level simulator based on Mobile WiMAX networks is described.

To explore the possibility of existing antenna systems, Chapter 4 presents two cooperative approaches using different simple antenna models with a "dumb" RS model. Here the controlled adjustment of the RF domain of the BSs is considered. Variable pilot signal strength and adaptive tilting are used. The dumb RSs are used in coordination with other network components to split the hotspot area based on the commands of BSs. As a simple RS model and existing antenna system are used, these approaches are very suitable for incremental deployment into existing cellular networks.

A fast deployment and self-organizing RS approach is described in Chapter 5. With the help of a phased array antenna and Passive DoA, RSs could be quickly added into a network similar to an ad hoc node. In this approach, RSs are more smart and powerful. They can learn from the environment during operation using an adaptive antenna, so that the radiation patterns generated provide coverage where and when it is needed. The coverage pattern of the phased array antenna can adapt if a node fails, a node moves, a MS moves, or simply if demand changes. It is suitable for emergency or temporary coverage situations, and a RBCN can be more efficient, more tolerant of environment changes because it can discover and re-discover its environment.

In chapter 6, practical issues appropriate to when cooperative control is applied in a realistic network are considered from radio coverage, handover and relevant standards aspects. Chapter 6 analyses the importance and difficulty in handling flexible radio coverage and responding to handover effects. A statistical radio coverage prediction model is proposed, where the cooperative control feature is used to support the learning of coverage prediction models. To mitigate excessive handover requirements caused if the optimal antenna configuration is applied in one step, a scheduling scheme is added into the cooperative control mechanism. The hard handover algorithm is recommended for urban area situations and FBSS is more suitable for rural area situations according to this analysis and if not require the diversity advantage of MDHO. If the BSs, RSs and MSs support diversity, MDHO is more suitable for rural areas as further diversity improvement could be achieved.

Chapter 7 proposes a novel approach to balance the traffic demand in a local cell when scalable QoS traffic is assumed. The objective is to maximise the network revenue and maintain predefined QoS constraints. The balancing adjustment of a real time service or non real time service can occur when the service request first arrives in a new cell or in handover to a neighbouring cell. The simulation results indicate that the proposed balancing scheme is flexible, efficient and extendable to more sophisticated wireless networks. Such adjustments can be made in conjunction with the adjustments offered by the system level cooperative resource allocation.

Apart from the above research, two other approaches for cooperative control with semi-smart antenna systems [17, 130] have also been investigated for cellular networks. If more autonomy and intelligence is placed at the BS, then cooperation between the BSs without the use of a central controller can be considered as a way of making the network more robust to failures and attacks. A completely decentralized approach is proposed where the intelligence is distributed to the BSs. The calculations required by the optimization algorithms can also be partitioned as a result of the distributed approach and so make the optimization more efficient. That work develops a novel technique for completely distributed colouring of a cellular network so that no two cells of the same colour are neighbours. Therefore, the optimization algorithms can be run in parallel in cells with the same colour.

In an emergency situation, the cooperative control not only can help resolve unbalanced traffic in the network. The fast deployment ability also has the potential to make rescue mission better supported. Another possibility of usage is to differentiate users/service to give better service to some specific users. This scheme can aid people who are more vulnerable in the environment, such as those dialling emergency numbers, or ensure a service such as video calling is given a better quality of service. More details about these two approaches can be found in [58, 131].

8.2 Future Work

In this thesis, the use of RF control in the physical layer is presented as the central method for providing adaptability in a relay based cellular network and analysed the performance in the presence of heterogeneous traffic and surrounding environment changes. This thesis allows the potential flexibility of the RF domain to be exploited. One area of future work is in combining the knowledge gained about cooperative RF control of network components in other wireless networks, e.g. mesh networks with limited hop capability. Another area is to apply the results studied here to network planning process improvement and to exploit other antenna systems such as MIMO or smart antennas in which environment awareness is an important problem.

- In this thesis, the hops between the BS and MSs are limited to 2 to minimize the relay delay, but the proposed approaches could be easily extended to multi-hop. Future work might concentrate on mesh network and sensor networks. While the use of cooperative control has proven to be useful in helping to understand how to obtain improved resilience and capacity, future research should move beyond looking at cooperative control alone and consider how it can be combined with mesh or sensor networks to further enhance performance on awareness-related tasks and activities.

- Additionally, research could be carried out into the area of network planning. As part of this work, the prediction model offers the possibility for accuracy radio coverage estimation and responds to the propagation environment changing. The network planning could be further improved and reduce the cost and time, if the ideas here are incorporated into a network planning tool. In this thesis, only existing antenna systems are considered. Further research extending to the use of MIMO and smart antennas will be interesting with cooperative control for large scale network optimization. The approach is in fact eminently suitable for application to smart antennas.

Appendix A-Author's publications

Refereed Journal Paper

---

1. Peng Jiang, John Bigham, and Anas Khan M., "Distributed Algorithm for Real Time Cooperative Synthesis of Wireless Cell Coverage Patterns," *IEEE Communications Letters,* vol. 12, no. 9, pp. 702-704, Sep 2008.

2. Peng Jiang, John Bigham, and Jiayi Wu, "Self-organizing Relay Stations in Relay Based Cellular Networks," *Computer Communications,* vol. 31, no. 13, pp. 2937-2945, 2008.

Refereed Conference Paper

---

3. Peng Jiang, John Bigham, and Jiayi Wu, "Distributed algorithm for cooperative coverage provisioning in mobile cellular networks," *Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference* on 19-21 Nov. 2008, pp.1565 – 1568.

4. Peng Jiang, John Bigham, and Jiayi Wu, "Scalable QoS Provisioning and Service Node Selection in Relay Based Cellular Networks," *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference* on 12-14 Oct. 2008.

5. Peng Jiang, John Bigham, and Jiayi Wu, "Cooperative Geographic Load Balancing in Hybrid Cellular Networks," *Signal Processing, Communications and Networking, 2008. ICSCN '08. International Conference* on 4-6 Jan. 2008, pp.288 – 294.

6. Peng Jiang, John Bigham, and Jiayi Wu, "Providing enhanced QoS differentiation to customers using geographic load balancing," *Wireless Technology, 2006. The 9th European Conference* on 10-12 Sept. 2006, pp.154 – 157.

7. Haris Pervaiz, John Bigham, Peng Jiang and Celana Mei., "A Game Theoretic Based Call Admission Control Scheme for Competing WiMAX Networks," *The 2nd IEEE International Conference On Computer, Control & Communication* on Feb 17-18. 2009.

8. Jiayi Wu, Peng Jiang, and John Bigham, "Adaptive Cellular Coverage for Radio Resource Management in Mobile Communications," *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference* on 12-14 Oct. 2008.

9. Jiayi Wu, John Bigham, Peng Jiang and Yapeng Wang, "Intelligent Control Of Cellular Network Coverage Using Semi-Smart Antennas For Load Balancing," *Signal Processing, Communications and Networking, 2008. ICSCN '08. International Conference* on 4-6 Jan. 2008, pp.295 – 301.

10. Jiayi Wu, John Bigham, Peng Jiang and Neophytou J. P., "Tilting and Beam-shaping for Traffic Load Balancing in WCDMA Network," Wireless Technology, 2006. The 9th European Conference on 10-12 Sept. 2006, pp. 63 – 66

References

[1] D. Gray, "*Mobile WiMAX-Part I: A Technical Overview and Performance Evaluation*", WiMAX Forum 2006.

[2] K. Kyun, and P. Ramjee, *4G roadmap and emerging communication technologies*, Artech House, 2006.

[3] H. Yanikomeroglu, "Fixed and mobile relaying technologies for cellular networks," *Second Workshop on Applications and Services in Wireless Networks*, pp. 75-81,2002.

[4] D. Schultz, B. Walke, R. Pabst *et al.*, "Fixed and Planned Relay Based Radio Network Deployment Concepts," *Wireless World Research Forum*, pp. 1-6,2003.

[5] Q. Chunming, and W. Hongyi, "iCAR: an integrated cellular and ad-hoc relay system," *Computer Communications and Networks, 2000. Proceedings. Ninth International Conference on*, pp. 154-161,2000.

[6] W. Hongyi, Q. Chunming, S. De *et al.*, "Integrated cellular and ad hoc relaying systems: iCAR," *Selected Areas in Communications, IEEE Journal on,* vol. 19, no. 10, pp. 2105-2115, 2001.

[7] H. Li, M. Lott, M. Weckerle *et al.*, "Multihop communications in future mobile radio networks," *Personal, Indoor and Mobile Radio Communications, 2002. The 13th IEEE International Symposium on*, pp. 54-58 vol.51,2002.

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

[9] H. Viswanathan, S. Mukherjee, and R. Ramjee, "Hybrid networks: cellular-relay architecture," *Wireless and Optical Communications, 2005. 14th Annual WOCC 2005. International Conference on*, pp. 47,2005.

[10] Y. Lin, and Y. Hsu, "Multihop cellular: A new architecture for wireless communications," *IEEE INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, pp. 1273 - 1282,2000.

[11] T. S. P. Yum, and W. S. Wong, "Hot-spot traffic relief in cellular systems," *Selected Areas in Communications, IEEE Journal on,* vol. 11, no. 6, pp. 934-940, 1993.

[12] A. Lindsay-Stewart, W. C. Y. Lee, M. A. Schulz *et al.*, "Incremental capacity gains for high blocking sites using dynamic channel sharing," *Vehicular Technology, IEEE Transactions on,* vol. 50, no. 1, pp. 1-11, 2001.

[13] Y. Argyropoulos, S. Jordan, and S. P. R. Kumar, "Dynamic channel allocation in interference-limited cellular systems with uneven traffic distribution," *Vehicular Technology, IEEE Transactions on,* vol. 48, no. 1, pp. 224-232, 1999.

[14] F. Delli Priscoli, N. P. Magnani, V. Palestini *et al.*, "Application of dynamic channel allocation strategies to the GSM cellular network," *Selected Areas in Communications, IEEE Journal on,* vol. 15, no. 8, pp. 1558-1567, 1997.

[15] D. Everitt, and D. Manfield, "Performance analysis of cellular mobile communication systems with dynamic channel assignment," *Selected Areas in Communications, IEEE Journal on,* vol. 7, no. 8, pp. 1172-1180, 1989.

[16] T. Dong-wan, K. Suk-yon, and J. Go-whan, “Load analysis of the soft handoff scheme in a CDMA cellular system,” *Selected Areas in Communications, IEEE Journal on,* vol. 19, no. 6, pp. 1147-1152, 2001.
[17] L. Du, J. Bigham, and L. Cuthbert, “Towards intelligent geographic load balancing for mobile cellular networks,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews,* vol. 33, no. 4, pp. 480-491, 2003.
[18] L. Du, J. Biaham, and L. Cuthbert, "A bubble oscillation algorithm for distributed geographic load balancing in mobile networks," *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, pp. 330-338,2004.
[19] A. Mishra, *Fundamentals of Cellular Network Planning and Optimisation: 2G/2.5 G/3G... Evolution to 4G*, Wiley-Interscience, 2004.
[20] A. Tanenbaum, *Computer networks*, Prentice Hall Professional Technical Reference, 2002.
[21] C. Murthy, and B. Manoj, *Ad hoc wireless networks: architectures and protocols*, Prentice Hall PTR Upper Saddle River, NJ, USA, 2004.
[22] D. Johnson, and D. Maltz, “Dynamic source routing in ad hoc wireless networks,” *KLUWER INTERNATIONAL SERIES IN ENGINEERING AND COMPUTER SCIENCE*, pp. 153-179, 1996.
[23] C. Lin, and J. Liu, “QoS routing in ad hoc wireless networks,” *IEEE Journal on selected areas in communications,* vol. 17, no. 8, pp. 1426-1438, 1999.
[24] R. Ananthapadmanabha, B. Manoj, and C. Murthy, “Multi-hop cellular networks: the architecture and routing protocols,” *Proc. IEEE PIMRC 2001*, pp. 78-83, 2001.
[25] L. Yeonwoo, S. Park, and S. Yun, “Resource Management for Fixed Relay Based Cellular Systems with Sectorization,” *IEICE Transactions on Communications*, pp. 625-628, 2006.
[26] IEEE. "IEEE 802.16's Relay Task Group," http://www.ieee802.org/16/relay/.
[27] IEEE, “Proposed Technical Requirements for IEEE 802.16 Relay TG,” *IEEE 802.16j-06/016*, 2006.
[28] H. Jeon, Y. Jung, B. Kwon *et al.*, "Analysis on coverage and capacity in adoption of repeater systemsin CDMA2000," *Broadband Communications, 2002. Access, Transmission, Networking. 2002 International Zurich Seminar on*, pp. 33-31,2002.
[29] R. Baines, "The Need for WiMAX picocell & Femtocells," *WiMax London 2007*, pp. 1-36,2007.
[30] V. Chandrasekhar, J. Andrews, and A. Gatherer, “Femtocell networks: a survey,” *Communications Magazine, IEEE,* vol. 46, no. 9, pp. 59-67, 2008.
[31] D. Lopez-Perez, A. Valcarce, G. De La Roche *et al.*, "Access methods to WiMAX femtocells: A downlink system-level case study," *Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference on*, pp. 1657-1662,2008.
[32] A. Mishra, *Fundamentals of cellular network planning and optimisation*, Wiley Hoboken, NJ, 2004.
[33] J. Schiller, *Mobile communications*, Addison Wesley Longman, 2003.
[34] L. Jaana, W. Achim, and N. Tomas, *Radio Network Planning and Optimisation for UMTS*, Wiley, 2002.

[35] J. Romero, O. Sallent, R. Agusti *et al.*, *Radio resource management strategies in UMTS*, John Wiley & Sons, 2005.

[36] N. L. Banerjee, K. Basu, and S. K. Das, "Adaptive resource management for multimedia applications in wireless networks," *World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. Sixth IEEE International Symposium on a*, pp. 250-257,2005.

[37] E. W. Knightly, and N. B. Shroff, "Admission control for statistical QoS: theory and practice," *Network, IEEE,* vol. 13, no. 2, pp. 20-29, 1999.

[38] F. Yuguang, and Z. Yi, "Call admission control schemes and performance analysis in wireless mobile networks," *Vehicular Technology, IEEE Transactions on,* vol. 51, no. 2, pp. 371-382, 2002.

[39] S. Das, S. Sen, and R. Jayaram, "A dynamic load balancing strategy for channel assignment using selective borrowing in cellular mobile environment," *Wireless Networks,* vol. 3, no. 5, pp. 333-347, 1997.

[40] IEEE. "The IEEE 802.16 Working Group on Broadband Wireless Access Standards," http://wirelessman.org/.

[41] IEEE, "IEEE Std 802.16e-2005 Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands.," *IEEE 802.16e*, 2006.

[42] IEEE, "Air Interface for Fixed Broadband Wireless Access Systems," *IEEE STD 802.16 - 2004*, 2004.

[43] J. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking (Prentice Hall Communications Engineering and Emerging Technologies Series)*, Prentice Hall PTR Upper Saddle River, NJ, USA, 2007.

[44] T. Rappaport, *Wireless communications: principles and practice*, Prentice Hall PTR Upper Saddle River, NJ, USA, 2001.

[45] J. Zander, "Radio resource management in future wireless networks: requirements and limitations," *Communications Magazine, IEEE,* vol. 35, no. 8, pp. 30-36, 1997.

[46] S. Das, S. Sen, and R. Jayaram, "A structured channel borrowing scheme for dynamic load balancing incellular networks," *Proceedings of the 17th International Conference on Distributed Computing Systems*, pp. 116-123, 1997.

[47] J. Li, N. Shroff, and E. Chong, "The study of a channel sharing scheme in wireless cellular networksincluding handoffs," *IEEE INFOCOM'99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, pp. 1179-1186,1999.

[48] E. Del Re, R. Fantacci, and G. Giambene, "Handover and dynamic channel allocation techniques in mobilecellular networks," *IEEE Transactions on Vehicular Technology,* vol. 44, no. 2, pp. 229-237, 1995.

[49] S. Anand, A. Sridharan, and K. Sivarajan, "Performance analysis of channelized cellular systems with dynamic channel allocation," *IEEE Transactions on Vehicular Technology,* vol. 52, no. 4, pp. 847-859, 2003.

[50] J. Blogh, P. Cherriman, and L. Hanzo, "Dynamic channel allocation techniques using adaptive modulation andadaptive antennas," *IEEE Journal on selected areas in communications,* vol. 19, no. 2, pp. 312-321, 2001.

[51] A. Jalali, "On cell breathing in CDMA networks," *IEEE International Conference on Communications,ICC 98. Conference Record*, pp. 985-988,1998.

[52] S. Hyukmin, L. Sanghoon, K. Soo-Chang *et al.*, “Soft Load Balancing Over Heterogeneous Wireless Networks,” *Vehicular Technology, IEEE Transactions on,* vol. 57, no. 4, pp. 2632-2638, 2008.

[53] S. Kundu, and S. Chakrabarti, "Call blocking in a mobile radio system with directed retry and priority handoff," *Personal Wireless Communications, 2002 IEEE International Conference on*, pp. 163-167,2002.

[54] M. J. Feuerstein, "Applications of smart antennas in cellular networks," *IEEE Antennas and Propagation Society International Symposium*, pp. 1096-1099,1999.

[55] T. Kitahara, Y. Kishi, Y. Imagawa *et al.*, "An adaptive load balancing in multi-hop mesh networks for broadband fixed wireless access systems," *Radio and Wireless Conference, 2004 IEEE*, pp. 463-466,2004.

[56] E. P. C. Jones, M. Karsten, and P. A. S. Ward, "Multipath load balancing in multi-hop wireless networks," *Wireless And Mobile Computing, Networking And Communications, 2005. (WiMob'2005), IEEE International Conference on*, pp. 158-166,2005.

[57] Y. Yang, J. Wang, and R. Kravets, “Interference-aware load balancing for multihop wireless networks,” *University of Illinois at Urbana-Champaign, Tech. Rep*, pp. 2005-2526, 2005.

[58] P. Jiang, J. Bigham, and A. K. M., “Distributed Algorithm for Real Time Cooperative Synthesis of Wireless Cell Coverage Patterns,” *IEEE Communications Letters,* vol. 12, no. 9, pp. 702-704, 2008.

[59] P. Behalf, "*Mobile WiMAX Part II: A Competitive Analysis*", WiMAX Forum, 2006.

[60] B. Zeigler, H. Praehofer, and T. Kim, *Theory of modeling and simulation*, Academic Press, 2000.

[61] R. Karp, and M. Luby, "Monte-Carlo algorithms for enumeration and reliability problems," *24th Annual Symposium on Foundations of Computer Science*, pp. 56-64,1983.

[62] B. Zeigler, *Multifacetted modelling and discrete event simulation*, Academic Press Professional, Inc. San Diego, CA, USA, 1984.

[63] J. Pitts, and J. Schormans, *Introduction to IP and ATM design and performance*, Wiley, 2000.

[64] P. L'Ecuyer, " Random Number Generation," *Handbook on Simulation*, pp. 93-137: Jerry Banks Ed., Wiley, 1998.

[65] G. Project. "BSD Random Number Functions," http://www.gnu.org/s/libc/manual/html_node/BSD-Random.html.

[66] Z. Gutterman, B. Pinkas, and T. Reinman, "Analysis of the Linux random number generator," *2006 IEEE Symposium on Security and Privacy*, pp. 15 pp.-385,2006.

[67] OenBSD.org. "Cryptography in Open BSD," http://www.openbsd.org/crypto.html.

[68] P. V. Orlik, and S. S. Rappaport, “Traffic performance and mobility modeling of cellular communications with mixed platforms and highly variable mobilities,” *Proceedings of the IEEE,* vol. 86, no. 7, pp. 1464-1479, 1998.

[69] F. Yuguang, I. Chlamtac, and L. Yi-Bing, “Portable movement modeling for PCS networks,” *IEEE Transactions on Vehicular Technology,* vol. 49, no. 4, pp. 1356-1363, 2000.

[70] C. Hoymann, and S. Goebbels, "Dimensioning Cellular WiMAX Part I: Singlehop Networks," *Proceedings of 13th European Wireless Conference*, pp. 150 - 157,2007.
[71] M. Ergen, S. Coleri, and P. Varaiya, "QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access systems," *IEEE Transactions on Broadcasting,* vol. 49, no. 4, pp. 362-370, 2003.
[72] R. Agrawal, R. Berry, J. Huang *et al.*, "Optimal scheduling for OFDMA systems," *Asilomar Conference on Signals, Systems, and Computers*, pp. 1347-1351, 2006.
[73] S. Grandhi, and J. Zander, "Constrained power control in cellular radio systems," *IEEE 44th Vehicular Technology Conference*, pp. 824-828,1994.
[74] J. Y. Jaehyeong Kim, "Downlink Power Control for OFDMA PHY," *IEEE 802.16 Broadband Wireless Access Working Group*, 2004.
[75] A. M. Jun Zhou, Linghang, "Downlink Power Control for WiMAX Femtocell in IEEE 802.16m," *IEEE 802.16 Broadband Wireless Access Working Group*, 2008.
[76] A. Law, and W. Kelton, *Simulation Modeling and Analysis*, McGraw-Hill Higher Education, 1997.
[77] R. wireless. "RDW-16e-OFDM-Enhanced White Paper- A WiMAX Compliant Technology," <http://www.rdwic.com/files/RDW-16e-OFDM-Enhanced%20-Whitepaper-09202007.pdf>.
[78] K. Taek Jin, and M. Gerla, "Clustering with power control," *Military Communications Conference Proceedings, 1999. MILCOM 1999. IEEE*, pp. 1424-1428 vol.1422,1999.
[79] D. Kim, Y. Chang, and J. Lee, "Pilot power control and service coverage support in CDMA mobile systems," *IEEE 49th Vehicular Technology Conference*, pp. 1464-1468,1999.
[80] I. Siomina, "Pilot power management in WCDMA networks: coverage control with respect to traffic distribution," *Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, pp. 276-282,2004.
[81] K. Valkealahti, A. Hoglund, J. Parkkinen *et al.*, "WCDMA common pilot power control with cost function minimization," *IEEE 56th Vehicular Technology Conference, 2002. Proceedings. VTC 2002-Fall*, pp. 2244- 2247,2002.
[82] H. Zhu, T. Buot, R. Nagaike *et al.*, "Load balancing in WCDMA systems by adjusting pilot power," *Wireless Personal Multimedia Communications, 2002. The 5th International Symposium on*, pp. 936- 940,2002.
[83] R. Love, K. Beshir, D. Schaeffer *et al.*, "A pilot optimization technique for CDMA cellular systems," *IEEE VTS 50th Vehicular Technology Conference*, pp. 2238-2242,1999.
[84] H. Yagoobi, "Scalable OFDMA Physical Layer in IEEE 802.16 WirelessMAN," *Intel Technology Journal,* vol. 08, August 2004.
[85] AWE. "WiMAX Network Planning Tool," http://www.awe-communications.com/Network/WiMAX/index.html.
[86] L. Zordan, and N. Rutazihana, "Niek Engelhart Capacity Enhancement of Cellular Mobile Network Using a Dynamic Electrical Down-tilting Antenna System," *Vehicular Technology Conference*, pp. 19-22,1999.

[87] D. Lee, C. Xu, and A. Inc, "Mechanical antenna downtilt and its impact on system design," *1997 IEEE 47th Vehicular Technology Conference*, pp. 447-451,1997.
[88] G. Wilson, "Electrical downtilt through beam-steering versus mechanicaldowntilt [base station antennas]," *IEEE 42nd Vehicular Technology Conference*, pp. 1-4,1992.
[89] J. Niemela, “Optimum antenna downtilt angles for macrocellular WCDMA network,” *EURASIP Journal on Wireless Communications and Networking,* vol. 2005, no. 5, pp. 816-827, 2005.
[90] J. Niemela, T. Isotalo, J. Borkowski *et al.*, "Sensitivity of optimum downtilt angle for geographical traffic load distribution in WCDMA," *IEEE VEHICULAR TECHNOLOGY CONFERENCE*, pp. 1202-1206,2005.
[91] S. C. Sungkyung Kim, Chulsik Yoon, “Group Handover on the Mobile RS,” in IEEE S802.16j-06/227, 2006.
[92] I. Akyildiz, and X. Wang, “A survey on wireless mesh networks,” *Communications Magazine, IEEE,* vol. 43, no. 9, pp. 23-30, 2005.
[93] L. Godara, “Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations,” *Proceedings of the IEEE,* vol. 85, no. 8, pp. 1195-1245, 1997.
[94] C. Balanis, *Antenna Theory: Analysis and Design*, Wiley-Interscience, 2005.
[95] R. Mailloux, *Phased array antenna handbook (Second Edition)*, Boston, MA, Artech House, 2005.
[96] B. Sklar, “Rayleigh fading channels in mobile digital communication systems. I. Characterization,” *Communications Magazine, IEEE,* vol. 35, no. 7, pp. 90-100, 1997.
[97] J. Eaves, and E. Reedy, *Principles of modern radar*, p.^pp. pp.567-597, Van Nostrand Reinhold New York, 1987.
[98] R. S. Kawitkar, and D. G. Wakde, "An approach for MUSIC algorithm in smart antenna system," *Antenna Technology: Small Antennas and Novel Metamaterials, 2005. IWAT 2005. IEEE International Workshop on*, pp. 403-406,2005.
[99] L. Correia, *Wireless Flexible Personalised Communications: COST 259, European Co-operation in Mobile Radio Research*, Wiley, 2001.
[100] H. Asplund, A. A. Glazunov, A. F. Molisch *et al.*, “The COST 259 Directional Channel Model Part I: Overview and methology,” *Wireless Communications, IEEE Transactions on,* vol. 5, no. 12, pp. 3421-3433, 2006.
[101] H. Asplund, A. A. Glazunov, A. F. Molisch *et al.*, “The COST 259 Directional Channel Model Part II: Macrocells,” *Wireless Communications, IEEE Transactions on,* vol. 5, no. 12, pp. 3434-3450, 2006.
[102] Y.-B. Lin, and V. W. Mak, “Eliminating the boundary effect of a large-scale personal communication service network simulation,” *ACM Trans. Model. Comput. Simul.,* vol. 4, no. 2, pp. 165-190, 1994.
[103] M. Nawrocki, H. Aghvami, and M. Dohler, *Understanding UMTS radio network modelling, planning and automated optimisation: theory and practice*, John Wiley & Sons, 2006.
[104] W. Stallings, *Wireless communications & networks*, Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 2004.
[105] R. Vaughan, and J. Andersen, *Channels, propagation and antennas for mobile communications*, Iet, 2003.

[106] J. Yin, Q. Yang, and L. Ni, "Adaptive temporal radio maps for indoor location estimation," pp. 85-94,2005.

[107] J. Yin, Q. Yang, and L. Ni, "Learning adaptive temporal radio maps for signal-strength-based location estimation," *IEEE Transactions on Mobile Computing,* vol. 7, no. 7, pp. 869-883, 2008.

[108] P. Bernardin, M. Yee, and T. Ellis, "Cell radius inaccuracy: a new measure of coverage reliability," *IEEE Transactions on Vehicular Technology,* vol. 47, no. 4, pp. 1215-1226, 1998.

[109] J. Parsons, and Knovel, *The mobile radio propagation channel*, Wiley New York, 2000.

[110] A. Annamalai, C. Tellambura, V. Bhargava *et al.*, "Simple and accurate methods for outage analysis in cellular mobile radio systems-unified approach," *IEEE Transactions on Communications,* vol. 49, no. 2, pp. 303-316, 2000.

[111] A. Kurniawan, "Prediction of Mobile Radio Propagation By Regression Analysis of Signal Measurements," *Magazine of Electrical Engineering (Indonesian,Majalah Ilmiah Teknik Elektro),* vol. 3, pp. 11-21, 1997.

[112] J. B. Andersen, T. S. Rappaport, and S. Yoshida, "Propagation measurements and models for wireless communications channels," *Communications Magazine, IEEE,* vol. 33, no. 1, pp. 42-49, 1995.

[113] B. Sklar, "Rayleigh fading channels in mobile digital communication systems part II: Mitigation," *IEEE Communications Magazine,* vol. 35, no. 9, pp. 148-155, 1997.

[114] G. Pollini, "Trends in handover design," *IEEE Communications Magazine,* vol. 34, no. 3, pp. 82-90, 1996.

[115] S. Su, J. Chen, and J. Huang, "Performance analysis of soft handoff in CDMA cellular networks," *IEEE Journal on selected areas in communications,* vol. 14, no. 9, pp. 1762-1769, 1996.

[116] R. Esmailzadeh, and M. Nakagawa, *TDD-CDMA for wireless communications*, Artech House, 2003.

[117] H. Holma, and A. Toskala, *WCDMA for UMTS: Radio access for third generation mobile communications*, John Wiley & Sons, Inc. New York, NY, USA, 2000.

[118] W. Honcharenko, J. Kruys, D. Lee *et al.*, "Broadband wireless access," *Communications Magazine, IEEE,* vol. 35, no. 1, pp. 20-26, 1997.

[119] A. Quintero, Y. Elalamy, and S. Pierre, "Performance evaluation of a broadband wireless access system subjected to heavy load," *Computer Communications,* vol. 27, no. 9, pp. 781-791, 2004.

[120] D. Schultz, R. Pabst, and T. Irnich, "Multi-hop based Radio Network Deployment for efficient Broadband Radio Coverage," *Wireless Personel Multimadia Communications,* vol. 2, pp. 377-381, 2003.

[121] D. Amzallag, T. Armarnik, M. Livschitz *et al.*, "Multi-Cell Slots Allocation in OFDMA Systems," *Mobile and Wireless Communications Summit, 2007. 16th IST*, pp. 1-5, 2007.

[122] H. Kim, Y. Han, and J. Koo, "Optimal subchannel allocation scheme in multicell OFDMA systems," *Vehicular Technology Conference, 2004. VTC 2004-Spring. 2004 IEEE 59th*, pp. 1821- 1825,2004.

[123] G. Li, and H. Liu, "Downlink dynamic resource allocation for multi-cell OFDMA system," *2003 IEEE 58th Vehicular Technology Conference, 2003. VTC 2003-Fall*, pp. 1698- 1702,2003.

[124] M. Ergen, S. Coleri, and P. Varaiya, "QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access systems," *Broadcasting, IEEE Transactions on,* vol. 49, no. 4, pp. 362-370, 2003.

[125] G. M. S Pfletschinger, J Speidel "Efficient subcarrier allocation for multiple access in OFDM systems," *7th International OFDM-Workshop 2002 (InOWo'02)*, pp. 21-25,2002.

[126] S. Kittipiyakul, and T. Javidi, "Resource Allocation in OFDMA with Time-Varying Channel and Bursty Arrivals," *Communications Letters, IEEE,* vol. 11, no. 9, pp. 708-710, 2007.

[127] N. F. Mehr, and H. Khoshbin, "Performance analysis of relay selection methods for IEEE802.16j," *Electrical and Computer Engineering, 2008. ICECE 2008. International Conference on*, pp. 484-487,2008.

[128] S. Shah, and H. Mouftah, "Scalable quality of service network design," *2000 IEEE International Conference on Communications, 2000. ICC* pp. 735-739,2000.

[129] H. Yaghoobi, "Scalable OFDMA physical layer in IEEE802. 16," *Intel Technol J,* vol. 8, no. 03, 2004.

[130] P. Nahi, C. G. Parini, S. Papadopoulos *et al.*, "A semi-smart antenna concept using real-time synthesis for use in a distributed load balancing scheme for cellular networks," *Antennas and Propagation, 2003. (ICAP 2003). Twelfth International Conference on (Conf. Publ. No. 491)*, pp. 168-171 vol.161,2003.

[131] P. Jiang, J. Bigham, and J. Wu, "Scalable QoS Provisioning and Service Node Selection in Relay Based Cellular Networks," *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on*, pp. 1-4,2008.