DIRECTIONAL CELL BREATHING - A FRAMEWORK FOR CONGESTION CONTROL AND LOAD BALANCING IN BROADBAND WIRELESS NETWORKS

by

KHALED AMER AHMED ALI

A thesis submitted to the
Department of Electrical and Computer Engineering
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
April 2009

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-65437-8
Our file Notre référence
ISBN: 978-0-494-65437-8

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.
.
The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Despite the tremendous bandwidth increase in 3^{rd} generation *(3G)* Broadband Wireless Networks *(BWNs)* such as Universal Mobile Telecommunication System *(UMTS)*, maintaining the mobile users' Quality of Service *(QoS)* requirements while maximizing the network operators' revenues is still a challenging issue. Moreover, spatial distribution of network traffic has a negative impact on the overall network performance where network resources are overutilized in parts of the network coverage area while such resources are underutilized in other network coverage areas. Therefore, network congestion and traffic imbalance become inevitable. Hence, efficient Radio Resource Management *(RRM)* techniques which release congestion and balance network traffic are of utmost need for the success of such wireless cellular systems. Congestion control and load balancing in *BWNs* are, however, challenging tasks due to the complexity of these systems and the multiple dimensions that need to be taken into consideration. Examples of such issues include the diverse *QoS* requirements of the supported multimedia services, the interference level in the system, which vary the mobile users and base stations allocated transmission powers and transmission rates to guarantee certain *QoS* levels during the lifetime of mobile users connections.

In this thesis, we address the problem of congestion control and load balancing in *BWNs* and propose efficient network coverage adaptation solution in order to deal with these issues, and hence enhance the *QoS* support in these systems. Specifically, we propose a directional coverage adaptation framework for *BWNs*. The framework is designed to dynamically vary the coverage level of network cells to release system congestion and balance traffic load by forcing mobile users handoff from a loaded cell to its nearby lightly loaded cell. The framework consists of three related components, namely directional coverage adaptation module, congestion control and load balancing protocol, and *QoS* provisioning module. These components interact with each other to release system congestion, balance network load, maximize network resource utilization, while maintaining the required *QoS* parameters for individual mobile users.

Co-Authors

◇ **Chapter 4**

- Khaled A. Ali, Hossam Hassanein, Abd-Elhamid M. Taha, and Hussein T. Mouftah“Directional cell breathing: a module for congestion control and load balancing in WCDMA networks”, International Conference On Communications And Mobile Computing, pages 317-324, July, 2006.
- Khaled A. Ali, Hossam Hassanein, and Hussein T. Mouftah“Directional Cell Breathing: Framework for Congestion Control and Load Balancing in *WCDMA* Systems”, Journal Paper, in preparation.

◇ **Chapter 5**

- Khaled A. Ali, Hossam Hassanein, and Hussein T. Mouftah“Directional Cell Breathing Based Reactive Congestion Control in WCDMA Cellular Networks”, IEEE Symposium on Computers and Communications, pages 685-690, July, 2007.
- Khaled A. Ali, Hossam Hassanein, and Hussein T. Mouftah“A Novel Dynamic Directional Cell Breathing Mechanism with Rate Adaptation for Congestion

Control in WCDMA Networks", IEEE Wireless Communications and Networking Conference, pages 2927-2932, March, 2008.

- Khaled A. Ali, Hossam Hassanein, and Hussein T. Mouftah"Directional Cell Breathing Based Congestion Control and Load Balancing Protocol", Journal Paper, in preparation.

◇ **Chapter 6**

- Khaled A. Ali, Hossam Hassanein, and Hussein T. Mouftah"Power-Controlled Rate and Coverage Adaptation for WCDMA cellular networks", IEEE Symposium on Computers and Communications, pages 194-200, July, 2008.
- Khaled A. Ali, Hossam Hassanein, and Hussein T. Mouftah"QoS provisioning in WCDMA cellular networks through rate and coverage adaptation", IEEE Conference on Local Computer Networks, pages 661-666, October, 2008.
- Khaled A. Ali, Hossam Hassanein, and Hussein T. Mouftah"DCB-based QoS provisioning in WCDMA cellular networks through rate and coverage adaptation", Journal Paper, in preparation.

Acknowledgments

I thank Allah *(SWT)*, the Holy, the Creator, the most Gracious, the most Merciful, and the Wise, whose help and support are unbounded and gave me patience and ability to reach this stage of knowledge.

I would like to thank Prof. Hossam S. Hassanein and Prof. Hussein T. Mouftah for their supervision of my Ph.D studies.

Prof. Hassanien, I would like to express my thankfulness and gratitude for a wonderful journey and experience. Your patience, experience, encouragement, and endless support are all appreciated beyond words limits.

Prof. Mouftah, I thank you for your wisdom, experience, and encouragement which helped me to carry this work to the stage I am in.

I would like to thank the examination committee members for their valuable remarks and considerable recommendations.

I would like to thank Ms. Bernice Ison and Ms. Debra Fraser for their support and kindness care since I arrived to Kingston.

I acknowledge with great appreciation the funding provided by the Libyan Education Ministry, Queen's University, and the Communication and Information Technology Ontario *(CITO)*.

I am thankful to Abd-Elhamid M. Taha who helped me during my Ph.D work. I thank you very much for your endless support and valuable advises during our long discussion with my endless questions.

Also, I would like to thank my Brother Hassan A. Ahmed for his encouragement and support. Hassan, I wish you a good luck with our wonderful supervisor; Prof. Hassanein.

I would like to thank all of the *TRL Lab* members for the joyful and worthwhile experience that made the course of my studies and stay in Kingston wonderful.

And last but never least, my wonderful family. My parents, my wife and children, my siblings, who continuously supported and encouraged me. You have given me so much, and I can never pay you back. I will owe you gratitude and thankfulness forever.

Contents

List of Tables

List of Figures

List of Acronyms

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	3rd Generation Partnership Project
AFA	Average Fair Adaptation
BCQ	Best Channel quality
BER	Bit Error Rate
BIP	Binary Integer Programming
BS	Base Station
BWNs	Broadband Wireless Networks
CAC	Call Admission Control
CB	Cell Breathing
CBM	Cell Breathing Management
CBR	Case Based Reasoning
CC	Congestion Control
CN	Core Network
CPICH	Common Pilot Channel
CS	Circuit Switched
DCB	Directional Cell Breathing

DCB-CC&LB	DCB-Based Congestion Control and Load Balancing
DS-CDMA	Direct-Sequence Code Division Multiple Access
eCommerce	Electronic Commerce
EDGE	Enhanced Data rate for GSM Evolution
EP	Elementary Procedure
FA	Fair Adaptation
FDD	Frequency Division Duplex
FIFO	First in First Out
FPP	Fixed Pilot Power
GPRS	General Packet Radio Service
GSM	Global System for Mobile communication
H-DCB	Heuristic DCB
IMT-2000	International Mobile Telephony 2000
IP	Internet Protocol
ITU	International Telecommunication Union
LS	Loaded Sector
MA	Minimum Adaptation
MRA	Minimum Rate Adaptation
MSC	Mobile Switching Center
NBAP	Node B Application Part
O&M	Operation and Management
PCRCA	Power Controlled Rate and Coverage Adaptation
PDCB	Proactive DCB
PS	Packet Scheduling

PSTN	Public Switching Telephony Network
QoS	Quality of Service
RA	Rate Adaptation
RAB	Radio Access Bearer
RAN	Radio Access Network
RDCB	Reactive DCB
RNC	Radio Network Controller
RNS	Radio Network Subsystem
RRM	Radio Resource Management
SIR	Signal-to-Interference Ratio
SPC	Selective Power Control
SRT	Sector Residence Time
SS	Supporting Sector
TDD	Time Division Duplex
TPC	Transmission Power Control
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
UTRAN	UMTS Terrestrial Radio Access Network
VoD	Video on Demand
VoIP	Voice over IP
WCDMA	Wideband Code Division Multiple Access
WCQ	Worst Channel quality
WWW	World Wide Web

Chapter 1

Introduction

The evolution of wireless communications started in the early 1980s when the first wireless cellular systems were commercially launched. These systems marked the beginning of a new era of personal commutations which enabled wireless voice communications. These systems utilized analog air interface technology and offered voice-centric applications only. These analog systems are widely known as first generation *(1G)* cellular networks. The limited bandwidth and low voice quality were the main drawbacks that prevented the wide spread of *1G* cellular systems. Therefore, to overcome the shortages of *1G* systems, digital wireless cellular systems, known as second generation *(2G)*, were introduced. *2G* systems utilized digital access air interference and provided a large bandwidth and higher voice quality. Although *2G* systems were designed to support voice communications, they had the capability to support limited data services enabled by their evolved systems such as the evolved systems of Global System for Mobile communication *(GSM)* of General Packet Radio Service *(GPRS)* and Enhanced Data rate for GSM Evolution *(EDGE)* which are known as *2.5G* systems [13], and [53]. The enhanced bandwidth capabilities and the higher voice quality have led to a worldwide deployment of the *2G* cellular systems and attracted a large number of subscribers.

The tremendous success of the *2G* systems and the widespread use of the Internet has increased user demand for wireless data services. This has motivated the development of Broadband Wireless Networks *(BWNs)* for third generation *(3G)* cellular systems. *BWNs* are characterized by higher bandwidth that enabled multimedia traffic multiplexing over a single channel and improved Quality of Service *(QoS)* support. Therefore, in addition to the legacy circuit switching voice services, these features enabled the *BWNs* to provide wireless multimedia packet services such as Voice over IP *(VoIP)*, video on demand *(VoD)*, mobile Internet browsing. One of the most well-known *3G* systems is Universal Mobile Telecommunications System (UMTS) which was developed by the 3rd Generation Partnership Project *(3GPP)* [30]. *UMTS* supports a transmission rate of up to 2 Mbps. In *UMTS*, multimedia service differentiation is made possible through the introduction of different *QoS* classes [1]. Such classes allow network operators to prioritize different users based on their *QoS* requirements.

UMTS systems are interference-limited due to the frequency reuse factor of *"one"* of the *WCDMA* air interface [20], and [24]. The consequences are reflected in the system capacity degradation in terms of the number of supported mobile users' calls and the offered bandwidth to such calls as well. This makes the *UMTS* base station capacity and cell[1] coverage have an inverse relationship; the base station capacity increases as the cell coverage area is decreased and vice versa. Therefore, for better system resource utilization different radio resource management *(RRM)* approaches are proposed by the *UMTS* standard [2]. These management entities mainly concern

---

[1] The term "cell" refers to the coverage area of a base station.

power control, admission control, congestion control, and packet scheduling [30]. The power control mechanism manages the mobile users' and base stations' transmission powers to minimize the interference level in the system, while guaranteeing the mobile users' minimum *QoS* requirements. Admission control is accomplished by Call Admission Control *(CAC)* schemes, which are responsible for admitting or rejecting new or handoff calls to the system. After admitting such calls to the system, congestion control is responsible for maintaining the system's operational state within the predefined performance levels at the system planning stage. Finally, packet scheduling is concerned with scheduling the data packet transmission of mobile users considering mobile users' channel quality conditions.

In this thesis, we investigate problems of congestion control and load balancing in *BWNs* and propose effective and intelligent solutions to maximize the radio resources and enhance the *QoS* support of these systems. Our solutions are designed to solve congestion and load imbalance scenarios of wireless communication systems and aim to achieve network wide congestion control and load balancing. In particular, we propose a connection level congestion control and load balancing framework, which consists of three related components. These components are a directional cell breathing *(DCB)* coverage adaptation module, a *DCB-Based* congestion control and load balancing protocol, and a *DCB-Based QoS* provisioning module.

1.1 Motivation and Objectives

The *RRM* implementation spans over the User Equipment *(UE)*[2], Node Bs, and the Radio Network Controller *(RNCs)*. Admission control and power control are broadly investigated in the literature. On the other hand, despite their prominent role in maximizing system resource utilization, the congestion control and load balancing mechanisms of current *BWNs* have not been given much attention. Hence, deep investigations of possible congestion control and load balancing mechanisms are needed since such mechanisms are a key complementing factor to the admission control and power control mechanisms in efficiently utilizing system radio resources, maximizing network profit, and increasing the number and satisfaction of mobile users admitted to the system.

As the demand for multimedia services in current and next generation *BWNs* increases, tremendous pressure is placed on the network operators to meet the mobile users' requirements. Therefore, without efficient and scalable radio resource management techniques, congestion and system load imbalance will certainly be a major issue in *BWNs*. When congestion and load imbalance scenarios are formed, network operators become incapable of maintaining the mobile users' assured *QoS* levels which then leads to mobile users' dissatisfaction and network operators' revenue loss. However, congestion control and load balancing in *BWNs* are challenging tasks due to the complexity of the *BWNs* in general and the different *QoS* requirements of the supported multimedia traffic. One of the most complex and performance degrading factors of the *BWNs* is interference, i.e., the received power of a mobile user is an interfering power to other mobile users in the system [24]. Therefore, as the number

---

[2] "User Equipment" is the name used by the *UMTS* standard for a "mobile user terminal". The two terms are interchangeably used with "mobile user" in this thesis.

of admitted mobile users to the system is increased, their received signal quality is degraded. The degradation level of mobile users' received signal quality increases their Bit Error Rate *(BER)* and makes the system incapable of correctly decoding the user information. Therefore, mobile users or base stations are required to increase their transmission power and retransmit the information. In an interference-limited congested system, increasing transmission power worsens the system congestion state and increases data retransmission requests which degrades the system performance level. Therefore, in order to support a large number of mobile users and provide multimedia services, more efficient congestion control and load balancing mechanisms that exploit the characteristics of the physical and network layers, as well as the recent advancements in antenna technology are essential for the operation of the interference limited *UMTS* systems.

Another important issue that must be considered is network traffic hotspots that are formed due to either non-uniform mobile users distribution, wireless multimedia traffic heterogeneity, or base station partial or full malfunction. In the case of spatially localized traffic scenarios caused by certain mobile users' mobility patterns and/or the heterogenous traffic requirements the base station radio resources are rapidly consumed, while such resources are underutilized in nearby base stations. Moreover, when base station hardware failure happens, it degrades its serving capability which increases call blocking and dropping rates. These scenarios negatively affect the *QoS* support capabilities of both the congested cell and its neighboring cells.

Currently, hotspot congestion in *3G* networks is handled through cell breathing in which the coverage area of a hotspot cell is contracted while the coverage areas of its adjacent cells are expanded. Therefore, the capacity of the hotspot cell is increased as its serving area is decreased to enable it to serve more traffic. Cell breathing management is made possible by the fact that mobile users of *WCDMA* systems are not always communicating with their nearest base station, but instead with the base station with the best received channel quality. However, the currently utilized omni-directional cell breathing approach negatively affects the system performance because of the unnecessary expansion of the supporting adjacent cells towards non congested neighboring cells. This may lead to a decrease in the number of served mobile users as a result of the increased interference levels in the system. In addition to the cell breathing mechanism, rate adaptation is also used to release congestion in wireless commutation systems through adapting mobile users' transmission rates to the system load conditions.

Also, in multimedia systems, preserving the provisioned *QoS* requirements to the higher priority traffic while maintaining at least the minimum service requirement level for low priority traffic is a key factor to the success of any proposed congestion control and load balancing schemes. Therefore, the required *QoS* requirements by mobile users need to be maintained by any proposed congestion control or load balancing framework. Also, fairness among different users in the same class and among users of different classes is one of the objectives of the *RRM* functionalities in general and the congestion control mechanism in particular.

Recent advances in wireless smart antenna technology combined with the fast and seamless handoff features of the current wireless cellular systems permit dynamic coverage adaptation of the network cells' coverage areas to meet the dynamic traffic requirements in each call [11][18]. Moreover, the directional property of smart antennas can be utilized to maximize the base station directional coverage range and practice congestion control and load balancing through directional coverage adaptation. These features enable base stations to adapt their serving areas based on their load conditions. Therefore, the utilization of the system limited resources can be increased, congestion can be released, and spatially varied system load conditions can be balanced among nearby base stations. Ultimately, the number of served mobile users and the network operators' revenue are maximized.

Currently, existing congestion control and load balancing schemes deal only with limited congestion and load imbalance scenarios which are mostly performed locally at a single cell level. Also, in some situations, these schemes are inefficient in releasing system congestion. This leads to poor radio resource utilization and hence, degrades the performance of *BWNs*. We, therefore, aim at considering all of the aforementioned issues in designing our framework. Since some of these issues may conflict with one another, e.g. varying the coverage areas of the congested and supporting cells while minimizing the interference level in the system, a trade-off needs to be made to compromise between different available options.

1.2 Thesis Contributions

BWNs are designed to support a large number of mobile users and provide a wide range of multimedia applications with different *QoS* requirements. In this thesis, we propose a congestion control and load balancing framework for *BWNs* composed of three novel modules to simultaneously achieve the following objectives:

1. Minimizing the interference level in the system;

2. Controlling congestion in both reactive and proactive manners;

3. Sharing system load through load balancing;
4. Maximizing the utilization of the system resources; and
5. Supporting multi-class services with users having different *QoS* and bandwidth demand requirements

The work of this thesis focuses on the uplink (i.e., from the mobile users to the base station) congestion control and load balancing in *BWNs* in general and *UMTS* systems in particular. The main contributions of this thesis are outlined in the following subsections:

1.2.1 Directional Cell Breathing

In Chapter 4, a novel congestion control and load balancing module for *BWNs* is proposed in order to release the system congestion and balance network load through nearby cell sectors load sharing. The module exploits the capabilities of smart directional antennas in online dynamic cell configuration. In this module, a controllable directional cell breathing is practised where a cell sectorized coverage is varied reactively in instances of congestion, and proactively in instances of exercising long-term congestion avoidance and load balancing. The concept, which is called Directional Cell Breathing *(DCB)*, overcomes the drawbacks of non-controllable cell breathing management in non-sectorized *WCDMA* network cells by optimizing the coverage levels within each cell sector under constraints minding sectoral traffic load and interference bounds as well as the nearby sector load state. We show that our coverage adaptation scheme suppresses the interference level in the system and increases the number of served mobile users in a loaded sector.

1.2.2 Congestion Control and Load Balancing Protocol

In Chapter 5, we introduce a novel *DCB-based* Congestion Control and Load Balancing *(DCB-CC&LB)* protocol to alleviate congestion and load imbalance problems in large-scale *BWNs* systems. *DCB-CC&LB* exploits the cell breathing properties which are practised in a directional manner to release spatially localized hotspot traffic and balance the system load. The *DCB-CC&LB* protocol is composed of both reactive and proactive schemes. The reactive *DCB-CC&LB* scheme evaluates the effect of practising coverage adaptation on the instantaneous network performance parameters, i.e., system load, average mobile users' transmission power and system average outage ratio. On the other hand, the proactive *DCB-CC&LB* scheme assesses the long-term effects of coverage adaptation on controlling congestion and balancing system load.

1.2.3 QoS Provisioning Module

In Chapter 6, a Power Controlled Rate and Coverage Adaptation *(PCRCA)* module is proposed and analyzed. A heuristic scheme that iteratively evaluates the performance of the proposed module is implemented. The proposed module evaluates the optimal allocation of nearby sectors' coverage levels which is locally, with respect to every cell sector, maximizes the average allocated transmission rates and minimizes the average transmission powers for active mobile users while maintaining the network base stations' loads below predefined thresholds. Moreover, a formula for multimedia traffic is derived to evaluate the system performance given traffic of multiple classes having different *QoS* requirements. We show that our proposed *QoS* provisioning module has the potential of supporting differentiated services based on different traffic requirements.

1.3 Thesis Outline

This thesis is organized as follows. Chapter 2 presents some background material and previous work that are necessary for understanding the discussions to follow. Also, the proposed framework is outlined in Chapter 2. The system model is outlined in Chapter 3. Chapter 4 introduces the proposed *DCB* Module that is designed to provide directional coverage adaptation management. Chapter 5 presents the *DCB-Based* congestion control and load balancing protocol, which utilizes the *DCB* module in releasing traffic congestion and balancing system load in *BWNs.* Chapter 6 introduces the power-controlled rate and coverage adaptation *(PCRCA)* module, which aims at optimizing the network sectors' coverage levels, and mobile users' transmission powers and rates given different system load scenarios. Chapter 6 also evaluates the ability of a *DCB-Enabled WCDMA* system to support multimedia services. Chapter 7 presents the conclusions drawn from the thesis and discusses possible future research directions.

Chapter 2

Background and Framework Overview

In this chapter, we present a background study explaining the evolution of Broadband Wireless Networks *(BWNs)*, expanding on areas relevant to the design and implementation of congestion control and load balancing. We also present a comprehensive literature review of the state-of-the-art of Radio Resource Management *(RRM)* techniques and their interaction in fulfilling the quality of service *(QoS)* requirements in wireless communications systems. Finally, an overview of our proposed Congestion Control and Load Balancing framework is outlined.

2.1 WCDMA for UMTS Cellular Networks

To achieve a unified multimedia radio access network, the International Telecommunication Union *(ITU)*, through the 3^{rd} Generation Partnership Project *(3GPP)*, proposed a universal radio access system for International Mobile Telephony 2000

(IMT-2000) mobile systems. The *3GPP* standardization body was formed in December of 1998 [30]. The *3GPP* adapted the Wideband Code Division Multiple Access *(WCDMA)* air interference access technology as the defacto air interface for *3G BWNs* systems [20] and [30]. The well known *3G* system that uses the *WCDMA* air interface is the Universal Mobile Telecommunication System *(UMTS)* [30]. *UMTS* is designed to achieve a high data rate capable of supporting multimedia communications ranging from the legacy circuit switching for voice services to the Internet Protocol *(IP)* packet switching communications such as Voice over IP *(VoIP)*, Electronic Commerce *(eCommerce)*, and wireless web browsing.

WCDMA has two operation modes; Frequency Division Duplex *(FDD)*, which uses two separate carriers for communication in the uplink (i.e. from mobile users to base stations) and the downlink (i.e. from base stations to mobile users), each of 5 MHz; and Time Division Duplex *(TDD)* consisting of a single 5 MHz carrier divided between the uplink and downlink [20][1]. *WCDMA* is allocated the frequency spectrum around the 2 GHz band. The radio communications principle of the *WCDMA* air interface is to use Direct-Sequence Code Division Multiple Access *(DS-CDMA)* to spread and despread mobile user information bits over a wide bandwidth channel. In the spreading operation, user data bits of rate R are multiplied by a sequence of X code bits called chips before transmitting them over the wireless channel as shown in Figure 2.1. In the despreading operation at the receiver side, the received spread signal is multiplied by the same code to recover the original data. The mobile users transmit over a single broadband frequency channel and are distinguished by their unique spreading codes. The use of the same frequency channel by all mobile users negatively affects the mobile users' received signals. This is because other active

---

[1] The proposed Framework in this thesis assumes FDD operation mode.

mobile users' signals act as an interference power with respect to the received signal of an active mobile user. Therefore, *WCDMA* systems are known as interference limited systems [24].

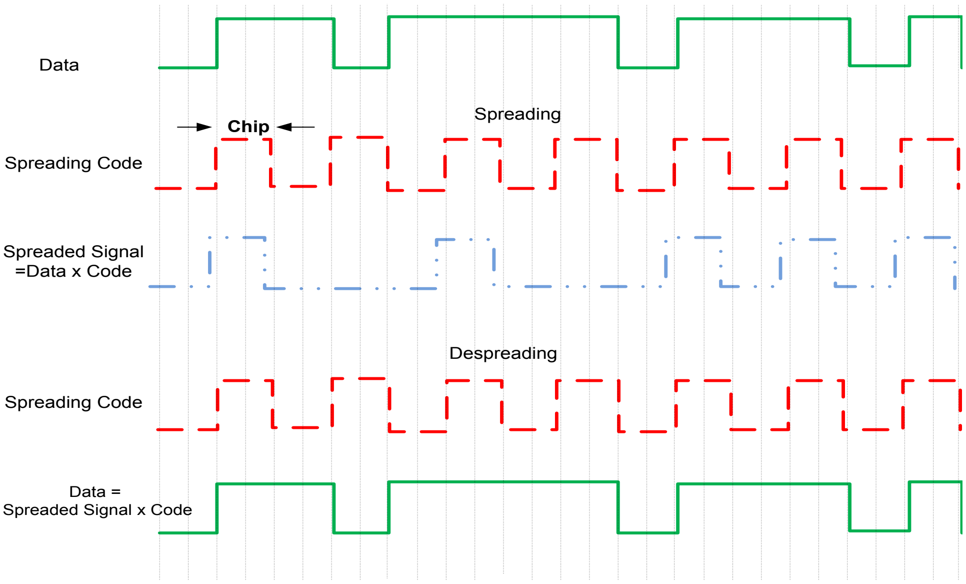


Figure 2.1: WCDMA Spreading and Despreading Principles

The design objectives of the *UMTS* systems are to be able to provide transmission rates up to 2 Mbps, support variable bit rate transmission for bandwidth on demand applications, multiplex multiple services on a single connection, provide support for delay-sensitive real-time traffic to best-effort packet date transmission. Also, support of asymmetric uplink and downlink traffic transmission, worldwide roaming and efficient spectrum utilization [30]. To achieve these objectives, a set of *QoS* classes were defined for *UMTS*: *Conversational, Streaming, Interactive, and Background.* The Conversational class is used for real time applications, e.g. voice, video conferencing. The requirements of the Conversational class traffic include bounded transmission delay and low jitter. The Streaming class is used for real time streaming applications; such as audio or video streaming. Nevertheless, the transmission delay requirements are less stringent in this class compared to the Conversational class. The timing relation between data packets have to be preserved and therefore the jitter must be low. The Interactive class is characterized by the request-response communication pattern of end users. Examples of such services are Hyper Text Transfer Protocol *(HTTP)* traffic and telnet. The *QoS* requirements of the interactive class include low packet error rate and low round trip time. Finally, the Background class is used for best-effort data service without delay requirements such as file transfer and e-mail. The provisioning of these *QoS* requirements is controlled by the *UMTS* Radio Resource Management *(RRM)* through a set of procedures which will be explained in a later section.

2.2 UMTS Radio Access Network Architecture

The *UMTS* Radio Access Network *(RAN)* architecture is made up of the Node Bs and the Radio Network Controllers *(RNCs)* which are contained in the *UMTS Terrestrial Radio Access Network (UTRAN)* as shown in Figure 2.2 [30] and [20]. The *RNC* overlooks the functionalities for one or more Node Bs. Although typical system design separates the implementation of the *RNC* and the Node Bs, they can be jointly implemented [31]. The *RNC* and its corresponding Node Bs are collectively known as the Radio Network Subsystem *(RNS)* where the *UTRAN* can be composed of more than one *RNS*. The *UTRAN* interfaces the core network *(CN)* and the User Equipment *(UE)* to connect a mobile user to another mobile user, to the Public Switched Telephony Network *(PSTN)*, or to the Internet through a set of defined network interfaces.

There are a number of logical interfaces defined for the *UTRAN* components: the

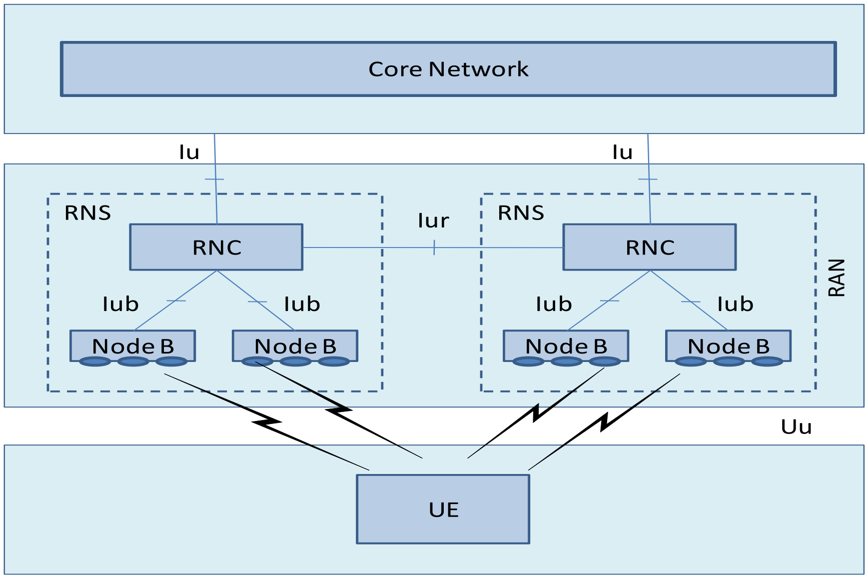


Figure 2.2: UMTS Terrestrial Radio Access Network (UTRAN)

Iub interface between the *RNC* and its Node Bs, the *Iu* which connects the *RNC* to the *CN*, the Uu which connects the Node B with the mobile user, and the *Iur* which connects two *RNCs* together. Through these interfaces, a set of dedicated and shared channels are used for uplink and downlink control and data communications [30], and [20].

The initial deployment of *UMTS* was in Finland in 1999, followed by Spain and Sweden in 2000 [30]. Since then, more than 150 commercial *WCDMA* systems have been deployed worldwide and *UMTS* mobile subscribers nowadays exceed 400 million [78].

2.3 RRM in Wireless Cellular Systems

Radio Resource Management *(RRM)* plays a major role in Quality of Service *(QoS)* provisioning for wireless cellular networks. Specifically, *RRM* techniques along with the network planning and air interface design determine the *QoS* performance at both the mobile users level and the network level. *RRM* policies are composed of frequency and time channels, transmit power, and access to base stations in order to efficiently control the allocated resources to different mobile users with the objective of maximizing a specific function such as network resource utilization, or total network revenue, subject to some constraints such as the maximum call blocking and dropping rates, and/or the minimum signal to interference ratio. The efficient design of *RRM* techniques has a direct impact on the performance of the mobile users in general and on the overall network performance in particular. For instance, the allocated transmission power for a mobile user has a positive effect on the *QoS* for that user, but it negatively affects the interference level for other mobile users which as a result degrades their *QoS* level. *RRM* strategies are not subject to the standardization process, so that they can be a differentiation feature among equipment vendors and operators.

Different *RRM* techniques are proposed for managing the radio resources for wireless communication systems. These techniques can be classified into three sets: Frequency and time allocation, which concerns channel allocation, scheduling, and bandwidth reservation schemes; power allocation, rate allocation and congestion control, which manage the allocation of mobile users' transmission powers and their allocated transmission rates, and reacts to system congestion scenarios; call admission control *(CAC)*, base station assignment, and handoff algorithms, which handle the mobile users' access to the wireless communication system [8]. These different sets interact together as shown in Figure 2.3 to maintain the mobile users' provisioned *QoS* parameters, taking the network loading conditions, network revenue, and other network performance objectives into consideration. In the following subsections, *CAC*, transmission power control *(TPC)*, handoff management *(HM)*, and congestion control *(CC)* techniques in modern wireless cellular networks are reviewed in depth.

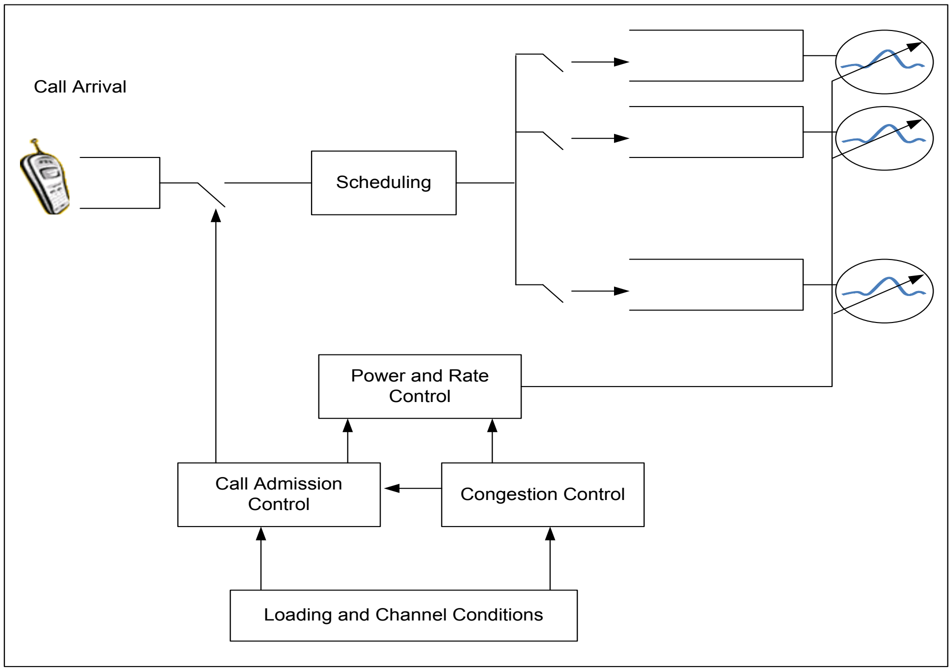


Figure 2.3: Components of Radio Resource Management

2.3.1 Call Admission Control

From the network perspective, admitting a call (new or handoff) has a reward. This reward comes from utilizing the network resources for a certain amount of revenue. However, for such a reward there is also a potential penalty, particularly at high congestion levels, which can be translated as a degradation in the offered *QoS* level to the existing mobile users' calls. Therefore, a balanced network state, which maximizes resource utilization and revenue while maintaining the provisioned *QoS* parameters is the objective of *CAC* schemes. *CAC* has been widely investigated in wireline networks as a *QoS* provisioning and congestion control tool. Different approaches of *CAC* design and performance analysis, particulary in the context of broadband integrated service digital network *(B-ISDN)* based on asynchronous transfer mode *(ATM)* technology, have been investigated in [62]. However, *CAC* in wireless networks is more challenging due to the unique characteristics of wireless networks such as wireless channel quality variation, handoff requirements, and bandwidth limitations.

Due to the growing popularity of wireless communications, *CAC* has been receiving special attention during the last two decades since it plays a central role in *QoS* provisioning in terms of signal quality, call blocking and dropping probabilities, packet delay and loss rate, transmission power, and transmission rate. In first and second generation wireless cellular systems, *CAC* schemes have been generally developed for a single class of service which mainly considers voice application [32], and [46]. However, the introduction of data services and multimedia services in *2.5G* and *3G* wireless systems, respectively has necessitated the development of more sophisticated *CAC* schemes to cope with the stringent *QoS* requirements of these applications. Multi-class *CAC* schemes are more challenging than single-class *CAC* schemes due to the need for considering service prioritization, fairness, and resource sharing policies [10], [14], [15], [16], [28], [33], [38], [45], [44], [49], [58], [61], [62], [89], [88] and [89].

CAC schemes can be classified as centralized or distributed [58], [17], [32] and [89]. Each of these *CAC* design options has its advantages and disadvantages. For example, centralized *CAC* schemes are implemented by a central network unit such as the Mobile Switching Center *(MSC)* in 1G and 2G cellular systems or the *RNC* in *3G* cellular systems. This central unit manages the admission policy in the whole network. Centralized *CAC* schemes are more efficient than distributed *CAC* schemes due to the global information availability but their complexity is high which prevents their use in real systems. Distributed *CAC* schemes are simple, scalable, and more reliable since they are implemented in each network base station. The amount of available information for distributed *CAC* schemes is limited when information exchange between network base stations is not enabled.

CAC schemes are designed with different objectives. For instance, *CAC* schemes for interference-limited systems are designed to maintain the signal quality above a predefined threshold [28], [45], and [46]. The soft capacity of these systems; e.g. *CDMA* systems, is limited by the received interference level [28]. Therefore, the more loaded the network is, the more degraded the signal quality for the existing mobile users' calls in terms of the interference level or the signal to interference ratio *(SIR)*. Hence, a *CAC* scheme is utilized to admit a mobile user's call only if a minimum signal quality is maintained for such call and existing calls as well.

From a mobile user's perspective, dropping an active call is more annoying than blocking a new call. Hence, *CAC* schemes are designed to minimize the mobile users' call dropping rate[2] by reserving some resources for handoff calls exclusively [15]. These approaches are known as guard-channel *CAC* schemes. Despite the mobile users' call dropping rate reduction, such schemes may lead to an increase in mobile users' call blocking rate[3]. Moreover, *CAC* schemes can also be deployed at the packet level. Such schemes permit or reject mobile users' packets to the system based on the system load state. The objectives of such schemes are to minimize packet transmission delay and delay jitter, and packet dropping rate [38]. Therefore, the number of active

---

[2] Call dropping rate is the ratio of unsuccessful handoff calls to the total handoff calls.

[3] Call blocking rate is the ratio of blocked new calls to the total arrived new calls.

calls, network available resources, and an estimate of packet-level *QoS* parameters can be used as a criterion for call admission.

Also, the offered transmission rate for data calls can be considered as an admission criterion for *CAC* schemes to ensure the minimum possible transmission rate of mobile users' data calls [41]. Moreover, service priority is considered in different *CAC* schemes to differentiate mobile users' calls based on their assigned service classes. For example, voice services have been given higher priority than data services due to the common belief among network operators and researchers that voice calls are more rewarding than data calls [45]. This leads to the problem of fairness among users of different classes. Therefore, *CAC* schemes also can be employed to assure fairness among users of different classes such that no calls from a certain class dominate the system resources utilization.

A number of previously proposed *CAC* schemes from the literature are summarized herein. An uplink *CAC* based on the received interference power is proposed in [28]. The uplink capacity formula is used to estimate the average number of mobile users per cell for a given blocking probability. In this scheme, the *Erlang B* formula is extended to consider the system soft blocking. The obtained results show an increased system capacity when soft blocking is considered compared to a system capacity with hard blocking.

A load-based *CAC* scheme is proposed in [45]. In this scheme, the mobile users' signal quality in terms of *SIR* is maintained by controlling the cell load level. Herein, traffic is differentiated into voice and data. Two thresholds are defined to control the number of voice and data calls in the system. The voice traffic is controlled at the admission level, while data traffic is regulated at both admission and packet levels. Once admitted to the system, voice traffic is transmitted without incurring any further delay. This scheme rejects data calls that otherwise experience long delay. Once admitted to the system, data users transmit with a predefined probability to minimize the packet error rate of real-time voice traffic. Despite its efficiency in maintaining the required *QoS* for voice traffic, the system performance with this scheme is degraded as the non-real-time data traffic becomes dominant which is the case in *3G* systems.

In [46], two *SIR-based CAC* schemes are proposed. The first scheme utilizes the residual cell capacity, $R_m = \lfloor \frac{1}{SIR_{th}} - \frac{1}{SIR_m} \rfloor$, as call admission criterion. In the R_m formula, SIR_m is the received uplink *SIR* in cell *m*, SIR_{th} is the *SIR* threshold, and $\lfloor x \rfloor$ is the largest integer smaller than *x*. R_m is periodically computed, and whenever a mobile user's call arrives at cell *m*, R_m is checked. If it is greater than zero, the call is admitted, otherwise the call is rejected. The second scheme is proposed to consider the effect of admitting a mobile user's call not only in the cell of arrival but on the adjacent cells as well. Therefore, the concept of residual capacity is modified to take into consideration such impact. In this scheme, $R_{m,j} = \lfloor \frac{1}{\beta} (\frac{1}{SIR_{th}} - \frac{1}{SIR_m}) \rfloor$, is the minimum residual capacity of the cell of arrival and the adjacent cells which is utilized as the call admission criterion. In $R_{m,j}$, β is the interference coupling between adjacent cells and *C(m)* is the set of adjacent cells to the target cell *m*. When the minimum residual capacity is positive, the arrived call is admitted, otherwise the call is rejected. The performance results of the first scheme outperform the ones of the second scheme for uniformly distributed traffic. For non-uniform traffic distribution, the second scheme shows better performance results.

An *SIR-based CAC* scheme for a combined *TDMA* and *CDMA (T/CDMA)* systems supporting voice and data applications is proposed in [14]. In this scheme, voice traffic is assigned a single code and a single time slot, while two approaches are used for data traffic; multicode or multislot *T/CDMA* allocations. In the multicode approach, a data call is admitted if there is at least one time slot of C codes which assure the minimum *SIR* parameters of the new and excising calls. Herein, C is the ratio of the transmission rate of data users to that of voice users. In the second approach of multislot allocation, the minimum *SIR* parameter of the new and existing calls must be satisfied in all of the C time slots using a single code per slot. In this approach, it is not mandatory to use the same code in the C time slots. The obtained performance results of the multicode *T/CDMA* system outperform the results of multislot *T/CDMA* in terms of system capacity , and also allows a tradeoff between the system capacity and the transmission rate.

2.3.2 Transmission Power Control

Transmission Power Control *(TPC)* is utilized to efficiently allocate base stations' transmission power and mobile users' transmission power in *3G* systems. *TPC* concerns the use of no more transmission power than required to meet the minimum signal-to-interference ratio *(SIR)* constraints for mobile users connections. Hence, using the *TPC* mechanism, the system interference level is minimized and the mobile terminal battery's expected life time is increased. In *WCDMA* systems, transmission power management is realized through the use of a closed loop power control mechanism proposed by the *UMTS* standard [30]. As Figure 2.4 shows, this mechanism comprises two algorithms; a fast inner loop and a slower outer loop. The inner loop aids a mobile user in the uplink to reach its *SIR* target and assists the base station at the downlink in maintaining the mobile received power at an acceptable level. This is achieved through a sequence of power up/down commands at a rate of 1500 Hz. On the other hand, the outer loop is located at the *RNC* to dynamically set the long term *SIR* target so that adequate performance in terms of frame error rate *(FER)* is achieved.

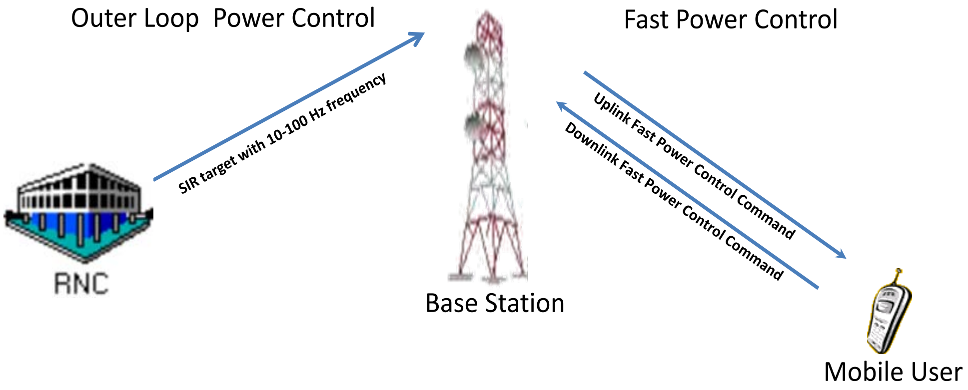
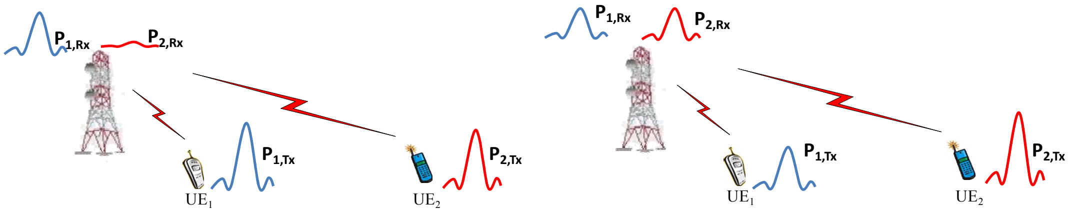


Figure 2.4: Power Control in WCDMA Systems

These power control algorithms are essential in the interference-limited *UMTS* networks (See Figure 2.5). In the uplink, the inner loop algorithm is used to solve the so-called near-far effect; a phenomenon by which strong interference at the base station overwhelms weak signals from distanced mobile users as shown in Figure 2.5(a). Therefore, by dynamically adjusting the transmission power of every mobile user, the base station receives more-or-less the same power from every mobile user as shown in Figure 2.5(b). In the downlink, mobile users near their base stations are typically unaware of the interference power of other cells, while the received signal quality of mobile users near cell edges will be greatly affected by the received interference power from adjacent cells. Therefore, without the downlink inner loop power control at the mobile user, the nearedge mobile users might be unable to decode their signals properly.



(a) No Power Control (b) Power Control

Figure 2.5: Near-Far Effect in WCDMA Systems

In the literature, *TPC* algorithms are extensively studied [36], [35], [39], [55], [56], [85], [23], [26], [25], [27], [12] and [84]. These algorithms investigate different methods for optimal power allocation to the active mobile users in the system. Generally, for N mobile users, a power vector $P = [p_1, p_2, \ldots, p_N]^T$ is defined to maintain the *SIR* level above a predefined target γ_n^{tar}. This value is achieved if the mobile users' power vector $\mathbf{P} > 0$ for all N mobile users exists. Therefore, the system is characterized by an $N \times N$ gain matrix whose entries are determined by mobile users' link gains to their base stations. Such system has been mathematically modeled in [55], [56], and [85].

Based on the $N \times N$ gain matrix, a number of iterative solutions have been developed. In [23], a downlink distributed power control algorithm has been proposed. The algorithm states that whenever there is a power vector that satisfies the *SIR* constraints for the N mobile users, the conversion speed to these values is exponential. Each mobile user iteratively resets its power to a value which is required to attain an acceptable performance level assuming that the interference power imposed by other mobile users are not going to change. Similarly, other mobile users in this algorithm follow the same approach.

In [12], a downlink power control algorithm combined with a transmission scheduling algorithm is proposed for supporting non real-time data traffic. A transmission scheduling interval, $0 \leq t \leq T$, is defined to guarantee a minimum average data rate, R_n, for a mobile user n. The power control algorithm objective is to minimize a power vector p(t), $\min_{p(t)} \sum_{n=1}^{N} \int_0^T p_n(t)dt$, subject to the constraints $\frac{1}{T}\int_0^T r_n(t)dt = R_n$ and $\sum_{n \subset C(m)} p_n(t) = P_m$, where $r_n(t)$ is an instantaneous mobile user n's transmission rate, *C(m)* is the set of connected mobile users to base station m, and P_m is the maximum transmission power of base station m. Moreover, the proposed scheme minimizes the system interference level by intracell transmission scheduling which permits only one mobile user transmission in a cell in every transmission interval T. Performance results show that the same amount of data can be delivered with lower transmission power as compared to the results when only power control algorithm is utilized. This scheme requires the search of transmission rate vectors, which minimize a transmission power vector, which might be not feasible when the number of active mobile users is large.

A distributed power control algorithm that supports multirate transmission is studied in [36]. The algorithm is called Selective Power Control *(SPC)* and it dynamically adapts the data rate transmission to meet the mobile users' transmission power constraints. In this scheme, the mobile users' transmission power is constrained by a maximum value. When this value is exceeded because of congestion, the mobile user's *UE* is turned off. The scheme is distributed and utilizes the locally measured *SIR* parameter to maximize the total effective data rates with the minimal power usage. The obtained results show considerable system throughput improvement.

In [27], an uplink distributed power control *(DPC)* algorithm based on a previously proposed power control algorithm for satellite systems in [51] is proposed. In this algorithm, a mobile user n adjusts its transmission power at the i^{th} time interval using the formula $p_n^i = c\frac{p_n^{i-1}}{\gamma_n^{i-1}}$ formula, where $1 \leq n \leq N$, $i \geq 1$, p_n^i is the transmitted power by mobile user n at the i^{th} iteration, γ_n^{i-1} is the current *SIR* of mobile user n, and c is some positive constant. The performance of the *DPC* algorithm is evaluated using a cellular system of 50 cells and compared to a Distributed Balancing Algorithm *(DBA)* proposed in [86]. Results show that the *DPC* has a faster converging rate than the *DBA* scheme with very few iterations being required for the minimum *SIR* to approach the optimal value, γ^*.

2.3.3 Handoff Management

Whenever a mobile user having an active call crosses a cell boundary, the active call necessitates transfer from its current base station to another one in order to avoid call discontinuation. This process is known as a handoff and concerns the transfer of the current communication channel of an active call from one base station to a new one. If the new base station has unoccupied channels the handoff call is assigned to one of them. When all channels in the new base station are in use, there are two possibilities: handoff failure which leads to call dropping, or handed off call queuing while waiting for possible channel availability by another call completion or handoff in the new cell. Generally, a handoff scheme's objectives are to minimize call dropping rate while not significantly increasing call blocking rate [22] and [50]. A handoff process composes of two main stages: handoff initiation and handoff decision [22]. These stages are detailed as follows.

Handoff Initiation

The initiation of a handoff procedure requires knowledge of different base stations' received signals' strengths at a mobile user. A mobile user examines the Received Signal Strengths *(RSSs)* of its current base station *(BS1)* and of one or more neighboring base stations *(BS2)* as shown in Figure 2.6. As the mobile user moves away from *BS1*, the *BS1*'s *RSS* degrades while the *RSS* of *BS2* becomes stronger as a result of signal propagation characteristics. Therefore, the handoff procedure is initiated considering these dynamic variations of *RSSs*.

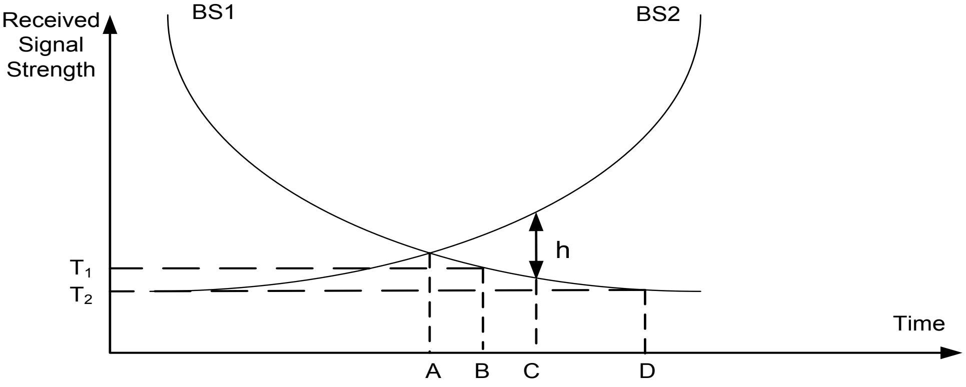


Figure 2.6: Mobile user Handoff in Cellular Networks

Basically, there are four handoff initiation techniques: *Relative signal strength, relative signal strength with threshold, relative signal strength with hysteresis*, and *relative signal strength with hysteresis and threshold*, see Figure 2.6 [22] and [63]. In the relative signal strength technique, the *RSSs* are measured overtime and the base station of the strongest received signal is selected for handoff. As shown in Figure 2.6, when *RSS* of *BS2* exceeds the *RSS* of *BS1* at point *A*, the handoff procedure is activated. Since the base stations' received signals strength fluctuate overtime, unnecessary handoffs may be requested while *BS1*'s *RSS* is still strong enough to serve the mobile user. These unnecessary handoffs are known as the *ping-pong* effect and have a negative impact on network load and call dropping probability.

To overcome the *ping-pong* effect, a threshold value is introduced in the relative signal strength with threshold technique *(T1 in Figure 2.6)*. In this technique, a handoff is only initiated when the *RSS* of *BS1* degrades below a predefined threshold. Therefore, instead of a handoff initiation at point *A* for the relative signal strength technique, the handoff is initiated at point *B*, where the relative signal strength with threshold technique is used. This reduces the number of unnecessary handoffs.

The relative signal strength with hysteresis technique uses a hysteresis value *h*. In this technique, when *BS2*'s *RSS* stronger than the *BS1*'s *RSS* by the defined hysteresis value, handoff is initiated as indicated by point *C* in Figure 2.6. Also, to minimize the number of handoffs, this technique is enhanced by combining the hysteresis and the threshold values. Hence, handoff is initiated only when *BS1*'s *RSS* is lower than the threshold value *T1* shown in Figure 2.6 and *BS2*'s *RSS* is stronger than *BS1*'s *RSS* by the hysteresis value *h*.

All of these handoff methods need to initiate handoff before point *D*, which is the minimum acceptable *RSS* value by a receiver for an active call continuation. If the *RSS* drops below this value, the active call is dropped. The time interval between the handoff request and the minimum acceptable *RSS* enables the system to delay the handoff request at a cell with an unavailable channel.

Handoff Decision

After the initiation of the handoff procedure, a handoff decision needs to be made. There are three handoff decision protocols proposed for cellular systems: *Network*

Controlled Handoff, Mobile Assisted Handoff, and Mobile Controlled Handoff [57], [50], [75], and [77]. The Network Controlled Handoff *(NCHO)* protocol is responsible for the overall handoff decision. In this protocol, the network performs the required *RSS* measurements and issues handoff decisions. The handoff execution time is on the order of many seconds since the network handles all of the processes [57].

To reduce the network work load, mobile users are involved in the handoff decision process in the Mobile Assisted Handoff *(MAHO)* protocol [50], and [77]. In *MAHO*, a mobile user measures its *RSSs* and periodically reports them to the network. Based on these measurements, the network controls when the mobile user's active call should be handed off. *MAHO* takes about 1 second for the active call to be transferred from one cell to another [57], and [77].

In the Mobile Controlled Handoff *(MCHO)* protocol, mobile users control all of the handoff process [50]. In this protocol, the network and mobile users make the necessary measurements. The network periodically reports its measurements to the mobile user which has to decide when to handoff based on these measurements. The handoff execution time of the *MCHO* protocol is 100-500 ms [57], and [77].

Handoff Types

Handoffs are classified into three main types: *Hard handoff, soft handoff, and softer handoff* [22], [42], [50], [77], [76], [82], [81] and [87]. The hard handoff, which is also known as *break-before-make*, involves only two base stations as shown in Figure 2.6. In a hard handoff, a mobile user's active call channel is released before its new channel is acquired from the new base station. The hard handoff degrades the quality of service since there is service interruption when the handoff occurs. It is implemented by *1G* and *2G* cellular systems.

Another handoff type is soft handoff, which is mainly used by *CDMA* systems such as *UMTS*. Also, soft handoff is termed as *make-before-break* since it can establish multiple connections to different neighboring base stations simultaneously. Therefore, active call interruption is eliminated as a mobile user moves from one cell to another [42], and [87]. In soft handoff, each mobile user maintains an *Active* set and *Neighbor* set. The base stations in the *Active* set form a soft handoff connection, while the base stations in the *Neighbor* set are the monitored base stations where either their *RSSs* are not strong enough to be added to the *Active* set or the *Active* set is full.

According to the *UMTS* standard [2], a soft handoff example is illustrated in Figure 2.7. Before explaining this example we need to introduce some parameters along with their definitions. The measured *RSS* of a base station is referred to here as $Pilot\_E_c/I_o$. The strongest and weakest measured pilot signals in an *Active* set are $Best\_Pilot\_E_c/I_o$ and $Worst\_Old\_Pilot\_E_c/I_o$ respectively. $Best\_candidate\_Pilot\_E_c/I_o$ is the strongest measured pilot signal in a *Neighbor* set, and $Reporting\_range$ is the threshold for soft handoff. $Hysteresis\_event1A$, $Hysteresis\_event1B$, and *Hysteresis_event1C* are the addition, removal, and replacement hystereses respectively. $Reporting\_range - Hysteresis\_event1A$ is the $Window\_add$ and $Reporting\_range + Hysteresis\_event1B$ is the $Window\_drop$. The required time to trigger the handoff procedure is defined by ΔT. In this example, the used *Active* set size is “2” which allows the mobile user to be simultaneously connected to maximum of “2” base stations.

In Figure 2.7, first, a mobile user is connected to base station 1 *(BS1)* which has the strongest $Pilot\_E_c/I_o$. As the mobile user moves away from *BS1* and gets

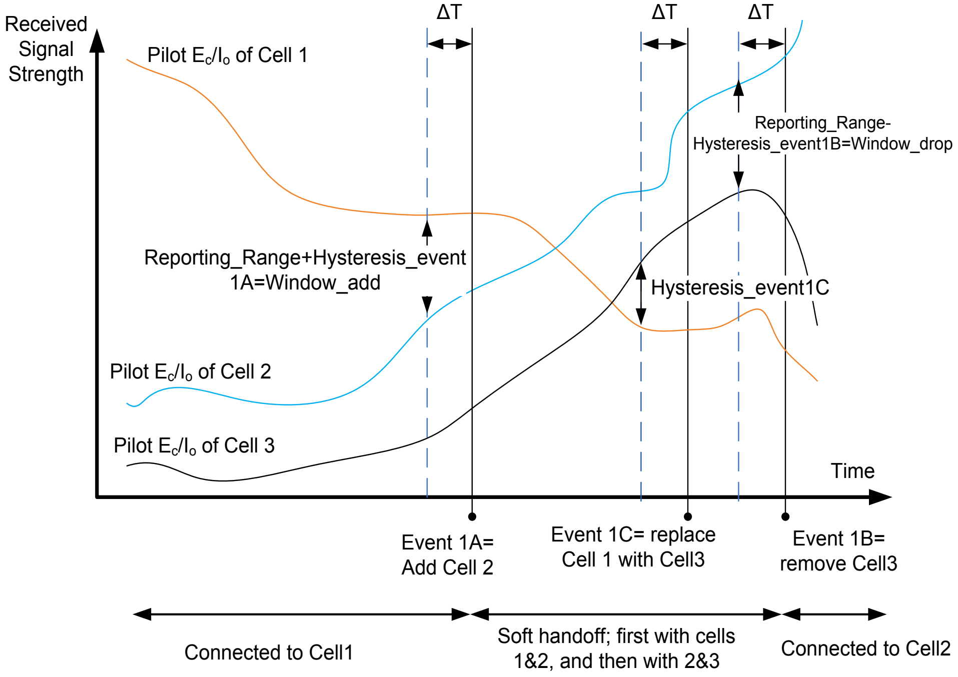


Figure 2.7: Soft Handoff in UMTS Systems [30]

closer to base station 2 *(BS2)*, the $Pilot\_E_c/I_o$ of *BS2* becomes stronger. When the $Pilot\_E_c/I_o$ of *BS2* becomes strong enough to be added to the active set as defined by: $Pilot\_E_c/I_o > Best\_Pilot\_E_c/I_o - Reporting\_range + Hysteresis\_event1A$, for a time period of ΔT and the *Active* set is not full, *BS2* is added to the *Active* set. This event is called *Event 1A* or *Radio Link Addition.*

As the mobile user moves further from *BS1*, the $Pilot\_E_c/I_o$ of base station 3 *(BS3)* of the *Neighbor* set becomes stronger. Since the *Active* set size is set to “2”, the base station of the weakest received signal needs to be replaced by the best received signal of *BS3*. Therefore, as the $Pilot\_E_c/I_o$ of *BS3*, $Best\_candidate\_Pilot\_E_c/I_o$, becomes stronger and the following condition is met for the ΔT time period, *BS1* is replaced by *BS3* in the *Active* set:

$$Best\_candidate\_Pilot\_E_c/I_o > Worst\_Old\_Pilot\_E_c/I_o + Hysteresis\_event1C$$

This event is called *Event 1C* or *Combined Radio Link Addition and Removal*. As time passes, the mobile user gets closer to *BS2* which leads to a degradation of the received signal quality of *BS3*. At some point when $Pilot\_E_c/I_o$ of *BS3* is:

$$Pilot\_E_c/I_o < Best\_Pilot\_E_c/I_o - Reporting\_range - Hysteresis\_event1B$$

for a period of time defined by ΔT, *BS3* is removed from the *Active* set. This event is called *Event 1B* or Radio Link Removal.

Therefore, when the mobile user is in the range of more than one base station, it will be connected to more than one base station of the strongest *RSSs* while monitoring the *RSSs* of other base stations in the *Neighbor* set. As the mobile user moves out of the handoff region, it will be communicating via one base station.

The last handoff type is the softer handoff, which occurs between sectors of one cell. This type of handoff is implemented by *CDMA* systems where the network base stations have some degree of independence from their *RNCs* [42].

Existing handoff schemes belong to one of these handoff types. In [37], a multi-level threshold handoff algorithm is proposed. This algorithm considers the co-existing of high and low speed mobile users in a cellular system and uses 8 handoff thresholds. Each mobile user is assigned a handoff threshold based on its estimated speed. Low speed mobile users are assigned a high threshold since they spend more time in the handoff zone, while high speed mobile users are assigned low threshold. The obtained results in terms of call blocking and dropping probabilities of this algorithm outperform the results of a handoff scheme of one threshold.

Other handoff schemes which also consider the mobile users' speeds are proposed in [65], and [54]. A multilayered network architecture composed of microcells overlaid by a macrocell is used to evaluate these schemes. The mobile users are assigned to each layer based on their speeds; low speed mobile users are assigned to microcells while high speed mobile users are assigned to macrocells. Since slow moving mobile users are assigned to microcells and high speed mobile users are assigned to macrocells, the total number of handoff requests is decreased. When a new or handoff call of a slow moving mobile user finds microcell's channels fully occupied, it is assigned to the macrocell layer. As channels become available in the microcell, the call is assigned a channel of the available ones. This is called a take-back handoff.

A guard channel handoff scheme is proposed in [7]. In this scheme, guard channels are reserved for handoff calls only while other channels are used by both new and handoff calls. The number of guard channels is dynamically changed based on the expected number of handoff mobile users from neighboring cells. The coverage area of a network cell is classified into two zones; pre-handoff zone *(PHZ)* and handoff zone *(HZ)*. Each base station estimates the number of mobile users in the *PHZ* and periodically informs its adjacent base stations. The corresponding base stations reserve a number of guard channels for handoff calls based on the received information from their neighboring base stations. When a handoff call arrives at a cell with no available channels, the call is queued. A newly arrived call is served only if there are no handoff calls in the queue. This scheme decreases the call dropping probability while increasing the call blocking probability and minimizes the system utilization.

Similar guard channel handoff schemes are proposed in [88][48]. These schemes are adaptive which dynamically change the number of reserved handoff channels based on the call dropping probability of the system. When the call dropping probability exceeds a predefined threshold, the number of reserved guard channels is increased to decrease the probability. On the other hand, when the base station does not utilize the guard channels significantly, the number of guard channels is decreased to decrease the blocking probability.

2.3.4 Congestion Control

Despite its broadband capacity, congestion in *BWNs* still exist as in *1G* and *2G* narrowband cellular systems. This is due to the tremendous growth in the number of wireless subscribers and the diversity in wireless multimedia applications which demand more resources. Moreover, the existence of a network utilization pattern in specific periods of the day is a major factor that influences congestion. Therefore, different congestion control *(CC)* methods, to meet the mobile users' *QoS* requirements and to better utilize the limited radio spectrum of wireless networks, are investigated. Generally, *CC* is accomplished through: Rate Adaptation *(RA)* and network cells coverage adaptation through Cell Breathing Management *(CBM)*. In this section, we investigate the advantages and disadvantages of these congestion control mechanisms. A literature review is provided as well.

Rate Adaptation

Rate Adaptation *(RA)* plays an important role in multimedia wireless networks. It utilizes the adaptive nature of multimedia applications to the benefit of both the mobile user and the network. The basic operation of *RA* is based on varying the allocated resources of existing mobile users' calls depending on demand intensity and network conditions. *RA* is invoked by the *CAC* to guarantee the admission of an arrived call if the currently available network resources cannot sustain such demand, and is utilized by the *CC* scheme to react to severe network conditions such as increased interference level or partial system malfunction.

In a multimedia wireless system, traffic is classified into different classes of service. Each class has a set of *QoS* requirements. For the consideration of *RA*, each class of service is characterized by a set of discrete resource allocations, Φ, where $\Phi = \{\phi_1, \cdots, \phi_i, \cdots, \phi_{max},\}$. The resource allocations are increasing in the order of the subscript i, that is $\phi_i < \phi_{i+1}$, for all i with ϕ_1 and ϕ_{max} being the minimum and maximum possible allocatable values, respectively. Usually, a certain allocation denoted ϕ_{target}, is considered as a default or reference allocation. *RA* and other *RRM* schemes attempt to provide ϕ_{target} for mobile users as much as possible. ϕ_{target} can hold the value of either the minimum or maximum possible class allocations, or a value in between. During a mobile user's call lifetime, if the call is given a resource allocation more than ϕ_{target}, the call is said to be upgraded. When the allocated resources are less than ϕ_{target}, the call is considered as downgraded. Generally, the operation of increasing and decreasing the resource allocations are known as upgradation and downgradation, respectively.

In *UMTS*, RA is based on a set of Radio Access Bearers *(RABs)* provided to support various services using different transmission rates [6]. Each *RAB* defines a certain peak transmission rate on both the uplink and the downlink to transfer user data across the radio access network between mobile users and the Core Network *(CN)*. The key function of *RAB* is to dynamically adapt the mobile users' transmission rates during the connection time according to the mobile users' activity and the currently available bandwidth resources in the system. Data rate adaptation is realized by performing *RAB* upgrades to increase the data rate and *RAB* downgrades to decrease the data rate.

In the literature, a wide range of proposals have investigated the *RA* mechanism. In [72], a rate adaptation scheme that investigates the tradeoff between network overhead and fairness in wireless networks is proposed. This scheme is composed of three algorithms; Minimum Adaptation *(MA)*, Fair Adaptation *(FA)*, and Average-Fair Adaptation *(AFA)*. The objective of the *MA* algorithm is to minimize the number of downgraded mobile users to obtain the required resources for a certain call request. The *FA* algorithm concerns fairness in allocating and re-allocating the resources to the active calls. The last algorithm, *AFA* is proposed to balance the two extremes of minimizing network overhead of *MA* and maximizing fairness of *FA*. The *AFA* achieves its objectives by considering the temporal average of bandwidth allocations.

In the *MA* algorithm, two lists are defined; *availlist* and *downgradedlist*. These lists are used to order existing calls' releasable and required bandwidth, respectively. The releasable bandwidth is the network resources that can be released by mobile users before reaching their minimum threshold, while the required bandwidth is the bandwidth needed by existing calls to reach their maximum possible allocations. When the *MA* is triggered by the *CAC* or *CC* schemes, a minimum number of mobile users from these lists are selected to satisfy the required demand or re-allocate the available bandwidth. This results in unfairness of mobile users' rate adaptation.

The *FA* is proposed to divide the allocatable bandwidth between all users. Achieving farness in *FA* comes at the increasing cost of system overhead. This problem is solved by using *AFA* to balance between the unfairness of *MA* and the higher cost of *FA*. *AFA* utilizes the *MA* modules and adds the average bandwidth allocation to the rate adaptation selection mechanism.

Another rate adaptation scheme is studied in [43]. In this scheme, the received power of each active data user is maintained constant by using transmission power control to compensate for channel fading and path loss, while the transmission rate is dynamically adjusted to guarantee a target bit energy-to-equivalent noise spectral density $\frac{E_b}{N_o}$ when interference varies. Analysis and simulation results show that rate adaptation outperforms conventional SIR-based power control in terms of lower average power and average delay for data users to achieve same network throughput.

A downlink power-based congestion control mechanism is proposed in [67] and [66]. In this scheme, mobile users are prioritized according to Best Channel Quality *(BCQ)*, Worst Channel Quality *(WCQ)*, or First in First Out *(FIFO)* order. As the downlink traffic channels' transmission powers exceed a predefined threshold for a specified time period, the congestion control algorithm is invoked to reduce the mobile users' transmission rates starting with lower priority calls first. As congestion is released, mobile users' transmission rates are restored starting with high priority calls first. The obtained results outperform the results of a system with CAC algorithms only.

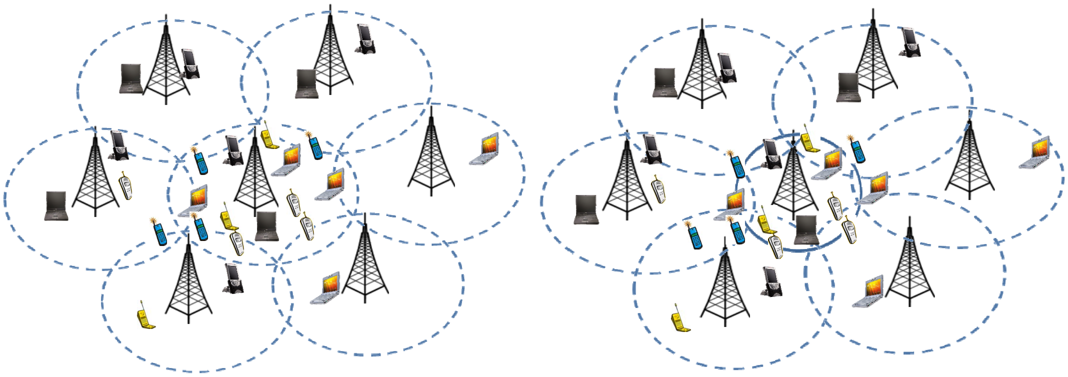
In [52], a RA scheme that considers a Multiple Access Interference *(MAI)* evaluation is proposed. First, an MAI model considers the burstiness of data traffic and the time-varying fading of a mobile user's channels is proposed. Then, a rate adaptation algorithm that utilizes the proposed MAI model is devised. In this algorithm, each base station measures the desired user received uplink signal quality and its corresponding MAI power. Based on such information, the base station determines the next transmission rate to be used by the mobile user for its next packet transmission. The mobile user adjusts its transmission rate according to the feedback information received from its base station. A reliable signal estimation is required for such algorithm.

Similar utility functions based *RA* schemes are proposed in [19], [47], and [70]. In these schemes, network traffic is classified into different classes; real-time non-adaptive multimedia traffic, real-time adaptive multimedia traffic, and non-real-time data traffic. The execution of these schemes are classified into upgradation and downgradation steps. The upgradation is triggered by call departure while downgradation is triggered by call arrival. In the case of congestion, the required resources are obtained by downgrading the transmission rates of the adaptable traffic. The main objective of these schemes is to maximize system utility while keeping the call blocking and dropping rates low as possible. A large volume of signaling traffic to upgrade and downgrade individual mobile users' transmission rates is required which is considered as a drawback of such schemes.

Generally, these *RA* schemes are associated with a high cost due to their deterministic triggers. To reduce such cost, a Stochastically Triggered Bandwidth Adaptation Algorithm *(ATBAA)* is proposed in [71]. This algorithm replaces the deterministic manner in which *ATBAA* adaptation is initiated with a probabilistic trigger. Therefore, despite the availability of downgradable or upgradable resources, the invocation of *STBAA* algorithm is controlled via a probabilistic trigger. Therefore, mobile users are protected against severe resource degradation and the number of adaptations a mobile user undergoes in a short period of time is reduced. Hence, network stability is achieved and *RA* cost is reduced.

Cell Breathing Management

In *WCDMA* systems, the Common Pilot Channel *(CPICH)* transmission power level is used to define the coverage area of a network cell while traffic channels carry mobile users' traffic [20], and [30]. Therefore, the base station transmission power is allocated to the *CPICH* and other traffic channels based on their needs. When the cell load is moderate, more power is allocated to the *CPICH*, hence, increasing cell coverage area. On the other hand, when the cell load is high, more power is allocated to traffic channels which leads to cell coverage contraction. This behavior makes the cell coverage and capacity have an inverse relationship; (i.e. cell capacity increases as its coverage level is decreased and vice versa). This is known as Cell Breathing *(CB)* in *UMTS* terminology. A simplified sketch of *CDMA* cell breathing mechanism is shown in Figure 2.8. Figure 2.8(a) shows the network cells without *CB*. In Figure 2.8(a), the loaded center cell is contracted (when *CB* is used) and its adjacent lightly loaded cells are expanded. Therefore, to release loaded cell congestion nearedge mobile users are forced to handoff towards the expanded adjacent cells. Proper management of the *CB* concept increases the utilization of the limited radio resources when traffic is spatially localized in *3G* and beyond systems.



(a) No Cell Breathing (b) Cell Breathing

Figure 2.8: Cell Breathing Mechanism

A number of congestion control and load balancing schemes have investigated the congestion control and load balancing problem in interference limited *3G* systems utilizing Cell Breathing Management *(CBM)*. A congestion control and load balancing scheme that uses *CBM* is proposed in [40]. It is based on the concept of reinforcement Q-learning method to dynamically adapt the coverage area of a loaded cell and its adjacent lightly loaded cells in an omni-directional manner. The limitation of this scheme is the possible formation of coverage gaps, since each cell independently changes its coverage area based on the received interference power in the uplink. Also, this scheme requires an infinite search space of possible coverage level which greatly increases its complexity.

In [83], a Case Based Reasoning *(CBR)* scheme based on the concept of dynamic cell breathing is proposed for releasing system congestion. It recalls a solution from a database, which has previously been used to resolve similar congestion scenarios. The scheme requires database maintainability and its complexity increases as the database size increases.

A hybrid network architecture is proposed in [68]. In this architecture, *CDMA* and *TDMA* networks cooperate to balance their load. When the *CDMA* system is congested, its cells are contracted and mobile users are forced to handoff to the *TDMA* system. As congestion is released, *CDMA* cells are expanded for the mobile users to take advantage of the higher bandwidth and speed of the *CDMA* system. The management of two different platforms and the co-allocation of cell sites are needed in this scheme.

In [21], the *CBM* is utilized to balance network traffic. It is a cooperative approach in that the loaded cell and its adjacent lightly loaded cells cooperatively work towards balancing their load. In this approach, the loaded cell coverage area is contracted while the coverage areas of its neighboring cells are expanded. This enforces nearedge loaded cell mobile users to handoff towards the expanded lightly loaded cells. Thus, congestion can be released and traffic can be balanced.

All of the previously proposed *CBM* congestion control and load balancing schemes consider uniform mobile users' distribution over the coverage area of a hotspot cell and adapt cells coverage areas in omni-directional manner. However, expanding all adjacent cells to assist a loaded cell may not be required since in real scenarios congestion maybe spatially localized at the side of one or two nearby cells. Furthermore, and even if traffic is uniformly distributed, omni-directional expansion of the adjacent cells may negatively affect the serving capabilities of other neighboring cells. Therefore, other congestion control and load balancing approaches that consider the disadvantages of the previously proposed schemes are needed. To this end, we propose a directional coverage adaptation mechanism called Directional Cell Breathing *(DCB)* in this thesis to release spatially localized congestion and balance system load in interference-limited wireless communication systems.

2.4 DCB Framework for BWNs

The main objectives of the proposed Directional Cell Breathing *(DCB)* framework are to maximize radio resource utilization and efficiently manage congestion and load imbalance scenarios in *3G* and beyond systems. This is achieved through a directional cooperative coverage adaptation method, which involves every two nearby sectors of two adjacent cells. The framework utilizes recent advances in smart directional antenna capabilities, which can be programmed to form the required sector coverage level. Also, the proposed scheme needs to maintain the provisioned *QoS* parameters and support service differentiation to the mobile users according to their associated class of service. *DCB* also needs to consider the system limitations in meeting such requirements. For example, the constraints on mobile users' power budget and the the received interference level in the system need to be considered, while practising the framework actions within the system.

The proposed framework can be integrated within the already existing hardware and software system components of the *BWNs*. For example, directional antennas are already installed in cellular systems which may require minimal upgrade requirements to support the functionality of our proposed framework. Also, the existing *Node B* Application Part *(NBAP)* signaling protocol requires software-based upgrades to include the required signaling messages proposed by our framework. In terms of operation management, the operation is reduced to the sector level instead of the *Node B* level as in previous cell breathing coverage adaptation mechanisms.

2.4.1 Overview of Framework Components

The framework, shown in Figure 2.9, comprises three major components

- Directional Cell Breathing *(DCB)* Coverage Adaptation Module
- DCB-Based Congestion Control and Load Balancing
- DCB-QoS Provisioning

These components are detailed in as following.

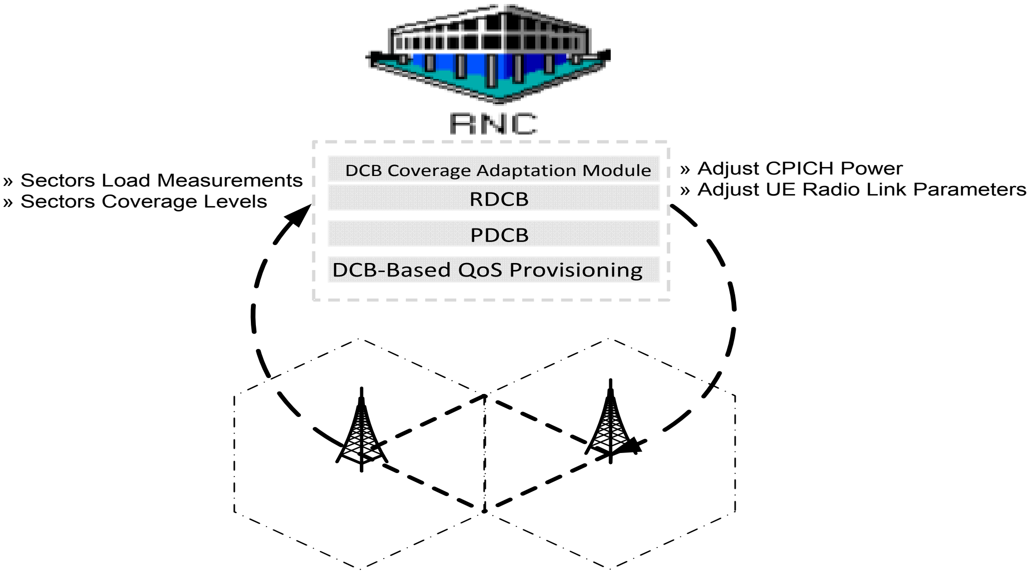
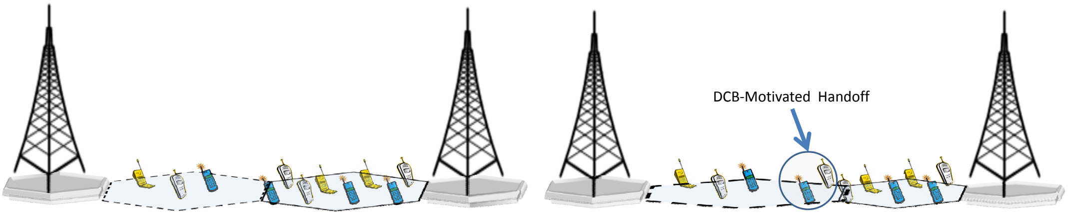


Figure 2.9: The DCB Framework Components

Directional Cell Breathing Coverage Adaptation Module

The *DCB* coverage adaptation module is the main component of our framework. The *DCB* coverage adaptation module considers the dynamic and directional variation of base station coverage area at the granularity of cell sectors' coverage levels. The coverage adaptation is a joint cooperation process of nearby cell sectors of every two adjacent base stations (definition of nearby sectors is given in Section 2.9). It aims at releasing spatially localized traffic congestion, maintaining the system load balance, and maximizing the radio resource utilization. However, unlike omni-directional cell breathing, avoiding unnecessary expansion of other adjacent base stations, and expansion of the adjacent base station coverage is only practised towards the congested adjacent base station sector.

DCB coverage adaptation is based on predefined combined coverage levels of every two nearby sectors of every two adjacent cells. These combined coverage levels correspond to pre-set pilot channel transmission power levels of every base station sector. In a spatially localized traffic load situation (See Figure 2.10(a)), the overloaded sector is contracted while the lightly loaded supporting sector is expanded (See Figure 2.10(b)). Hence, nearedge mobile users are motivated to handoff towards the expanded sector.



(a) Before DCB (b) After DCB

Figure 2.10: Directional Cell Breathing Mechanism

The *DCB* coverage adaptation module examines the capabilities and limitations of the base station of a network sector in directionally expanding and contracting the sector coverage level towards an adjacent nearby sector considering different load scenarios at both sectors. This behavior stimulates a hotspot congestion or load imbalance management in interference-limited *UMTS* systems. In practice, the *DCB* module is invoked to order the corresponding base station, by *DCB*-specific Node B Application Part *(NBAP)* protocol signaling messages, to activate certain coverage levels based on a predefined *CPICH* transmission power levels at each *Node B* as shown in Figure 2.10. As congestion is released or load is balanced, the *DCB* action is reversed to restore the normal coverage of each base station sector.

DCB-Based Congestion Control and Load Balancing

Despite the challenges imposed by the *DCB-Based* Congestion Control and Load Balancing *(DCB-CC&LB)* module on the *UTRAN* system, it can be exploited to the benefit of operators and mobile users. For example, due to the daily network traffic patterns, spatially localized congestion and load imbalance can be handled by dynamically configuring the coverage area of congested or traffic imbalanced sector along with its nearby supporting sector. Therefore, congestion is released in the loaded sector and resource utilization is increased in the supporting sector. Moreover, the need for installing extra base stations to increase system capacity is avoided.

DCB-CC&LB includes two modules– a reactive module and a proactive module. The reactive module is intended for urgent situations, which are in need of quick action, while the proactive module is used to achieve long-term objectives. For example, the reactive *DCB-CC&LB* module can support the admission control and handoff *RRM* components when the arrival rate of new or handoff calls is suddenly increased in a sector. Another important application of the reactive approach is when the serving base station of a sector partially or completely malfunctions. This requires an action to handoff the existing traffic in that sector to the nearby expanded sector.

The proactive approach of the *DCB-CC&LB* is designed to evaluate the system behavior on long-term basis. Proactive *DCB-CC&LB* periodically evaluates the received base station measurement reports with respect to every sector. Whenever a predefined traffic imbalance or congestion threshold parameters are violated in a sector, the proactive *DCB-CC&LB* module acts to balance network traffic, and prevent or release system congestion. Network sectors coverage levels converge in the long term to a state which meets the traffic profiles of each sector.

DCB-Based QoS Provisioning

The success of any congestion control and load balancing mechanism depends on meeting the stringent requirements of multimedia traffic. Therefore, the ability of the proposed framework to support multimedia applications with different *QoS* requirements is examined in the *DCB-QoS Provisioning* component. In this component, the *DCB* module is integrated with Transmission Power Control *(TPC)* and Rate Adaption *(RA)* to support mixed traffic of different transmission rates and different *QoS* demands.

The *DCB-Based QoS Provisioning* module adapts mobile users' transmission rates to achieve a certain coverage combination of loaded and supporting sectors. Such coverage combination would not be feasible without rate adaptation because of the increased interference level in the system. Therefore, for the sake of maximizing the support level, this component includes a module to perform near-optimal sector coverage level and transmission rate and power allocations to the mobile users in the loaded and supporting sectors. Moreover, this proposed module supports service differentiation when multiple classes of services are supported by the system. Hence, the framework capabilities in fulfilling the *QoS* requirements of mixed traffic are examined with respect to different coverage combinations of every two nearby sectors.

2.4.2 The Framework Architecture and Operation

In *BWNs*, radio resources are managed at four levels; namely: sector, *Node B*, *RNC*, and *RNS* levels which are shown in Figure 2.2. At the sector level, the radio resources are managed to maintain the mobile users' connectivity to the upper levels. Therefore, the required transmission power to guarantee the planned coverage of each sector need to be maintained by the controlling *Node B*.

At the *Node B* level, the *Node Bs* cooperation is required to guarantee the elimination of coverage gaps between the adjacent *Node Bs* and the efficient utilization of their limited resources. For example, an optimal coverage level of two nearby sectors requires the optimal allocation of uplink and downlink transmission powers at both sectors. Since these sectors are overlooked by different *Node Bs*, such optimization needs to be done at the *Node Bs* level which requires the involvement of the controlling *RNC*.

The upper level of the network hierarchy is the *RNC* which manages the logical resources of the *Node Bs* in an *RNS* system. The main functionalities of the proposed framework are located at the *RNC*. Therefore, two objectives are followed in this level. The first objective concerns the involvement of two *Node Bs* of the same *RNS* system. In this case, the framework action requires communication only over the *Iub* interfaces. Whenever the framework action requires the involvement of two *RNS* systems because the involved sectors are controlled by two adjacent *Node Bs* of different *RNS*, *inter-RNS* communication is required over the *Iur* interface. In this thesis, we implement the proposed framework in a single *RNS* system; therefore, *inter-RNS* communication is not required.

The implementation and the effect of the proposed framework spans this network architectural levels via the interaction among the framework components which is shown in Figure 2.11. The framework is implemented at the *RNC* and relies on information provided by the lower levels. This information is communicated over the *Iub* interface using the *Node B* Application Part *(NBAP)* signaling protocol. The framework operations can be classified into four phases; namely: Measurement, Reporting, Evaluation, and Action. In the measurement phase, the *Node B* periodically monitors the sector resources utilization and reports the requested information to its controlling *RNC*. This information may include interference and load levels in the uplink, and used transmission power in the downlink. The information is reported based on a predefined manner by the *RNC*. For example, the *Node B* may report the measured information periodically or the reporting process may be triggered when a predefined threshold, such as uplink load factor is reached.

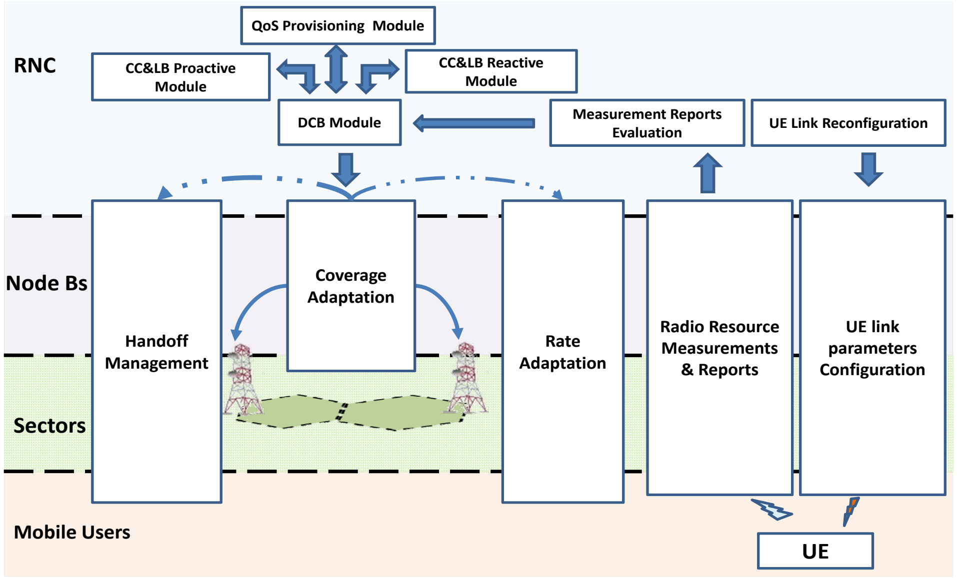


Figure 2.11: The Framework components and their interaction

The received reports with respect to every sector in all *Node Bs* are evaluated in the *RNC*. Hence, the *RNC* is aware of the sectors' conditions. Whenever a sector is defined as congested or network traffic is imbalanced, the *DCB* module is involved to evaluate the possible coverage levels of such sector and its nearby sector. The *DCB* module requires the involvement of the *DCB-CC&LB* and *QoS- Provisioning* modules.

If the activation of a coverage combination of two nearby sectors releases sector congestion or balances network load, an action is conveyed to the corresponding *Node Bs* to reconfigure the nearby sectors, proactively handoff nearedge mobile users and reconfigure the uplink and downlink radio resources. On the other hand, if there is no possible coverage combination that can be provisioned to release congestion and balance network load, the classical *RRM* approaches are used.

2.5 Summary

In this chapter, we reviewed the architecture and *RRM* protocol design for 3G and beyond wireless cellular systems. First, the design objectives and the architecture of *WCDMA* is explained in detail. The *UMTS* system architecture and its proposed *QoS* classes to support multimedia services are reviewed. The radio resource management approach, emphasizing on call admission control, transmission power control, handoff management, and congestion control techniques, has been investigated in depth. Also, a wide range of previous work proposals of these *RRM* techniques are summarized. The Cell Breathing Management *(CBM)* concept and its role in releasing congestion and balancing load in interference limited *WCDMA* system was discussed. The drawbacks of the *CBM* are outlined. The objectives of the proposed framework and it components were explained, including the framework schematic architecture and its components' interaction.

Chapter 3

System Model

In this chapter, we introduce the system model on which we base the mathematical formulation of the proposed framework. The model is based on the *UMTS* standard proposed by the *3GPP* in Release 99 [2]. The *DCB-enabled UMTS* network architecture and the concept of Coverage Levels *(CL)* and Supporting Levels *(SLs)* are introduced in Section 3.1. The adapted *UMTS* Node B Application Part *(NBAP)* signaling protocol that communicates the required signaling messages between the Node Bs and the *RNCs* is explained in Section 3.2. The mobile users' signal propagation models are presented in Section 3.3. The chapter is summarized in Section 3.4.

3.1 DCB Network Architecture

Generally, the cellular system architecture is composed of M cells as shown in Figure 3.1. Each cell is served by a base station which is the radio access interface for the mobile users in that cell. To reduce the interference level at a base station, the coverage area of each cell is divided into Q sectors. The network cells are divided into a number of manageable groups. Each group is controlled by one *RNC* which logically manages the radio resources of these cells and it also acts as a gateway to other communication systems such as the Public Switching Telephony Network *(PSTN)* and the Internet. Mobile users move freely over the network coverage area. When a mobile user moves out of its current serving cell coverage area and enters the coverage area of an adjacent cell, a handoff protocol is executed to transfer the mobile user connection to the base station of the moved to cell.

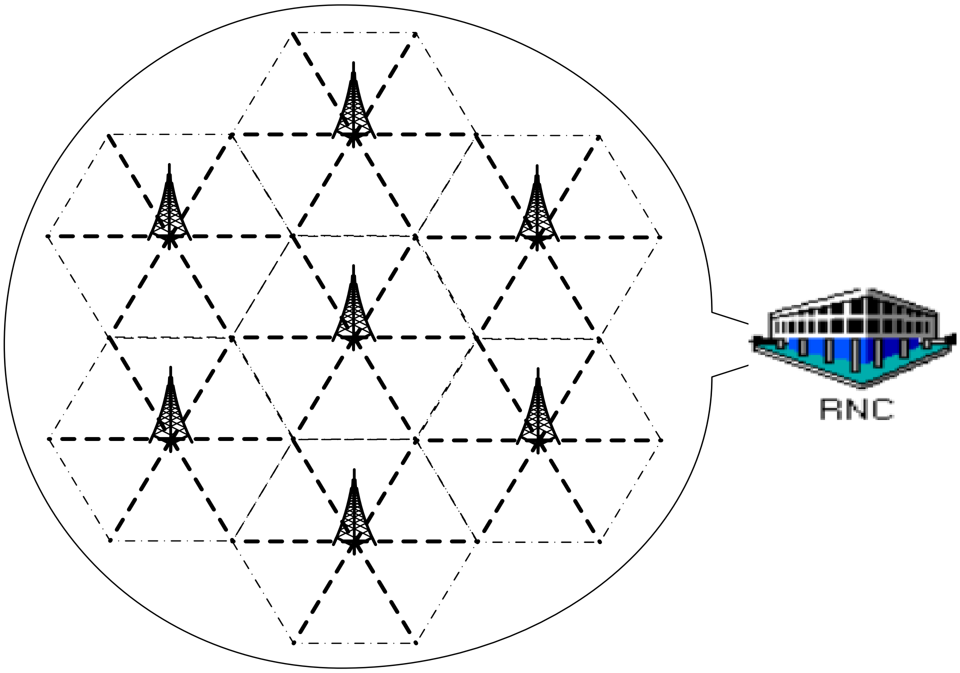


Figure 3.1: Cellular Network Architecture

The proposed *UMTS* network architecture that enables nearby sectors' coverage adaptation by the Directional Cell Breathing *(DCB)* framework of this thesis is shown in Figure 3.2. In this architecture, the coverage area of each sector is virtually partitioned into L coverage levels of equal width (See Figure 3.2). Each sector has multiple coverage levels (3 in Figure 3.2). These coverage levels correspond to different pre-set downlink Common Pilot Channel *(CPICH)* transmission power levels of the base stations serving these sectors. Each cell sector is served by a directional smart antenna, which has the capability to expand the sector coverage area towards the nearby loaded sector and to contract it towards the cell center in case of sector congestion. Therefore, the expanded sector acquires nearedge mobile users of the contracted sector to ease its congestion. The *DCB* mechanism manages the coverage levels of every two nearby sectors; viz sectors q and $\hat{q}$ in Figure 3.2, as one set. In this set, the normal coverage of a sector occurs when its activated coverage levels and the activated coverage levels of its nearby sector are equal.

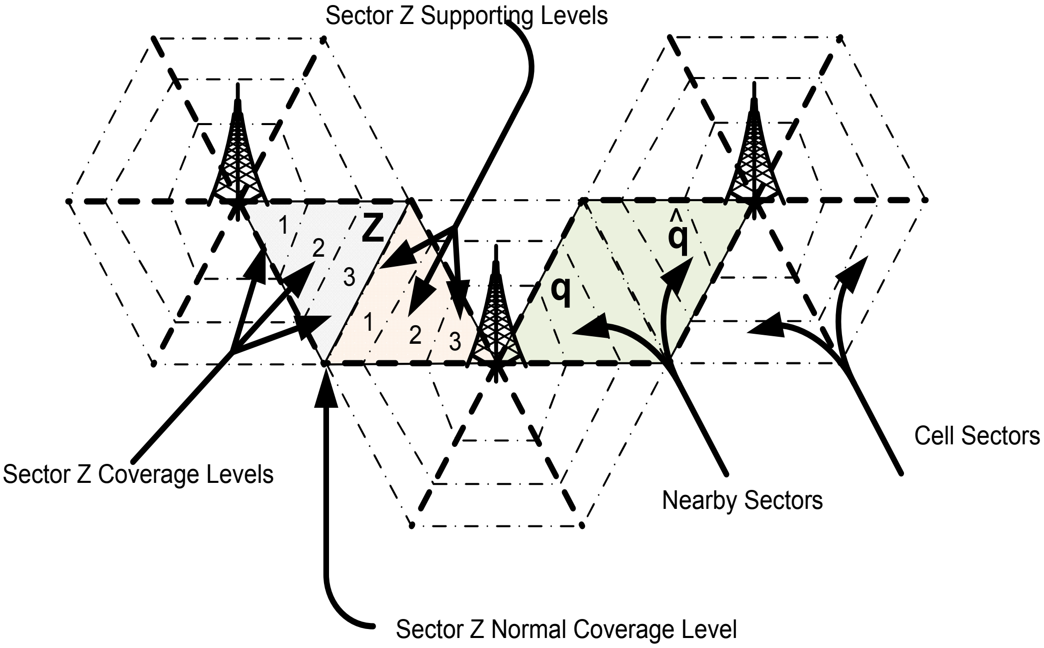


Figure 3.2: DCB Network Architecture

To elaborate more on the concept of the proposed coverage and supporting levels of a cell sector, we use the following example. With respect to a sector Z, shown in

Figure 3.2, its coverage levels and the coverage levels of its nearby sector are defined into two groups: sector Z Coverage Levels *(CL)* and sector Z Supporting Levels *(SL)*. Based on the traffic load of sector Z and the traffic load of its nearby sector, the coverage area of sector Z can be in one of three states; normal coverage; i.e coverage level 3 which also includes coverage levels 1 and 2, contracted coverage; i.e. coverage level 1 or 2 of sector Z in the figure, or expanded coverage; i.e. supporting levels 1, 2, or 3. When the coverage area of sector Z is normal, the load of sector Z and the load of its nearby sector are balanced. As sector *Z's* load increases given that the load of its nearby sector is moderate, the coverage area of sector Z is contracted and the coverage area of its nearby sector is expanded. Hence, nearedge mobile users of sector Z are handed off towards the nearby sector. When the coverage area of sector Z is expanded, it implies that its nearby sector is overloaded and nearedge mobile users of the nearby sector are handed off towards the expanded Z sector. The *CL* and *SL* definitions are valid with respect to every sector in the network except when for a given sector there is no nearby sector.

The aforementioned expansion and contraction processes of every two nearby sectors requires joint management of the pilot transmission power to assure the activation of the complementing coverage levels of every two nearby sectors. Hence, network coverage gaps are eliminated and a loaded sector congestion is released. This requires the involvement of the controlling *RNC(s)* to manage the coverage area of every two nearby sectors of two adjacent Node Bs. Therefore, an efficient signaling protocol needs to be in place such as the *NBAP* protocol defined by the *UMTS* standard [5].

3.2 NBAP Signaling Protocol

In *UMTS*, Node B Application Part *(NBAP)* protocol provides a set of functions such as *Cell Configuration Management, Radio Link Management, Measurements on Common Resources, Measurements on Dedicated Resources*, etc. [5]. These functions are mapped to *NBAP* Elementary Procedures *(EPs)*, where an Elementary Procedure is a unit of interaction between the *RNC* and the *Node B*. An *EP* consists of an initiating message and possibly a response message. The *NBAP's EPs* are divided into common procedures and dedicated procedures. *NBAP* common procedures are procedures that request initiation of a *Node B* Communication Context for a specific mobile user in *Node B* or are not related to a specific mobile user such as the ***Cell Reconfiguration Procedure***. Also, *NBAP* common procedures incorporate logical Operation and Management *(OAM)* procedures. On the other hand, *NBAP* dedicated procedures are procedures that are related to a specific mobile user.

Since the *DCB* coverage adaptation mechanism involves two nearby sectors managed by two different *Node Bs*, the *NBAP* protocol is utilized to communicate the required sectors' information to the *RNCs* and conveys the *DCB* action requests issued by the *RNCs* to the *Node Bs*. The enhanced functionalities of *NBAP* for the *DCB* technique are explained in Chapter 5.

3.3 Mobile User Signal Propagation Model

In this thesis, we use the widely accepted lognormal attenuation propagation model for *WCDMA* networks [24][80]. In this model, the received signal power of a mobile user at the base station is given by:

$$P_r = P_{tr} d^{-\sigma} 10^{\zeta/10} \tag{3.1}$$

where P_{tr} and P_r are the transmitted and received powers respectively, d is the distance between the transmitter and the receiver, σ is the path-loss exponent with a typical value of 4 and ζ represents the shadowing effects. This model is used to mathematically calculate the average received mobile user signal at a base station of a network cell.

On the other hand, when the *DCB* is evaluated heuristically and mobile users mobility is considered, the path loss calculation of a mobile user transmitted signals is modeled based on an outdoor pedestrian radio propagation model as used in [60] and [59]. In this model, the mobile user signal path loss is defined as:

$$PL = (40 log_{10} d) + (30 log_{10} freq) + 49 \tag{3.2}$$

where PL is the path loss for a transmitted signal, d is the distance between transmitter and receiver, and *freq* is the system center frequency. Accordingly, the received power of a signal transmitted from a transmitter at distance d is expressed as:

$$P_r = P_t * 10^{txGain} * 10^{rxGain} * (\frac{1}{PL}) * x \tag{3.3}$$

where P_r and P_t are the received and transmitted powers respectively, *txGain* is the transmitting antenna gain and *rxGain* is the receiving antenna gain, PL is the path loss between the transmitter and the receiver, and x is the compensation term for orthogonality factor which computed as

$$x = 1 - \frac{1}{PG} \tag{3.4}$$

where PG is the spread spectrum processing gain which is calculated by dividing the

WCDMA system bandwidth W by the mobile user transmission rate R [30].

In *WCDMA* systems, an acceptable call quality has to have a Bit Error Rate (BER) above a minimum threshold value, which is maintained by keeping the bit energy to interference ratio (E_b/I_o) of such a call above a certain threshold [24]. Practically, (E_b/I_o) for a mobile user is obtained from its received signal to interference ratio *(SIR)* [29][24]. The *SIR* for a mobile user n is defined as the ratio of the desired signal of this mobile user to the undesired signals of other mobile users which is mathematically defined by

$$SIR_n = \frac{g_{nb} p_n^{tr}}{\sum_{\hat{n} \neq n}^{N} g_{\hat{n}b} p_{\hat{n}}^{tr} + I_{inter} + X_o} \qquad (1 \leq n, \hat{n} \leq N) \tag{3.5}$$

where g_{nb} and p_n^{tr} are respectively the channel gain to the base station b and the transmission power of mobile user n. $g_{\hat{n}b}$ and $p_{\hat{n}}^{tr}$ are respectively the channel gain and the transmission power of other mobile users $\hat{n}$ in the same cell, I_{inter} is the inter-cell interference and X_o is the background noise power at the base station. Hence, the (E_b/I_o) for a mobile user is calculated by dividing the mobile user's received signal power by its transmission rate (R_n) and dividing the received interference power by the total system bandwidth *(W)*

$$\left(\frac{E_b}{I_o}\right)_n = \frac{W}{R_n} * SIR_n \tag{3.6}$$

To maintain the required quality of service *(QoS)* parameters for all calls, their received (E_b/I_o) values must be higher than or equal to a predefined threshold value τ_{th}.

The (E_b/I_o) QoS parameter is used to compute the uplink load factor of a cell sector in *WCDMA* systems. In this thesis, we utilize the following uplink load factor formula [30]:

$$\eta_{UL} = (1 + I_{ratio}) * \sum_{n=1}^{N} \frac{1}{1 + \frac{W}{(\frac{E_b}{I_o})_n * R_n}} \tag{3.7}$$

where η_{UL} is the uplink load factor, W is the system bandwidth, $\left(\frac{E_b}{I_o}\right)_n$ is the energy per bit to noise ratio of mobile user n, and I_{ratio} is the inter-cell interference to intra-cell interference ratio.

3.4 Summary

The system model to evaluate the proposed framework is explained in this chapter. First, the general network architecture is explained. Then, the *DCB* network architecture of a *DCB-Enabled UMTS* system is introduced and the concepts of sector Coverage Levels *(CL)* and Supporting Levels *(SL)* are explained in detail. Then, the *NBAP* signaling protocol which communicates *DCB* signaling messages between network Node Bs and their controlling *RNC(s)* is explained. The signal propagation models for mobile users are explained which, in the following chapters, will be used to calculate the average and actual received mobile users signals at a network base station. Also, the mathematical models for computing a mobile user energy per bit-to-interference noise ratio and a cell sector uplink load factor are presented.

Chapter 4

Directional Cell Breathing

Directional coverage adaptation positively impacts the performance of the interference-limited *BWNs* due to its ability to release hotspot congestion and balance the system load while maximizing the utilization of the limited system resources. A distinctive feature of directional coverage adaptation is the cooperation of *Node Bs'* sectors in managing congestion as well as balancing the system load. This, however, adds a new dimension to the congestion control and load balancing problem as it requires the simultaneous management of resources of different *Node Bs.* Also, the coverage variation, if mismanaged, can negatively affect the performance of interference-limited systems; e.g. coverage gaps, increased interference, etc. Therefore, to efficiently release congestion and maximize the limited system resources utilization, the coverage adaptation of a *Node B* service area should not affect the performance of other *Node Bs* in the system.

In this chapter, we introduce our cooperative directional coverage adaptation technique for congestion control and load balancing in *BWNs.* We propose controllable directional cell breathing *(DCB)* where a network cell's sectorized coverage is varied reactively in instances of congestion, and proactively in instances of exercising long-term congestion avoidance and load balancing. *DCB* overcomes the drawbacks of non-controllable cell breathing in non-sectorized *WCDMA* network cells by optimizing the coverage levels within a cell sector under constraints, minding sectoral traffic load and interference bounds and the nearby sector load state as well.

The remainder of this chapter is organized as follows. The motivation and problem formulation of the proposed *DCB* technique are described in Section 4.1. The *DCB* mechanism is explained in Section 4.2. Then, the mathematical analysis, interference calculation and system capacity quantification of the *DCB* system model is detailed in Section 4.3. The proposed *DCB* optimization model and its numerical results are presented in Section 4.4. The chapter is summarized in Section 4.5.

4.1 Motivation and Problem Formulation

In *WCDMA* systems, coverage and capacity have an inverse relationship; capacity is increased when the coverage area of a cell is decreased and vice versa. Therefore, a trade-off between coverage and capacity needs to be carefully considered to maximize the system capacity while guaranteeing access to mobile users in the whole planed coverage area. Currently, online cell configuration is performed in an omni-directional manner based on the cell breathing concept. Despite the ability of omni cell breathing in increasing the cell capacity or coverage, it negatively impacts the capacity of the adjacent cells, increases the chance of coverage gaps, and degrades the overall network performance. Also, wireless networks tend to have a certain daily utilization pattern; i.e. in the morning more traffic is localized at city centers as people head towards their work. Then, in the afternoon, such traffic is shifted towards the residential and recreational areas, which renders city centers network cells resources underutilized, while such utilization is increased in the residential and recreational areas network cells. Therefore, in spatially localized traffic scenarios system performance can be enhanced if online cell configuration can be performed in a directional manner; expanding cell coverage directionally only towards a congested directionally contracted cell to release its congestion or balance the load among these cells and preventing the unnecessary omni-directional expansion and contraction towards other adjacent cells.

The recent advancement in antenna technology is a promising avenue for congestion control and load balancing in current and next generation *BWNs*. Also, instead of exercising cell coverage adaptation in an omni-directional manner, the directional properties of smart directional antennas can be utilized to expand and contract cell coverage towards the hotspot directions only. Accordingly, we assert that directional coverage adaptation in nodal-distribution-aware *WCDMA* systems increases system capacity and minimize interference level in the system.

4.2 Directional Cell Breathing Mechanism

The *DCB* mechanism is illustrated by considering two nearby sectors; q and $\hat{q}$ of *Node B1* and *Node B2* respectively as shown in Figure 4.1. Each sector coverage area is partitioned into L concentric coverage levels[1](4 in Figure 4.1). Theoretically, based on these L coverage levels each *Node B* sector has the capability to cover the total service area of both sectors when the maximum number of its Supporting Levels *(SLs)* is used. Each coverage level in one sector has a complementary coverage

---

[1] The CLs and SLs of a cell sector here are dealt with as a one set and referred to as coverage levels unless they are mentioned specifically by CLs or SLs

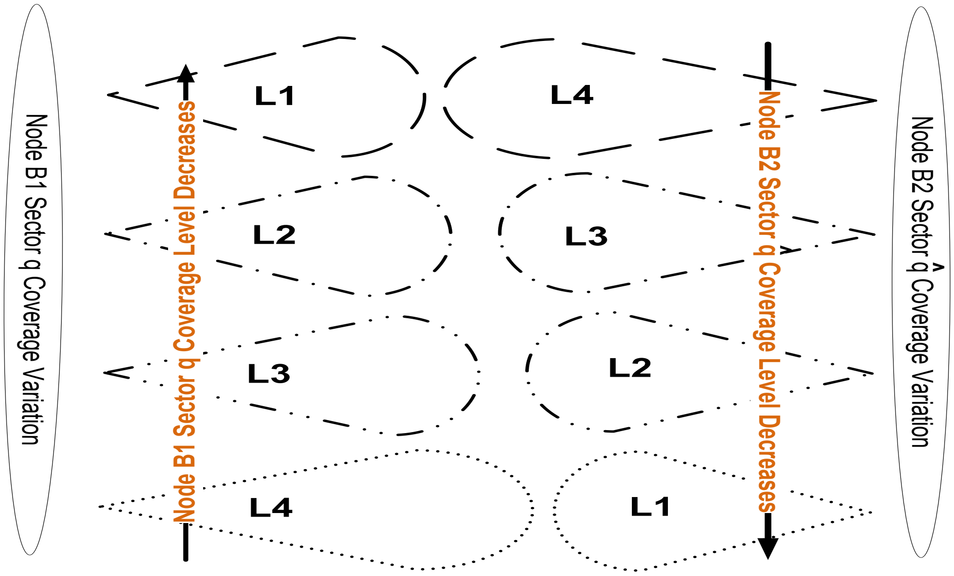
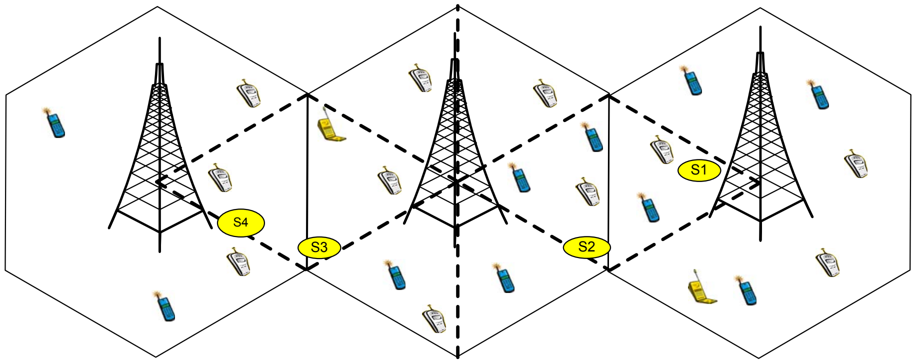


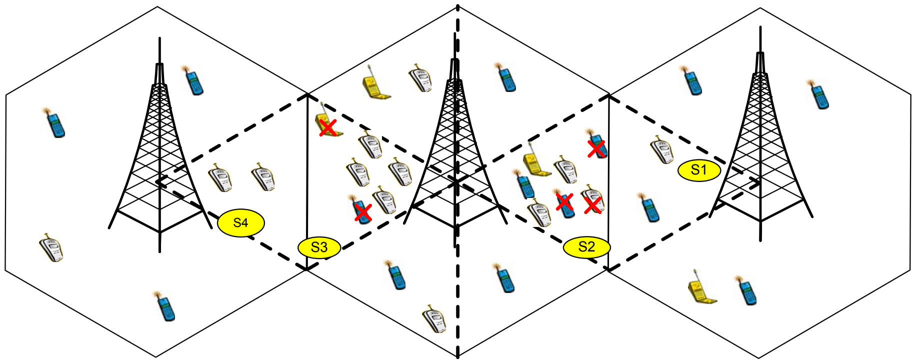
Figure 4.1: Illustration of nearby sectors coverage combination in Directional Cell Breathing

level in the other sector; coverage levels of the same line pattern in Figure 4.1. The complementary coverage levels have to be activated simultaneously to prevent network coverage gaps. Therefore, as the coverage level of sector q is increased, the one for sector $\hat{q}$ is decreased and vise versa. This sectoral coverage variation helps release congestion and balance network load by motivating nearedge mobile users of the loaded sector to handoff towards the expanded lightly loaded sector.

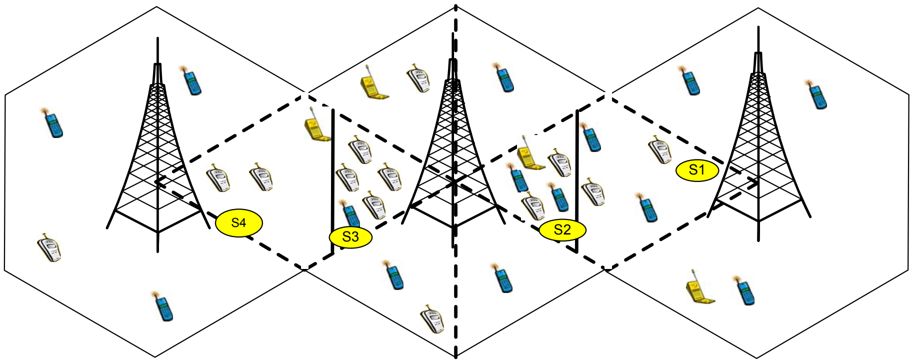
Based on these coverage levels, an example showing the interaction of every two nearby sectors to release hotspot congestion using the *DCB* mechanism is illustrated in Figure 4.2. In this figure, when the network load is balanced each sector has a normal coverage level as shown in Figure 4.2(a). Without the *DCB* mechanism, as traffic becomes unbalanced and concentrated at certain sectors such as S_2 and S_3 of the center cell in 4.2(b), hotspots are created which renders the serving base stations unable to support the increased load. Therefore, call dropping and blocking



(a) Mobile Users Mobility Trajectory



(b) Without DCB Mechanism



(c) With DCB Mechanism

Figure 4.2: Traffic HotSpots W/O DCB Mechanism

become inevitable as shown in Figure 4.2(b). Moreover, other network sectors may be negatively affected due to the increased interference level in the system.

Utilizing the proposed *DCB* technique as shown in Figure 4.2(c) can benefit the whole system by shifting some traffic of the loaded sector towards the lightly loaded sectors, i.e., expanding sectors S_1 and S_4 while contracting S_2 and S_3 sectors. The combined coverage adaptation of nearby sectors makes the coverage level of a *Node B* sector dependent on the coverage level of its nearby sector but it is independent of the coverage levels of other sectors of the same *Node B*. Therefore, coverage gaps between neighboring cells are eliminated.

The *Node Bs* and mobile users may have some limitations which limits the expanding of a cell sector above a certain coverage level. Examples of such limitations are the maximum tolerable inter-cell interference power, and the maximum *Node B* and mobile devices transmission powers. Therefore, a maximum support level L_{max} and a minimum coverage level L_{min} need to be defined for each sector. Whenever the complement coverage level of two nearby sectors reaches those levels, the *DCB* mechanism cannot be exercised and classical radio resource management is used to release network congestion.

4.3 DCB Mathematical Formulation

The uplink capacity of *WCDMA* systems is limited by the total received power at a base station which is mainly a combination of the received powers of the mobile users in the corresponding sector and the received powers from the mobile users in neighboring sectors. These components are known as intra- and inter-cell interferences respectively [24], and [80]. With respect to a certain mobile user, its received signal can be decoded correctly if its signal to interference ratio *(SIR)* is above a pre-set threshold. Therefore, in interference-limited systems such as *UMTS*, maintaining the *SIR* level of every mobile user above a certain threshold is mandatory to minimize the error probability in decoding such received signals [24].

In a *DCB-Enabled WCDMA* system, a mobile user's average received power at its base station and its average interference power on other network base stations vary based on the activated coverage level of its cell sector. Therefore, for every coverage level of a cell sector shown in Figure 4.1 the average received power of mobile users at the base station of its corresponding sector and the average received interference power of that mobile user at the base stations of other sectors need to be mathematically quantified. Then, the quantified values are used to compute the nearby sectors capacity for different combined coverage levels. Before proceeding with the mathematical interference analysis for the *DCB* module, we first give the following definitions:

- M: total number of cells
- m and $\hat{m}$: interfered and interfering cells, m and $\hat{m} \in \{0, 1, 2 \ldots, M-1\}$
- l: Loaded sector coverage level index, $l \in \{0, 1, 2 \ldots, L_{max}\}$
- $\hat{l}$: Supporting Sector coverage level index, $\hat{l} \in \{0, 1, 2 \ldots, L_{max}\}$
- Q: total number of sectors in a cell
- q: Loaded sector ID, $q \in \{0, 1, 2 \ldots, Q-1\}$
- $\hat{q}$: Supporting sector ID, $\hat{q} \in \{0, 1, 2 \ldots, Q-1\}$

- $N_l^{m,q}$: number of mobile users in the loaded sector q of cell m
- $\hat{N}_{\hat{l}}^{\hat{m},\hat{q}}$: number of mobile users in the interfering sector $\hat{q}$ of cell $\hat{m}$
- $A_l^{m,q}$: activated coverage area l of loaded sector q in cell m
- $A_{\hat{l}}^{\hat{m},\hat{q}}$: activated coverage area $\hat{l}$ of supporting sector $\hat{q}$ in cell $\hat{m}$
- $A_{tot} = (A_l^{m,q} \cup A_{\hat{l}}^{\hat{m},\hat{q}})$: total serving area of both sectors
- $r_l^{m,q}$: radius of loaded sector q in cell m given coverage level l
- $r_{\hat{l}}^{\hat{m},\hat{q}}$: radius of supporting sector $\hat{q}$ in cell $\hat{m}$ given coverage level $\hat{l}$
- $d_n^{m,q}$: mobile user distance to the interfered base station of cell m
- $d_n^{\hat{m},\hat{q}}$: mobile user distance to its base station of cell $\hat{m}$

4.3.1 Inter-Cell Interference Analysis

The directional antenna property limits the number of cells interfering with a cell sector [24], [80], and [79]. Also, the effect of the received signal from beyond first tier cells is negligible which simplifies the inter-cell interference calculation [24]. Therefore, for every cell sector the inter-cell interference is imposed by the mobile users in the sectors of the adjacent cell as shown in Figure 4.3. The inter-cell interference analysis is identical for every sector and the analysis herein is with respect to one sector and its adjacent cell sectors. The inter-cell interference is composed of Q components where Q is the number of sectors in the neighboring interfering cell. Each component has L_{max} values which correspond to the maximum possible coverage level of a cell sector. With the assumption of perfect power control, the signals from mobile users are received with power P_r at their corresponding base station. These signals are propagated to the neighboring *BSs* and received with power given by:

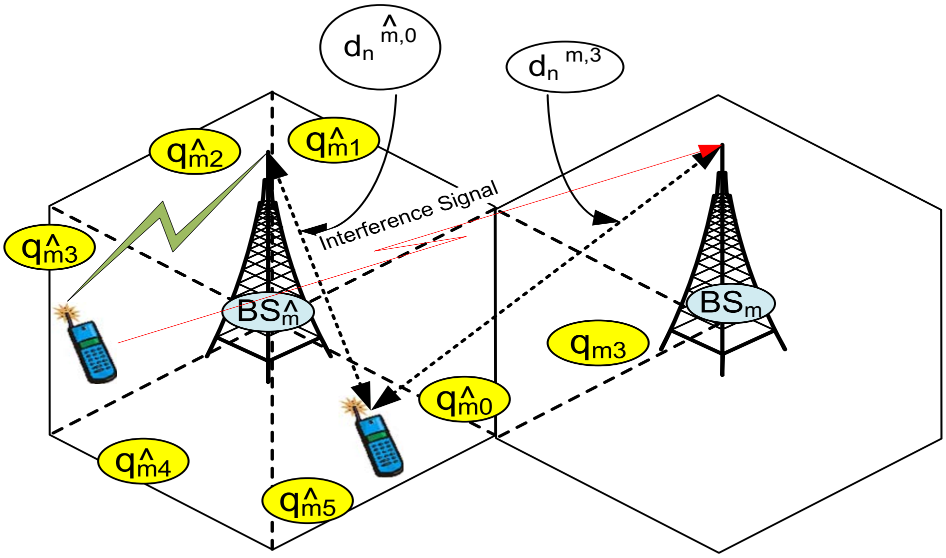


Figure 4.3: Inter-Cell Interference on Cell m from users in Cell $\hat{m}$

$$P_r^{inter} = P_r \left(\frac{d_n^{\hat{m},\hat{q}}}{d_n^{m,q}} \right)^{\sigma} \tag{4.1}$$

where P_r is defined in Equation (3.1) of Chapter 3, $d_n^{\hat{m},\hat{q}}$ is the distance from the interfering mobile user to its base station and $d_n^{m,q}$ is its distance to the interfered base station. To mathematically derive the average interference on a base station of a cell sector caused by a mobile user located in a neighboring cell sector, the distance ratio of $d_n^{m,q}$ and $d_n^{\hat{m},\hat{q}}$ needs to be defined and integrated over the corresponding sector area then multiplied by the number of active mobile users in that sector. The average received interference power from $N_{\hat{l}}^{\hat{m},\hat{q}}$ interfering mobile users with respect to the activated coverage level $\hat{l}$ of each interfering sector is given in the following formula:

$$I_{inter}^{\hat{m},\hat{q}} = \frac{\hat{N}_{\hat{l}}^{\hat{m},\hat{q}} P_r}{A_{\hat{l}}^{\hat{m},\hat{q}}} \int_{\theta_1}^{\theta_2} \int_{0}^{r_{\hat{l}}^{\hat{m},\hat{q}}} \frac{r^{m+1}}{\left(r^2 + R^2 - 2Rr\cos(\theta)\right)^{\frac{m}{2}}} d_r d_\theta \tag{4.2}$$

Since the total received interference power is composed of Q values of $I_{inter}^{\hat{m},\hat{q}}$, the total average received interference power at the corresponding base station i can be computed as follows:

$$I_{inter}^{avg} = \frac{\sum_{q=0}^{Q-1} I_{inter}^{\hat{m},\hat{q}}}{Q} \tag{4.3}$$

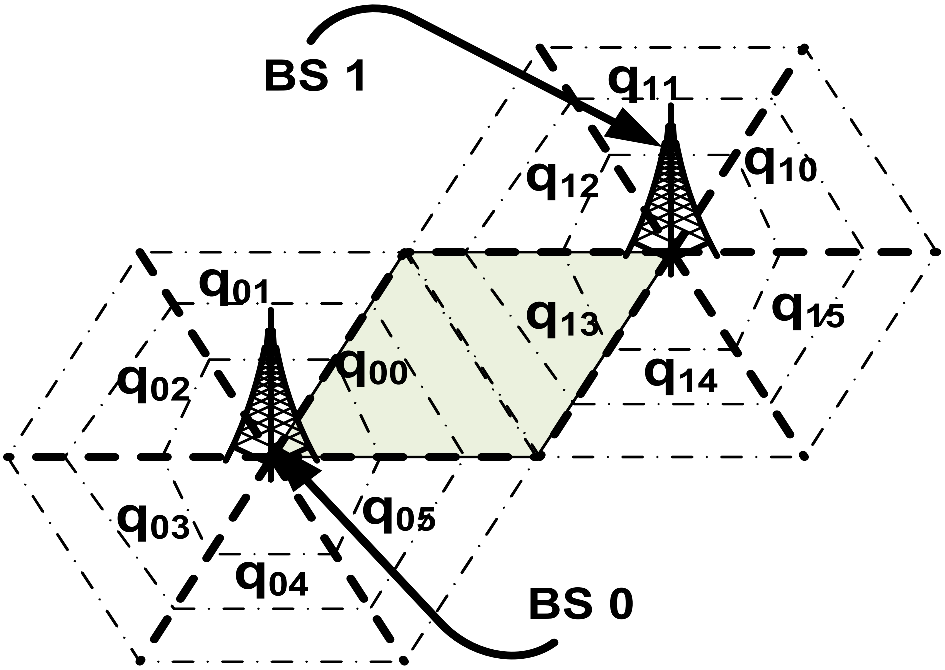


Figure 4.4: Interfering Cell '0' and Interfered Cell '1' of a DCB-Enabled WCDMA System

The system architecture in Figure 4.4 is used to illustrate the computation of these interference expressions. The average inter-cell interference is computed at the base station of sector (q_{13}) which is imposed by each mobile user in every interfering sector,

$(q_{00} - q_{05})$, of the adjacent cell. The combined sectors' coverage areas is partitioned into 40 coverage levels. Generally, each sector is allowed to extend its coverage area by 10 *SLs* beyond its normal coverage level of 20. For simplicity of analysis and results representation, we only vary the coverage area of the nearby interfering sectors (q_{00}) and (q_{13}) while maintaining the normal coverage area of other interfering sectors $(q_{01} - q_{05})$, $(q_{10} - q_{12})$, and $(q_{13} - q_{15})$. The normalized interference results obtained from such analysis are shown in Figure 4.5. The X-axis represents the coverage level of the interfering sector (q_{00}) and the Y-axis is for the average inter-cell interference imposed on sector (q_{13}) from the six sectors on the adjacent cell $(q_{00} - q_{05})$

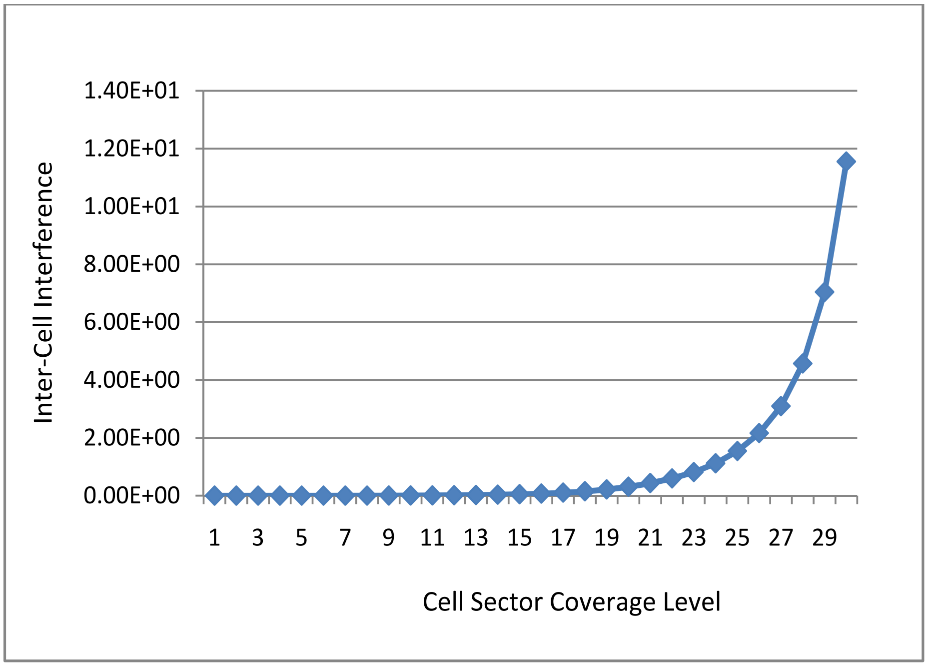


Figure 4.5: Average Normalized Inter-Cell Interference

As the figure shows, as the coverage level of a sector is contracted, its inter-cell interference on its nearby sector becomes negligible. Therefore, serving a large number of mobile users in a small coverage area has a negligible interference effect on other sectors. On the other hand, as the coverage level of a sector increases the inter-cell interference on the nearby sector becomes dominant. Despite the increased interference level of the expanded sector, the actual interference value will not be significant because of the smaller number of served mobile users in the expanded sector. Therefore, in the case of load imbalance varying the coverage area of the two nearby sectors will release congestion and balance network traffic load while maximizing network resource utilization. Hence, directionally managing the coverage levels of every two nearby sectors will maximize the number of served mobile users and increase the utilization of the limited wireless resources.

4.3.2 Intra-Cell Interference Analysis

With respect to a certain mobile user n in a cell sector, the most dominant interference is the received power from other mobile users $\hat{n}$ in the same sector which is widely known as intra-cell interference. For a power-controlled system such as *UMTS* systems, intra-cell interference is computed by the following formula [24], and [80]

$$I_{intra} = \sum_{\hat{n}=0 \neq n}^{N-1} P_{r,\hat{n}} \tag{4.4}$$

where N is the total number of mobile users which are uniformly distributed over the sector coverage area and $P_{r,\hat{n}}$ is the received interference power of mobile user $\hat{n}$ at the serving base station.

In a *DCB-enabled WCDMA* system, the number of mobile users served by a sector is increased as the sector coverage level is increased which leads to an increase of the intra-sector interference power. Therefore, for each sector coverage level, the *DCB* mechanism uses Equation (4.4) to numerically evaluate the intra-sector interference for the loaded and supporting sectors. The results obtained are used with the values of the inter-cell interference in computing the system capacity with respect to various coverage levels.

4.3.3 System Capacity Quantification

Herein, for a DCB-enabled WCDMA system, the computed inter-cell interference values and Equation (4.4) of the intra-cell interference are used to quantify the maximum capacity in terms of the number of mobile users a cell sector can support based on the number of mobile users in its nearby sector and other interfering sectors. Constant Bit Rate *(CBR)* traffic such as a voice application service is considered. The network architecture shown in Figure 4.4 is used to quantify such capacity. In this architecture, each network cell has 1 km, mobile user transmission rate R of 12.2 kbps, minimum required E_b/I_o of 5 dBm and a system bandwidth W of 3.84 Mcps are used. All of the calculations are performed in the uplink direction for a single time slot.

In a *DCB-enabled WCDMA* system, the received $\frac{E_b}{I_o}$ of a mobile user in a cell sector varies as the activated coverage levels of the sector and its nearby sectors are varied. Therefore, Equation (3.6) is modified to consider the effect of such variation in computing mobile users $\frac{E_b}{I_o}$ parameters. The modified formula is shown in Equation (4.5).

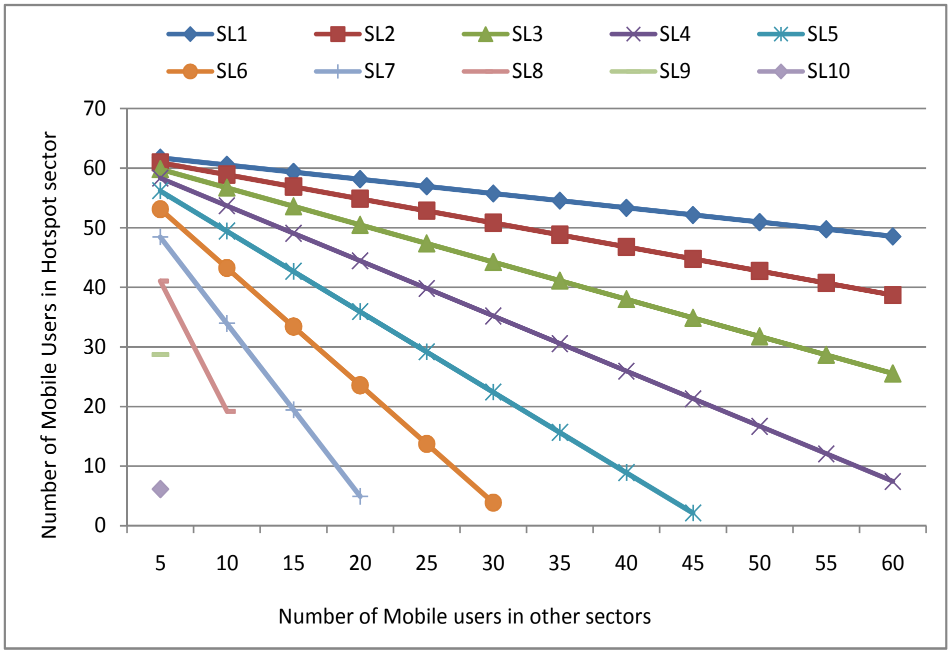


Figure 4.6: Capacity Quantification of a Hotspot Sector

$$\left(\frac{E_b}{I_o}\right)_n^l = \tau_n \geq \tau_{th} \tag{4.5}$$

$$R_n \geq R_n^{th} \tag{4.6}$$

$$P_{r,n}^l \geq P_r^{min} \tag{4.7}$$

$$L_{min} \leq l \leq L_{max} \tag{4.8}$$

$$n = 1, 2, \ldots, N_l^{m,q}$$

where R_n^{th} is the minimum acceptable transmission rate of a mobile user n, $P_{r,n}^l$ and P_r^{min} are the received power of mobile user n given coverage level l and the minimum acceptable received power of any call respectively. The received power is lower bounded, which is required by the receiver circuits to decode the received signal properly, l is the current coverage level of a cell sector, and L_{min} and L_{max} are the minimum and maximum sector coverage levels, respectively. These constraints need to be maintained to guarantee the requested QoS parameters of every call.

Equation (4.5) is manipulated to compute the maximum number of mobile users N for every possible coverage level of a sector[2]. Therefore, the maximum number of mobile users in each sector can be obtained by using the following formula:

$$N_l^{m,q} = \left(\frac{W}{R_n \tau_n} - I_{inter}^l N^{NC} + 1 \right) \tag{4.9}$$

where I_{inter}^l is the average inter-cell interference imposed on the corresponding sector given that its activated coverage level is l. N^{NC} is the total number of the interfering calls in the interfering sectors.

For a sector capacity quantification using Equation (4.10), we vary the number of mobile users in the supporting sector, sector q_{00} in Figure 4.4, while maintaining the number of mobile users in other interfering sectors fixed at 10 mobile users per sector. For each load scenario in the supporting sector, the maximum number of mobile user in the hotspot sector, sector q_{13} in Figure 4.4, is computed. In this analysis, the number of interfering mobile users in the supporting sector is increased from 5 to 60 in step of 5 mobile users. The obtained results for a uniform[3] call distribution in every sector are shown in Figure 4.6.

Figure 4.6 shows that, as the average number of interfering mobile users increases, the average number of acceptable calls with respect to every supporting level (SL_l) is decreased. This is because of the increased inter-cell interference power, which negatively affects the hotspot sector capacity. Despite the capacity decrease, users

---

[2] The system background noise has a negligible value as compared to the interference level. Therefore, it is omitted in this analysis.

[3] Herein, traffic distributed in each cell sector is uniform, but network traffic distribution is non-uniform.

density is increased in the loaded sector as the support level is increased.

The *WCDMA* system's interference and capacity analysis through the proposed *DCB* technique shows that the unbalanced system traffic due to mobile users spatial distribution can be solved by directionally varying the network base stations coverage levels. On the other hand, increasing the coverage area of a cell sector beyond certain coverage levels drives the system interference power to an unacceptable level which negatively affects the *DCB* performance. Therefore, there is a tradeoff between increasing the coverage level of a cell sector for supporting a loaded nearby sector and the interference level increase in the system. Hence, a mechanism should be in place which prevents the activation of a sector supporting level which negatively affects the overall system performance.

4.4 DCB Optimization Model

Based on the interference results and the capacity quantification for the *DCB* module, a *DCB* optimization model is proposed. The objective of the model is to maximize the support level given by the supporting sector to the nearby loaded sector. In doing so, the interference level of the supporting sectors needs to be maintained below a predefined threshold to avoid reaching an overloaded state. The formulation of the optimization model is:

$$F = \max_{\hat{l}} \sum_{m,\hat{m}} \sum_{q,\hat{q}} \sum_{\hat{l}=0}^{L_{max}} \hat{l} * g(\hat{l}) \tag{4.10}$$

s.t.

$$\begin{aligned}
& 0 \leq \hat{l} \leq L_{max} \\
& g(\hat{l}) = \hat{l}_0 * b_0 + \hat{l}_1 * b_1 + \ldots \hat{l}_{L_{max}} * b_{L_{max}} \leq 1 \\
& I_{LS}^{l} = I_{intra}^{LS} + I_{inter}^{LS} \leq I_{thrshold} \\
& I_{SS}^{\hat{l}} = I_{intra}^{SS} + I_{inter}^{SS} \leq I_{thrshold} \\
& N_{\hat{l}}^{\hat{m},\hat{q}} \leq N_{l}^{m,q}
\end{aligned}$$

where $\hat{l}$ is the active supporting level of the supporting sector, and $g(\hat{l})$ is a binary function which defines that a maximum of one and only one sector coverage level to be active. Note that selecting coverage level $\hat{l}$ implies the selection of all coverage levels below $\hat{l}$. I_{LS}^{l} and $I_{SS}^{\hat{l}}$ are the total interference seen by the base station of the loaded and supporting sectors given that their activated coverage levels are l and $\hat{l}$ respectively. Finally, $N_{l}^{m,q}$ and $N_{\hat{l}}^{\hat{m},\hat{q}}$ are the number of mobile users in the loaded and supporting sectors for candidate coverage levels l and $\hat{l}$. To prevent coverage gaps between the nearby sectors, l and $\hat{l}$ coverage levels need to complement each other.

To evaluate the performance of the proposed *DCB* optimization model, the average number of proactively handed off mobile users $\overline{N_{ho}^{l}}$ from sector q to a nearby sector $\hat{q}$ needs to be quantified. To simplify analysis, we consider a uniform traffic distribution over the coverage area of each sector. Given the coverage area, $A_{l}^{m,q}$, of sector q, the expected number of handed off mobile users, $\overline{N_{ho}^{l}}$, towards nearby sector $\hat{q}$, with respect to every coverage level l, can be mathematically defined by:

$$\begin{aligned}
\overline{N^{l}_{ho}} &= N^{m,q}_{L} - \frac{N^{m,q}_{L} * A^{m,q}_{l}}{A^{m,q}_{L}} \qquad (4.11) \\
\text{where} & \\
A^{m,q}_{l} &= \int_{0}^{r^{m,q}_{l}} \int_{\theta_1}^{\theta_2} r dr d\theta \\
A^{m,q}_{L} &= \int_{0}^{r^{m,q}_{L}} \int_{\theta_1}^{\theta_2} r dr d\theta
\end{aligned}$$

where $N^{m,q}_{L}$ is the total number of mobile users in sector q of normal coverage level L, $A^{m,q}_{L}$ is the normal coverage area of sector q, $A^{m,q}_{l}$ is the coverage area of sector q when its activated coverage level is l, $r^{m,q}_{l}$ is sector q radius given that supporting level l is activated, and θ is the sector angle. Therefore, as the coverage level of sector q is decreased, the expected number of handed off mobile users is increased.

The procedure for *DCB* optimization is performed in a number of steps which are summarized as following:

1. Define the number of mobile users for a normal coverage of the loaded and supporting sectors
2. Quantify the number of mobile users to be handed off from loaded sector q towards the supporting sector $\hat{q}$ using Equation (4.11)
3. Run the DCB optimization model of Equation (4.10)
4. Activate the maximum SL which satisfies the constraints of the objective function in Equation (4.10) is activated

The Binary Integer Programming *(BIP)* optimization model of the Matlab 7.0 Optimization ToolBox is used to solve the proposed optimization model. The reason for using *BIP* is that the activation of a sector coverage level is based on the binary function of $g(\hat{l})$ which defines that a maximum of one and only one sector coverage level to be active. The activation of a sector coverage level implies the activation of other coverage levels below it. The following section presents numerical examples to evaluate the performance of the *DCB* mechanism. For each scenario, different load values are used to test the *DCB* mechanism in providing both load balancing and congestion control in interference-limited *WCDMA* systems.

4.4.1 Performance Evaluation and Numerical Example

The *DCB* optimization model is evaluated using the network model shown in Figure 4.4. The coverage area of a cell sector can be extended towards the nearby sector by up to 10 supporting levels above its normal coverage level. Homogenous mobile users are used to evaluate the *DCB* optimization model. The system parameters in this analysis are the same parameters used in the capacity quantification analysis in Subsection 4.3.3. The mobile users are uniformly distributed over the coverage area of each sector but their density differs from sector to sector for the purpose of creating hotspot congestion. In this analysis one sector is considered as a loaded sector, sector (q_{13} in Figure 4.4), while its nearby sector, sector q_{00} in Figure 4.4, is considered as a supporting sector. Two different scenarios are used in evaluating the optimization model:

- Dynamic Supporting Sector Load and Fixed Loaded Sector Load.
- Fixed Supporting Sector Load and Dynamic Loaded Sector Load.

For both testing scenarios, the number of mobile users in the interfering sectors $(q_{01} - q_{05})$, $(q_{10} - q_{12})$ and $(q_{14} - q_{15})$ in Figure 4.4, are fixed at 10 per sector. This implies that such sectors are lightly load. The *DCB* optimization model uses the interference results and capacity quantification defined in the previous section.

Dynamic Supporting Sector Load and Fixed Loaded Sector Load

In this scenario, the loaded sector load is maintained at a fixed value of 60 mobile users which creates a hotspot in that sector. The number of mobile users in the supporting sector varies from 0 to 30 in steps of 5 mobile users. For each load scenario, the *DCB* optimization model is invoked to search for the optimal coverage levels of the loaded and supporting sectors which balances the load and releases the congestion of the spatially localized traffic at q_{13}.

Numerical results are presented in Table 4.1. In the table, SL_1 to SL_{10} are the candidate supporting levels which can be given to the loaded sector, N_{13} and N_{00} are the number of mobile users served by the loaded and supporting sectors with respect to every candidate supporting level. The shaded cells of Table 4.1 correspond to the optimal supporting level found by the *DCB* optimization model of such load scenario.

As the table shows that, the level of support depends on the load of the supporting sector. As the supporting sector load increases, the support level is decreased. Therefore, increasing the number of mobile users in the normal coverage of the supporting sector has a negative impact on the number of mobile users to be handed off and limits the level of support given to the loaded sector. The optimal values in Table 4.1 comply with the quantified capacity shown in Figure 4.6.

Table 4.1: Number of mobile users in Loaded Sector, N_{13}, and Supporting Sector, N_{00}, (Dynamic Supporting Sector Load and Fixed Loaded Sector Load Scenario)

	SL1	SL2	SL3	SL4	SL5	SL6	SL7	SL8	SL9	SL10
N13	54	49	43	38	34	29	25	22	18	15
N00	6	11	17	22	26	31	35	38	42	45
N13	54	49	43	38	34	29	25	22	18	15
N00	11	16	22	27	31	36	40	43	47	50
N13	54	49	43	38	34	29	25	22	18	15
N00	16	21	27	32	36	41	45	48	52	55
N13	54	49	43	38	34	29	25	22	18	15
N00	21	26	32	36	41	46	50	53	57	60
N13	54	49	43	38	34	29	25	22	18	15
N00	26	31	37	42	46	51	55	58	62	65
N13	54	49	43	38	34	29	25	22	18	15
N00	31	36	42	47	51	56	60	63	67	70
N13	54	49	43	38	34	29	25	22	18	15
N00	36	41	47	52	56	61	65	68	72	75

Fixed Supporting Sector Load and Dynamic Loaded Sector Load

In addition to congestion control, *DCB* can be used as a load balancing mechanism for *UMTS* cellular systems. Therefore, given a nearly loaded sector and lightly loaded supporting sector, the activation of the *DCB* mechanism balances load among those sectors. The *DCB* optimization model is evaluated for 10 mobile users over the normal coverage area of the supporting sector while the number of mobile users in the loaded sector varies from 30 to 70 mobile users in steps of 5 mobile users. Other network sector retain their mobile users as in the previous scenario. Such load variation in the loaded sector drives the sector load from a nearly loaded state, in which the *DCB* mechanism acts as a load balancing mechanism, to an overloaded state, in which the *DCB* mechanism acts as a congestion control mechanism as in the previous scenario. Numerical results are shown in Table 4.2.

We can see different behaviors of the optimization model from the one in previous

Table 4.2: Number of mobile users in Loaded Sector, N_{13}, and Supporting Sector, N_{00}, (Dynamic Loaded Sector Load and Fixed Supporting Sector Load Scenario)

	SL1	SL2	SL3	SL4	SL5	SL6	SL7	SL8	SL9	SL10
N13	27	24	22	19	17	15	13	11	9	8
N00	13	16	18	21	23	25	27	29	31	32
N13	32	28	25	22	20	17	15	13	11	9
N00	13	17	20	23	25	28	30	32	34	36
N13	36	32	29	26	22	20	17	14	12	10
N00	14	18	21	24	28	30	33	36	38	40
N13	41	36	33	29	25	22	19	16	14	11
N00	14	19	22	26	30	33	36	39	41	44
N13	45	40	36	32	28	25	21	18	15	12
N00	15	20	24	28	32	35	39	42	45	47
N13	50	45	40	35	31	27	23	20	17	14
N00	15	20	25	30	34	38	42	45	48	51
N13	54	49	43	38	34	29	25	22	18	15
N00	16	21	27	32	36	41	45	48	52	55
N13	59	53	47	42	37	32	27	23	20	16
N00	16	22	28	33	38	43	48	52	55	59
N13	**63**	**57**	**51**	**45**	**39**	**34**	**30**	**25**	**21**	**17**
N00	**17**	**23**	**29**	**35**	**41**	**46**	**50**	**55**	**59**	**63**

scenario. For example, when the loaded sector load is moderate, the level of support is maintained at SL_3. This level balances the load of both sectors. As the loaded sector load increases, the level of support is increased to SL_4. When the loaded sector becomes overloaded, the level of support is decreased due to the increasing number of handed off mobile users. Therefore, the optimization model becomes unable to find more support to the loaded sector. As the load increases, the support sector becomes unable to provide any support to the loaded sector as seen in the last two rows of the table.

The efficiency of the *DCB* mechanism can be increased by increasing the granularity of the sectors' coverage levels. This decreases the average number of handed off mobile users per every activated coverage level which increases the probability of providing more support to an overloaded sector. However, increasing the granularity of sectors' coverage levels requires dynamic sectors' coverage areas partitioning. Such dynamic cell sectors' partitioning increases the complexity of the proposed *DCB* mechanism.

4.5 Summary

A novel coverage adaptation module was introduced in this chapter. This module, which is called Directional Cell Breathing *(DCB)*, dynamically changes the coverage area of a cell sector to release hotspot congestion and balance network traffic. *DCB* utilizes recent advancements in antenna technology by using directional configurable smart antennas. The average interference level imposed by each mobile user in every sector of the adjacent cell is computed with respect to every coverage combination of every two nearby sectors. The *DCB* interference results are used to compute the *DCB-Enabled UMTS* system capacity. Such capacity quantification results are used in an optimization model to evaluate the optimal coverage levels of loaded and supporting sectors using different system load scenarios. Numerical results demonstrate the *DCB* ability to reduce hotspot congestion and balance system load. The efficiency of the *DCB* module, however, is decreased as the load of the loaded sector is largely increased. This is because of the increased mobile users' density in the loaded sector which increases the number of handed off mobile users for increased coverage level of the supporting sector. Therefore, a possible solution to such a scenario is to increase the granularity of the coverage levels which will decrease the number of handed off mobile users. On the other hand, dynamically changing the number of coverage levels increases the design complexity of such a module.

Chapter 5

Congestion Control and Load Balancing

In this chapter, we introduce a novel *DCB-based* Congestion Control and Load Balancing *(DCB-CC&LB)* protocol to alleviate network congestion and load imbalance. *DCB-CC&LB* is the first such protocol that exploits the cell breathing properties which are practiced in a directional manner. The *DCB-CC&LB* protocol is composed of reactive (static) and proactive (dynamic) schemes. The reactive *DCB-CC&LB* scheme evaluates the effect of practicing cell sectors coverage adaptation on the instantaneous network performance parameters; system load, average mobile users transmission powers and mobile users average outage ratio. The proactive *DCB-CC&LB* scheme, on the other hand, considers the long-term effects of directional coverage adaptation on congestion control and load balancing of interference limited wireless systems such as *UMTS*.

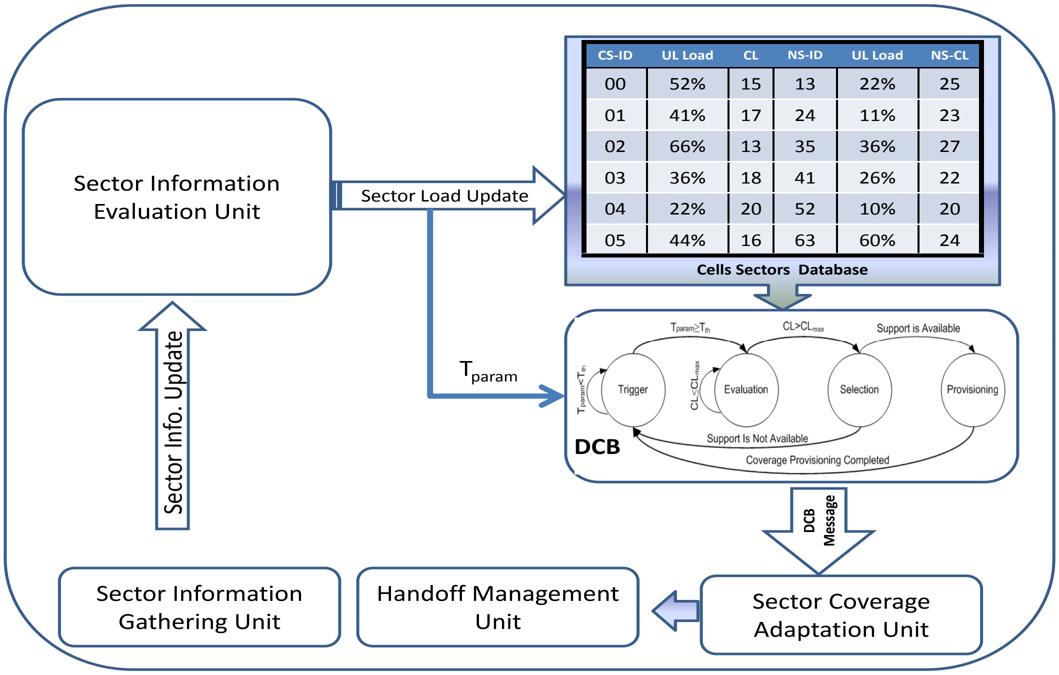
The chapter is organized as follows. The proposed protocol components and their interactions are outlined in Section 5.3. The design stages of the *DCB-CC&LB* protocol are explained in Section 5.2. The *DCB-CC&LB* heuristic schemes to evaluate the performance of a *DCB-enabled* wireless communication system are explained in Section 5.4. The system performance evaluation and results analysis are detailed in Section 5.5. The chapter is summarized in Section 5.6.

5.1 DCB-CC&LB Protocol Architectural Design

The *DCB-CC&LB* protocol is a generic protocol which can be implemented by any interference-limited wireless communication system such as *UMTS*. The architectural design of the *DCB-CC&LB* protocol includes five main parts: information gathering, information evaluation, *DCB*, coverage adaptation, and handoff management units. These units are integrated within different existing hardware and software components of the targeted wireless communication system. A block diagram of the *DCB-CC&LB* protocol is shown in Figure 5.1. Detailed functionalities of the units shown in this diagram are outlined in the following subsections.

5.1.1 Information Gathering

The *DCB-CC&LB* protocol requires some information to be available for proper functioning. This information is obtained from the base stations of cell sectors through the Information Gathering unit and concerns the current sectors' congestion state. This information can be in the form of an average value, which is computed periodically, or an instantaneous value. The information may include one element such as *SIR* value, or multiple information elements such as the values of *SIR* and dropping



CS-ID	UL Load	CL	NS-ID	UL Load	NS-CL
00	52%	15	13	22%	25
01	41%	17	24	11%	23
02	66%	13	35	36%	27
03	36%	18	41	26%	22
04	22%	20	52	10%	20
05	44%	16	63	60%	24

Figure 5.1: DCB-CC&LB Protocol Block Diagram Design

probability parameters. These depend on the design of the *Trigger* stage which may require one or multiple trigger parameters.

5.1.2 Information Evaluation

The received congestion information from every cell sector is averaged every S samples by the Information Evaluation unit. The average value is maintained in a database of M elements; where M is the number of cells in the network. One element is shown in Figure 5.1 corresponding to cell "0" of six sectors. The database is accessed by the Information Evaluation unit for information update and by the *DCB* unit to obtain sector load state information. Each element has six parameters; Cell Sector ID *(CS-ID)*, trigger parameter (Uplink load in Figure 5.1), sector Coverage Level *(CL)*, Nearby Sector ID *(NS-ID)*, Nearby Sector trigger parameter, and Nearby Sector Coverage Level *(NS-CL)*. In addition to updating the corresponding sector's information element in the database of its cell, this information element is also updated into the database of other cells where such sector is a nearby sector. Simultaneously, the updated information is propagated to the *DCB* unit for possible *DCB* technique trigger when certain criteria are met.

5.1.3 DCB

The *DCB* unit implements the *DCB* mechanism proposed in Chapter 4 and represents the heart of the proposed *DCB-CC&LB* protocol. The *DCB* unit utilizes the outcome of the Information Evaluation unit as a trigger parameter to judge if the invocation of the *DCB* coverage adaptation technique is necessary or not. When the *DCB* unit is activated, it utilizes the information stored in the cell sectors database to evaluate the *DCB-CC&LB* protocol stages. Based on the *DCB* unit outcome, specific messages are issued to the Coverage Adaptation units of the involved nearby sectors. The explained protocol design stages of the next section are implemented by the *DCB* unit.

5.1.4 Coverage Adaptation

After the selection of a nearby sectors' coverage combination, which releases loaded sector congestion and balances system load, the *DCB* unit issues two different signaling message to the *Node Bs* of the involved nearby sectors. The loaded sector message concerns the contraction of the sector service area, while the other message orders the supporting sector to expand its coverage area. Hence, the handoff area is shifted towards the loaded sector which enforces the load sector's nearedge mobile users to handoff towards the expanded supporting sector. These messages are received by the Coverage Adaptation unit which is responsible for reconfiguring the sectors' base stations radio parameters and manages mobile users' migration towards the supporting sector.

5.2 DCB-CC&LB Protocol Stages Design

In this section, we detail the considerations involved in the design of the proposed *DCB-CC&LB* protocol. The protocol is classified into four execution stages, namely: Trigger, Evaluation, Selection, and Provisioning. These stages interact with each other based on the received load information from *Node Bs'* sectors. The transition from a current stage to its preceding stage is governed by the results of the evaluated information at the current stage. When triggered, these stages evaluate the potential coverage combinations of the loaded and supporting sectors. Then, the coverage combination whose activation releases loaded sector load, while maintaining supporting sector load below limits, is selected. These stages are shown in Figure 5.2 and their explanations are as following.

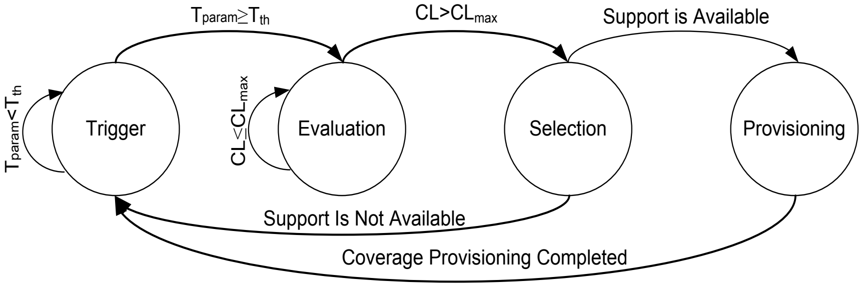


Figure 5.2: DCB-CC&LB State diagram

5.2.1 The Trigger Stage

In the Trigger stage, given the provisioned coverage combination of every two nearby sectors, the *DCB-CC&LB* protocol monitors the trigger parameters for every network sector. The trigger parameter can be the uplink load factor, uplink received interference power level, received *SIR*, etc. Whenever a sector's trigger parameter exceeds its threshold, the *DCB-CC&LB* is invoked to evaluate other coverage combinations with its nearby sector to release congestion.

5.2.2 The Coverage Combinations Evaluation Stage

Whenever congestion is detected in a sector and there is a nearby sector, the Coverage Combinations Evaluation stage is activated. In this stage, other coverage combinations of the nearby sectors are evaluated for possible loaded sector congestion release. For every candidate coverage combination, mobile users are virtually handed off from the loaded sector to the nearby sector and their connection level parameters are estimated for evaluating nearby sectors' congestion. After examining all candidate coverage combinations, the Selection stage is invoked.

5.2.3 The Coverage Combination Selection Stage

The obtained results of the Coverage Combinations Evaluation stage are examined in the Coverage Combination Selection stage. With respect to every nearby sectors' candidate coverage combination, the congestion parameters for each sector are compared to a predefined congestion threshold value. Whenever the sectors' congestion parameters with respect to the given coverage combination are found to be lower than their predefined congestion threshold values, the coverage combination is selected to be activated in the next stage. On the other hand, when there is no candidate coverage combination that can be activated, the currently active coverage combination is maintained. In this case, classical *RRM* schemes can be used to release loaded sector congestion.

5.2.4 The Coverage Combination Provisioning Stage

When a candidate coverage combination is selected, the Coverage Combination Provisioning stage is activated. In this stage, *DCB* signaling messages are involved to convey the required information for adapting the coverage areas of the nearby sectors. Such information includes the Common Pilot Channel *(CPICH)* transmission power values to be provisioned by the *Node Bs* controlling those sectors for adapting the nearby sectors coverage levels.

The design of the *DCB-CC&LB* protocol can be centralized or distributed. In a centralized protocol design, the network global information needs to be communicate to the network central unit such as *(RNC)* and the protocol decisions need to be delivered to the network base station. On the other hand, when the information exchange between network base stations is enabled, the protocol can be implemented by every base station. In the following section, practical consideration for a centralized *DCB-CC&LB* protocol for a *UMTS* system is introduced.

5.3 Practical Consideration for UMTS

Practical design of a centralized *DCB-CC&LB* protocol for *UMTS* requires the integration of the protocol units within *Node Bs* and *RNC* hardware and software components as shown in Figure 5.3. *Node Bs* implement the functionalities of the Information Gathering and the Coverage Adaptation units, where as the *RNC* implements the functionalities of the Information Evaluation and the *DCB* units. In the following, *Node B* and *RNC* design for the *DCB-CC&LB* protocol units and the enhancement of the *NBAP* signaling protocol are explained. Then, an example demonstrating the interaction of the *RNC* and *Node Bs* of a *DCB-enabled UMTS* system through the enhanced *NBAP (ENBAP)* protocol signaling messages follows.

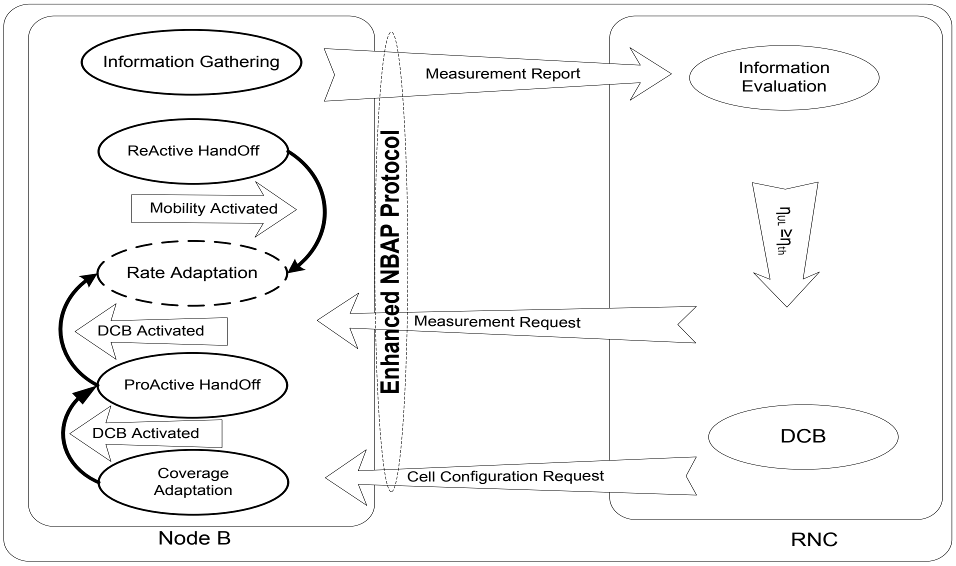


Figure 5.3: The DCB-CC&LB Protocol Integration with UMTS

5.3.1 Enhancement of Node B

We amend *Node B*'s design to support the *DCB-CC&LB* protocol's functionalities. In this amendment, the Information Gathering and Coverage Adaptation units of the *DCB-CC&LB* protocol are integrated within *Node Bs* of a *UMTS* system as shown in Figure 5.3. Herein, we consider the sectors received interference power as the information element which is periodically measured by the base station of every cell sector. The measured interference power is composed of intra- and inter-cell interference components which are used to compute the uplink load of the corresponding cell sector. The intra-cell interference is the total uplink received power of the active mobile users N in sector q of cell m which is computed using the formula shown in Equation (5.1)

$$I_{intra}^{m,q} = \sum_{n=1}^{N} p_{n,recev}^{m,q} \tag{5.1}$$

where $p_{n,recev}^{m,q}$ is the received power from mobile user n.

The other interference component, inter-cell interference, is the total received power from active mobile users which are served by the sectors of the interfering cells included in the set *C(m)*. This value is computed using Equation (5.2)

$$I_{inter}^{m,q} = \sum_{\hat{m} \in C(m)} \sum_{\hat{q}=1}^{Q} \sum_{\hat{n}=1}^{\hat{N}} p_{\hat{n},recev}^{m,q} \tag{5.2}$$

where $\hat{q}$ is the serving sector of the interfering mobile user $\hat{n}$ in the interfering cell $\hat{m}$, Q is the number of sectors in each network cell, $\hat{N}$ is the total number of mobile users in the interfering sector $\hat{q}$, and $p_{\hat{n},recev}^{m,q}$ is the received power of mobile user $\hat{n}$ at the interfered sector q of network cell m.

The periodically measured intra- and inter-cell interference values are used by Equation (5.3) to calculate the interference ratio. For network stability, this ratio has to be between "0" and "1" [30]. The computed $I_{ratio}^{m,q}$ value is periodically updated to the Information Evaluation unit implemented by the *RNC*

$$I_{ratio}^{m,q} = \frac{I_{inter}^{m,q}}{I_{intra}^{m,q}} \tag{5.3}$$

The other *DCB-CC&LB* protocol unit which is implemented by the *Node B* is the Coverage Adaptation unit. This unit reacts to the received *ENBAP* signaling message from the *DCB* unit to adapt the coverage area of a *Node B* sector. The received message includes the required *CPICH* transmission power value which needs to be provisioned by the receiving *Node B*. Therefore, the sector coverage level increases in the case of increased power value or decreases in the case of a decreased power value. When a sector coverage level increases, nearedge mobile users of its nearby sector receive its stronger pilot signal and handoff towards it. This handoff is called a Proactive Handoff since it is triggered as a result of sectors coverage adaptation, while the mobility handoff is performed through the Reactive Handoff mechanism. Moreover, the rate adaptation mechanism can be implemented to support the *DCB* technique in the case of severe loaded sector severe congestion.

5.3.2 RNC Design Considerations

The *RNC* is the network unit which implements the central part of the *DCB-CC&LB* protocol. Since the *RNC* oversees the control functionalities of different *Node Bs* and periodically receives the status of *Node Bs* radio resources utilization, global network information becomes available which simplifies the design and implementation of our proposed protocol. Therefore, the *RNC* implements the Information Evaluation and the *DCB* units.

The Information Evaluation unit is designed to compute the average of the periodically received $I_{ratio}^{m,q}$ of every cell sector for every T samples, where T is a network design parameter. The formula used for averaging such samples is given in Equation (5.4)

$$I_{Avg}^{m,q} = \frac{1}{T}\sum_{1}^{T} I_{ratio}^{m,q} \tag{5.4}$$

The averaged value, $I_{Avg}^{m,q}$, is utilized in Equation (3.7) to compute the uplink load factor of the corresponding sector. Then, the computed uplink load value is written into the *DCB-CC&LB* database and used by the *DCB* unit as a trigger parameter.

The second *DCB-CC&LB* protocol unit located at the *RNC* is the *DCB* unit which concerns the *DCB-CC&LB* protocol stages design introduced in Section 5.2. The *DCB* unit utilizes the periodically evaluated uplink load factor in its trigger stage. The successful outcome of the *DCB* unit leads to nearby sectors coverage levels adaptation. On the other hand, when the *DCB-CC&LB* protocol fails in releasing loaded sector congestion, the classical congestion control mechanisms are activated.

5.3.3 Enhanced NBAP Protocol

[1] Currently, for managing *Node B* radio resources in *UMTS* systems, the *Node B* Application Part *(NBAP)* protocol communicates the required information between *Node Bs* and *RNC* with respect to cell configuration, radio resource management, Operation and Maintenance *(OAM)*, etc. [5]. This is achieved through the *NBAP*'s common and dedicated Elementary Procedures (EPs), where an *EP* is a unit of interaction between *RNC* and *Node B*. The practical visibility and backward compatibility of the *DCB-CC&LB* protocol in *UMTS* systems require the enhancement

---

[1] In the following, the modified ENBAP procedures are capitalized, italized and in bold. UMTS messages are listed in all upper case italized letters.

of the *NBAP* protocol to support the *DCB-CC&LB* protocol functionalities. The enhancement is called Enhanced *NBAP (ENBAP)* and involves the *NBAP*'s functions of *Measurements on Dedicated Resources* and *Cell Configuration Management* [5]. More specifically, the enhancement modifies the ***Dedicated Measurement*** **EPs** and the common ***Cell Reconfiguration*** **EP** of *NBAP* protocol.

The modified dedicated resources measurement procedures of the *ENBAP* are the ***Dedicated Measurement Initiation*** and ***Dedicated Measurement Reporting*** [5]. In the ***Dedicated Measurement Initiation*** procedure, we modify the *DEDICATED MEASUREMENT INITIATION REQUEST* message which is usually issued by the *RNC* to request *Node B's* cell sector measurement of the uplink received interference ratio, I_{ratio} Information Element *(IE)*. The *DEDICATED MEASUREMENT INITIATION RESPONSE* message of the *NBAP* protocol is unchanged in the *ENBAP* protocol.

If the requested measurement reporting criteria are met, the *Node B* uses the ***Dedicated Measurement Reporting*** procedure to report the result of the requested measurements by the *RNC*. The *DEDICATED MEASUREMENT REPORT* message is modified to incorporate the *Node B* cell sectors' I_{ratio} measured values. The reporting method of the measured values is defined in the *DEDICATED MEASUREMENT INITIATION REQUEST* message, which can be "On Demand", "Periodic", etc. as defined in [5].

Based on the results of the reported measurements, the *RNC* may initiate the common ***Cell Reconfiguration*** procedure. This procedure is composed of a *CELL RECONFIGURATION REQUEST* message sent from the *RNC* to *Node Bs*, and its replay message of either *CELL RECONFIGURATION RESPONSE* or *CELL*

RECONFIGURATION FAILURE. In the *ENBAP* protocol, we modify the *Primary CPICH Information IE* of the *CELL RECONFIGURATION REQUEST* message to have a discrete set of transmission power values instead of a value from a continuous power range [5]. Each discrete power value corresponds to a *Node B* cell sector coverage level. The *CELL RECONFIGURATION REQUEST* message delivers this value to the Coverage Adaptation unit for actual provisioning of the pilot transmission power of the base station of the corresponding sector. The message has either a *CELL RECONFIGURATION RESPONSE* for indicating successful sector coverage reconfiguration, or a *CELL RECONFIGURATION FAILURE* when the operation is failed [5].

5.3.4 Node Bs and RNC Interaction

A successful example of the *ENBAP* scenario shows the interaction of two *Node Bs* and their *RNC* with respect to the nearby sectors q and $\hat{q}$ which is presented in Figure 5.4. In this scenario, first *RNC* issues a *DEDICATED MEASUREMENT INITIATION REQUEST* message to q and $\hat{q}$ of *Node B1* and *Node B2* respectively at time t_1. Upon receiving these requests, the *Node Bs* establish the required measurement procedure and reply with the *DEDICATED MEASUREMENT INITIATION RESPONSE* messages at t_2. Whenever the reporting criterion included in the previously received *DEDICATED MEASUREMENT INITIATION REQUEST* message is met, the *Node Bs* issue the *DEDICATED MEASUREMENT REPORT* message as in t_3, t_4, and t_7, which includes the value of the requested measurement of the I_{ratio} IE.

Upon receiving T reported values, the *Information Evaluation* unit evaluates their

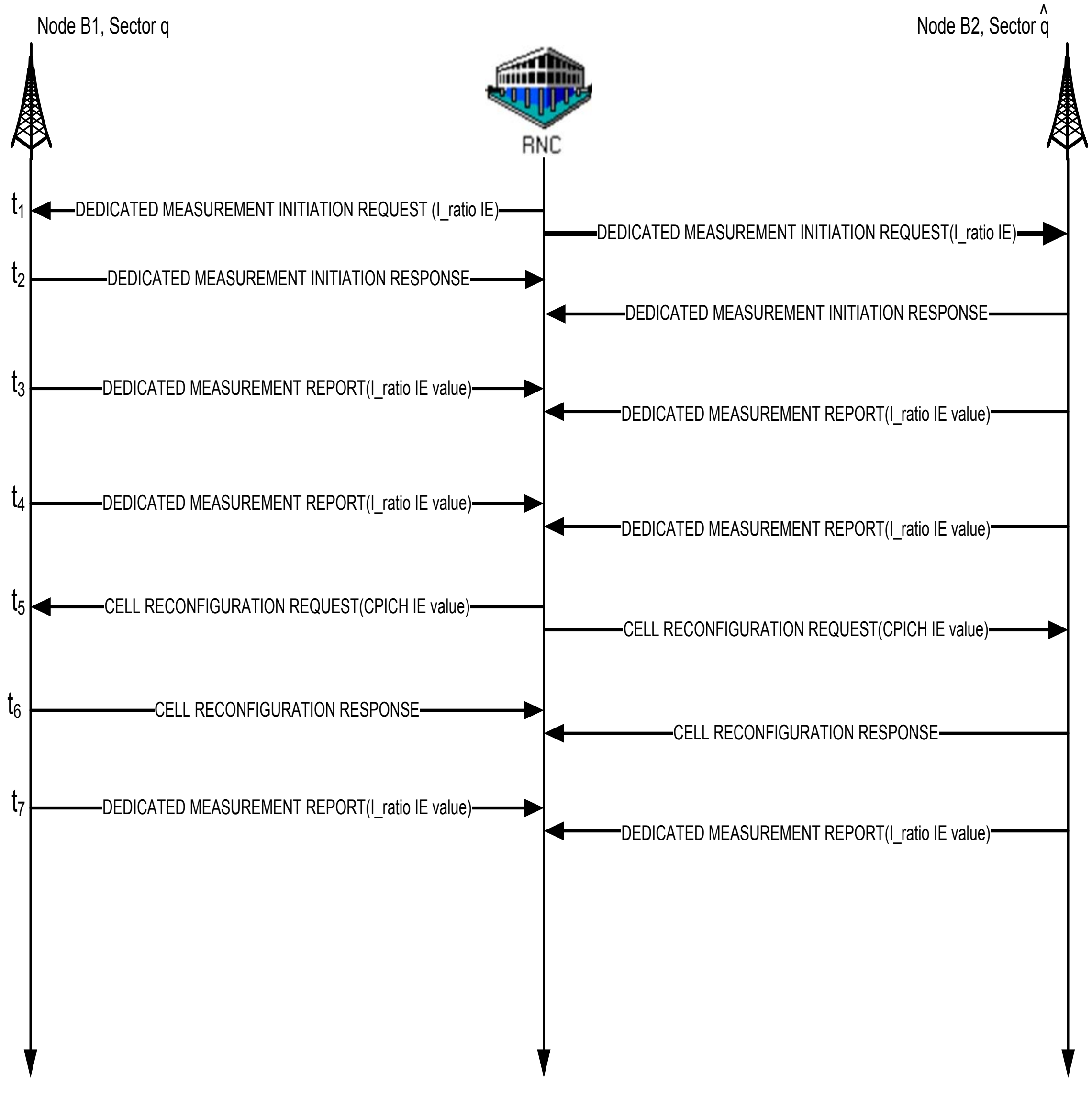


Figure 5.4: A Successful ENBAP Protocol Sequence Diagram

average and computes the uplink load factor for the corresponding sector. The computed value is updated to the *DCB-CC&LB* protocol database and simultaneously fed to the *DCB* unit to be triggered when the computed uplink load factor exceeds a predefined threshold. The invocation result of the *DCB* unit is two *CELL RECONFIGURATION REQUEST(CPICH IE)* messages sent to the Coverage Adaptation unit of each *Node B*. These messages include the required *CPICH* transmission power values for sectors q and $\hat{q}$. Based on the received *CPICH* transmission powers, each *Node B* provisions its sector coverage level based on the selected coverage combination by the *DCB* unit. The successful sectors configuration results in the *CELL RECONFIGURATION RESPONSE* messages from both *Node Bs*. Then, the *DEDICATED MEASUREMENT REPORT* messages resume again from both *Node Bs*.

In the following section, we introduce the design of the Reactive and Proactive *DCB-CC&LB* schemes, which implement the *DCB-CC&LB* protocol in a heuristic manner.

5.4 Heuristic DCB Schemes

Two heuristic *DCB* schemes for *DCB-CC&LB* protocol are proposed to evaluate the performance of a *DCB-enabled UMTS* system. These schemes are called the Reactive *DCB (RDCB)* and the Proactive *DCB (PDCB)*. The interference ratio parameter, I_{ratio}, of a cell sector is utilized for evaluating the uplink load factor. The computed value is used as a trigger parameter of the *DCB* technique. These schemes can be used for load balancing when the system load is moderate, and as a congestion control schemes when the system is congested. This requires the adaptation of the trigger threshold, which is lowered when the system is not congested while the threshold is increased when the system is loaded. Before introducing these schemes, we define the following (in addition to the definitions in Section 4.3):

- η_{UL}: uplink load factor of a sector
- $\eta_{UL}^{m,q,l}$: uplink load factor of loaded sector q of cell m given coverage level l
- $\eta_{UL}^{\hat{m},\hat{q},\hat{l}}$: uplink load factor of supporting sector $\hat{q}$ of cell $\hat{m}$ given coverage level $\hat{l}$
- η_{UL}^{th}: uplink load factor threshold

5.4.1 Reactive DCB-CC&LB Scheme

The Reactive *DCB (RDCB)* is a heuristic scheme for congestion control and load balancing in a *DCB-enabled UMTS* system. The *RDCB* is a *measurement-based* scheme which monitors the load state changes of the system due to mobile users' call arrival, call departure, and/or mobility. Moreover, the *RDCB* scheme can be used in the case of partial or full network base station failure to migrate mobile users from the cell sector of a malfunctioning base station to the base station of a nearby supporting sector. When congestion is formed or nearby sectors' load is unbalanced, the scheme is triggered to adapt the coverage area of the loaded and supporting sectors which leads to releasing congestion and balancing system load. The *RDCB* scheme is composed of three procedures; *InformationGathering(m,q)*, *InformationEvaluation*(I_{ratio}, m, q, $N_l^{m,q}$, l, L_{Max}, L_{Min}), and *DCB*(m, q, $N_l^{m,q}$, l, L_{Max}, L_{Min}). This scheme is sketched in Algorithm 1 with respect to a loaded sector q of cell m and its nearby supporting sector $\hat{q}$ of adjacent cell $\hat{m}$.

Algorithm 1 Reactive DCB Scheme

```
$NetworkInitialization(M, Q)$
while $i \leq I$ do
    procedure InformationGathering($m, q, l$)
        $ReactiveIniti(m, q, N_l^{m,q}, l, L_{Max}, L_{Min})$
        $I_{ratio} \leftarrow InterferenceMeasurement(m, q, l)$
        $InformationEvaluation(I_{ratio}, m, q, N_l^{m,q}, l, L_{Max}, L_{Min})$
    end procedure
    procedure InformationEvaluation($I_{ratio}, m, q, N_l^{m,q}, l, L_{Max}, L_{Min}$)
        $\eta_{UL} \leftarrow (1 + I_{ratio}) \star \sum_{n=1}^{N_l^{m,q}} \frac{1}{1+\frac{W}{(\frac{E_b}{I_o})_n \star R_n}}$
        $DCBDataBase(\eta_{UL}, m, q, l)$
        if $(\eta_{UL} \geq \eta_{UL}^{th})$ then                              ▷ DCB Trigger Stage
            $DCB(m, q, l, L_{Max}, L_{Min})$
        end if
    end procedure
    procedure DCB($m, q, l, L_{Max}, L_{Min}$)
        for $cl \leftarrow l - 1, L_{Min}$ do                          ▷ Coverage Evaluation Stage
            if $d_n^{m,q} > r(cl)$ then
                $VirtualHandoff(\hat{m}, \hat{q}, \hat{l})$
            end if
            $ComputeTxRxPowers(m, q, l, \hat{m}, \hat{q}, , \hat{l})$
            $I_{ratio}^{m,q} \leftarrow InformationGathering(m, q)$
            $I_{ratio}^{\hat{m},\hat{q}} \leftarrow InformationGathering(\hat{m}, \hat{q})$
            $\eta_{UL}^{m,q,l} \leftarrow (1 + I_{ratio}) \star \sum_{n=1}^{N_l^{m,q}} \frac{1}{1+\frac{W}{(\frac{E_b}{I_o})_n \star R_n}}$
            $\eta_{UL}^{\hat{m},\hat{q},\hat{l}} \leftarrow (1 + I_{ratio}^{\hat{m},\hat{q}}) \star \sum_{\hat{n}=1}^{\hat{N}_{\hat{l}}^{\hat{m},\hat{q}}} \frac{1}{1+\frac{W}{(\frac{E_b}{I_o})_{\hat{n}} \star R_{\hat{n}}}}$
        end for

        for $cl \leftarrow l - 1, L_{Min}$ do                 ▷ Selection and Provisioning Stage
            if $(\eta_{UL}^{m,q,l} \leq \eta_{UL}^{th}) \&\& (\eta_{UL}^{\hat{m},\hat{q},\hat{l}} \leq \eta_{UL}^{th})$ then
                $CoverageAdaptation(m, q, l, )$
                $CoverageAdaptation(\hat{m}, \hat{q}, \hat{l})$
            end if
        end for
    end procedure
    $i = i + 1$
end while
```

The execution of the *RDCB* scheme starts by initializing the network and mobile users parameters using *NetworkInitialization(M,Q)* function. Then, the algorithm iterates I times. The mobile users' distribution over the coverage area of their sectors varies with respect to every iteration i. In each iteration the *RDCB* procedures are executed. First, the *InformationGathering*(m, q, l) procedure is invoked with respect to every sector q of network cell m for a given sector's coverage level l. At the start of each iteration, the function *ReactiveIniti(m,q,* $N_l^{m,q}$*, l,* L_{Max}*,* L_{Min}*)* is invoked to redistribute mobile users in each sector to vary the received interference power in the network. Then, mobile users parameters such as transmission and receiving powers, *SIR* parameters, etc. are recomputed. The *InterferenceMeasurement(m,q, l)* function is invoked to estimate the received intra- and inter-cell interference ratio, I_{ratio}, of the sector q of cell m. The computed interference ratio is passed to the *InformationEvaluation(*I_{ratio}*, m, q, N, l,* L_{Max}*,* L_{Min}*)* procedure.

Upon receiving the I_{ratio}, the *InformationEvaluation()* procedure computes the uplink load factor, η_{UL}, of sector q for current coverage level l. The computed value is written into the *DCBDataBase(*η_{UL}*, m, q, l)* and then evaluated against a predefined uplink load threshold, η_{UL}^{th}. A successful evaluation indicates that the sector is congested and leads to the *DCB(m, q, l,* L_{Max}*,* L_{Min}*)* procedure activation to evaluate other coverage combinations with the nearby sector for releasing such congestion. For every candidate coverage combination, the expected mobile users to be handed off towards the supporting sector are virtually handed off and the new transmission and receiving powers of mobile users at the loaded and supporting sectors are re-evaluated. Then, the *InformationGathering()* procedure is invoked and the uplink load factor is computed with respect to both sectors. The evaluated uplink load factors with respect to every coverage combination are maintained in $\eta_{UL}^{m,q,l}$ and $\eta_{UL}^{\hat{m},\hat{q},\hat{l}}$ parameters. After the evaluation of the uplink load factors of every possible coverage combination, the Selection and Provisioning stages are invoked to select a coverage combination that can release the loaded sector congestion and maintain the supporting sector load below limits. Then, the *DCB* decision is relayed to the Coverage Adaptation units of the loaded and supporting sectors through the *CoverageAdaptation(m, q, l)* and *CoverageAdaptation*($\hat{m}$, $\hat{q}$, $\hat{l}$) respectively for actual sectors coverage adaptation.

5.4.2 Proactive DCB-CC&LB Scheme

For long-term congestion relief and system load balancing a Proactive *DCB (PDCB)* scheme is introduced. In *PDCB*, the *DCB* technique is invoked based on some predictive assessment of the *QoS* parameters such as the received interference level, *SIR*, call dropping probability, etc. Herein, the *PDCB* also utilizes the uplink load factor of cell sectors to invoke the *DCB* technique in the case of congestion.

There are four events in the *PDCB* scheme shown in Algorithm 2; namely: *callArrival*, *callDeparture*, *handOff*, and *informationEvaluation*. The occurrence of each event invokes its corresponding procedure; *CallArrival()*, *CallDeparture()*, *MobileUserHandOff()*, and *InformationEvaluation()* respectively. The *CallArrival()*, and *CallDeparture()* procedures overlooks the arrival and departure processes of mobile users' calls to the system, while the *MobileUserHandOff()* procedure manages mobile users' mobility. The *InformationGathering* unit of the *DCB-CC&LB* protocol is embedded in these procedures to update the system load state parameters after call arrival, call departure, and mobile users' handoff events.

Algorithm 2 Proactive DCB Algorithm

```
procedure PDCB-CC&LB
    $NetworkInitialization(M, Q)$
    $TotalSimulationTime \leftarrow Time$
    $CurrentTime \leftarrow 0$
    $NumberOfSamples \leftarrow 0$
    while $CurrentTime \leq TotalSimulationTime$ do
        $CurrentTime =$
        $Min(ArrivalTime, DepartureTime, HandOffTime, DCBTime)$
        if $CurrentTime = ArrivalTime$ then
            $CallArrival()$
        else if $CurrentTime = DepartureTime$ then
            $CallDeparture()$
        else if $CurrentTime = HandOffTime$ then
            $MobileUserHandOFF()$
        else if $CurrentTime = InformationEvaluation$ then
            $InformationEvaluation()$
        end if
    end while
end procedure
```

The *InformationEvaluation()* procedure shown in Algorithm 3 evaluates the periodically updated network sectors' load information. Herein, the updated information is the sectors' received interference ratio, $I_{ratio}^{m,q,l}$, which varies based on mobile users' calls arrivals and departures, and mobile users mobility events. In this procedure, for every sector's S received $I_{ratio}^{m,q,l}$ samples, the average value, $I_{avg}^{m,q,l}$, is computed. This computed average is used to calculate the uplink load factor, $\eta_{UL}^{m,q,l}$ of that sector. Then, the obtained $\eta_{UL}^{m,q,l}$ is updated into the system database and it is evaluated against a predefined threshold, η_{UL}^{th}. The successful evaluation invokes the *DCB(m, q, l,* L_{Min}*)* procedure shown in Algorithm 4.

When it is activated, the *DCB(m, q, l,* L_{Min}*)* procedure decreases the coverage area of the load sector while increasing the coverage area of the nearby supporting sector through the *CoverageAdaptation()* function. Changing the coverage levels of the loaded and supporting sectors proactively handoff nearedge mobile users of the contracted loaded sector towards the expanded supporting sector through the *MobileUsersHandOff()* procedure. The uplink load factors of both sectors are then re-evaluated and updated into the system database. Executing this scheme for long enough time, converges the network sectors coverage levels to a stable state which sustain the network sectors' traffic demands while minimizing system's call blocking and dropping probabilities. The performance of the proposed heuristic schemes is evaluated in the following section using a representative *UMTS* cellular system.

5.5 Performance Evaluation

In this section, we evaluate the performance of a *DCB-enabled UMTS* system through the proposed *H-DCB* schemes. Using simulation, we examine the effect of directional

Algorithm 3 InformationEvaluation Algorithm

```
procedure InformationEvaluation
    for $m \leftarrow 1, M$ do
        for $q \leftarrow 1, Q$ do
            $I_{total}^{m,q,l} \leftarrow I_{total}^{m,q,l} + I_{ratio}^{m,q,l}$
            $NumberOfSamples \leftarrow NumberOfSamples + 1$
            if $(NumberOfSamples \geq S)$ then
                $I_{avg}^{m,q,l} \leftarrow \frac{I_{total}^{m,q,l}}{S}$
                $\eta_{UL}^{m,q,l} \leftarrow (1 + I_{avg}^{m,q,l}) \star \sum_{n=1}^{N_l^{m,q}} \frac{1}{1+\frac{W}{(\frac{E_b}{I_o})_n \star R_n}}$
                $DCBDataBase(\eta_{UL}, m, q, l)$
                if $(\eta_{UL}^{m,q,l} > \eta_{UL}th)$ then
                    $DCB(m, q, l, L_{Min})$
                end if
                $I_{total}^{m,q,l} \leftarrow 0$
                $NumberOfSamples \leftarrow 0$
            end if
        end for
    end for
end procedure
```

Algorithm 4 DCB Algorithm

```
procedure DCB($m, q, l, L_{Min}$)
    if $((l-1) \geq L_{Min})$ then
        $CoverageAdaptation(m, q, (l-1),)$
        $CoverageAdaptation(\hat{m}, \hat{q}, (\hat{l}+1))$
        for $n \leftarrow 1, N_l^{m,q}$ do
            if $d_n^{m,q} > r(l-1)$ then
                $MobileUserHandOFF(\hat{m}, \hat{q})$
            end if
        end for
        $I_{ratio}^{m,q,(l-1)} \leftarrow InterferenceMeasurement(m, q, (l-1))$
        $I_{ratio}^{\hat{m},\hat{q},(\hat{l}+1)} \leftarrow InterferenceMeasurement(\hat{m}, \hat{q}, (\hat{l}+1))$
        $\eta_{UL}^{m,q,(l-1)} \leftarrow (1 + I_{ratio}^{m,q,(l-1)}) \star \sum_{n=1}^{N_l^{m,q}} \frac{1}{1+\frac{W}{(\frac{E_b}{I_o})_n \star R_n}}$
        $\eta_{UL}^{\hat{m},\hat{q},(\hat{l}+1)} \leftarrow (1 + I_{ratio}^{\hat{m},\hat{q},(\hat{l}+1)}) \star \sum_{\hat{n}=1}^{\hat{N}_{\hat{l}}^{\hat{m},\hat{q}}} \frac{1}{1+\frac{W}{(\frac{E_b}{I_o})_{\hat{n}} \star R_{\hat{n}}}}$
        $DCBDataBase(\eta_{UL}^{m,q,(l-1)}, m, q, (l-1))$
        $DCBDataBase(\eta_{UL}^{\hat{m},\hat{q},(\hat{l}+1)}, \hat{m}, \hat{q}, (\hat{l}+1))$
    end if
end procedure
```

coverage adaptation on the QoS metrics such as outage ratio, call blocking and dropping probabilities and mobile users' transmission powers and compare these results to the results of the Fixed Pilot Power *(FPP)* scheme where the coverage area of network sectors remains fixed.

5.5.1 Simulation Model

To evaluate the performance of the proposed *H-DCB* schemes, the 7-cell network model shown in Section 3.1 is used. Each cell is divided into Q sectors where each sector is served by a smart directional antenna. The coverage area of each sector is partitioned into L supporting levels of equal width. Each sector has a minimum coverage level and maximum coverage level defined by L_{min} and L_{max} respectively. The system bandwidth is W. The energy per bit to interference noise ratio for a mobile user is defined by $(\frac{E_b}{I_o})$. The minimum and maximum mobile users transmission powers are P_{min} and P_{max}, respectively, and the cell sector uplink load threshold is η_{UL}^{th}. The system parameters used in the simulation are shown in Table 5.1. The used mobile users' signal propagation loss model and the received power model given in Equations (3.4), and (3.4) are used to evaluate mobile users' transmission and receiving powers. In our simulation, we assume that the effect of fast fading is mitigated by several other techniques such as a RAKE receiver, channel coding, bit interleaving, etc.

In a dynamic system environment, users mobility which leads to handoff from one cell sector to another, needs to be implemented. Herein, we adapt a simple and well known mobility model– the Random Walk model [69]. In this model, a mobile user pauses in a cell sector q for a random period of time as shown in Figure 5.5. When the time period expires, the mobile user hands off towards one of the adjacent sectors,

Table 5.1: DCB-enabled UMTS Network Simulator Parameters

Network Parameters	**Parameters Values**
number of cell (M)	7
Cell Radius	1 km
number of sectors (Q)	6
System Bandwidth (W)	3.840 MHz
Uplink base Frequency	1920 MHz
η_{UL}^{th}	80%
Antenna gains	18dB
Thermal noise	-104 dBm
Minimum Sector Coverage Level (L_{min})	10
Maximum Sector Coverage Level (L_{max})	30
UE Parameters	**Parameters Values**
Maximum transmitted power (P_{max})	27 dB
Minimum transmitted power (P_{min})	-50 dB
Thermal noise	-100 dBm
E_b/N_o	5 dB

q_j, with probability $P(q_j)$, where j is the mobile user destination sector index with respect to its current sector. Herein, the mobile user residence time in a cell sector is defined by an average sector residence time *(SRT)* provisioned value. This value is a reciprocal to the mobile users handoff rate from the sector.

Whenever a mobile user hands off from a sector to a neighboring sector, its new location in the new sector is randomly generated. Moreover, if the handed off mobile user has a call in progress, its new transmission and receiving powers, intra- and inter-cell interference, *SIR*, and the uplink load factor of its former and current sectors are re-evaluated. If the current resources of the new sector do not permit the admission of the new handed off call and rate adaptation is not enabled, the call is dropped. When rate adaptation is enabled, the transmission rate of active mobile users and the newly arrived one are adapted for the system to be able to admit the new call.

The Minimum Rate Adaptation scheme proposed in [72] is used for mobile users'

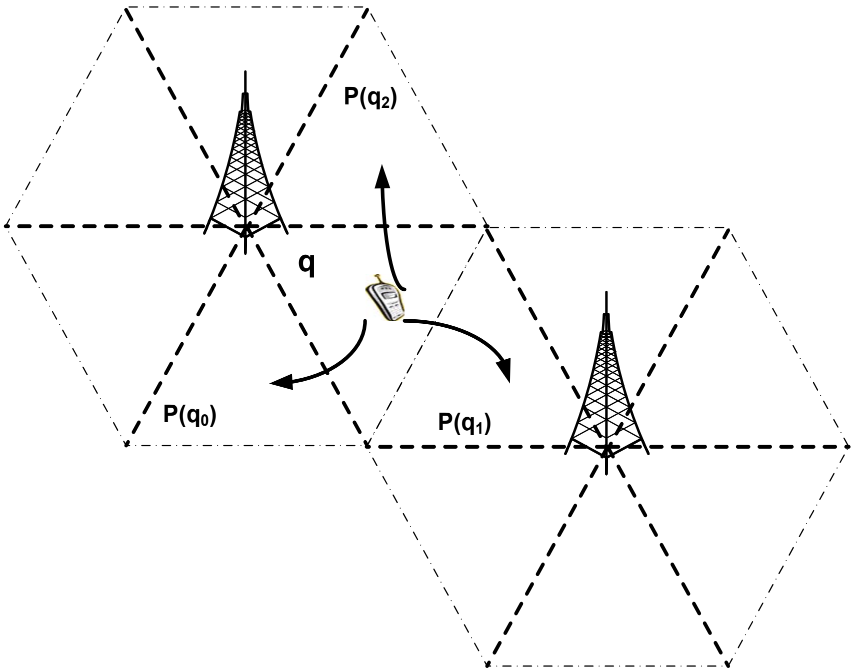


Figure 5.5: PDCB Mobility Model

rate adaptation in the *PDCB* scheme. The scheme is utilized for adapting the transmission rates of the handed off mobile users to maximize the supporting level given to a loaded sector and minimize call dropping rates as well.

5.5.2 RDCB Performance Evaluation

The performance of the *RDCB* scheme is evaluated using a snapshot simulation. Each mobile user's transmission rate R is set and fixed at 25 kb/s. Mobile users remain stationary during simulation. For a given number of mobile users in each sector, the received interference level at the sectors' base stations is varied through different mobile users distribution realizations. The mobile users' distribution realizations correspond to the *RDCB* scheme iterations. The instantaneous performance parameters defined in this section are evaluated with respect to every realization.

Then, the average values of such parameters are calculated and analyzed. First, the tested performance parameters and the simulation scenarios are outlined. Then, the simulation results and their analysis are discussed.

RDCB Performance Metrics

- *Average Uplink Load Factor:* The uplink load factor is a measure of the instantaneous uplink load in a *WCDMA* system. The η_{UL}^{bef} and η_{UL}^{aft} parameters are defined to hold the uplink load factor before and after the *RDCB* scheme execution.

- *Average transmission power:* The average transmission power of mobile users in the loaded and supporting sectors. An increased mobile users' transmission powers indicate the sector is loaded.

- *Outage Ratio:* The outage ratio of a sector indicates the ratio of the number of mobile users' calls turned off, because of exceeding their power budget due to the increased interference level in the system, to the total number of mobile users in the sector.

Simulation Scenarios

The daily network utilization pattern follows a certain behavior in which network resources at certain coverage area are more utilized during a certain period of time while such resources become underutilized during other time periods. For example, city centers are more crowded in the morning but less crowded in the evening. On the other hand, residential and recreational areas are less likely to be congested in the morning while more mobile users will be located in such areas at the afternoon and evening time. To evaluate the network performance using the proposed *RDCB* scheme and considering this network utilization pattern, first, 420 mobile users are distributed over the network coverage area. Hence, each cell sector will have 10 mobile users uniformly distributed over its coverage area. Then, for the purpose of creating a hotspot congestion, 120 extra mobile users are uniformly distributed over the coverage area of the center cell sectors. Therefore, network traffic imbalance is created and drives center cell sectors to the congestion state. These mobile users are redistributed over the coverage area of their sectors for I =*100* iterations. In each iteration, the instantaneous performance parameters are collected. At the end of these iterations, the average values of the collected performance parameters are computed and the obtained results are maintained for analysis. Then, a step-wise approach for increasing hotspot congestion is used. In each step, 5 extra mobile users are uniformly distributed over each hotspot sector and the *RDCB* scheme is invoked again for another I iterations. This process continues until the number of mobile users in each loaded sector reaches 50 mobile users. The simulation results for these scenarios are discussed next.

Simulation Results

Since the obtained results with respect to every loaded and supporting sectors pairs are similar, herein we show the results of one pair. The averages of the uplink load factor, mobile users' transmission powers and mobile users' outage ratio are shown in Figures 5.6-5.8. These averages are with respect to 100 different traffic distribution realizations. In each figure, the level 0 results correspond to the *FPP* scheme, while the results of levels $SL_1 - SL_{10}$ correspond to the activated supporting levels of the supporting sector with respect to the *RDCB* scheme.

The results in Figure 5.6 show the average uplink load factor of the loaded and supporting sectors with respect to different supporting levels from the support sector. The results of the *FPP* scheme shows a higher loaded sector load and low supporting sector load. This is due to the traffic imbalance of these sectors. Then, for every load scenario, the *RDCB* scheme is executed to evaluate all possible coverage combinations of the loaded and supporting sectors. In each scenario, the first coverage combination that maintains the loaded and supporting sectors' loads below η_{UL}^{th} is activated. As shown in the figure, for the first load scenarios of 30 and 35 mobile users in the loaded sector, the predefined uplink load threshold of 80% has not been exceeded. Therefore, the normal coverage levels of both sectors can be maintained.

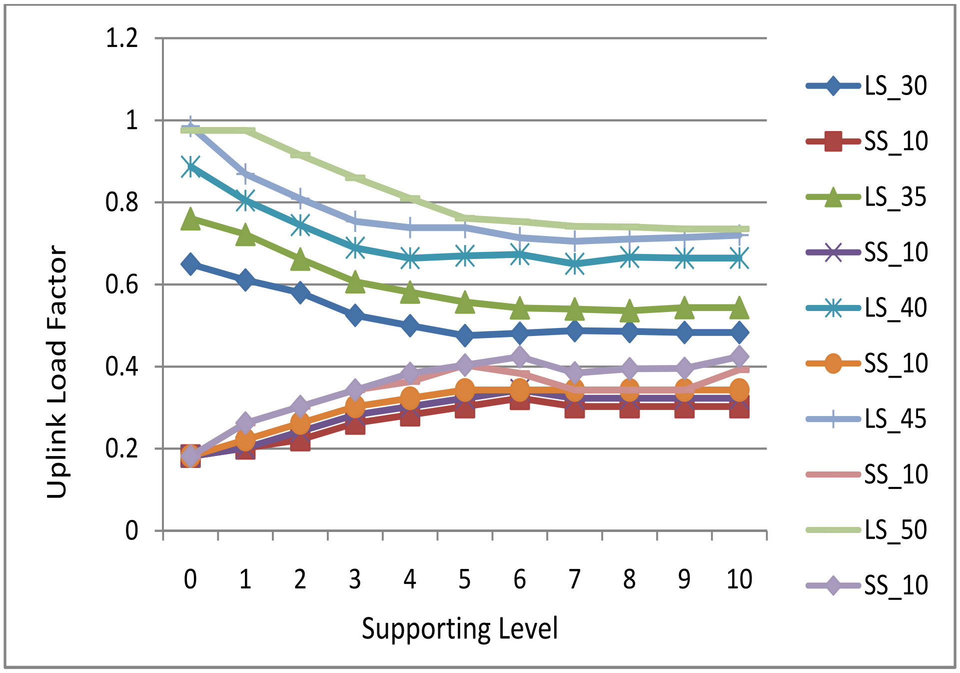


Figure 5.6: Load and Supporting Sectors uplink load Factors

On the other hand, as the loaded sector load exceeds η_{UL}^{th} due to the higher number of served mobile users, the coverage adaptation of the nearby sectors needs to be practiced to enforce nearedge mobile users of the loaded sector to handoff towards the supporting sector. As can be seen from the figure, different supporting levels need to be activated for these load scenarios. For example, in the case of 40 mobile users in the loaded sector activating supporting level SL_2 brings the loaded sector load below 80%. For the 45 and 50 mobile users load scenarios, supporting levels SL_3 and SL_5 are required to be activated respectively. As the figure shows, the loaded sector load is decreased as the supporting level is increased up to level SL_5. This is due to the coverage adaptation of the loaded and supporting sectors practiced by the *RDCB* scheme.

Starting of support level SL_6, the uplink load factor of the loaded and supporting sectors are nearly constant. This is due to the increased mobile users' transmission powers, especially the proactively handed off mobile users. Therefore, the interference level in the system is increased as the support level is increased. This drives some mobile users to exceed their maximum available transmission power and they are turned off.

The average mobile users transmission powers with respect to the provisioned supporting levels are shown in Figure 5.7. As the figure reveals, the average combined transmission power varies from one supporting level to another. This variation is based on the load level at the loaded sector and the activated supporting level of its nearby sector. When the loaded sector load is moderate, the average transmission power is decreased as the support level is increased. As the sector load is increased, the average transmission power is increased. This is because of the increased number of mobile users requesting access to the system, as well as the increased interference

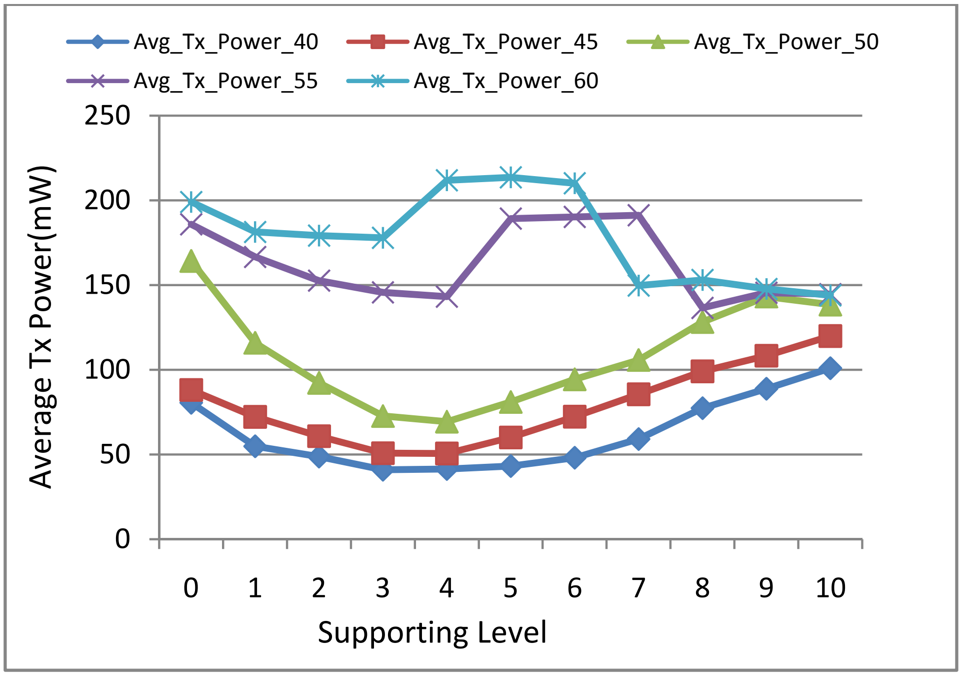


Figure 5.7: Mobile Users Average Transmission Power

level imposed by the proactively handed off mobile users towards the supporting sector. Therefore, the mobile users need to increase their transmission powers to maintain their provisioned parameters.

The transmission powers of the mobile users in the loaded and supporting sectors are increased as the system load and the given support are increased. At some point, the *SIR* of a mobile user cannot be maintained because its transmission power exceeds the maximum provisioned value. This is because of the increased distance of handed off mobile users towards the base station of the supporting sector. Therefore, these mobile users are turned off. Figure 5.8 shows the average combined outage ratio. As the figure shows, this ratio is increased as the support level and loaded sector load are increased.

After the evaluation of every possible coverage level of the loaded and supporting

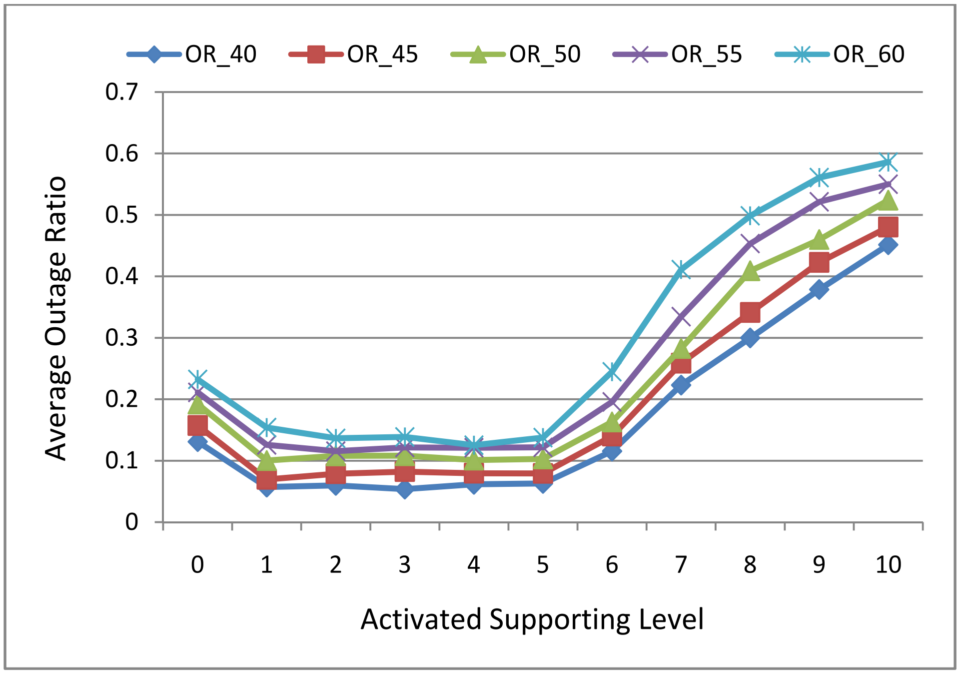
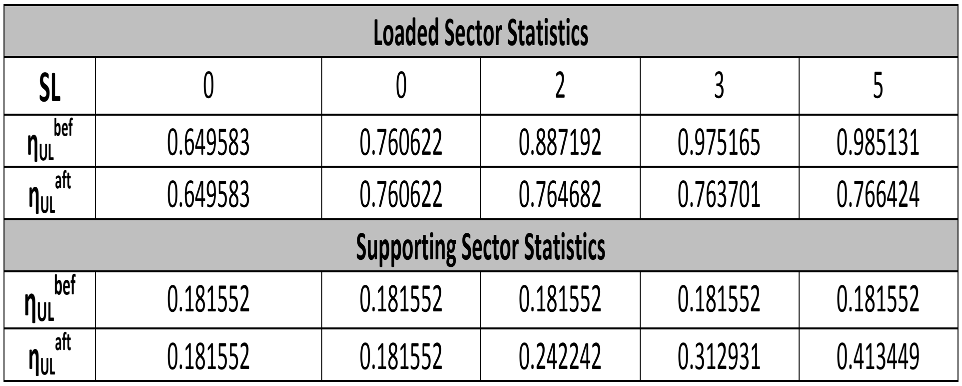


Figure 5.8: Mobile Users Outage Ratio



Loaded Sector Statistics					
SL	0	0	2	3	5
η_{UL}^{bef}	0.649583	0.760622	0.887192	0.975165	0.985131
η_{UL}^{aft}	0.649583	0.760622	0.764682	0.763701	0.766424
Supporting Sector Statistics					
η_{UL}^{bef}	0.181552	0.181552	0.181552	0.181552	0.181552
η_{UL}^{aft}	0.181552	0.181552	0.242242	0.312931	0.413449

Table 5.2: Provisioned Coverage levels and Load Statistics

sectors, the *RDCB* algorithm activates the minimum coverage level which releases hotspot congestion and maintains the supporting sector load below the provisioned threshold. Table 5.2 shows the provisioned supporting levels for every load scenario along with the sectors' loads before; η_{UL}^{bef}, and after; η_{UL}^{bef}, *RDCB* execution.

Based on the presented results, the *RDCB* scheme reacts well to release system congestion by adapting the coverage levels of the loaded and supporting sectors. The given support to the loaded sector depends on the load state of the supporting sector.

5.5.3 PDCB Performance Evaluation

A dynamic simulator for a *DCB-enabled UMTS* system is implemented to evaluate the system performance using *PDCB* scheme. In this simulation experiment, mobile users' call arrivals and departures follow random processes which are defined by the average call arrival and departure rates. Moreover, mobile users' mobility and rate adaptation for mobile users' active calls are enabled. The uplink load factor of a cell sector is sampled periodically every t seconds. Every S load samples, the average uplink load factor is computed and whenever the computed value exceeds a predefined threshold, the *DCB* procedure of the *PDCB* scheme is invoked to practice nearby sectors coverage adaptation. The simulation parameters, performance parameters, and simulation scenarios and results analysis are introduced next.

Simulation Setup

In addition to the simulation parameters given in Section 5.5.1, herein, the base stations use a measurement interval of 2s to sample their received uplink interference. Also, the *RNC* computes the average interference ratio every T =10 samples. In each simulation scenario, initially 100 mobile users are uniformly distributed over network sectors and remain in the system until the end of the simulation. However, during simulation the mobile users' distribution varies from one sector to another based on the average exponentially distributed *Sector Residence Time (SRT)* parameter. Users' mobility is controlled by the *SRT* of each sector which is inversely related to the users' handoff rate. In the simulation scenarios, we vary *SRT* to vary sectors' loads.

Herein, a single class of service is considered and the presented results are for multi rate voice traffic of 8, 9.6 and 12.2 kbps. Calls are generated according to a Poisson process with an average call arrival rate of 10 calls/h/user and exponentially distributed call holding time of an average of 180 s. Upon its arrival, a call is assigned a predetermined transmission rate, $r_{requetsed}$, where $r_{Min} < r_{requested} \leq r_{Max}$. In this chapter, $r_{requested}$ is assigned the maximum transmission rate of 12.2kbps.

PDCB Performance Metrics

- *Call Blocking Rate:* The ratio of the system denied new calls to the total arrived new calls. It is an important system *QoS* parameter which needs to be minimized.

- *Call Dropping Rate:* The ratio of the system denied handoff calls to the total handoff calls. It is also an important *QoS* parameter which needs to be as low as possible. This parameter has greater importance than the call blocking rate parameter since, with respect to the mobile user, blocking a new call is less annoying than dropping an active call.

- *Activated Supporting Level:* The activated supporting level with respect to every load scenario. This parameter shows the dynamic behavior of the cell sectors' coverage levels based on the load level at each sector.

Simulation Scenarios

A delay-based hotspot [34] is used to evaluate the performance of the proposed *PDCB*. The center cell of the 7-cell network model is the hotspot cell while the surrounding cells are the supporting cells. Different approaches have been used to examine the proposed algorithm, namely:

- Delay-Based Hotspot Cell and Lightly Loaded Supporting Cells
- Delay-Based Hotspot Cell and Nearly Loaded Supporting Cells

The results of these simulation scenarios are explained next.

Delay-Based Hotspot Cell and Lightly Loaded Supporting Cells

In this scenario, a delay based hotspot cell of 6 sectors is formed. This cell is supported by lightly loaded adjacent cells. The scenario is implemented for *FPP* and *PDCB* schemes. The rate adaptation scheme in [72] is applied for both. In this scenario, the average *SRT* of all network sectors except the loaded center cell sectors is set and fixed at 5 seconds. To create a hotspot at the selected loaded cell sectors, the average *SRT* is varied from 10-30 seconds in steps of 5s. Since the results obtained for all loaded and supporting sectors are similar, the call blocking and dropping results of a single loaded sector *(LS)* and its nearby supporting sector *(SS)* with and without rate adaptation are depicted in Figures 5.9-5.12[2]. These results are the average of 10

---

[2] In the figures, *LS_FPP*, *LS_DCB*, *LS_DCB_RA*, *SS_FPP*, *SS_DCB*, *SS_DCB_RA* correspond to the dropping and blocking probabilities of the loaded and supporting sectors for the FPP and DCB schemes. *RA* stands for rate adaptation

runs where each run lasts for one simulation hour.

As can be inferred from Figure 5.9, the average call blocking rate for both loaded and supporting sectors remains the same for *FPP* and *PDCB* protocols up to 20s of the loaded sector *SRT* parameter value. As the *SRT* value of the loaded sector exceeds 20s, the blocking rate for the *PDCB* protocol becomes slightly higher than that of the *FPP* scheme. Therefore, as the loaded sector load increases, the support level is increased and the number of proactively handed off mobile users is also increased. The increased number of handed off mobile users leads to an increase in the interference level in the system. This negatively impacts the loaded sector by increasing the inter-cell interference, and the supporting sector due to the increased intra-cell interference level. Therefore, the blocking probability increases at the loaded and supporting sectors because of the higher transmission power of the proactively handed off mobile users, which prevents admission of new call arrivals to the system.

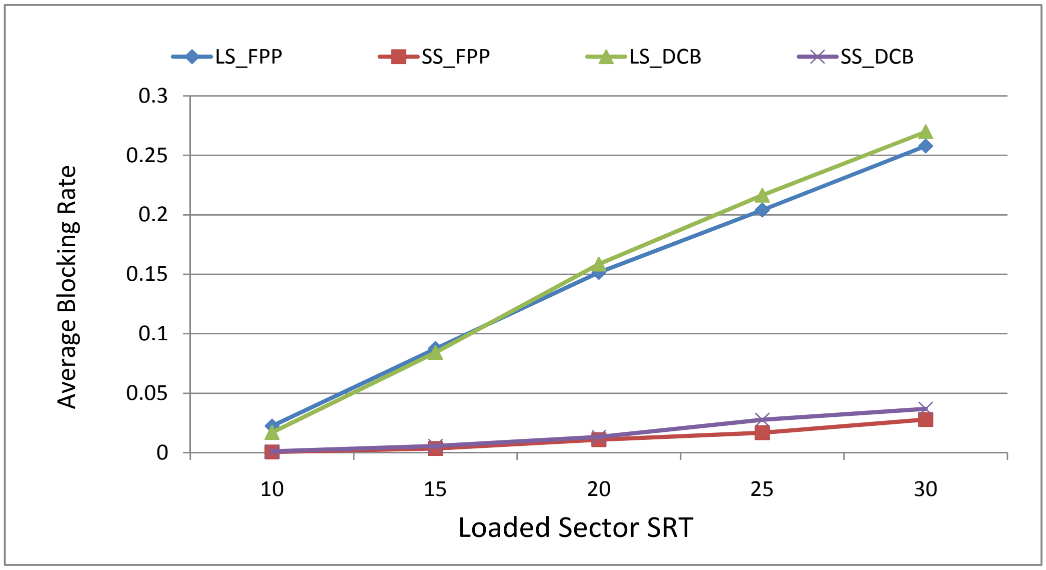


Figure 5.9: Average call blocking rate for different LS' SRT values and fixed SS' SRT value

With respect to the call dropping rate, it is significantly improved under the *PDCB-CC&LB* as shown in Figure 5.10. For the *FPP* scheme, the increase in the call dropping rate is related to the increased number of mobile users at he loaded sector, which increases the loaded sector load. Therefore, as the *DCB* procedure is activated the dropping rate is decreased significantly due to the handoff of nearedge mobile users of the loaded contracted sector towards the supporting lightly loaded expanded sector. Despite their distance increase towards their new sector, these handed off mobile users use lower transmission power because their new sector is not loaded as their former sector. Therefore, the load distribution between the loaded and supporting sectors becomes more balanced and the congestion of the loaded sector is released. The supporting sector's call dropping rate increase is due to the sector's load increase caused by the handed off mobile users from the loaded sector. The figure shows that this increase is not significant as compared to the call dropping rate decrease of the loaded sector.

Figures 5.11 and 5.12 respectively show the call blocking and dropping rates when the rate adaptation algorithm is enabled. As can be seen from the figures, the call blocking and dropping rates are maintained at lower values with respect to the values in Figures 5.9 and 5.10. On the other hand, the lower values of the blocking and dropping rates come at the cost of low users transmission rates, which in multimedia systems decrease mobile users' satisfaction and service providers' profitability. Therefore, to guarantee a minimum mobile users' transmission rate which maintains their provisioned *QoS* parameters a minimum mobile users' transmission rate level should be exercised.

The *PDCB* scheme adapts the coverage areas of the loaded and supporting sectors

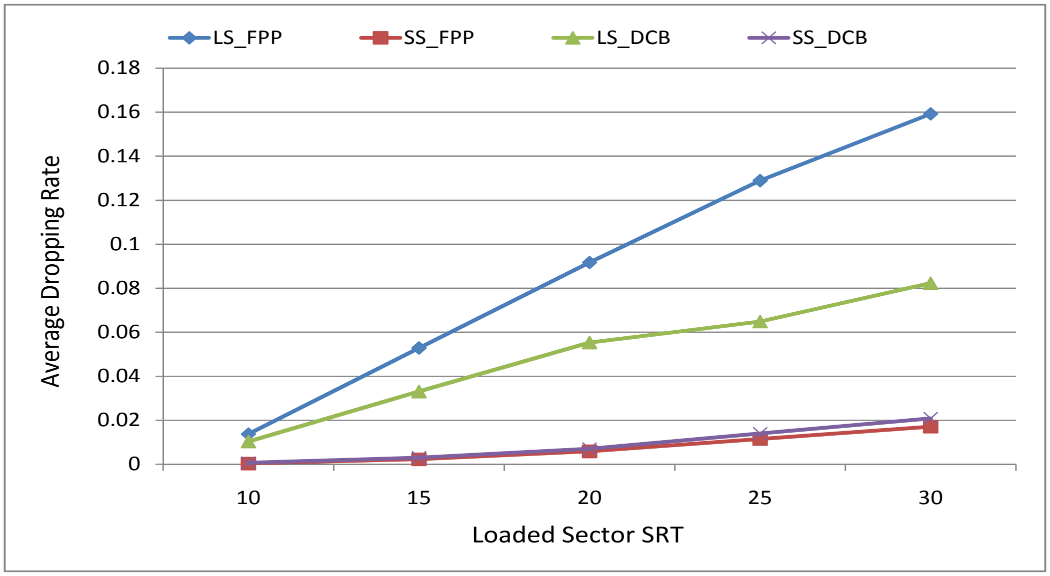


Figure 5.10: Average call dropping rate for different LS' SRT values and fixed SS' SRT value

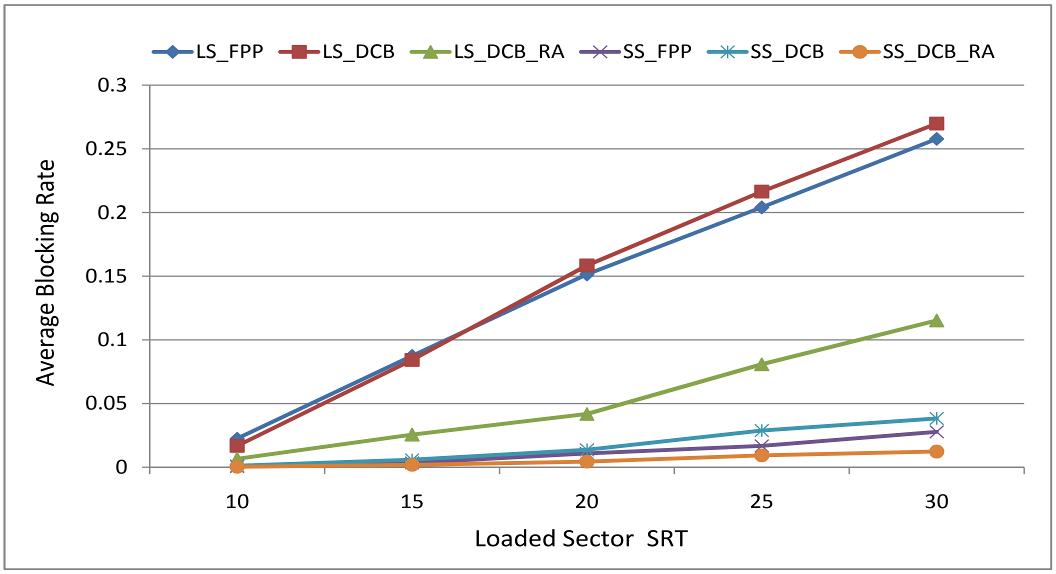


Figure 5.11: Average call blocking rate for different LS' SRT values and fixed SS' SRT value - Rate Adaptation Enabled

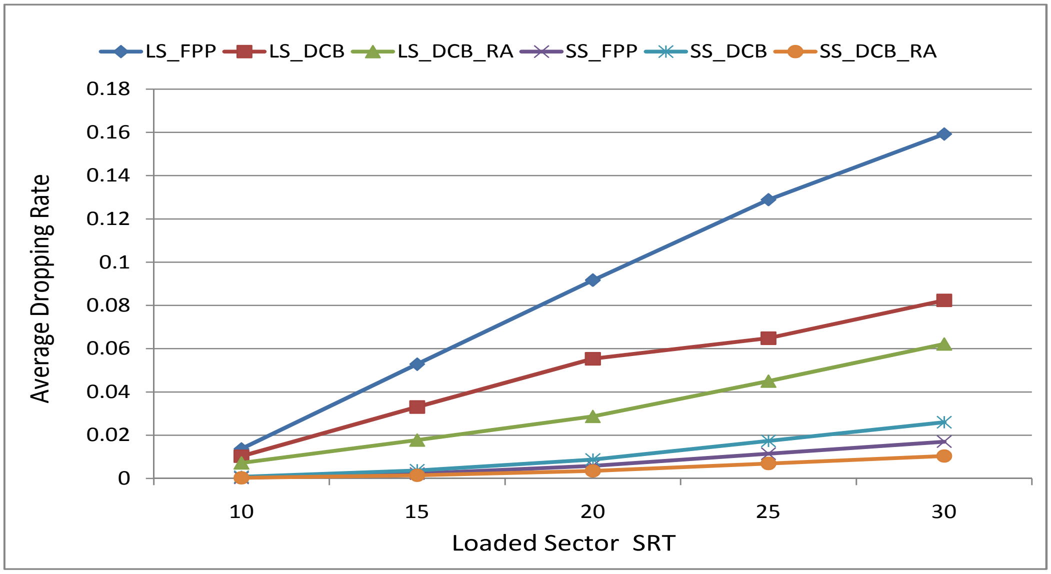


Figure 5.12: Average call dropping rate for different LS' SRT values and fixed SS' SRT value - Rate Adaptation Enabled

to sustain the spatially increased traffic load. The combined coverage adaptation of the loaded and supporting sectors is shown in Figure 5.13. The figure shows that the coverage of the loaded sector decreases as its load increases, while the coverage of the supporting sector increases to acquire the loaded sector nearedge mobile users.

Delay-Based Hotspot Cell and Nearly Loaded Supporting Cells

In this scenario, the ability of a supporting sector to provide support to a loaded sector is examined. The loaded sector load is maintained at a higher load level while the supporting sector load is gradually increased. To achieve this system load scenario, the center cell loaded sectors' SRT parameters are assigned a fixed value of 30s, while the SRT parameters of non-support sectors are also assigned fixed value of 5s for the whole simulation time. On the other hand, the supporting sectors' SRT parameters are increased from 10s to 30s in steps of 5s. The rate adaptation scheme is not applied

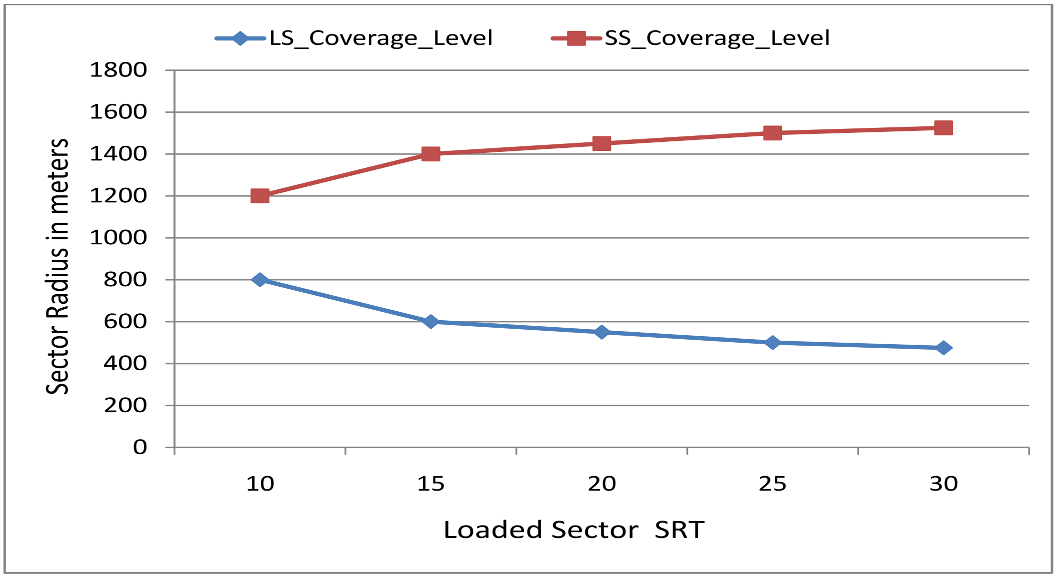


Figure 5.13: Combined Coverage Adaptation of Loaded and Supporting Sectors

in this scenario. The results with respect to only one loaded sector and its nearby supporting sector are shown in Figures 5.14 and 5.15.

The system behavior as shown in Figure 5.14 explains the effect of increasing the supporting sector load on the call blocking rate of both loaded and supporting sectors. As the *SRT* parameter value of the supporting sector increases, the mobile users' handoff rate out of this sector is decreased due to the increase average sector residence time. Hence, the supporting sector load is increased which decreases the supporting level provided to the loaded sector. The interesting phenomena here is that as the *SRT* value increases the loaded sector call blocking rate of *FPP* and *PDCB-CC&LB* schemes decreases while the call blocking rate for the supporting sector increases. This can be explained as follows. As the average *SRT* value of a sector is increased, mobile users spend more time in this sector before handing off towards another sectors. Therefore, the load of the supporting sector is increased since its *SRT* parameter is gradually increased. This leads to an increase of the

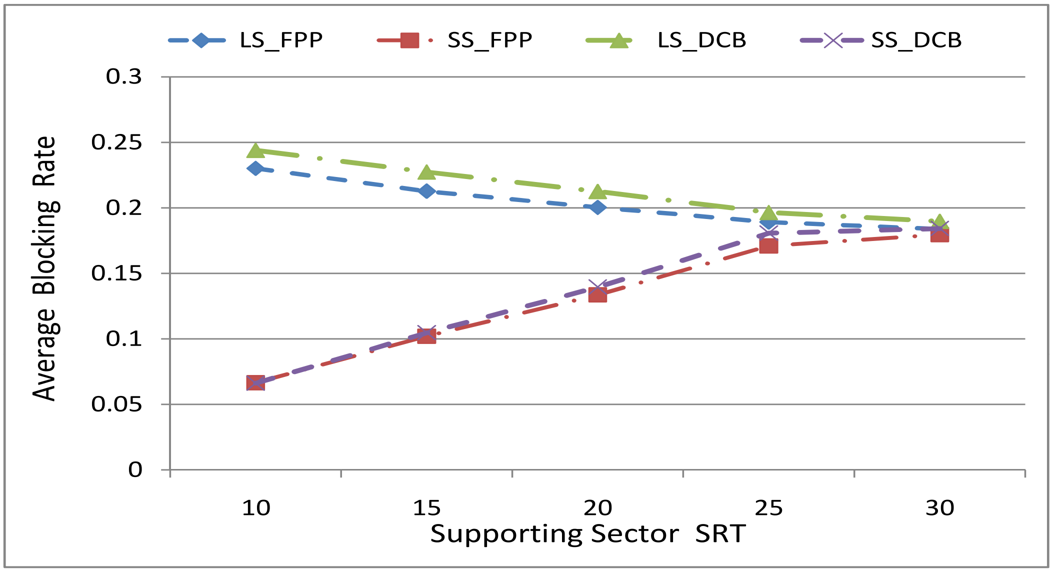


Figure 5.14: Average call blocking rate for fixed LS' SRT value and different SS' SRT values

blocking and dropping rate at the supporting sector while such parameters' values are decreased in the loaded sector. At the point where the *SRT* parameters have the same values at the loaded and supporting sectors, the blocking rate becomes comparable for both sectors and the *PDCB-CC&LB* scheme converges to the *FPP* protocol scheme as no support can be provided by any sector.

Also, the call dropping rate for both schemes is affected as the supporting sector load increases, as seen in Figure 5.15. As the figure shows, the call dropping rate for the loaded sector is decreased while it is increased for the supporting sector. As the load of both sectors become nearly balanced, their dropping rates converge to nearly equal vales.

Therefore, the system behavior depends on the load level at the loaded and supporting sectors. When a sector is loaded, the expected support depends on the load

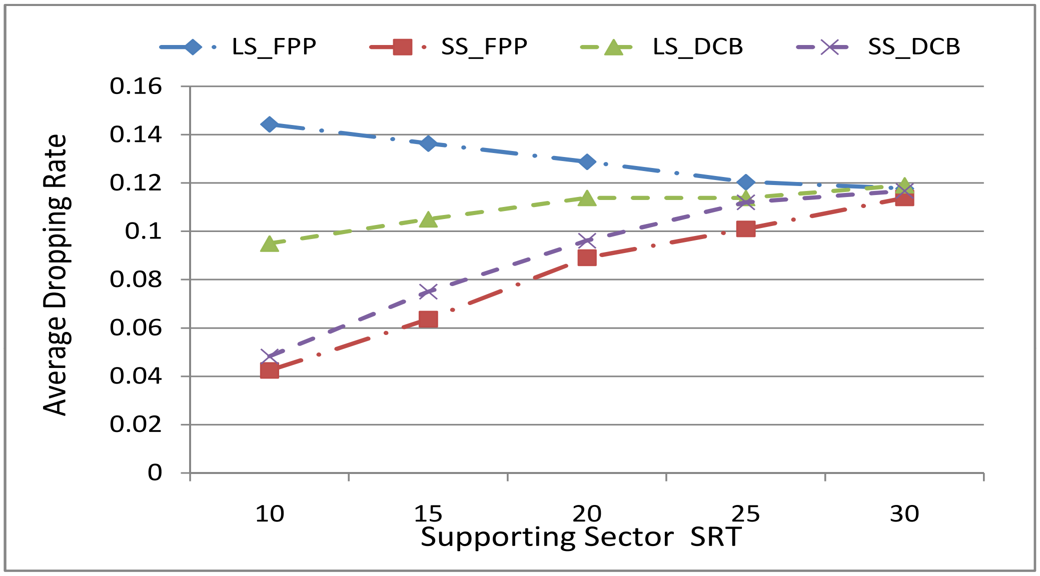


Figure 5.15: Average call dropping rate for fixed LS' SRT value and different SS' SRT values

level at its nearby sector; the support level increases as the supporting sector load decreases. Therefore, the *RDCB* and *PDCB* schemes effectively manages the spatially localized traffic in *DCB-enabled UMTS* networks.

In a *UMTS* system, the online evaluation of the proposed *DCB* module is computationally expensive. Lookup tables can be used to speedup the execution of the proposed heuristic schemes. For example, these Lookup tables entries may compose of different levels of the average uplink load factor of every two nearby sectors along with the candidate coverage levels to be activated by each sector. Hence, when the nearby sectors' computed load values match the Lookup table values the associated coverage levels of the nearby sectors are activated. Therefore, the heuristic schemes execution time can be dramatically decreased which will increase the efficiency of the proposed *DCB* mechanism.

5.6 Summary

A novel congestion control and load balancing mechanism for interference-limited *WCDMA* cellular networks is proposed in this chapter. The mechanism is based on the concept of Directional Cell Breathing *DCB* and it is called the Directional Cell Breathing based Congestion Control and Load Balancing *(DCB-CC&LB)* protocol. Two schemes are proposed to evaluate the performance of the *DCB-CC&LB* protocol, namely reactive and proactive *DCB-CC&LB*. The reactive scheme considers no mobile users' mobility and utilizes a measurement-based approach to invoke the *DCB* technique. On the other hand, the proactive scheme is a parameter-based approach which measures the average sectors' congestion parameters for a predefined period of time before it examines a *DCB* trigger parameter. Results demonstrate the ability of the proposed scheme to manage unbalanced load scenarios in situation-aware *WCDMA* systems. As sector congestion becomes severe and the support level from the supporting sector or the load level in the supporting sector are increased, the call blocking and dropping rates, and mobile users transmission powers are increased. Therefore, there should be a limit on the maximum support that can be given to a loaded sector after which the classical radio resource management schemes can be used instead.

Chapter 6

DCB-Based QoS Provisioning

The support of quality of service *(QoS)* is a key component for the success of *BWNs*. Therefore, a number of *QoS* classes were defined for *UMTS* networks, where each class contains a set *QoS* requirements that define its traffic. *UMTS QoS* classes were briefly explained in Section 2.1. Real-time traffic requires strict delay requirements but it is tolerant to some bit error rate *(BER)*. Therefore, transmission power control is utilized to maintain the required *QoS* parameters at or above the provisioned limits. Since non-realtime traffic is tolerant to transmission delay but is restricted on *BER*, transmission rate adaptation of non-realtime traffic is used to cope with congestion scenarios in *BWNs*. Therefore, combining the power control and rate adaptation mechanisms with *DCB* coverage adaptation can further enhance the overall performance of the interference limited wireless multimedia systems.

The benefits of coverage adaptation are twofold: First, reducing the loaded sector coverage area increases its capacity, which prevents a severe transmission rate degradation to its mobile users. Secondly, since the sector is lightly loaded, the handoff mobile users towards the supporting sector will use less power to communicate with their new sector. Also, the variations in sectors' coverage areas and mobile users' transmission rates increase the processing gain of each mobile user, which is the ratio of the system bandwidth W to the mobile transmission rate R. Such gain is beneficial to the signal recovery at the receiver station.

In this chapter, we expand the *DCB* module proposed in Chapter 4 to accommodate the heterogeneous nature of traffic in future *BWNs*. First; a *Power-Controlled Rate and Coverage Adaptation (PCRCA)* module for a single class multi-rate traffic is proposed to maximize the number admitted mobile users to the system, assuring their *QoS* requirements, and balancing the network load through transmission power control, and transmission rate and cell sectors' coverage adaptation. A heuristic algorithm is implemented to solve an optimization model of the proposed *PCRCA* scheme. Then, a derivation of mathematical formulas to involve more realistic considerations for inter-cell interference in a *DCB-Enabled* multimedia system is derived . The model is used to quantify the degradation level in terms of *QoS* support of low priority traffic to preserve the *QoS* level of high priority traffic. Achievable data rates and *BER* are determined for every possible nearby cell sectors' coverage combination. The effect of rate and coverage adaptation on the transmission powers of mobile users is also analyzed.

The chapter is organized as follows. The *PCRCA* model and its performance evaluation is detailed in Section 6.1. The *DCB-QoS Provisioning* module is explained in Section 6.2. The performance evaluation of the *DCB-QoS Provisioning* module is given in Section 6.3. The chapter is summarized in Section 6.4.

6.1 Power-Controlled Rate and Coverage Adaptation Module

The Power-Controlled Rate and Coverage Adaptation *(PCRCA)* module is an adaptation mechanism which searches for optimized coverage of nearby sectors of two adjacent cells for which mobile users' transmission powers are minimized, their transmission rates are optimally allocated, and network congestion is released in both sectors. Although, it can be easily extended for a general coverage adaptation scheme, herein, the proposed module utilizes the *DCB* coverage adaptation scheme. Moreover, the *PCRCA* module is generalized to work with any power or rate adaptation schemes. The proposed module and its optimization model are detailed in Section 6.1.1. The heuristic algorithm for solving the optimization model is provided in Section 6.1.2.

6.1.1 PCRCA Optimization Model

Given the traffic of two nearby sectors, the *PCRCA* module optimally and dynamically determines their optimized coverage levels with respect to the optimal transmission powers and rates of their existing mobile users. Achieving this goal leads to system congestion release and traffic load balancing while maximizing the utilization of the network resources. The optimization of the *PCRCA* module is defined in two steps; namely *Local* and *Global*. These two steps are detailed next.

Local Optimization

With respect to *PCRCA* local optimization, given N mobile users distributed over a given sector q of service area A_q , we define a function w_{A_q} which optimally allocates transmission powers and rates for such mobile users as follows:

$$
\begin{aligned}
w_{A_q} &= \max_k \sum_{n=1}^{N_{A_q}} f\left(p_n, \gamma^k\right) \qquad (6.1)\\
\text{s.t.}& \\
& 0 \leq p_n \leq P_{max} \\
& \gamma^1 \leq \gamma^k \leq \gamma^K \qquad k \in (1, 2 \ldots K)
\end{aligned}
$$

where N_{A_q} is the number of mobile users in sector q given that its current serving area is A_q, and $f(.,.)$ is a power controlled rate adaptation scheme which is a function of mobile users' transmission powers and their target *SIR*(γ^k). The outcomes of $f(.,.)$ are the optimal power and rate vectors to the mobile users of sector q. Herein, since A_q is dynamic and depends on the load of sector q as well as nearby sector $\hat{q}$, the values of N_{A_q} vary and so do their power and rate allocations. This approach is only performed locally for every cell sector in the network.

Global Optimization

The global module of *PCRCA* searches the optimal coverage of nearby sectors which gives the optimal joint allocation for transmission powers and rates for mobile users in nearby sectors of a *WCDMA* network. Recall that, the service area of each cell sector is partitioned into L concentric coverage levels. These L coverage levels define L service areas in each sector. Therefore, the number of users in each sector depends on its activated coverage area. Then, the *DCB* module is used to vary the coverage levels of the nearby sectors to balance unevenly distributed traffic, and hence release system congestion.

For two given loaded and supporting nearby sectors, q and $\hat{q}$, the objective of

PCRCA-G is to maximize the support level given to the loaded sector q by the supporting sector $\hat{q}$. Therefore, *PCRCA-G* proposes coverage levels combination for sectors q and $\hat{q}$. Then, *PCRCA-L* locally searches for optimal resource allocation to mobile users of sectors q and $\hat{q}$. After evaluating all possible combinations, the optimal coverage combination, which maintains predefined system objectives, is activated. The *PCRCA-G* module is represented mathematically in (6.2):

$$F = \max_{k,sl} \sum_{q,\hat{q}} \left\{ \sum_{sl=0}^{L} \left(w_{A_q} + w_{A_{\hat{q}}} \right) \cdot f(sl) \right\} \tag{6.2}$$

s.t.

$$0 \leq L \leq L_{max}$$

$$f(sl) \leq 1$$

where $f(sl) = sl_0 * b_0 + sl_1 * b_1 + \ldots + sl_L * b_L$ and $b_{sl} = 1$ only for the supporting level which is currently activated. In cases where no possible support can be provided, all b_{sl} will be assigned a binary value of "0". Note that the optimal solution of this scheme is computationally expensive. Therefore, a heuristic solution is devised and is detailed herein.

6.1.2 PCRCA Heuristic Algorithm

A *PCRCA* iterative algorithm is proposed to heuristically evaluate the proposed local and global optimization modules. The scheme efficiently allocates mobile users' transmission power and transmission rate with respect to possible coverage combinations for every two nearby sectors. Then, the coverage combination which provides maximum support to the loaded sector is activated. The *PCRCA* is mainly composed of three procedures, namely: network initialization procedure; *Initialization()*, local *PCRCA* procedure; *PCRCA-L()*, and global *PCRCA* procedure; *PCRCA-G()*. The detailed functionalities of these procedures are explained in the following subsections.

Initialization()

The *Initialization* procedure oversees the settings of the number of algorithm iterations *(I)*, and the initial settings of network parameters, such as base stations X_m and Y_m locations, their provisioned maximum transmission powers P_m, and the initial coverage level l of each cell sector. Also, the initial settings of mobile users' parameters are performed by this procedure. Such settings are mobile users' distance $d_n^{m,q}$ to their base stations, maximum transmission power P_n^{max}, initial transmission and receiving powers, $P_{tx}^{m,q,n}$ and $P_{rx}^{m,q,n}$ respectively, and initial transmission rates r_n. The pseudocode of this procedure is shown in Algorithm 5

Algorithm 5 PCRCA Initialization Algorithm

```
procedure Initialization(M, Q, N)
    $I \leftarrow MaxNumberOfIterations$                    ▷ Number of iterations
    for $m \leftarrow 1, M$ do
        for $q \leftarrow 1, Q$ do
            InitializeNetworkParameters(m, q)
        end for
    end for
    for $m \leftarrow 1, M$ do
        for $q \leftarrow 1, Q$ do
            for $n \leftarrow 1, N_{A_q}$ do
                InitializeMobileUserParameters(m, q, n)
            end for
        end for
    end forreturn (0)
end procedure
```

PCRCA-L()

After the initialization of mobile users and network parameters, the *PCRCA-L()* procedure shown in Algorithm 6 is locally invoked with respect to every loaded cell sector$q$ and its nearby supporting sector $\hat{q}$. In this procedure, transmission powers and rates parameters of mobile users in the loaded and supporting sectors are allocated using the power controlled rate adaptation module defined in Equation (6.1). The *PCRCA-L()* iteratively evaluates the possible mobile users' transmission powers and rates allocations with respect to every possible sector coverage level defined by currently activated coverage area A_q and $A_{\hat{q}}$ of the loaded and supporting sectors respectively. If, for a mobile user, no such feasible allocation exists, it is turned off and an outage is reported with respect to the evaluated coverage level. Finally, the allocated transmission powers and rates, and the *SIR* values for active mobile users, as well as the number of outages are reported with respect to every coverage combination.

Algorithm 6 PCRCA-L Algorithm

```
procedure PCRCA-L(q, $\hat{q}$)
    for $i \leftarrow 1, I$ do
        $w_{A_q}()$
    end for
    for $i \leftarrow 1, I$ do
        $w_{A_{\hat{q}}}()$
    end forreturn ($w_{A_q}()$ and $w_{A_{\hat{q}}}()$)
end procedure
```

PCRCA-G()

The execution of *PCRCA-L()* is based on the nearby sectors' coverage levels allocated by the *PCRCA-G()* procedure, as shown in Algorithm 7. Therefore, whenever the *PCRCA* is invoked, the *PCRCA-G()* procedure evaluates all possible coverage scenarios of the loaded and supporting sectors using the *PCRCA-L()* procedure. Then, Equation (6.2) is used to evaluate the obtained results of the *PCRCA-L()* procedure and then activate the coverage combination of the best obtained combined results of the two nearby sectors. The main *PCRCA* procedure is sketched in Algorithm 8.

Algorithm 7 PCRCA-G Algorithm

```
procedure PCRCA-G(q, $\hat{q}$)
    for $sl \leftarrow 1, L_{Max}$ do
        $PCRCA - L(q, \hat{q})$
    end for
    $F = \max_{k,sl} \sum_{q,\hat{q}} \left\{ \sum_{sl=0}^{L_{Max}} \left( w_{A_q} + w_{A_{\hat{q}}} \right) \cdot f(sl) \right\}$
     return (F)
end procedure
```

Algorithm 8 PCRCA Algorithm

```
procedure Main
    $Initialization()$
    for $m \leftarrow 1, M$ do
        for $q \leftarrow 1, Q$ do
            $PCRCA - G(q, \hat{q})$
        end for
    end forreturn (0)
end procedure
```

6.1.3 Transmission Power and Transmission Rate Control

The proposed *PCRCA* algorithm requires power and rate control algorithms to efficiently allocate the limited network resources to the active mobile users in the system. Such an allocation needs to maintain the interference level in the system at the minimum level possible. Therefore, in a congested system mobile users are allowed to use only the minimum transmission power required for achieving their required *QoS* parameters.

A distributed power control and rate adaptation algorithm, known as Selective Power Control *(SPC)*, is proposed in [36]. It is an iterative algorithm which determines the next transmission power used by mobile user n based on the power vector *P(i)* of iteration i, γ^k target, where k is the transmission rate level utilized by mobile user n, and $\gamma_n(P(i))$ is the current *SIR* parameter value of mobile user n. The mobile user's transmission power is constrained by lower and upper values

$$p_n(i+1) = \max_k \left(\frac{p_n(i) \cdot \gamma^k}{\gamma_n(P(i))} \cdot I_{\left(0 < \frac{p_n(i) \cdot \gamma^k}{\gamma_n(P(i))} \leq p_{max}\right)} \right) \tag{6.3}$$

where $p_n(i+1)$ is the computed transmission power for mobile user n in iteration *i+1*, $I_{(E)}$ is the indicator function of the event *E*.

In this chapter we utilize the *SPC* algorithm in the *PCRCA* module for controlling mobile users' transmission powers and transmission rates keeping in mind the system load conditions.

6.1.4 Simulation Model

The simulation model used to evaluate the performance of a *DCB-enabled UMTS* system implementing the proposed *PCRCA* module is similar to the one in Section 5.5.1. The proposed *RDCB* scheme in Section 5.4 is extended to consider multi-rate traffic and iterative *SPC* transmission power and rate control algorithms.

Simulation Setup

We extended the *UMTS* network simulator used in Chapter 5 to evaluate the performance of the *PCRCA* module. In addition to the network parameters shown in Table 5.1, we define the parameters of Mobile users' transmission rate levels; K, maximum transmission rate; R_{max}, mobile user's used transmission rate; r_m^k, and the target *SIR*; γ_m^k along with their used values in Table 6.1. Initially, the coverage areas of network sectors are equal to 20 coverage levels each, and mobile users are associated with their nearest base station.

Table 6.1: PCRCA-enabled UMTS Network Simulator Parameters

Network Parameters	**Parameters Values**
Number of iterations/Simulation run (I)	500
UE Parameters	**Parameters Values**
E_b/N_o (Γ)	5 dB
Number of transmission rate levels (K)	4
Max transmission rate (R_{Max})	25 kb/s
Mobile user transmission rate (r_m^k)	$(R_{Max} \cdot \frac{1}{2^{k-1}})$
Mobile user target SIR (γ_m^k)	$\frac{(r_m^k \cdot \Gamma)}{W}$

Mobile users are uniformly distributed over the cell sectors but their distribution varies from one sector to another for the purpose of creating traffic hotspots. Mobility is not considered in this simulation. Only a single class of service is considered and each mobile user n utilizes the transmission rate r_n^k where $k \in \{1, 2, 3, 4\}$. For each utilized r_n^k the mobile user received signal has to meet the $SIR_n \geq \gamma_n^k$ constraint to be decoded properly. The maximum utilized transmission rate is 25 kb/s which corresponds to $k = 1$. Each mobile user n is initially allocated the maximum transmission rate R_{Max} and its initial transmission power is a randomly selected value within the range of $P_{Min} \leq p_n \leq P_{Max}$. The number of *SPC* iterations, *I*, is 500 and the presented results are the average of 100 simulation runs of the *PCRCA* heuristic algorithm. The results are compared to the results of a *WCMDA* system of fixed pilot power *(FPP)* in which the coverage area of the cell sectors remains fixed at 20 coverage levels during the simulation time.

Simulation Scenarios and Numerical Results

A delay-based hotspot [34] is modeled to evaluate the performance of the proposed *PCRCA* module. The center cell of the 7-cell network model is the hotspot cell while the surrounding cells are the supporting ones. The following simulation scenarios are used to evaluate the performance of the *PCRCA* module. In each scenario, *N* mobile users are distributed over each cell sector. The transmission powers, transmission rates, and the received *SNR* for mobile users in every two nearby sectors are iteratively evaluated for every possible coverage combination. If no feasible values exist for a mobile user, its state is set to inactive and an outage is reported.

Scenario 1: Hotspot Sector and Lightly Loaded Supporting Sector

In this scenario, the number of active mobile users in the loaded sector is varied from 40-80 users in steps of 10 users, while the number of mobile users in the supporting

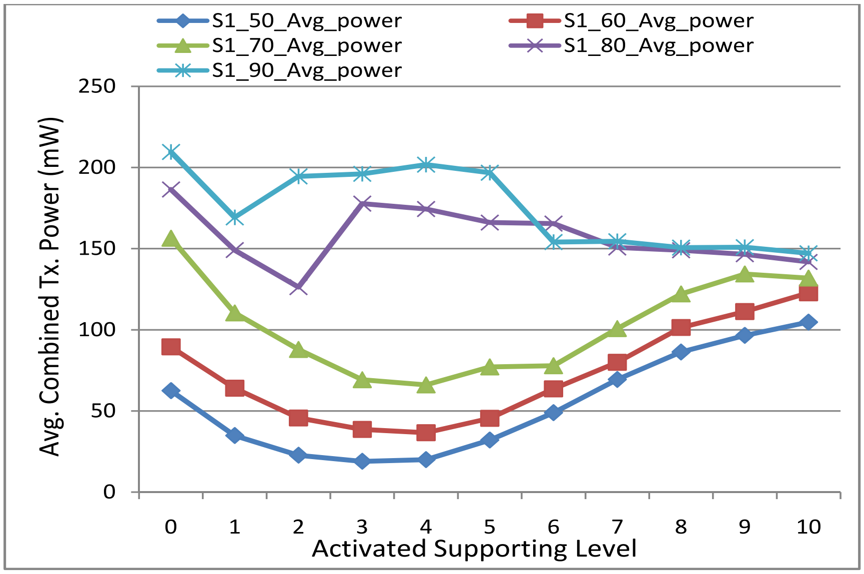


Figure 6.1: Scenario 1: Combined Average Transmission Power

sector as well as other network sectors is maintained fixed at 10 mobile users per sector. The average combined results of transmission power, transmission rate, and outage ratio with respect to every coverage combination of the loaded and supporting sectors are shown in Figures 6.1-6.3.

Average Combined Transmission Power

The average combined transmission power of mobile users of the loaded and supporting sectors varies from one coverage level to another[1] as shown in Figure 6.1[2]. First, the average combined transmission power for the *FPP* algorithm has a large value due to the increased load in the loaded sector which require the mobile users to increase their transmission power. Such increases in transmission power also affect the mobile users in the adjacent cells.

---

[1] In this figure and subsequent ones, supporting level"0" results correspond to the *FPP* algorithm.

[2] S1, and later S2, shown in the results figures correspond to simulation scenarios 1 and 2, respectively.

When the *PCRCA* algorithm is invoked, the average combined transmission power is decreased due to the handoff of nearedge mobile users of the loaded sector towards the expanded nearby lightly loaded sector. Therefore, the intra cell interference of the loaded sector is decreased, which leads to a decrease in the transmission power of its mobile users. Also, since the nearby sector is lightly loaded, the handed off mobile users require less transmission power to communicate with their new base station then with the loaded sector even if the communication distance is increased.

Given the total number of mobile users (80 and 90) in the nearby sectors, as the support level is increased beyond level 1 and 2, the average combined transmission power is dramatically increased. This is because of the increased number of handed off mobile users and the high load of their corresponding sector. Therefore the handed off mobile users are required to increase their transmission power to compensate for interference from the loaded sector and the increased distance to their new sector. As the support level is further increased, some mobile users exceed their transmission power budget so they are turned off and the average transmission power is decreased due to those turned off mobile users.

For the other combination of mobile users (50, 60, and 70), the system is not overloaded. Therefore, as the support level is increased traffic becomes more balanced and the loaded sector load is decreased. As Figure 6.1 shows, the average combined transmission power is decreased due to the decrease in intra-cell interference of the loaded sector. The decrease in the mobile users' transmission power in the loaded sector results in a capacity gain in the loaded and supporting sectors. Therefore, despite the distance increase to their new sector, the handed off mobile users require lower transmission power then was required when they were associated with the loaded sector. As the support level is increased beyond supporting level 5, the inter cell interference imposed on the loaded sector by the handed off mobile users becomes dominant and leads to an increase of the transmission power of the mobile users in both loaded and supporting sectors. Therefore, as the support level is increased, some mobile users, especially the handed off ones, exceed their maximum transmission power and are turned off. This can be concluded from the decrease in the mobile users average transmission power in Figure 6.1.

Average Combined Transmission Rate

The average combined transmission rate of the loaded and supporting sectors depends on the load level in the loaded sector and varies from one supporting level to another (See Figure 6.2). For the overload scenarios of 80 and 90 mobile users cases, the transmission rate is maintained almost at the same level for the *FPP* and the *PCRCA* up to supporting level 5. Therefore, the increase in transmission power shown in Figure 6.1 is just to maintain the minimum possible transmission rate because of the higher interference level in the system. As the support level exceeds support level 5, the average combined transmission rate is decreased due to the increased mobile users' outage ratio in the system.

On the other hand, the average combined transmission rate of 50, 60, and 70 mobile users' load scenarios is increased as the support level is increased. This can be explained as follows: as the support level is increased, the number of handed off mobile users is gradually increased which smoothes the load balance and permits the existing mobile users in the loaded sector to lower their transmission power. Therefore, the system interference level is decreased, which allows the mobile users to increase their

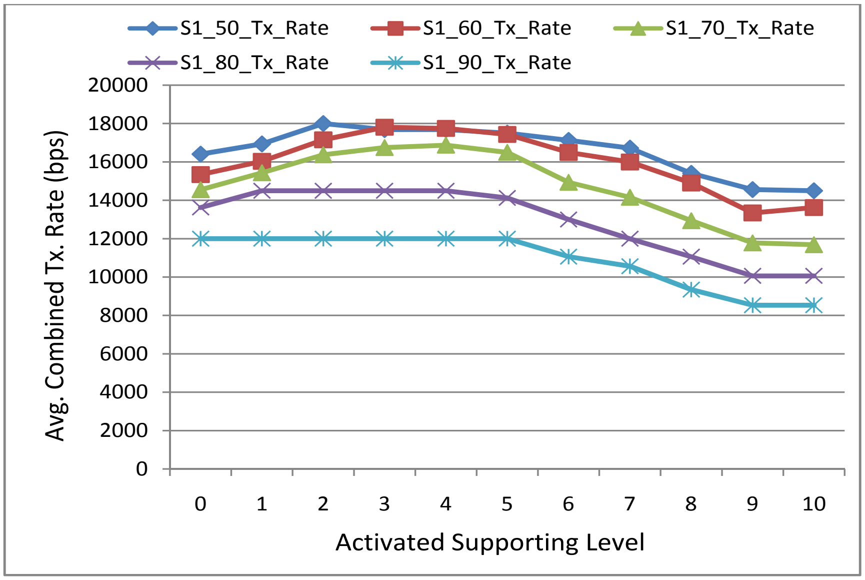


Figure 6.2: Scenario 1: Combined Average Transmission Rate

transmission rates by utilizing the decrease in the mobile users transmission power. Similarly, as the support exceeds level 5, the average combined transmission rate is decreased due to the increased interference level caused by the handed off mobile users.

Average Combined Outage Ratio

The average combined outage ratio against the activated supporting levels is shown in Figure 6.3. Note that, the average outage ratio for the *FPP* has a high value. As the *PCRCA* scheme is invoked, the ratio, however, is decreased for all load scenarios. This ratio is maintained at lower values at the cost of increased transmission power as shown in Figure 6.1. As the support level exceeds level 5, the outage ratio increases sharply for all load scenarios. This increase is because of the increased number of mobile users exceeding their maximum transmission power due to the high interference level in the system. Therefore, mobile users are turned off to give an opportunity to

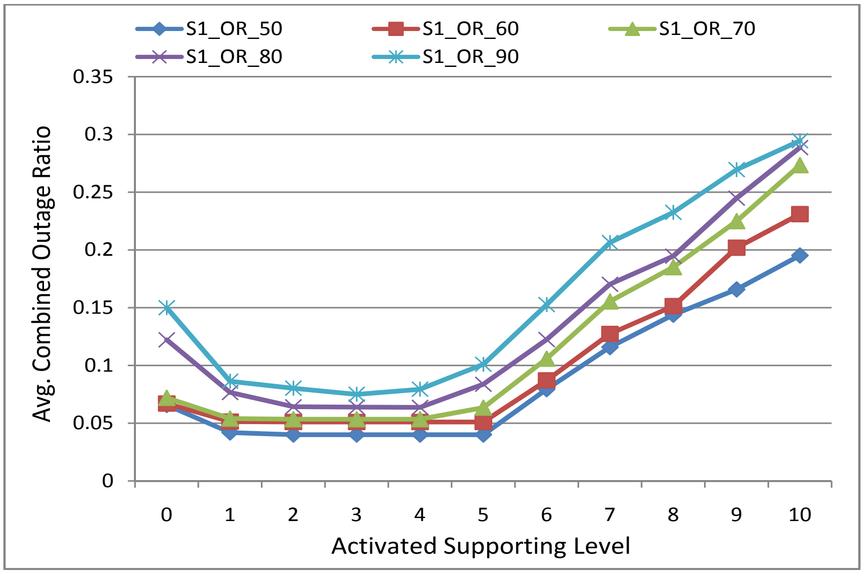


Figure 6.3: Scenario 1: Combined Average Outage Rate

other mobile users to maintain their connectivity.

In this scenario, the load level at the loaded sector has a great impact on the support level given by the supporting sector. In the case when the loaded sector is overloaded, the support given to the loaded sector negatively affects the system performance due to the increased interference level in the system. On the other hand, when the system is loaded, activating certain supporting levels balances the system load, increases the system capacity, and releases the system congestion.

Scenario 2: Hotspot Sector and Nearly Loaded Supporting Sector

In this scenario, the number of mobile users of the loaded sector is maintained at a fixed value of 60 mobile users while the number of mobile users of the supporting sector varies from 10-30 mobile users. For each mobile users' combination, the performance of the *PCRCA* and *FPP* algorithms are evaluated to test the ability of a nearly loaded sector to give support to a loaded sector. The performance results of the *PCRCA*

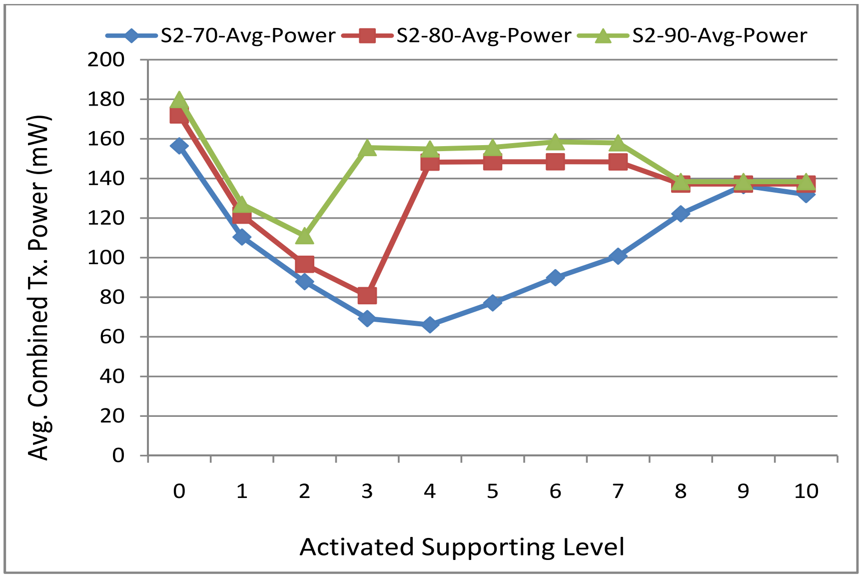


Figure 6.4: Scenario 2: Combined Average Transmission Power

and *FPP* are shown in Figures 6.4-6.6.

Average Combined Transmission Power

Since the number of mobile users in the loaded sector remains fixed, as the support level is increased the intra cell interference is decreased and the average required transmission power of the mobile users in the loaded sector is decreased. Therefore, despite the increased number of mobile users in the normal coverage area of the supporting sector the handed off mobile users require less transmission power. Hence, the average combined transmission power is much lower then in the previous scenario. As the support level is increased, the handed off mobile users increase their transmission power to compensate for their increased distance from their new base station. This leads to high interference levels in the system and drives other mobile users in the loaded and supporting sectors to increase their transmission power. This is shown as an increase in the average combined transmission power in Figure 6.4. When the

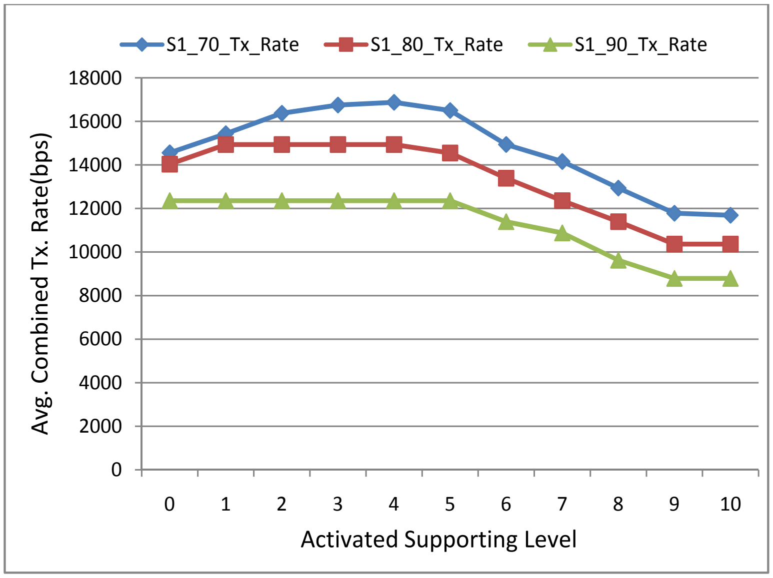


Figure 6.5: Scenario 2: Combined Average Transmission Rate

maximum transmission power is exceeded by a mobile user, it is turned off, hence a decrease in the average combined transmission power is achieved as shown in the figure.

Average Combined Transmission Rate

In comparison to the average combined transmission rate shown in the previous scenario, the average combined transmission rate for the mobile users in both sectors is increased (see Figure 6.5). This is because of the decrease in the average mobile users' transmission power. Similarly, as the support level is increased beyond supporting level 5, the average combined transmission rate is decreased due to the increased interference level in the system.

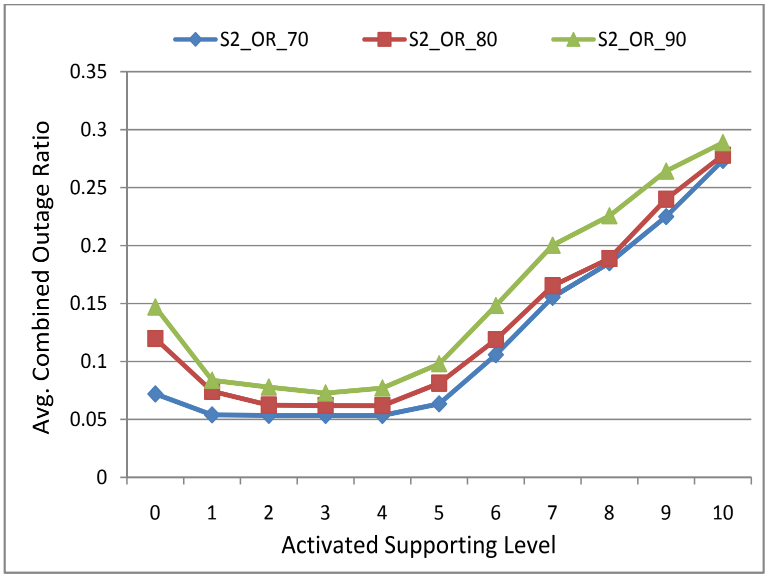


Figure 6.6: Scenario 2: Combined Average Outage Ratio

Average Combined Outage Ratio

The average combined outage ratio of the loaded and supporting sectors is also decreased because of the decreased interference level in the system. As Figure 6.6 shows, the interference level suppression gives mobile users better chance to increase their transmission power to maintain their maximum possible transmission rates. When the given support exceeds level 5, the average combined outage ratio is increased since the mobile users exceeding their maximum transmission power are turned off.

In this scenario, the ability of a nearly loaded sector to support a loaded sector is examined. Results indicate that the supporting sector has the ability to support a nearby loaded sector to a certain extent. Since the loaded sector load is initially maintained fixed as the support level is increased the average transmission power is decreased due to the decrease of intra cell interference level in the loaded sector. Moreover, the traffic of the supporting sector is mostly located in its normal coverage area which requires less transmission power. This leads to an increase in the mobile users' transmission rate and a decrease in the mobile users' outage ratio. As the support level is increased, most supporting sector traffic becomes located near the sector edge which requires more transmission power and drives other mobile users of the loaded and supporting sectors to increase their transmission power. Hence, the transmission rate is decreased and the mobile users outage ratio is increased.

6.2 QoS provisioning in DCB-Enabled WCDMA Systems

The ability of a *DCB-Enabled WCDMA* system in providing multimedia services is examined in this section. The mobile users' signal quality is defined by their received E_b/I_o, which needs to be maintained above a certain value for correct signal decoding by the receiver's electronics. In this system, the E_b/I_o of calls from multiple classes is defined by:

$$\left(\frac{E_b}{I_o}\right)_C^l = \frac{WP_{r,C}^l/R_C}{P_{tot}^l - P_{r,C}^l} \tag{6.4}$$

where $\left(\frac{E_b}{I_o}\right)_C^l$, $P_{r,C}^l$, R_C are the bit energy to interference ratio, the received power and transmission rate of a mobile user of class C respectively, and P_{tot}^l is the total received power. All of these values are with respect to the coverage level l[3] of the serving sector. Therefore, *QoS* provisioning requires the E_b/I_o of all calls of a certain class C to be higher than or equal to the threshold value τ_C^{th}. To maintain such a requirement in a dynamically configured system, certain conditions such as maximum

---

[3] The coverage level l of the serving sector implies the complement coverage level $\hat{l}$ in the nearby sector

transmission power, transmission rate, maximum sector coverage level, and maximum received interference power have to be satisfied. Equation (4.5) is extended to support multiclass traffic as follows:

$$\begin{aligned} \left(\frac{E_b}{I_o}\right)_C^l &= \tau_C \geq \tau_C^{th} && (6.5)\\ R_C &\geq R_C^{th} && (6.6)\\ P_{r,C}^l &\geq P_r^{min} && (6.7)\\ L_{min} &\leq l \leq L_{max} && (6.8)\\ C &= 1, 2, \ldots \end{aligned}$$

where R_C^{th} is the minimum acceptable transmission rate of class C calls, P_r^{min} is the minimum received power of any call which is required by the receiver circuits to decode the received signal properly, l is the current coverage level of a cell sector, and L_{min} and L_{max} are the minimum and maximum supporting levels, respectively. These constraints need to be maintained to guarantee the requested *QoS* requirements by every call of each class.

In *UMTS* multimedia system design, the *QoS* requirements of calls need to be handled in a joint manner to maintain priority and fairness among calls of different classes. Therefore, the *QoS* constraints need to be utilized to provide the required *QoS* level of each individual call. By examining Equation (6.5), all of its parameters are defined by the required *QoS* level of each class except P_{tot}^l, which is a function of the received power of other mobile users inside and outside the corresponding cell. Such power is classified into intra- and inter-cell interference in the wireless communication terminology. Therefore, P_{tot}^l can be defined as:

$$P^l_{tot} = \sum_{C=1}^{C_{max}} N_C I_C + I_{iner} \tag{6.9}$$

where N_C and I_C are respectively the class C total number of calls and their average received power at the corresponding cell, I_{inter} is the inter-cell interference power and C_{max} is the maximum number of classes supported in the system.

In this analysis, the inter-cell interference from calls of different classes is considered. Therefore, the number of mobile users of each class in every cell is quantified and is taken into consideration in the calculations of different *QoS* values. Therefore, with respect to the combined coverage levels of every two nearby sectors, the denominator of Equation (6.4) is redefined to consider the actual number of mobile users of each class in every sector as shown in Equation (6.10).

$$P^l_{tot} - P^l_{r,C} = P^l_{r,C}\left[(N_C - 1) + \sum_{m=1}^{M} N^m_C I^m_C\right] + \sum_{i \neq C}^{C_{max}} P^l_{r,i}\left[N_i + \sum_{m=1}^{M} N^m_i I^m_i\right] \tag{6.10}$$

where N_C is the number of calls from the same class in the same cell, N^m_C and I^m_C are the number of calls and the average inter-cell interference of a single call of class C from the interfering cell m respectively, N_i represents the number of calls from class i in the same cell, and N^m_i and I^m_i are the number of calls of class i and the average inter-cell interference of a single call from class i residing in the interfering cell m respectively. Since perfect power control is assumed, the received powers from mobile users of the same class in the same cell are equal. The mobile users' received powers are multiple of the minimum required received power P^{min}_r. The average inter-cell interference (I^m_x) of a single call of class x in a neighboring cell m for different coverage levels of *WCDMA* nearby sectors followed by the system capacity quantification has been introduced in Section 4.3.

From Equation (6.10), for a given combination of the number of calls from each class in every cell, the *QoS* constraints can be satisfied for every call by adjusting the sectors' coverage levels, varying the required received signal power $P^{l}_{r,C}$, adapting mobile users' transmission rates R_C and/or varying the E_b/I_o class threshold, τ_C, of such calls. The following section presents a numerical analysis of a multiclass *WCDMA* system utilizing two traffic classes, *C1 and C2*.

6.3 Performance of DCB-Based QoS Provisioning

We utilize the mathematical model defined in Equation (6.10) to study the effect of varying mobile users' transmission rates and cell sectors coverage levels in providing the required *QoS* of multimedia calls in *WCDMA* systems. When the network cells' sectors are lightly loaded, their coverage areas are equal and the requested transmission rates by their mobile users are granted. As the load of a sector reaches a saturation point and a hotspot is formed such as the end of a game in a football stadium where call arrival rate suddenly increases, which makes such a load not possible to be accommodated by the base station serving that area, the network can dynamically change the coverage area of the loaded and supporting sectors using the DCB mechanism, and adapt the transmission rates and powers of their mobile users to maintain the required QoS requirements of the ongoing calls.

We study the *QoS* provisioning for two classes; C1 (high priority traffic) and C2 (low priority traffic). The utilized transmission rate of C1 calls is a fixed value which can take any value from 12.2 kbps to 23.2 kbps. Fixing the transmission rate of C1 calls, the adaptable transmission rate of low priority C2 calls varies between 0 kbps and 12.2 kbps. We assume that system traffic load is unevenly distributed over the network service area making a sector highly loaded while its nearby sector is lightly loaded. As can be inferred from Equation (6.10), to maintain higher transmission rate and low *BER* of higher class traffic, higher received power is required. Therefore, the capacity of the system is limited by the quality of the higher class traffic. For the rest of this section, we mathematically analyze the effect increasing the transmission rate of higher priority calls on the transmission power of low and high priority calls, transmission rates degradation of lower priority calls, and the degradation level of E_b/I_o for lower class calls. The quantified capacity in Section 4.3 is utilized to define the number of mobile users of each class. The results and their analysis follow.

6.3.1 The Effect of Rate and Coverage adaptation

Herein, we analyze the effect of increasing the transmission rate of *C1* calls on both the transmission rate of *C2* calls and the ratio of the received power of *C1* to that of *C2* for every possible coverage combination of the loaded and supporting sectors. In this analysis, both classes are provisioned with the same *BER* of ($\tau_C = \tau_C^{th} = 5$ dBm). The transmission rate of *C2* is degradable with a threshold value of ($R_2^{th} = 3$ Kbps), while *C1* calls have a fixed transmission rate ($R_1 = R_1^{th}$). The ratio of the number of *C1* calls to that of *C2* calls is varied and has the same value in every sector. Based on this ratio, we change R_1 for every coverage combination of the loaded and supporting sectors and study the transmission rate degradation of *C2* calls in order to maintain the required transmission rate of *C1* calls. Results for the first 4 coverage levels are plotted in Figure 6.7.

The increase in the transmission rate of *C1* calls is coupled with an increase to the received powers from such calls which raises the interference level in the system. Therefore, to be able to decode their signals correctly, the transmission rate of *C2* calls are degraded. This decrease in the transmission rate increases the processing gain to *C2* calls and makes them more tolerant to an increase in interference level. Similar behavior is observed for different ratios of number of *C1* calls to *C2* calls. As shown in Figure 6.7, when the transmission rate of *C1* decreases, the achievable transmission rate of *C2* calls is increased until the curves of all coverage levels intersect at the point where there is no adaptation. At this intersection point, calls from both classes for every coverage level are granted the same transmission rate. On the other hand, as the curves cross the threshold value of *C2* transmission rate which is represented by the dotted line in Figure 6.7, no more degradation is performed. Therefore, this point represents the maximum transmission rate which can be granted to *C1* calls.

The degradation level of the *C2* transmission rate slightly increases as the given support to the loaded sector increases. This is because of the increased number of handed off mobile users towards the expanded supporting sector. These handed off mobile users increase the average transmission power of *C1* calls, which forces *C2* calls to decrease their transmission rates to maintain both classes *QoS* requirements. As the ratio of *C1* calls to *C2* calls decreases, the degradation level of *C2* transmission decreases. The reason behind that is the decrease of the number of higher quality calls which need to increase their transmission power to maintain their *QoS* requirements.

6.3.2 Transmission Power Variation

As mentioned before, in order to maintain higher quality for *C1* calls, their received power ratio to that of *C2* calls has to be increased ($P_{r,1}/P_{r,2}$). The increase in ratio is

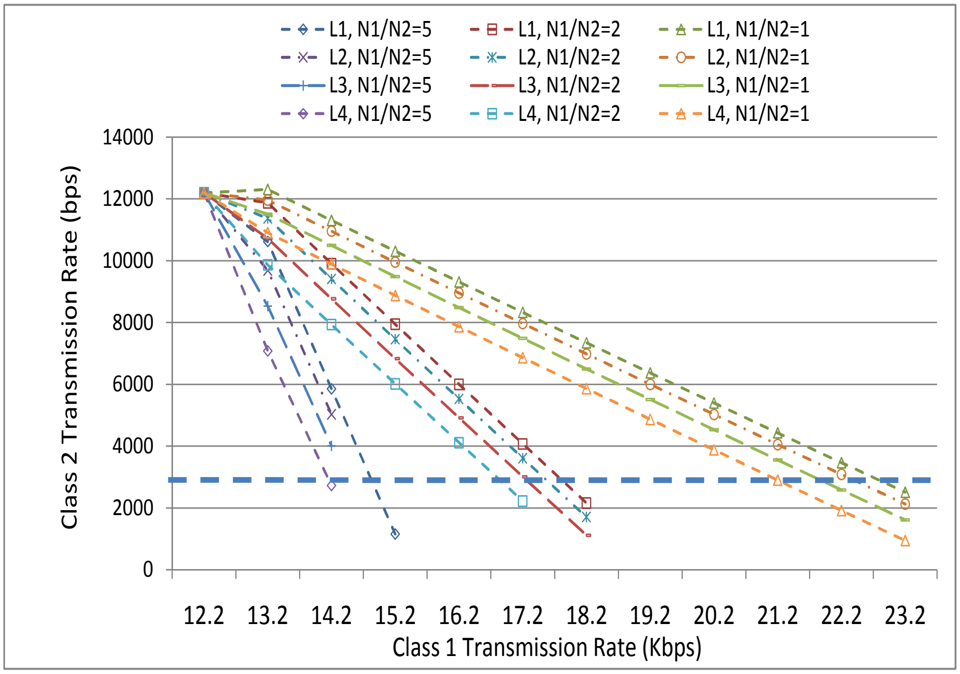


Figure 6.7: Effect of Increasing Class 1 Transmission Rate on Class 2 Transmission Rate Adaptation

plotted in Figure 6.8 against the increase in *C1* transmission rate. From this figure, it can be noted that the increase of *C1* transmission rate comes at the cost of increasing its transmission power. Moreover, as the ratio of *C1* calls to *C2* calls increases, such cost increases dramatically because of the increased number of *C1* calls. This increase in transmission power increases the interference level which forces *C2* calls to lower their transmission rates to maintain their quality. Also, this leads to an increase in the inter-cell interference on the adjacent cells which drives higher class mobile users to increase their transmission powers and lower the transmission rates of the lower class calls. Similar trends are observed for different supporting levels with more increase of the power ratio as the support level is increased. The reason behind this change is the increased inter-cell interference on the loaded sector and the intra-cell interference of the supporting sector as the number of handed off users towards the supporting sectors increases.

Since mobile equipment is limited by its power supply, the increased transmission power needs to be bounded at the level of the mobile equipment capability. Also, the transmission power needs to be efficiently managed to suppress the system interference level and to extend the lifetime of the mobile power supply. Hence, an upper threshold on the transmission power of *C1* calls needs to be determined to satisfy these goals. To satisfy these constraints, the provisioned transmission rates of *C1* calls may be lowered to comply with these restrictions.

6.3.3 Effect of BER Adaptation

Based on the requested application, different *QoS* requirements can be defined. For example, data transmission is tolerant to transmission delay as long as its *BER* is

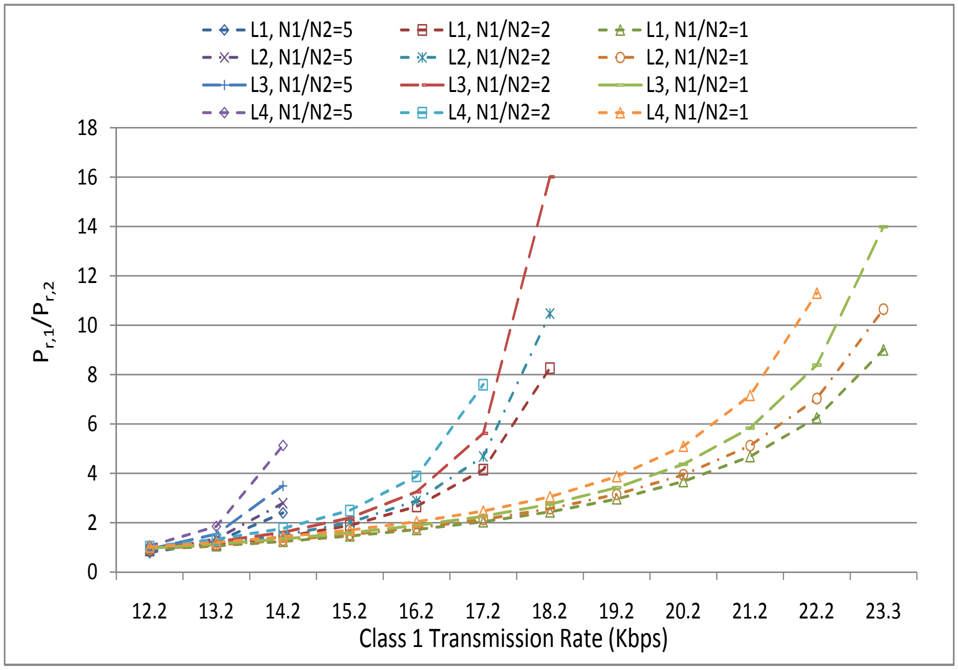


Figure 6.8: Effect of Increasing Class 1 rate on transmission power ratio

maintained below a certain threshold. On the other hand, voice and video services are tolerant to limited data loss but they are intolerant to transmission delay. Therefore, given the fixed transmission rate of *C2*, we calculate the achievable E_b/I_o of *C2* calls to maintain the required transmission rate of *C1* calls. In Figure 6.9, the results of the achievable E_b/I_o of *C2* calls are plotted against the transmission rate of *C1* with respect to different supporting levels and different ratios of the number of *C1* calls to that of *C2* calls. As shown in Figure 6.9, the minimum acceptable E_b/I_o of *C2* is set to 4.

As shown in the figure, the *BER* of *C2* calls is increased as the required transmission rate of *C1* calls is increased. Also, as the ratio of *C1* calls to that of *C2* calls increases, the *BER* of *C2* calls increases. Finally, as the transmission rate of *C1* calls increases, the *BER* of *C2* calls exceeds the threshold value marked on the figure with

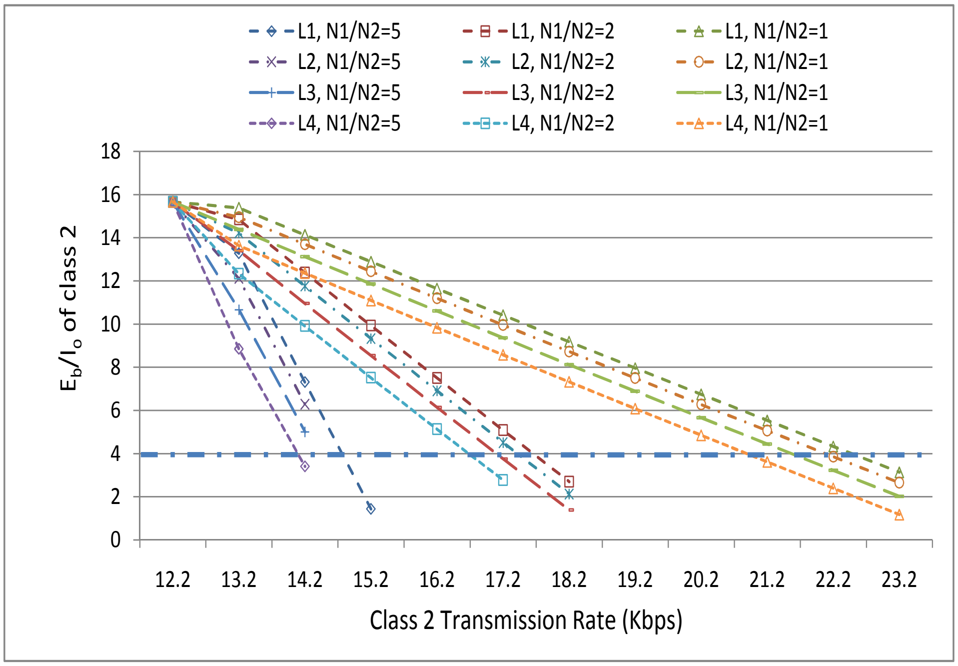


Figure 6.9: Effect of C1 Transmission rate on E_b/I_o of C2 on transmission power ratio

the dotted line. Again, this behavior varies with different activated supporting levels because of the increased number of handed off calls which leads to an increase in the interference level of the system.

The results presented in this section show a strong relationship between the transmission rate and the transmission power of calls from different classes with respect to different coverage levels of two nearby sectors when their traffic is unbalanced. As the transmission rate of *C1* increases, their transmission power is increased which negatively affects the transmission rate of *C2* calls. Also, as the supporting level increases, the inter-cell interference on the loaded sector increases and becomes dominant as the number of calls in the supporting sector increases. Also, the *BER* of the lower class can be negatively affected in the same manner with the increase of higher class transmission rate.

6.4 Summary

In this chapter we integrate the *DCB* module with power-controlled rate and coverage adaptation to maximize the number of admitted mobile users into the system and maximize the support given to a loaded sector. A Power Controlled Rate and Coverage Adaptation *PCRCA* module is proposed and analyzed. A heuristic scheme which iteratively evaluates the proposed *PCRCA* module is proposed. Adaptable traffic of four different transmission rate levels was used to evaluate the performance of the proposed module. The *PCRCA* is evaluated for different load scenarios; a varying loaded sector load with a static lightly loaded sector load, and static loaded sector load with a varying load of nearly loaded supporting sector. Results of such scenarios show the ability of the supporting sector to provide efficient support to the loaded sector up to a certain supporting level. As this level is exceeded, increasing the support level degrades the overall system performance due to the increased interference level mainly caused by the handed off mobile users towards the supporting sector.

A *DCB-Based QoS Provisioning* module is proposed. This module is used to evaluate the performance of a *DCB-enabled UMTS* system for two given traffic classes with different *QoS* requirements. The higher class, *C1*, required a fixed transmission rate while the lower class, *C2*, traffic was adaptable to the system load conditions. As the required transmission rate of the higher class was increased, the required transmission power to maintain the transmission rate was increased too. The increase in the transmission power increased the interference level in the system which negatively affected the transmission rate of the adaptable traffic. Numerical results also show that the received power ratio of *C1* and *C2* traffic increased as the transmission rate of *C1* traffic is increased. The E_b/I_o of the lower class traffic *C2* is adaptable given that its transmission rate is fixed. As the transmission rate of the higher class traffic *C1* was increased, the achievable *BER* of the lower class traffic was decreased.

Numerical results show the ability of the *DCB-Enabled WCDMA* system in managing congestion and providing differentiated service in a multimedia system. The maximum level of support given to a loaded sector needs to be provisioned to prevent degrading the system performance due to the increased inter cell interference caused by the handed off mobile users towards the expanded sector. Also, when multiclass traffic is supported in the system, the maximum achievable transmission rate needs to be upper bounded to maintain the mobile users' transmission powers below a certain limit.

Chapter 7

Conclusions and Future Work

Broadband Wireless Networks *(BWNs)* are increasingly deployed worldwide because of their convenient use and their ability to support multimedia services of different *QoS* requirements. These systems enable wireless data communications in addition to enhancing the traditional voice services of *1G* and *2G* systems. Despite their tremendous bandwidth increase, as compared to the *1G* and *2G* narrowband cellular systems, *BWNs* still experience congestion and load imbalance problems due to the rapid increase of wireless subscribers, which is motivated by the economical growth and the reduction of the communication cost per bit, and the diversity of the offered wireless applications. Therefore, congestion control and load balancing schemes become necessity to efficiently utilize the valuable radio resources and to meet the mobile users' *QoS* requirements. On the other hand, the problem of congestion control and load imbalance in *BWNs* is more complex as compared to the congestion control of *1G* and *2G* cellular systems. This is because of the divers *QoS* requirements of the provided applications and the increasing demand for such applications given that theses systems are interference-limited such as *UMTS*. Hence, congestion and load imbalance rapidly develops.

This thesis studied the problem of designing and developing efficient congestion control and load balancing mechanisms for *BWNs* and made solid contributions to the ongoing research in this area. In this chapter, we summarize and discuss the conclusions from this thesis and provide directions for future research.

7.1 Summary of Contributions

We define and classify the Directional Cell Breathing *(DCB)* congestion control and load balancing framework into three related components, namely the *DCB* coverage adaptation module, DCB-based Congestion Control and Loaded Balancing *(DCB-CC&LB)* protocol, and *DCB-Based QoS* Provisioning Module. The framework has been designed to simultaneously achieve and balance the following objective:

- Releasing hotspot congestion and balancing system load through dynamic network coverage adaptation.
- Minimizing the interference level in a network sector of spatially localized traffic.
- Efficiently utilizing the system resources by sharing traffic among network sectors.
- Preventing network coverage gaps.
- Supporting QoS differentiation for admitted traffic of varying *QoS* requirements.

In Chapter 4, we have introduced a novel coverage adaptation module, called Directional Cell Breathing *(DCB)*, which utilizes the capabilities of smart directional antennas to dynamically vary the coverage area of a cell sector to meet its coverage and capacity requirements. In this module, the coverage area of each cell sector is partitioned into L coverage levels that correspond to different transmission power levels of the Common Pilot Channel *(CPICH)*. The coverage adaptation of every two nearby sectors are practiced simultaneously to prevent network coverage gaps. When a traffic hotspot is formed in a cell sector, reducing its coverage area while expanding the coverage area of its nearby sector maximizes the loaded sector capacity and releases its congestion. This is achieved because nearedge mobile users are forced to handoff towards the expanded supporting sector, which results in a transmission power reduction of the other mobile users in the loaded sector. Therefore, the proposed directional coverage adaptation minimizes the system interference level and maximizes the system capacity. Herein, Cell Breathing Management *(CBM)* is practiced at the sector level and avoids unnecessary expansion towards other stable cell sectors, as opposed to the omni-directional cell breathing approach.

The *DCB* module is used to quantify the average interference power imposed on a cell sector by mobile users of adjacent cells' sectors. The interference is analyzed with respect to different coverage combinations, which correspond to different complementary coverage levels of every two nearby sectors. The calculated interference values are used to quantify the *DCB-Enabled WCDMA* system capacity. The capacity of a hotspot sector is quantified with respect to different coverage levels and for different load scenarios in the nearby supporting sector.

Based on the *DCB* interference analysis and the capacity quantification, a *DCB* optimization model is proposed. The objective of this model is to maximize the support level given by a supporting sector to its nearby loaded sector. A mathematical formula to compute the expected number of proactively handed off mobile users from a contracted load sector towards an expanded lightly loaded sector with respect to every coverage level is devised. The proposed *DCB* optimization model has been evaluated using different network load scenarios.

In Chapter 5, a congestion control and load balancing protocol utilizing the *DCB* coverage adaptation module is proposed for *BWNs*. This protocol is named DCB-based Congestion Control and Load Balancing *(DCB-CC&LB)*. *DCB-CC&LB* is composed of four components; namely Information Gathering, Information Evaluation, *DCB*, and Coverage Adaptation. The Information Gathering and Evaluation components are responsible for collecting and evaluating the network load state information, while the *DCB* component implements the proposed coverage adaptation techniques introduced in Chapter 4. The outcome of the *DCB* component is utilized by the Coverage Adaptation component which manages the actual cell sectors' coverage adaptation task. The implementation of the *DCB-CC&LB* protocol is composed of *Trigger, Evaluation, Selection,* and *Provisioning* stages and relies on the periodically updated load information of each cell sector. A *DCBDataBase* is defined to hold recent updated network cell sectors' state information for use by the *DCB-CC&LB* protocol stages.

The design approach of the *DCB-CC&LB* protocol for *UMTS* systems is also introduced in Chapter 5. In this design, the *Node B* implements the Information Gathering and Coverage Adaptation units with respect to its cell sectors, while the *RNC* implements the Information Evaluation and *DCB* units. The *NBAP* signaling protocol over the *Iub* interference is enhanced to communicate the required *DCB-CC&LB* protocol information between the *Node Bs* and *RNCs*. Such enhancement is called Enhanced *NBAP (ENBAP)* protocol.

Two heuristic *DCB* schemes based on the *DCB-CC&LB* protocol are proposed to evaluate the performance of a *DCB-enabled UMTS*. These schemes are Reactive *DCB (RDCB)* and Proactive *DCB (PDCB)*. The *RDCB* is a *measurement-based* scheme which is invoked when a measured sector's congestion parameter value exceeds a predefined sector load threshold. This scheme reacts to rapidly increased load in cell sectors due to a sector's base station partial or full failure, or higher call arrival rate increase in a certain network coverage area. On the other hand, the *PDCB* scheme is a *parameter-based* scheme which is periodically invoked to evaluate certain congestion sector parameters and acts upon these evaluated results. This scheme prevents and releases system congestion on a long-term basis.

The performance of a *DCB-enabled UMTS* is evaluating using *RDCB* and *PDCB* schemes. A snapshot simulator is implemented for the *RDCB* scheme. Mobile users' mobility is not considered in this scheme. On the other hand, a dynamic simulator is implemented for the *PDCB* scheme where mobile users' mobility is considered. Also, rate adaptation has been implemented in the *PDCB* scheme to support the *DCB* technique in maximizing sectors' support levels and minimizing call blocking and dropping probabilities. The results of both schemes showed an improved system performance in comparison to the results of a *UMTS* system without *DCB*.

In Chapter 6, a Power Controlled Rate and Coverage Adaptation *(PCRCA)* module for a *DCB-Enabled WCDMA* system is proposed. The *PCRCA* module objectives are to locally optimize nearby sectors' mobile users' transmission powers and transmission rates allocation while maximizing the support level given to a loaded sector. This module is composed of two parts, namely *PCRCA* Local *(PCRCA-L)* and *PCRCA* Global *(PCRCA-G)*. *PCRCA-L* concerns the optimal allocation of mobile users' transmission powers and transmission rates within a given sector while the *PCRCA-G* proposes different coverage combinations for the *PCRCA-L* evaluation.

The *PCRCA* module is evaluated heuristically. The obtained *PCRCA* results of the mobile users' average transmission power, transmission rate, and outage ratio outperform the performance results of the fixed pilot power *(FPP)* system. As the supporting sector coverage level exceeds a certain supporting level, the average transmission power and average outage ratio are increased sharply. Also, the average transmission rate is decreased. This is because of the increased interference level in the system due to the handed off mobile users towards the supporting sector. These handed off mobile users use higher transmission power to compensate for their distance increase to their new base station. Therefore, a maximum possible supporting level needs to be provisioned to prevent system performance degradation.

The second proposed module in Chapter 6 is the *DCB-Based QoS Provisioning* module which evaluates the *QoS* support in a multiclass *DCB-enabled UMTS* system. This module differentiates mobile users based on their traffic classes. The *DCB-Based QoS Provisioning* model considers the actual number of mobile users of each class within each sector. Two traffic classes, *C1* and *C2* defining high and low priority traffic respectively, are used to evaluate the performance of this module. The effect of increasing the transmission rate of high priority traffic on the transmission rate and $\frac{E_b}{I_o}$ degradation of low priority traffic is evaluated. Also, the transmission power ratio increase is evaluated. These evaluations are performed for different mobile users' ratios of these traffic classes. As the ratio increases, the average mobile users' transmission rate and $\frac{E_b}{I_o}$ of low priority traffic are decreased until they exceed their minimum provisioned level. This decrease slightly differs as the support sectors' coverage levels are increased. Also, the required transmission power ratio of high priority to low priority traffic is increased as the transmission rate of mobile users of high priority traffic is increased. The transmission power increase is required to compensate for the interference level in the system. Therefore, as the transmission rate and high priority to low priority mobile users' ratios are increased, the high priority to low priority mobile users transmission power ratio is increased. The increase slightly changes as the support sectors' coverage levels are increased. An upper threshold on mobile users' transmission power increase need to be provisioned to meet the maximum available transmission power of mobile users devices.

7.2 Future Research Directions

There are several directions that can be investigated based on the proposed work in this thesis. In this section, we highlight some of these directions.

We have shown that the proposed coverage adaptation module in Chapter 4 varies the service area of a cell sector based on the sector traffic intensity and the load of its nearby sector. We remark though that sector coverage adaptation is only optimized on the radial direction where the radius of a cell sector is varied to extend or contract its service area towards a nearby cell sector. To further enhance the performance of the proposed coverage adaptation module, sector coverage adaptation can be also based on varying the angle of a cell sector. Considering such an enhancement will increase coverage adaptation options.

The congestion control and load balancing protocol of Chapter 5 utilizes a basic Call Admission Control *(CAC)* where mobile users' calls are admitted to the system as long as their admission maintains the sector uplink load factor below a predefined threshold. A *CAC* scheme which considers the activated coverage combination of every two nearby sectors would increase the robustness and efficiency of the proposed protocol. Such a *CAC* scheme would integrate the activated coverage combination of every two nearby sectors in its call admission policy.

Elevating the proposed framework to the packet level would give the system operators a broad understanding of the effect of dynamic coverage adaptation on the system's packet level performance such as throughput, delay, etc. Therefore, the coverage adaptation module can be utilized to enhance the mobile users' channel conditions which are required for efficient data transmission scheduling especially in High Speed Packet Access *(HSPA)* systems which are also known as of 3.5G and beyond [3], [4] and [9].

Another promising research direction is the integration of the proposed coverage adaptation framework with the concept of Multi-hop Cellular Networks *(MCN)* [64], [74] and [73]. In such direction, the concept of traffic relaying of *MCN* can be invoked to support the coverage adaptation framework functionalities when system load is extremely high. Therefore, commanding some mobile users to switch to mutli-hop communication mode leads to system interference level suppression and hence motivates the nearby supporting sector to provide more support to the loaded sector. On the other hand, a tradeoff needs to be investigated with respect to the added complexity and cost of the *MCN* architecture.

Bibliography

[1] 3GPP TR 23.107. "Technical Specification Group Services and System Aspects, QoS Concept and Architecture". *Release 1999*, 3(9), February 2002.

[2] 3GPP TS 21.101. "Technical Specification Group Services and System Aspects; Technical Specifications and Technical Reports for a UTRAN-based 3GPP system". *Release 99*, 3(15), March 2005.

[3] 3GPP TS 25.308. "High Speed Downlink Packet Access (HSDPA)". *3GPP*, Release 5, March 2003.

[4] 3GPP TS 25.309. "High Speed Uplink Packet Access (HSUPA)". *3GPP*, Release 6, April 2006.

[5] 3GPP TS 25.433. "UTRAN Iub interface NBAP signalling Release 99". *3GPP*, Release 99, September 2006.

[6] 3GPP TS 34.108. "Technical Specification Group Terminals; Common Test Environments for User Equipment (UE) Conformance Test". *Release 99*, 3(16), June 2004.

[7] Agrawal, P., Anvekar, D., and Narendran, B. “Channel management policies for handovers in cellular networks”. *Bell Labs Technical Journal*, 1(1):97–110, Autumn 1996.

[8] Ahmed, H. “Call Admission Controll in Wireless Networks: A Comprehensive Survey”. *IEEE Communications Surveys*, 7(1):50–69, First Quarter 2005.

[9] Al-Manthari, B., Hassanein, H., and Nasser, N. “Packet Scheduling in 3.5G High-Speed Downlink Packet Access Networks: Breadth and Depth”. *IEEE Network*, 21(1):41–46, January/Febrauray 2007.

[10] AlQahtani, A. and Mahmoud, S. “Uplink Call Admission Schemes for 3G and Beyond Mobiles Wireless Networks Supporting Heterogeneous Traffic”. *Information and Communication Technologies*, 2:2657–2662, 2006.

[11] Balanis, C., and Ioannides, P. *“Intorduction to Smart Antennas”*. 2007.

[12] Berggren, F., Kin, S., Jantti, R., and Zander, J. “Joint Power Control and Intracell Scheduling of DS-CDMA Nonreal Time Data”. *IEEE Journal on Selected Areas in Communications*, 19(10):1860–1870, October 2001.

[13] Cai, J., and Goodman, D. “General packet radio service in GSM”. *IEEE Communications Magazine*, 35(10):122–131, October 1997.

[14] Casoni, M., Immovilli, G., and Merani, M. “Admission Controll in T/CDMA Systems Supporting Voice and Data Applications”. *IEEE Transaction on Wireless Communications*, 1(3):540–548, July 2002.

[15] Chao, C., Chen, W. “Connection admission control for Mobile Multiple-class Personal Communications Networks”. *IEEE Journal on Selected Areas in Communications*, 15(8):1618–1626, October 1997.

[16] Chatziperis, S., Koutsakis, P., and Paterakis, M. “Highly Efficient Call Admission Control for Wireless Multimedia Traffic”. *International Conference on Networking*, pages 675–680, April 2008.

[17] Choi, S., and Sohraby, K. “Analysis of a Mobile Cellular Systems with Hand-off Priority and Hysteresis Control”. *IEEE INFOCOM*, 1:217–224, March 2000.

[18] Chryssomallis, M. “Smart Antennas”. *IEEE Antennas and Propagation Magazine*, 42(3):129–136, June 2000.

[19] Curescu, C., and Nadjm-Tehrani, S. “Time-Aware Utility-Based Resource Allocation in Wireless Networks”. *IEEE Transactions on Parallel and Distributed Systems*, 16(7):624–636, July 2005.

[20] Dahlman, E., Beming, P., Knutsson, J., Ovesjo, F., Persson, M., and Roobol, C. “WCDMA-The Radio Access Interface for Future Mobile Multimedia Communications”. *IEEE Transactions on Vehicular Technology*, 47(4), November 1998.

[21] Din, L., Bigham, J., Cuthbert, L., Nahi, P., and Parini, C. “Intelligent cellular network load balancing using a cooperative negotiation approach”. *IEEE Wireless Communications and Networking Conference*, 3:1675–1679, March 2003.

[22] Ekiz, N., Salih, T., Kkner, S., and Fidanboylu, K. "An Overview of Handoff Techniques in Cellular Networks". *World Academy of Science, Engineering and Technology*, 6:1–4, June 2005.

[23] Foschini, G., and Miljanic, Z. "A simple distributed autonomous power control algorithm and its covergence". *IEEE Transactions on Vehicular Technology*, 42(4):641–646, November 1993.

[24] Gilhousen, K., Jacobs, I., Padovani, R., Viterbi, A., Weaver, L., and Wheatley, C. "On the Capacity of a Cellular CDMA System". *IEEE Transaction on Vehicular Technology*, 40(2):303–312, May 1991.

[25] Grandhi, S., and Zander, J. "Constrained Power Control in Cellular Radio Systems". *IEEE Vehicular Technology Conference*, 2:824–828, June 1994.

[26] Grandhi, S., Vijayan, R., and Goodman, D. "Distributed power control in cellular radio systems". *IEEE Transactions on Communication*, 42(2/3/4):226–228, FEBRUARY/MARCH/APRL 1994.

[27] Grandhi, S., Vijayan, R., and Goodman, D. "Distributed Power Control in Cellular Radio Systems". *IEEE Transactions on Communications*, 42(2/3/4):226–228, February 1994.

[28] Holma, H., and Laakso, J. "Uplink Admission Control and Soft Capacity with MUD in CDMA". *IEEE Vehicular Technology Conference*, 1:431–435, September 1999.

[29] Holma, H., and Toskala, A. *"W-CDMA for UMTS"*. 2000.

[30] Holma, H., and Toskala, A. *"WCDMA for UMTS- HSPA Evolution and LTE"*, volume 4. 2007.

[31] Holma, H., Toskala, A., Ranta-aho, K., and Pirskanen, J. "High-Speed Packet Access Evolution in 3GPP Release 7". *IEEE Communications Magazine*, 45(12):29–35, December 2007.

[32] Huang, Y., and Yates, D. "Call admission in power controlled CDMA systems". *IEEE Vehicular Technology Conference*, 3:1665–1669, May 1996.

[33] Imre, S. "Dynamic Call Admission Control for Uplink in 3G/4G CDMA-Based Systems". *IEEE Transactions on Vehicular Technology*, 56(5):2617–2629, Septmber 2007.

[34] Jobin, J., Faloutsos, M., Tripathi, S., and Krishnamurthy, S. "Understanding the Effects of HotSpots in Wireless Cellular Networks". *IEEE INFOCOM*, 1:660–671, March 2004.

[35] Kim, S., and Lee, Y. "Combined Rate and Power Control Adaptation in DS/CDMA Communication over Nakagami Fading Channels". *IEEE Transaction On Communications*, 48(1):162–168, January 2000.

[36] Kim, S., Rosberg, Z., and Zander, J. "Combined Power Control and Transmission Rate Selection in Cellular Networks". *IEEE Vehicular Technology Conference*, pages 1653–1657, September 1999.

[37] Kim, Y., Lee, K., and Chin, Y. "Analysis of multi-level threshold handoff algorithm". *IEEE Global TeleCommunications Conference*, 2:1141–1145, November 1996.

[38] Koutsakis, P., Paterakis, M., and Psychis, S. "Call admission control and traffic policing mechanisms for the wireless transmission of layered videoconference traffic from MPEG-4 and H.263 video coders". *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 5:2155–2159, September 2002.

[39] Kyatanavar, N., Patil, S., Patil, S., and Zope, G. "WLS Channel Prediction for Power Control in DS-CDMA Cellular Communication Systems". *International Conference on Wireless Communication and Sensor Networks*, pages 95–99, December 2007.

[40] Laio, C., Yu, F., Leung, V., Chang, C. "A Novel Dynamic Cell Configuration Scheme in Next-Generation Situation-Aware CDMA Networks". *IEEE Journal on Selected Areas in Communications*, 24(1):16–25, January 2006.

[41] Larijani, P., Hafez, R., and Lambadaris, I. "Adaptive Access Control for Mutlimedia Traffic in a CDMA Cell with Imperfect Power Control". *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2:729–733, September 1997.

[42] Lee, C., and Steele, R. "Effect of soft and softer handoffs on CDMA system capacity". *IEEE Transactions on Vehicular Technology*, 47(3):830–841, Auguest 1998.

[43] Liang, X., Xuemin, S., and Mark, J. "Performance analysis of rate adaptation scheme for data traffic in DS-CDMA systems". *IEEE International Conference on Communications*, 5:3372–3376, April/May 2002.

[44] Liu, J. and Liang, Z. “Efficient Call Admission Control Algorithm for the Next Generation Mobile Communications Systems”. *International Conference on Wireless Communications, Networking and Mobile Computing*, pages 922–925, Septmber 2007.

[45] Liu, T., and Silvester, J. “Joint Admission/Congestion Control for Wireless CDMA Systems Supporting Intergated Services”. *IEEE Journal On Selected Areas In Communications*, 16(6):845–857, August 1998.

[46] Liu, Z., and El Zarki, M. “SIR-Based Call Admission Control for DS-CDMA Cellular Systems”. *IEEE Journal on Selected Areas in Communications*, 12(4):638–644, May 1994.

[47] Lu, N., and Bigham, J. “Utility-maximization bandwidth adaptation for multi-class traffic QoS provisioning in wireless networks”. *International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems*, pages 136–143, 2005.

[48] Madan, B., Dharmaraja, S., and Trivedi, S. “Combined Guard Channel and Mobile-Assisted Handoff for Cellular Networks”. *IEEE Transactions on Vehicular Technology*, 57(1):502–510, January 2008.

[49] Manaffar, M., Bakhshi, H., and Pilevari, M. “A New Dynamic Pricing Scheme with Call Admission Control to Reduce Network Congestion”. *International Conference on Advanced Information Networking and Applications*, pages 347–352, March 2007.

[50] Marichamy, P., Chakrabati, S., and Maskara, S. “An Overview of handoff schemes in cellular mobile networks and their comparative performance evaluation”. *IEEE Vehicular Technology Conference*, 3:1486–1490, September 1999.

[51] Meyerhoff, H. “Method for computing the optimum power balance in multibeam satellites”. *COMSAT Technical Review*, 4(1):139–146, 1974.

[52] Ming, H., and Junshan, Z. “Rate adaptation for bursty data transmission in CDMA networks”. *Asilomar Conference on Signals, Systems and Computers*, 2:1718–1722, November 2001.

[53] Molkdar, D., Featherstone, W., and Larnbotharan, S. “An overview of EGPRS: the packet data component of EDGE”. *Electronics and Communication Engineering Journal*, 14(1):21–38, February 2002.

[54] Naslund, J., Carneheim, C., Johansson, C., Jonsson, S.-O., Ljungberg, M., Madfors, M., and Skold, J. “An evolution of GSM”. *IEEE Vehicular Technology Conference*, 1:348–352, June 1994.

[55] Nettleton, R., and Alavi, H. “Power control for a spread-spectrum cellular mobile radio system”. *IEEE Global Telecommunication Conference*, pages 84–88, November 1982.

[56] Nettleton, R., and Alavi, H. “Power control for a spread-spectrum cellular mobile radio system”. *IEEE Vehicular Technology Conference*, 33:242–246, May 1983.

[57] Noerpel, A. and Lin, Y. “Handover Arrangement for a PCS Network”. *IEEE Personal Communications*, 4(6):18–24, December 1997.

[58] Nuaymi, L., Godlewski, P., and Mihailescu, C. "Call admission control algorithm for cellular CDMA systems based on best achievable performance". *IEEE Vehicular Technology Conference*, 1:375–379, May 2000.

[59] Ojanpera, T. and Prasad, R. *"WCDMA: Towards IP Mobility and Mobile Internet"*. 2000.

[60] Opnet Network Simulator. "UMTS Model User Guide 21.1". May 2007.

[61] Park, J., Huang, L., and Kuo, C. "Computationally Efficient Dynamic Code Assignment Schemes With Call Admission Control (DCA-CAC) for OVSF-CDMA Systems". *IEEE Transactions on Vehicular Technology*, 57(1):286–296, January 2008.

[62] Perros, H., and Elsayed, K. "Call Admission Control Schemes: A Review". *IEEE Communications Magazine*, 34(11):82–91, November 1996.

[63] Pollini, P. "Trends in handover design". *IEEE Communications Magazine*, 34(3):82–90, March 1996.

[64] Radwan, A. and Hassanein, H. "On the Capacity of Multi-hop CDMA Cellular Networks". *IEEE Symposium on Computers and Communications*, pages 409–414, July 2007.

[65] Ramsdale, A., and Harrold, B. "Techniques for cellular networks incorporating microcells". *IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, pages 169–173, October 1992.

[66] Rodrigues, E., Cavalcanti, F., Lima, C., and Sousa, V. "(Re)Active Load Control Based on Radio Link Quality for UMTS/WCDMA Forward Link". *IEEE Vehicular Technology Conference*, 1:363–367, June 2005.

[67] Rodrigues, E., Lima, C., Sousa, V., and Cavalcanti, F. "Power-Based Congestion Control Framework for WCDMA Systems". *IEEE Vehicular Technology Conference*, 3:1870–1874, June 2005.

[68] Romero, J., Sallent, O., Agusti, R.,Garcia, N., Wang, L., Aghvami, H. "Network Controlled Cell-Breathing for Capacity Improvement in Heterogeneous CDMA/TDAM Scenarios". *IEEE Wireless Communications and Networking Conference*, pages 36–41, April 2006.

[69] Ross, S. "Introduction to Proabability Models". *Academic Press*, 2003.

[70] Stanisic, V., and Devetsikiotis, M. "Dynamic utility-based bandwidth allocation policies: the case of overloaded network". *IEEE International Conference on Communications*, 4:1958–1962, June 2004.

[71] Taha, A., Hassanein, H., and Mouftah, H. "On reducing the operational cost of bandwidth adaptation algorithms". *International Conference on Broadband Networks*, 2:781–783, October 2005.

[72] Talukdar, A., and Badrinath, B. "Rate Adaptation Schemes in Networks with Mobile Hosts". *International Conference on Mobile Computing and Networking*, pages 169–180, October 1998.

[73] Tam, H., Benkoczi, R., Hassanein, H. and Akl, S. "Optimal Channel Assignment in Multi-Hop Cellular Networks". *IEEE Global Telecommunications Conference*, 6:731–735, November 2007.

[74] Tam, H., Hassanein, H., Akl, S. and Benkoczi, R. "Optimal Multi-hop Cellular Architecture for Wireless Communications". *IEEE Conference on Local Computer Networks*, pages 738–745, November 2006.

[75] Tanenbaum, S. *"Computer Networks"*. 2003.

[76] Tekinay, S., Jabbari, B. "Handover and Channel Assignment in Mobile Cellular Networks". *IEEE Communications Magazine*, 29:42–46, November 1991.

[77] Tripathi, D., Reed, J., and VanLandinoham, H. "Handoff in Cellular System". *IEEE Personal Communications*, 5:26–37, December 1998.

[78] UMTS Forum. "More than 400 million 3G customers around the world". *http://www.umts-forum.org/content/view/2522/174/*, October 2008.

[79] Wacker, A., Laiho-Steffens, J., Sipila, K., and Heiska, K. "The Impact of The Base Station Sectorisation On WCDMA Radio Network Performance". *IEEE Vehicular Technology Conference*, 5:2611–2615, September 1999.

[80] Wang, L., Huang, S., and Tseng, Y. "Interference Analysis and Resource Allocation for TDD-CDMA Systems to Support Asymmetric Services by Using Directional Antennas". *IEEE Transactions on Vehicular Technology"*, 54(3):1056–1069, May 2005.

[81] Xiaomin, M., Yun, L., and Trivedi, S. "Design and Performance Analysis of a New Soft Handoff Scheme for CDMA Cellular Systems". *IEEE Transactions on Vehicular Technology*, 55(5):1603–1612, Septmber 2006.

[82] Xiaomin, M., Yun, L., and Trivedi, S. "Modeling and performance analysis for soft handoff schemes in CDMA cellular systems". *IEEE Transactions on Vehicular Technology*, 55(2):670–680, March 2006.

[83] Yao, N., Cuthbert, L. "Reducing Congestion over Hotspot Clusters in WCDMA Networks". *IEEE Wireless Communications and Networking Conference*, pages 3734–3738, March 2007.

[84] Yates, R. "A framework for uplink power control in cellular radio systems". *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1347, September 1995.

[85] Zander, J. "Performance of optimum transmitter power control in cellular radio systems". *IEEE Transactions on Vehicular Technology*, 41(1):57–62, February 1992.

[86] Zander, J. "Performance of optimum transmitter power control in cellular radio systems". *IEEE Transactions on Vehicular Technology*, 41(1):57–62, February 1992.

[87] Zhang, N., and Holtzman, J. "Analysis of a CDMA soft handoff algorithm". *IEEE Transactions on Vechilure Technology*, 47(2):710–714, May 1998.

[88] Zhang, Y., and Liu, D. “An Adaptative Algorithm for Call Admission Control in Wireless Networks”. *IEEE Global TeleCommunications Conference*, 6:3628–3632, November 2001.

[89] Zhao, D., Shen, X., and Mark, W. “Call admission control for heterogeneous services in wireless networks”. *IEEE International Conference on Communications*, 2:964–968, June 2000.