Resource Management in Multi-hop Cellular Networks

BY

Yik Hung Tam

A thesis submitted to the School of Computing

in conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

January 2009

Abstract

In recent years, mobile communications have become affordable and popular. High cellular capacity in terms of number of users and data-rates is in need. As the available frequency spectrums for mobile communications are limited, the utilization of the radio resources to achieve high capacity without imposing high equipment cost is of utmost importance. Recently, multi-hop cellular networks (MCNs) were introduced. These networks have the potential of enhancing the cell capacity and extending the cell coverage at low extra cost. However, in a cellular network, the cell or system capacity is inversely related to the cell size. In MCNs, the cell size, the network density and topology affect the coverage of source nodes and the total demands that can be served and, thus, the system throughput. Although the cell size is an important factor, it has not been exploited for maximizing throughput. Another major issue in MCNs is the increase in packet delay because multi-hopping is involved. High packet delay affects quality of service provisioning in these networks.

In this thesis, we propose the Optimal Cell Size (OCS) and the Optimal Channel Assignment (OCA) schemes to address the cell size and packet delay issues for a time division duplex (TDD) wideband code division multiple access (W-CDMA) MCN. OCS finds the optimal cell sizes to provide an optimal balance of cell capacity and coverage to maximize the system throughput, whereas OCA assigns channels optimally in order to minimize packet relaying delay. Like many optimized schemes, OCS and OCA are computationally expensive and may not be suitable for large real-time problems. Hence, we also propose heuristics for solving the problems. For the cell size problem, we propose two heuristics: Smallest Cell Size First (SCSF) and Highest Throughput Cell

Size First (HTCSF). For the channel assignment problem, we propose the Minimum Slot Waiting First (MSWF) heuristic. Simulation results show that OCS achieves high throughput compared to that of conventional (single-hop) cellular networks and OCA achieves low packet delay in MCNs. Results also show that the heuristics, SCSF, HTCSF and MSWF, provide good results compared to the optimal ones provided by OCS and OCA, respectively.

Acknowledgements

First, I would like to thank God for guiding me in this wonderful journey that I have never thought of.

To my two supervisors, Dr. Selim Akl and Dr. Hossam Hassanein, I would like to express my deepest gratitude. Dr. Akl is an excellent and exceptional supervisor. He is a humble, gentle, knowledgeable, vigorous, and efficient researcher. He inspires me by showing me what an excellent researcher should be like. His expertise in algorithm design and problem solving techniques broadens my intellectual horizon. He always gives me valuable advice for my research work. Dr. Hassanein is also an excellent supervisor. He not only provided excellent guidance and support on my research work, but also gave valuable advice, support, and encouragement for my professional and personal development. He is a very considerate supervisor with a wonderful sense of humor. His expert knowledge in the networking field helped me tackle a number of obstacles during this work. I am very lucky to have them as my supervisors. I would also like to thank Dr. Robert Benkoczi, who is my collaborator, friend, and mentor, for his great help and patience. What an excellent researcher and an exceptional mentor he is. Not only has he broadened my knowledge of linear programming techniques, but also helped me to excel in my research work. I would like to thank Dr. Juergen Dingel for his advice, help, support and encouragement.

To my parents, Mom and Dad, thank you so much for your deepest and unconditional love for me so that I could carry on my journey. This journey has never been easy for me.

But I finally made it. I am proud of being your son. I miss you so much. Grandma and Grandpa, thank you for your unconditional and genuine love which was important to me when my parents left. Auntie King Ming So, thank you so much for your love and care and treating me like your son. Brian Cheuk Hung Siu, you are the best of my best friends. Without your unconditional support financially and spiritually, my life would be much more difficult and this accomplishment may not be possible. Winnie Man, thank you for being part of my life. Your unconditional love and care makes me feel rich. Chun-ho Chan and Angela Lau, thank you for caring, listening, and standing by me all the time especially in Canada where I do not have many friends. Auntie Pauline So, thank you for your love and care which makes me feel home in Canada. Benjamin Chan, you are such a wonderful person. You are always helpful and considerate. Ben Kam, thank you for being my buddy in this study journey. Without you, this would be a long and lonely journey. Dr. Wenying Feng, thank you so much for your tremendous help and guidance in my undergraduate studies. Thank you also for your care and support all the time. I would also like to thank Mrs. Ruth Kennedy, Carolina Cuthbert, Rosie, Emily Jansons and Mrs. Ruth Pester, who volunteers to help me improve my English. You are such wonderful persons.

To my relatives and friends including the members of So family, Priscilla Chan, Elaine So, Edmond Lo, Kevin Lai, SKC friends, Chi-Shun Chan, Ven Lee, FAS friends, PIE friends, Jenny Li, Ronnie Kwan, Ivy Lam, Connie Chan, Benny Ng, Agatha Wong, Aunt Lam Yuen Tai, Simon Yu, Chi-hung Ho, UI colleagues, Adrianne Wai, Mandy Lee, Simon Yam, Sharen Li, Marianne Lo, IN25 buddies, Patrick Lam, John Lau, Sebastian Hung, Ian, and Debby, thank you for your love, care and support all the time – I love you all.

I would like to thank the students, the staff, especially Debby Robertson, the faculty members of the School of Computing, and my colleagues, especially Dr. Abd-Elhamid Taha, in the Telecommunications Research (TR) Lab for their help, care, valuable advice, and fun times we spent together.

Finally, I would like to thank the financial support provided by Queen’s University, the Natural Sciences and Engineering Research Council of Canada (NSERC), the Ontario Graduate Scholarship program, and the Communication Information Technology of Ontario (CITO) and Bell University Labs (BUL).

Contents

Abstract i
Acknowledgements iii
Contents vi
List of Tables ix
List of Figures ix
List of Acronyms xi

1. Introduction 1
1.1. Cellular and Multi-hop Cellular Networks 1
1.2. Challenges of MCNs 6
1.2.1. Cell size 7
1.2.2. Channel assignment 8
1.3. Thesis Objectives 9
1.4. Thesis Contributions 10
1.5. Thesis Organization 11

2. Overview and Related Work 13
2.1. Network Architecture 13
2.2. Design Factors and Issues 16
2.2.1. Wireless technology 17
2.2.2. Relaying device 21
2.2.3. Wireless interface/ communication mode/ supporting technology 23
2.2.4. Routing 24
2.2.5. Load balancing 26
2.2.6. Other issues 27
2.3. Cell Size 29
2.3.1. Cell size in wireless networks 29
2.3.2. Cell size in MCNs 32
2.4. Channel Assignment 35
2.4.1. Channels in cellular systems 35
2.4.2. Channel assignment in cellular networks 36
2.4.3. Channel assignment in MCNs 38
2.5. Summary 42

3. System Model 44
3.1. System Features 44
3.2. System Components 50
3.3. Network Model 51

3.4. Assumptions 53
3.5. Summary 54
4. Channel Assignment in Multi-hop Cellular Networks 56
4.1. Channel Assignment Model 57
4.1.1. Channel assignment 58
4.1.2. Channel conflicts 58
4.1.3. Relaying delay 60
4.1.4. Delay bounds 61
4.2. Optimal Channel Assignment (OCA) 62
4.2.1. OCA formulation 62
4.3. Minimum Slot Waiting First (MSWF) 68
4.3.1. The MSWF Scheme 68
4.3.2. Illustration of MSWF algorithm 73
4.3.3. Worse case of MSWF algorithm 75
4.4. Complexity Analysis 76
4.5. Performance Evaluation 77
4.5.1. Simulation model and parameters 77
4.5.2. Performance metrics 81
4.5.3. Simulation results 82
4.6. Limitations 88
4.7. Summary 88
5. Cell Size in Multi-hop Cellular Networks 90
5.1. Cell Size 91
5.1.1. The importance of cell size 91
5.1.2. Optimal cell size - single-cell case (Capacity-demand model) 91
5.1.3. Optimal cell size - multi-cell case 93
5.2. Optimal Cell Size (OCS) 96
5.2.1. OCS formulation 96
5.3. Small Cell Size First (SCSF) 100
5.3.1. The SCSF scheme 100
5.3.2. Illustration of SCSF algorithm 103
5.4. Highest Throughput Cell Size First (HTCSF) 105
5.4.1. The HTCSF scheme 105
5.4.2. Illustration of HTCSF algorithm 108
5.5. Complexity Analysis 109
5.6. Performance Evaluation 110
5.6.1. Simulation model and parameters 110
5.6.2. Performance metrics 114
5.6.3. Simulation results 114
5.7. Limitations 119

5.8. Summary 120
6. Conclusions and Future Work... 121
BIBLIOGRAPHY 124

List of Tables

Table 2.1: DESIGN DECISION FACTORS AND RELATED ISSUES FOR MCNs 21
Table 4.1: SIMULATIONS PARAMETERS 80
Table 5.1: SIMULATIONS PARAMETERS 112
Table 5.2: CAPACITY CORRESPONDING TO THE CELL RANGE................................... 113
Table 5.3: UPLINK RANGE VERSUS CAPACITY 113

List of Figures

Figure 1.1: System architecture of a 3G cellular system 2
Figure 1.2: Network architecture of a MCN... 5
Figure 1.3: Reduction of number of BSs and cell size of a cellular system 6
Figure 2.1: An ad hoc network 15
Figure 2.2: Traffic is relayed through ARSs 17
Figure 2.3: Classification of major MCN proposals based on the design factors 20
Figure 2.4: A switching point in a TDD CDMA transmission frame. 37
Figure 3.1: Operation of FDD and TDD 45
Figure 3.2: Three multiple access techniques: a) FDMA, b) TDMA and c) CDMA 48
Figure 3.3: A TDD transmission frame of 3G UMTS.. 48
Figure 3.4: Interactions between the cell size and channel assignment components 50
Figure 3.5: Topology in a MCN with directional antennas 52
Figure 4.1: Topology and channel assignment in a MCN with directional antennas.................... 57
Figure 4.2: Time-slot waiting time (relaying delay) at relaying nodes on relaying path 60
Figure 4.3: Virtual points for the relaying paths corresponding to the topology in Figure 4.1 63
Figure 4.4: Collision graph corresponding to the co-channel conflicting points in Figure 4.1 64
Figure 4.5: Consecutive graph corresponding to the co-time-slot conflicting points in Figure 4.1 65
Figure 4.6: The MSWF Algorithm 71
Figure 4.7: Channel assignment for the virtual points corresponding to the scenario in Figure 4.1 73
Figure 4.8: A single-cell model 77
Figure 4.9: Per-hop Packet Delay (Scenarios 1 and 2).. 85
Figure 4.10: C.O.V. of Per-hop Packet Delay (Scenarios 1 and 2) 86
Figure 4.11: Packet Delay (Scenario 2)... 86
Figure 4.12: Cell Throughput (Scenario 2) 87

Figure 4.13: Packet Delay versus Cell Throughput (Scenario 2) 87
Figure 5.1: Capacity-demand model 92
Figure 5.2: Cell size in a multi-cell environment 94
Figure 5.3: Distance graph *G* for base station b_1 in Figure 5.2 96
Figure 5.4: The SCSF Algorithm 102
Figure 5.5: An example to illustrate the SCSF algorithm 104
Figure 5.6: The HTCSF Algorithm 107
Figure 5.7: 3-Cell model 111
Figure 5.8: Cell Throughput 117
Figure 5.9: Call Acceptance Ratio 117
Figure 5.10: Cell Size 118
Figure 5.11: Packet Delay 118
Figure 5.12: Packet Delay versus Cell Throughput 119

List of Acronyms

1G	1st Generation Wireless Communication Systems
2G	2nd Generation Wireless Communication Systems
3G	3rd Generation Wireless Communication Systems
3GPP	3rd Generation Partnership Project
4G	4th Generation Wireless Communication Systems
ACAR	A-Cell Adaptive Routing
A-Cell	Ad hoc-Cellular
A-GSM	Ad hoc Global System for Mobile
ALBA	A-Cell Load BAlancing
AMC	Adaptive Multi-hop Cellular
AODV	Ad hoc On-Demand Distance Vector
AP	Access Point
ARS	Ad hoc Relaying Station
BCR	Base-Centric Routing
BER	Bit error rate
BS	Base Station
CAMA	Cellular Aided Mobile Ad hoc Network
CAHAN	Cellular Ad hoc Augmented Network
CBM	Cellular Based Multi-hop
CBR	Call Blocking Ratio
CDMA	Code Division Multiple Access
cMCN	Clustered Multihop Cellular Network
DCF	Distributed Coordination Function
DSR	Dynamic Source Routing
DSSA	Delay-Sensitive Slot Assignment
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
GPS	Global Positioning System
GSM	Global System for Mobile

HTCSF	Highest Throughput Cell Size First
HMCN	Hierarchical Multi-hop Cellular Network
HWN	Hybrid Wireless Network
iCAR	integrated Cellular and Ad hoc Relay
IP	Internet Protocol
ISM	Industrial, Scientific, and Medical
LDPR	Location-Dependent Packet Relay
LCS	Large Cell Size
MAC	Medium Access Control
MADF	Mobile Assisted Data Forwarding
MANET	Mobile Ad Hoc NETwork
MCN	Multi-hop Cellular Network
MCN-p	Multi-hop Cellular Network – power reduction
MCN-b	Multi-hop Cellular Network – base station reduction
MRAC	Multi-hop Radio Access Cellular
MSWF	Minimum Slot Waiting First
MSC	Mobile Switching Centre
MT	Mobile Terminal
OCA	Optimal Channel Assignment
OCS	Optimal Cell Size
ODMA	Opportunity-Driven Multiple Access
OFDMA	Orthogonal Frequency Division Multiple Access
P2P	Peer to Peer
PSTN	Public Switching Telephone Networks
QoS	Quality of Service
PARCelS	Pervasive Ad hoc Relaying for Cellular System
RNC	Radio Network Controller
SCS	Small Cell Size
SCSF	Small Cell Size First
SOPRANC	Self-Organizing Packet Radio Ad hoc Networks with Overlay
TDD	Time Division Duplex

TDMA	Time Division Multiple Access
UCAN	Unified Cellular and Ad hoc Network
UMTS	Universal Mobile Telecommunication System
VCN	Virtual Cellular Network
W-CDMA	Wideband Code Division Multiple Access
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network

Chapter 1

Introduction

Wireless communication technology has made great gains in popularity over the past decade and will be playing a more important role in access networks, as evidenced by the widespread adoption of cellular networks, Wireless Local Area Networks (WLANs) and Worldwide Interoperability for Microwave Access (WiMAX) [Tane04]. A common feature of these wireless technologies is the presence of a base station (BS) and central control. Users of these wireless access networks expect high quality, reliability, and easy access to high-speed services anytime, anywhere, and in any form.

1.1. Cellular and Multi-hop Cellular Networks

Mobile communications are facilitated by cellular networks. These networks basically consist of mobile terminals, BSs, the radio network controller (RNC) and the core network. A region that is being served is divided into sub-regions, called cells. Each cell is covered by a BS and is allocated a number of channels for mobile terminals to communicate with the BS. A mobile terminal communicates with another mobile terminal, a landline phone of the Public Switched Telephone Network (PSTN), or the Internet through the BS and the RNC. RNC and the core network are developed for third generation (3G) cellular systems to facilitate both voice and data services and to provide better radio resource management, handoff, and security. In first generation (1G) and second generation (2G) cellular systems, the mobile switching centre (MSC) is used instead of the RNC. Figure 1.1 illustrates the system architecture of a 3G cellular system.

In 3G networks, the wideband code division multiple access (W-CDMA) technology is used. The technology allows higher frequency reuse and higher data-rates than that of 1G frequency division multiple access (FDMA) and 2G time division multiple access (TDMA) technologies. In 3G networks, frequency division duplex (FDD) and time division duplex (TDD) modes are available. More discussion on the multiple access technologies and the duplex schemes are in Chapter 3.

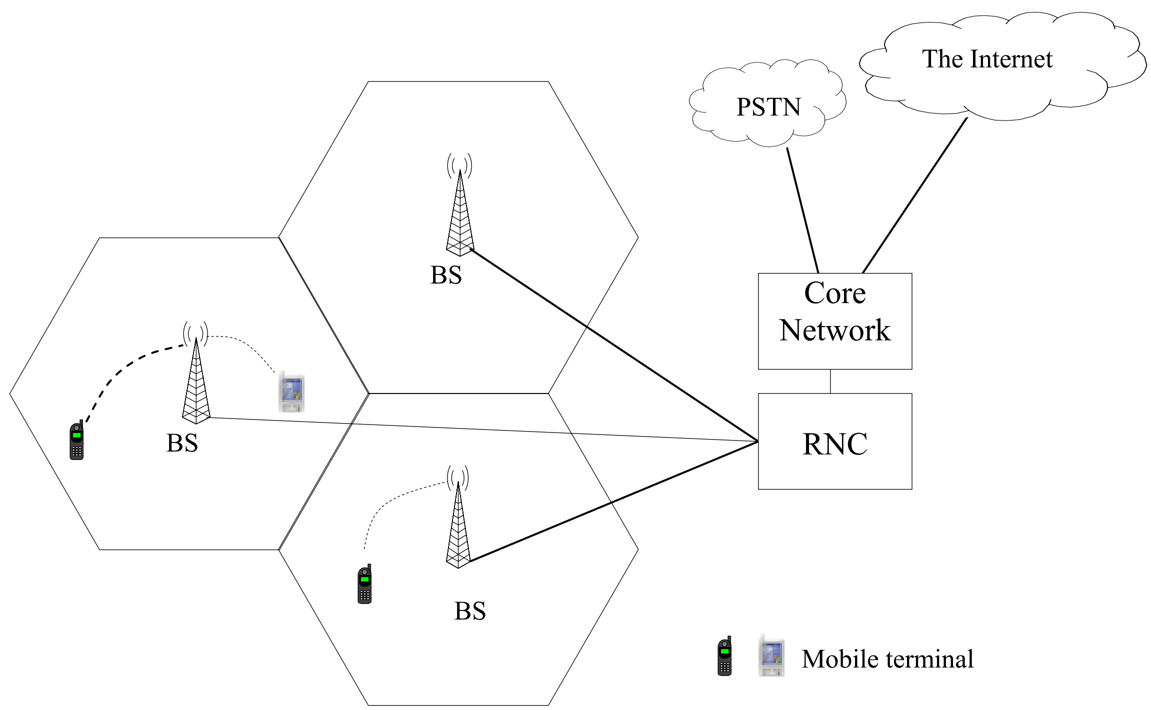


Figure 1.1: System architecture of a 3G cellular system

Limitations and problems of cellular networks

Cellular networks have inherent limitations on cell capacity and coverage. They also suffer from the dead spot problem. Limited capacity also raises the hot spot problem and the issue of radio resource utilization. We discuss these limitations, problems, and issues as follows.

Limited capacity – In a cellular network, the capacity of a cell is limited by the number of channels allocated to the cell. The larger the number of channels, the greater the number of users that can be served. The number of channels is limited by the available frequency spectrums and by the frequency reuse factor [Rapp02]. A smaller cell size allows higher frequency reuse and, thus, a higher capacity can be achieved. In a 3G system, the cell capacity is not only limited by the available frequency spectrums, but also by the interference among mobile nodes and the BSs. The higher the interference, the lower the cell capacity is.

The hot spot problem - Due to the limited capacity, in dense areas known as *hot spots*, such as downtown areas and amusement parks, mobile users tend to experience higher call blocking, i.e., call requests are denied. This is because, in hot spots, there are more mobile users than the number of available channels.

Radio resource utilization - The hot spot problem in turn raises the issue of radio resource utilization. While there are not enough channels or capacity in a hot spot for serving mobile users, the cells neighbouring the hot spot may still have available channels. In other words, the radio resources of neighbouring cells are under-utilized.

Limited coverage – the coverage of cells is limited by the communication range or transmission power of the BS. Mobile users, which are outside the coverage of the BSs, are not able to access the networks.

The dead spot problem - Even though mobile users are within the communication range of the BS, there are still some areas where coverage is not available. These areas are often referred to as *dead spots* such as indoor environments and underground areas.

A possible solution to extend cell coverage and alleviate the hot spot and dead spot problems is to install extra BSs or repeaters in the out-of-coverage region, congested areas and dead spots. However, such a solution is expensive and may not be flexible to adapt the dynamic traffic load conditions in the networks. A multi-hop cellular network (MCN) [Chan03, De02, Kwon02, Lin00, Safw03, Zhou02] can be an alternative complementary solution in cellular systems.

Benefits of multi-hop cellular networks

The idea of MCNs is based on multi-hop relaying. The source node signals are relayed through other intermediate nodes to the BS. The intermediate nodes can be fixed, mobile or ad hoc relays. In this way, the capacity can be enhanced, the coverage can be extended, the hot spot and dead spot problems can be alleviated and the radio resource can be better utilized. Figure 1.2 illustrates a general network architecture of a MCN which consists of source nodes, relaying nodes and the BSs. The use of MCNs

enhances capacity - By using the concept of MCN or multi-hop relaying, the cell size can be smaller, which allows higher system capacity in terms of higher frequency reuse [Rapp02]. A higher transmission rate (cell capacity) due to a shorter transmission range can also be achieved. This is further explained in Chapter 2.

alleviates the hot spot problem - With multi-hop relaying, congestions in hot spots can be alleviated by relaying the traffic from the hot spots to their neighbouring less-congested or non-congested cell through other mobile terminals or relaying devices. The radio resource or available channels in the neighbouring cells can also be utilized.

extends cell coverage – Mobile users, in areas where BS coverage cannot be attained, can relay their messages via one or more mobile terminals and/or special stationary devices that have a direct or indirect link to the BS.

alleviates the dead spot problem – By using multi-hop relaying, mobile users in dead spots are still able to reach the BS through other intermediate nodes.

is economically desirable – In MCNs, the number or the transmission power of the BSs can be reduced (see Figure 1.3). In other words, a simpler infrastructure or cheaper devices can be used. Therefore, MCNs can be more economically desirable. A study of the economic issues of MCNs can be found in [Li08].

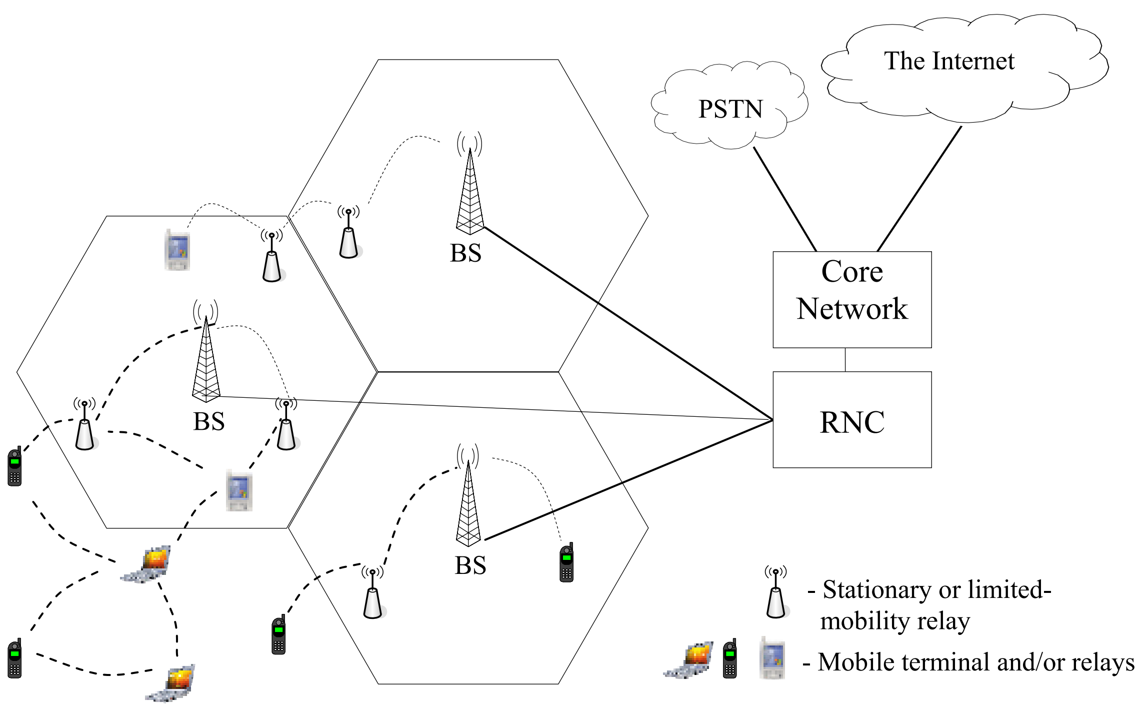


Figure 1.2: Network architecture of a MCN

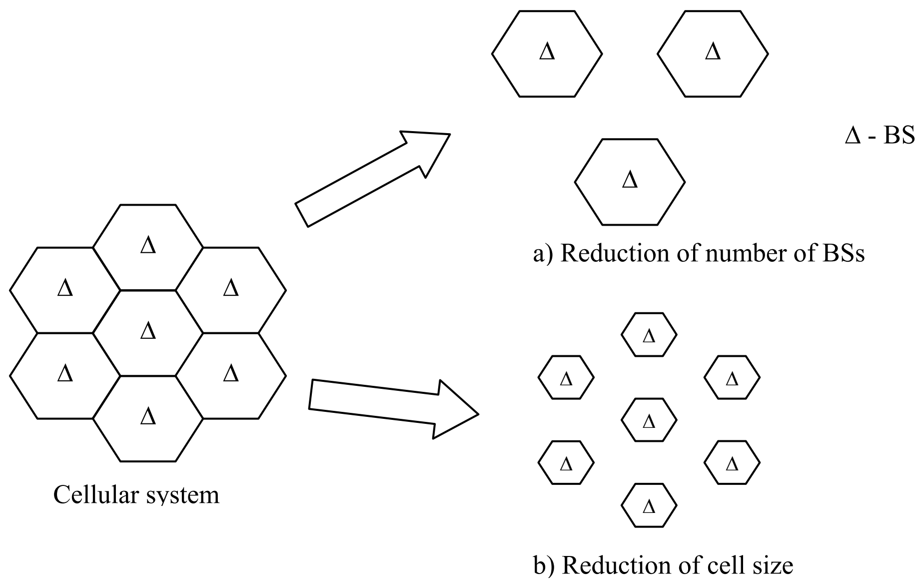


Figure 1.3: Reduction of number of BSs and cell size of a cellular system

1.2. Challenges of MCNs

Although the MCN concept has many benefits, MCNs raise a number of issues, linked to architectural design, the computation of the cell size, maximizing capacity, higher packet delay, channel assignment, routing, load balancing, power control, assuring quality of service (QoS), and providing security. In particular, the cell size is inversely related to the cell capacity and directly related to the coverage. The combined effect of cell size and cell capacity affects the reachability of source nodes and the demands that can be served, which in turn affect the radio resource utilization and system throughput. The demand can be expressed in terms of the number of calls or the amount of data-rates. Although the cell size is an important design factor, little work has been done in this area. Another important issue in MCNs is channel assignment which determines the packet delay and, thus, affects QoS provisioning in these networks.

1.2.1. Cell size

In a cellular network, the cell coverage is basically the cell size or the communication range of the BS. The cell size is inversely related to system or cell capacity. A large cell size decreases frequency reuse in the systems because fewer BSs which reuse the frequency can be installed. In order to avoid signal interference, neighbouring cells cannot use the same frequencies or channels. It also requires high transmission power of mobile nodes and BSs and high interference margins to compensate the interference level. In both cases, the cell capacity would be reduced. On the other hand, a small cell size allows high system or cell capacity because the frequency reuse can be increased and the transmission power as well as the interference level decreases.

In a MCN, the coverage (or network reachability) not only depends on the cell size, but also the node density and network topology. Although a small cell size has high cell capacity, if the node density is low, most distant mobile nodes do not have relaying paths to reach the BS to use the available and abundant capacity. The utilization of the radio resources and the system throughput are low. Thus, some MCN architectures assume a small cell size with high node density. However, in practice, high node density may not always be the case. In addition, the node density may vary during a day. Thus, the assumption may not be applicable all the time. Some MCN proposals assume a large cell size. But, the benefit of capacity gain of using small cell size cannot be achieved. Thus, finding a cell size to provide an optimal balance among the coverage, capacity and demand to maximize the system throughput and the radio resource utilization is important.

1.2.2. Channel assignment

In communications, a channel refers to the medium used to convey information from a sender (or transmitter) to a receiver. In a cellular network, a physical channel is the transmission medium to convey electromagnetic signals from a transmitter to a receiver. In 1G wireless systems, FDMA [Rapp02] is used. A channel is a frequency. As cellular technology evolves, a channel could be referred as a frequency, a time-slot, a code, or a time-slot code pair, depending on which cellular technology is used.

The objective of channel assignment in different contexts is different. In a cellular network, channel assignment involves assigning channels to the mobile nodes for communicating with the BS. It may also involve assigning channels to different cells in order to increase the channel reuse. In a MCN, channel assignment involves assigning channels to the source node and relaying nodes on a relaying path to the BS. In a TDD W-CDMA MCN, a channel is presented by a time-slot and a code pair. When a packet arrives at a relaying node, it has to wait for the time-slot allocated for transmission to the next hop node or the destination node. Improper channel assignment causes the collisions of transmission signals, packet losses, and high packet delay which greatly affect the QoS provisioning in MCNs.

Directional antennas

Directional antennas [Gyod00, Chen04, Dimo08] send signals in a specific direction with a specific beam angle. This helps increase channel reuse and reduce interference. However, when directional antennas are used, the interference patterns in the networks are quite different from those using omni-directional antennas. Channel assignment schemes need to take into consideration the directional antennas environment.

1.3. Thesis Objectives

The objective of this thesis is to design schemes to maximize the utilization of radio resources and to minimize the packet delay in TDD W-CDMA MCNs. In particular, we focus on finding the optimal cell sizes of these networks to maximize the system throughput and finding an optimal channel assignment to minimize the packet delay. W-CDMA is a current and prominent 3G cellular technology. The earlier generation cellular technologies, such as 2G TDMA and 2G TDD CDMA, can be considered as a special case of TDD W-CDMA in terms of channel assignment. This fact makes our channel assignment scheme also applicable to a MCN designed based on earlier generation cellular systems.

As we mentioned earlier in this Chapter, existing cellular systems have inherent limitations on capacity and coverage, have dead spot and hot spot problems, and suffer from the radio resource utilization issue. MCNs help ease the limitations, alleviate the problems and address the issues with an additional benefit of providing faster and cheaper deployment. Although MCNs have many benefits, they also bring along a number of issues. Among them the cell size and channel assignment issues are utmost important because they affect the radio resource utilization and packet delay. As radio resources are scarce, maximizing the radio resource utilization is a common goal of radio resource management which has impact on the profitability of the service provider. High radio resource utilization also implies resource availability to mobile users is higher.

Packet delay affects the QoS which could in turn affect the satisfactory level of mobile subscribers. Thus, it is important to address these issues.

1.4. Thesis Contributions

In this thesis, we propose the following schemes to address the cell size and packet delay issues in a TDD W-CDMA MCN where directional antennas are used. We seek optimized solutions. Note that the optimized schemes may be computationally expensive and inefficient for large real-time (on-line) problems. In this case, heuristic schemes which provide good results compared to the optimal solutions would be more suitable. To this end, we introduce the following schemes.

- Optimal Cell Size (OCS) [Tam08] finds the optimal cell sizes that maximize the system throughput in a TDD W-CDMA MCN. OCS is formulated as an Integer Linear Programming (Integer Programming) problem. OCS can determine the cell size dynamically and can be considered as a network planning aid in cellular networks and MCNs.

- Small Cell Size First (SCSF) is a heuristic for solving the cell size problem. Given a set of possible relaying paths for a source node, each path goes to a different nearby BS of the source node, SCSF chooses the path which requires the smallest cell size and the corresponding BS still have enough capacity to fulfill the demand of the source node. In this way, high system throughput can be achieved. Although SCSF has a better performance than a static cell size strategy when the network is sparse, its throughput performance could be dominated by the source nodes which require large cell size to reach.

- Highest Throughput Cell Size First (HTCSF) is another heuristic for solving the cell size problem. It addresses the limitation of SCSF. Given the same inputs as that of SCSF, HTCSF chooses the path which requires the smallest cell size and the corresponding BS still have enough capacity to fulfill the demand of the source node. It also can drop some connections, which require a large cell size to cover, for a new connection which requires a smaller cell size to cover. In these ways, high cell capacity and system throughput can be achieved.

- Optimal Channel Assignment (OCA) [Tam07] finds an optimal channel assignment that minimizes the system packet delay in a TDD W-CDMA MCN. OCA is also formulated as an Integer Programming problem. OCA can be considered as an un-biased benchmark tool to evaluate the performance of different network conditions, such as node density and network topology, and different networking schemes.

- Minimum Slot Waiting First (MSWF) [Tam06a] is a heuristic for solving the channel assignment problem. Given an input path for a connection of a source node, MSWF finds a channel assignment to minimize the packet delay for the connection along the path.

1.5. Thesis Organization

In the next chapter, the network architecture, design factors, and related issues of MCNs are discussed. Existing MCN proposals are classified based on the design factors. The related work on cell size and channel assignment in MCNs is also discussed. In Chapter 3, the system model is presented. The multiple access technologies and duplex modes used in cellular systems are described. The network model for our schemes and the assumptions for the model for TDD W-CDMA MCNs are presented. Chapter 4 presents the OCA and MSWF schemes which address the packet delay issue in MCNs. The problem definition, formulation of OCA, and the MSWF scheme illustrated with an example are provided. The simulation model, simulation parameters, and performance metrics are described. Simulation results are discussed. In Chapter 5, the OCS, SCSF, and HTCSF schemes are proposed to address the cell size issue. The importance of cell size is discussed. A capacity-demand model illustrating the optimal cell size in a single-cell case is presented. An example to illustrate the optimal cell size in a multi-cell case is provided. The problem definition, formulation of OCS, the SCSF and HTCSF schemes and examples for illustrating the schemes are provided. The simulation model and simulation results are presented and discussed as well. In the last chapter, conclusions of this research work are made and future work is discussed.

Chapter 2

Overview and Related Work

In this chapter, we provide an overview of the network architecture of MCNs and discuss the design factors of these networks, including wireless technology, relaying device, wireless interface, communication mode, and supporting technology. We also discuss existing challenges, such as packet routing, load balancing, power control and security. Then, we explain the importance of cell size with respect to cost, coverage, capacity, user mobility, and handoff, radio resource utilization, routing and packet delay in a cellular or multi-hop cellular systems. The problems and related work on channel assignment in cellular networks and MCNs are discussed. The importance of channel assignment with respect to packet delay in MCNs is also discussed.

2.1. Network Architecture

The architecture of MCNs basically consists of cellular and ad hoc relay components. Signals of mobile nodes are relayed through a relaying device to a gateway device. Signals are then sent through a network controller (a 2G MSC, or a 3G RNC with a core network) to the PSTN, the Internet, or other networks (see Figure 1.2 in Chapter 1). A gateway device is typically a BS. A relaying device can be a stationary dedicated repeater, a wireless router, an ad hoc relay station or a mobile terminal (MT). If the relaying devices are MTs, the MCNs are basically a hybrid of cellular networks and ad hoc networks. Note that the technology of the cellular component of a MCN is not limited to 2G or 3G wireless systems, but could be extended to infrastructure-based wireless technology such as WLAN and WiMAX [Tane04, Tao07]. In fact, there are MCN proposals designed based on WLAN technology or a combination of the cellular and WLAN technologies. Nevertheless, when a hybrid network architecture is used, MCNs gain the benefits and inherit the weaknesses of both infrastructure-based networks and ad hoc networks.

Cellular networks provide large coverage, medium data-rates, and high quality voice and data communications, but require high infrastructure cost and frequency band licensing cost, whereas WLANs provide small coverage, and are widely implemented based on contention-based medium access with no QoS guarantees, but the cost is low because the devices are inexpensive and no frequency band licensing cost is required. WLANs also allow a high data-rate if the number of users is small and interference from other WLANs is low. WiMAX has similar features to those of cellular networks, except the quality of voice service. However, WiMAX has lower equipment cost. For example, the BS of WiMAX is less expensive than the BS of a cellular network. In a 3G network, W-CDMA technology is used. This allows high frequency reuse, but requires power control to minimize the interference among cells and within cells to maintain a high level of cell capacity and to avoid mobiles, which are close to the BS, dominating the reception of the BS.

Ad hoc networks [Toh01], also called mobile ad hoc networks (MANETs), do not require any network infrastructures or central administrations. These networks are formed when mobile nodes are within transmission range of one another. Mobile nodes communicate with each other over the wireless air interface in a peer-to-peer fashion through other intermediate nodes. Figure 2.1 illustrates an example of an ad hoc network. In the figure, although node *C* is outside the transmission range of node *A*, node *A* can communicate with node *C* through the intermediate node *B* or node *D*. These networks have the advantage of flexibility and are cost efficient. They can be deployed anywhere, anytime, with no infrastructure cost. Military operations in a battlefield and emergency rescue operations are two typical examples of their usages. However, they are characterized by frequent network disconnections due to mobility and limited battery life of mobiles. If no route exists, a source node cannot communicate with a desired destination node. Multi-hopping also increases packet delay. Naturally, routing is a major issue in such networks. When designing a MCN, the characteristics, problems and issues, of cellular networks and ad hoc networks need to be considered.

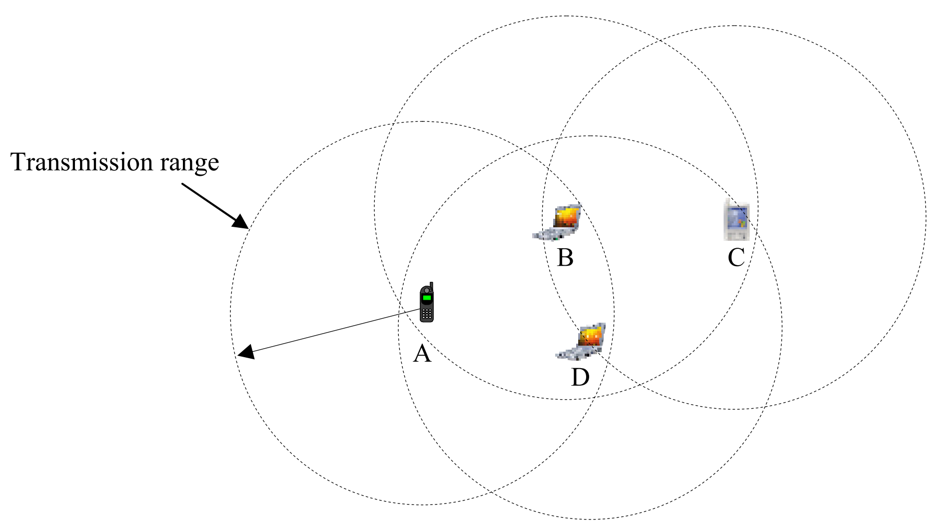


Figure 2.1: An ad hoc network

2.2. Design Factors and Issues

There are a number of decision factors that affect the design of MCNs. These factors include wireless technology, cell size, relaying device, wireless interface, communication mode and supporting technology such as directional antennas and Global Positioning System (GPS). Some of these factors in turn raise some other related issues such as cost and coverage. There are also other issues such as routing strategy, channel assignment, load balancing and security. Figure 2.3 shows the classification of major existing MCN proposals based on the design factors. Table 2.1 depicts the design factors and their related issues. In this section, we discuss the design factors and related issues, except the cell size and channel assignment which will be covered in detail in the next two sections.

Before we start the discussion, we use the Integrated Cellular and Ad hoc Relay (iCAR) [De02] architecture as an example to illustrate the classifications in Figure 2.3. iCAR is a relaying and load balancing scheme for cellular networks. The idea is to place a number of low cost limited mobility ad hoc relay stations (ARSs) in hot spots (congested cells) to relay excessive traffic from hot spots to their neighbouring less congested cells. The traffic is further relayed to outer non-congested cells so that congestion in hot spots is alleviated, the call blocking probability of these cells is reduced, and the load among cells is balanced. iCAR is designed for any cellular systems with the assumption that WLAN technology is used for the relaying component. A medium to large cell size is used. ARSs are equipped with two wireless interfaces: cellular and relaying. The cellular interface is used for communicating with the BS using licensed frequency bands whereas the relaying interface is used for communicating with ARSs using the Industrial, Scientific and Medical (ISM) bands. ISM bands are unlicensed frequency bands that can be used for unlicensed operation in wireless communications. The equipment using these bands needs to tolerate the interference from other ISM band devices. A centralized routing scheme, having a combination of hierarchical and flat routing strategy, is executed at MSCs. No channel assignment scheme is proposed for this architecture. Balancing load among BSs is the main propose of this scheme. No supporting technology, such as GPS or directional antennas, is assumed. Figure 2.2 illustrates the operation of iCAR. Traffic of mobile node *A* is relayed to adjacent BS through the ARSs.

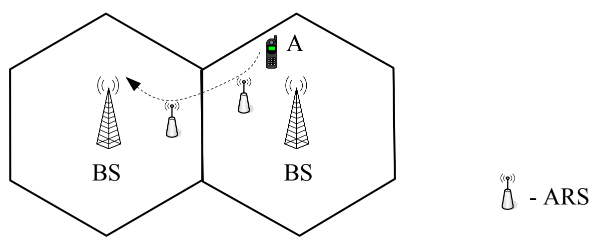


Figure 2.2: Traffic is relayed through ARSs

2.2.1. Wireless technology

Wireless technology is a main design factor for MCNs because it raises a number of important issues including cost, capacity, coverage, QoS, channel assignment, and power control (see Table 2.1).

Choosing cellular technologies for implementing MCNs implies high infrastructure cost, low to medium cell capacity, large to small cell size, good QoS provisioning, and reliable and secure services. A multiple medium access scheme is used and channel assignment is required for all source nodes and relaying nodes. Channel assignment is a challenging task. If 3G technology is chosen, power control issues also arise because a 3G network is based on an interference-limited technology, CDMA. The frequency bands also require licensing and might not be available due to already being used, political, or national security reasons. In addition, the licensing cost is often transferred to end users. Multi-hop Radio Access Cellular (MRAC) [Yama02], Location-Dependent Packet Relay (LDPR) [Li02], Opportunity-Driven Multiple Access (ODMA) [Anti99], Adaptive Multi-hop Cellular (AMC) [Tam06b], Ad hoc Cellular (A-Cell) [Safw03], and Mobile-Assisted Data Forwarding (MADF) [Wu00] are designed based on cellular technology (see Figure 2.3).

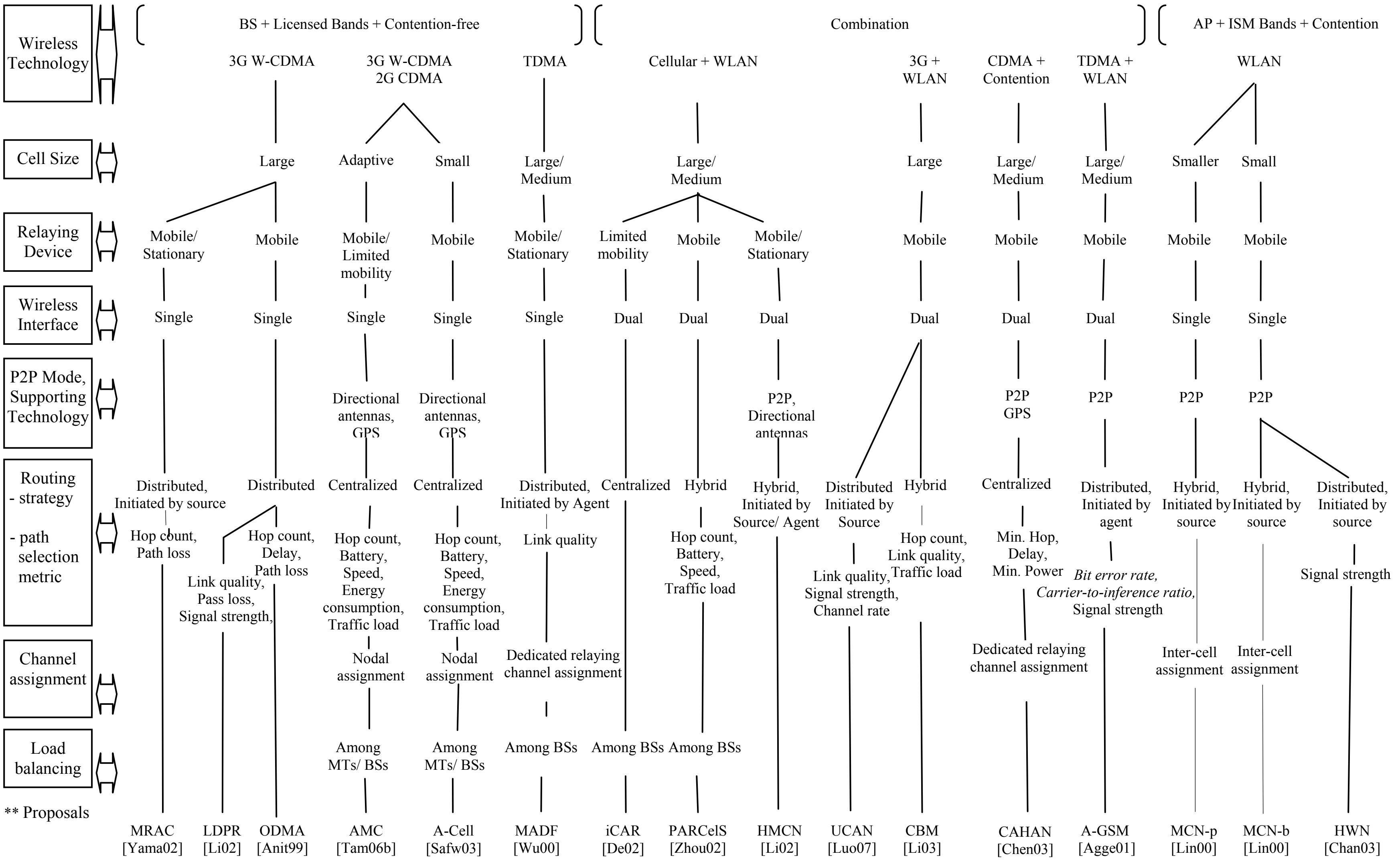
WiMAX is based on the IEEE 802.16 standard. Choosing WiMAX for MCNs has similar benefits and issues to that of cellular technology in terms of cell size, QoS provisioning (except voice quality), and channel assignment, but lower BS cost. WiMAX also provides the option of using license-free ISM bands. IEEE 802.16j [Tao07] is a mobile multihop relay proposal based on WiMAX technology.

WLANs or WiFi is based on the IEEE 802.11 standard. WLAN technology is a low-cost option for MCNs because the BSs (access points (APs)) of WLANs are cheap and no frequency licensing fee is required. Using WLAN technology implies low cost, high data-rates, but small cell size due to restricted transmission power, and more difficult to assure QoS due to the use of ISM bands. If contention-based 802.11 medium access control (MAC) protocol is used, no QoS guarantee can be provided. When there are signal collisions, the senders are required to back-off and retransmit their packets. If there are many users accessing (contending for) the medium, the overhead of medium contention greatly degrades the network performance in terms of throughput, delay and QoS assurance. Multi-hop Cellular Network - power reduction (MCN-p) [Lin00], Multi-hop

Cellular Network - base station reduction (MCN-b) [Lin00], and Hybrid Wireless Network (HWH) [Chan03] are examples of architecture designed based on WLANs technology.

A combination of cellular or WiMAX technology and WLAN technology is another option for a MCN. In this case, licensed frequency bands can be used together with ISM bands. This is especially convenient when extra licensed bands for relaying are not available. Licensed bands are used for cellular access whereas ISM bands are used for ad hoc relaying. An added advantage to such setting is that the signals of the relaying component do not interfere with the signals of the cellular communication. The disadvantage is that QoS of the relaying component, WLAN, cannot be assured due to the use of ISM bands. iCAR [De02], Pervasive Ad hoc Relaying for Cellular Systems (PARCelS) [Zhou02], Hierarchical Multi-hop Cellular Network (HMCN) [Li02], Unified Cellular and Ad hoc Network (UCAN) [Luo07], Cellular Based Multi-hop (CBM) [Li03], and Ad hoc Global System for Mobile (A-GSM) [Agge01] has been designed based on the combination of cellular technology and WLAN technology (see Figure 2.3).

Another option of combining technologies is to assign some dedicated cellular channels for contention-based medium access for the relaying component. In this way, the 802.11 WLAN MAC protocol can be used, but interference from ISM band users cannot be avoided. The Cellular Ad Hoc Augmented Networks (CAHAN) [Chen03] is an example based on this concept.



** Full names of proposals are in the List of Acronyms.

*** Some other proposals that are not listed in the table are SOPRANO [Zade02], CAMA [Bhar04], VCN [Kudo05], and cMCN [Li06].

Figure 2.3: Classification of major MCN proposals based on the design factors

Table 2.1: DESIGN DECISION FACTORS AND RELATED ISSUES FOR MCNs

Wireless Technology	Cell Size	Relaying Device	Wireless Interface	P2P mode	Supporting Technology	Routing Strategy	Channel Assignment	Load Balancing
Cost, Capacity, Coverage, QoS, Channel assignment, Power control	Cost, Capacity, Coverage, Utilization, Throughput, Routing, Mobility, Handoff, Packet delay	Cost, Reliability, Flexibility, Security, Node placement	Cost, Complexity, Interference	Complexity	Cost, Reliability	Delay, Throughput, QoS	Channel reuse, Interference, Delay, Throughput.	Cost, Flexibility, Overhead

2.2.2. Relaying device

A relaying device helps forward the signals of a source node to a BS, an AP or other MTs. The device can be carrier-owned or user-owned.

Carrier-owned devices can be stationary dedicated repeaters, APs, or limited mobility ARSs [De02]. A device has limited mobility means that the device can be set-up and relocated easily to accommodate network access needs for special events, such as parades. Choosing these devices implies that considerable infrastructure, administration, and maintenance cost and node placement planning are required. The flexibility is low. However, more reliable and secure services can be provided.

User-owned devices can be stationary wireless enabled desktops, wireless routers, or MTs. Choosing these devices allows high flexibility at no extra infrastructure cost. No node placement is required. However, the reliability and security diminishes especially when mobile terminals are used because link failures due to users' mobility and/or battery drainage become more frequent and the relaying host may not be trusted.

To decide which types of relaying devices should be used, the wireless network environment should be considered. In 3G networks, users can access services with a wide range of data-rates. The traffic patterns are no longer proportional to the number of users in the cell. Users may have several ongoing connections, each corresponding to a different data-rate for a different QoS class. Such type of users can collectively cause hot spots anywhere, anytime. Carrier-owned relaying devices are inflexible and cannot deal with these hot spots, unless the traffic patterns and network topologies are known a priori. An alternative approach is to utilize user-owned devices such as MTs. In fact, most MCN proposals such as ODMA [Anti99], A-Cell [Safw03], AMC [Tam06b], and PARCelS [Zhou02] assume mobile nodes as relaying devices.

Carrier-owned relaying devices and user-owned relaying devices can probably co-exist. The former can be used to serve the areas where traffic patterns and network topologies are known a priori and/or are predictable. Using these devices allows high reliability which is important for “always on” service provisioning. The latter could be used in a dynamic load network environment, dealing with an unexpectedly high call demand in emergency situations, and providing services that are not sensitive to delay. In fact, user-owned wireless stationary (wall-plugged) devices have great potential because they are more reliable in terms of energy supply and availability. Indeed, many such devices are readily available in cities and residential areas. Note that when user-owned relaying devices are used, incentive schemes, such as credit schemes, may be needed to encourage users to offer signal relaying service. This also raises the issue of cheating for credit [Sale03].

2.2.3. Wireless interface/ communication mode/ supporting technology

The kind of wireless interface, communication mode, and supporting technology impacts on the design of MCNs in several ways. The issues related to these decision factors are cost, complexity, interference, and reliability.

In MCNs, a mobile device may be equipped with a single- or dual- wireless interface. Dual-interface requires two frequency bands, one for cellular access and the other for relay access [De02]. The trade-off is between cost and complexity. Using a single-interface has no equipment cost impact, but signals for relaying cause interference to the cellular access. Using dual-interface reduces system complexity and avoids interference from the relaying component, but equipment cost increases and two frequency bands are required. Sometimes, an extra frequency band may not be available. If free ISM bands are used, interference from other ISM bands users may exist. Both choices are commonly used in MCN proposals. iCAR [De02], PARCelS [Zhou02], HMCN [Li02], UCAN [Luo07], CAHAN [Chen03] and A-GSM [Agge01] are examples of MCNs using dual-wireless interface (see Figure 2.3).

In HMCN [Li02], the peer-to-peer (P2P) communications mode is proposed to help reduce the load of a cell. The trade-off is an increase in the complexity of the system. This mode is useful when the source node and destination node are not too far away from each other in terms of number of hops. Although not many MCN proposals have P2P capability, it could easily be added by slightly modifying existing MCN routing schemes.

Supporting technologies, such as directional antennas and GPS, are assumed in some recent MCN proposals, such as A-Cell [Safw03] and AMC [Tam06b]. Directional antennas [Gyod00, Chen04, Dimo08] help reduce interference and power consumption, increase spatial (channel) reuse and decouple multi-path routes. GPS provides location information of mobile nodes, which reduces routing overhead for obtaining the network topology information [Ko00]. Both technologies require more expensive mobile terminals. The technology of directional antennas for mobile terminal is still in the development stage. The reliability of this technology is still unknown. GPS technology is usually augmented with a wireless sensing component to alleviate the line-of-sight problems.

2.2.4. Routing

Routing is one major issue in MCNs because it affects packet delay and system throughput. When designing a routing protocol, the control strategy and path selection metric need to be decided. Most routing protocols or strategies for MCNs use hybrid control to take advantage of the presence of the BS.

2.2.4.1. Control strategy

As a MCN contains a central controller, such as a BS or an AP, the routing protocol can be centralized, distributed or a combination of the two. In centralized routing, BSs utilizes their unlimited power supply and high computational power to discover and maintain routes. It also helps avoid consuming the precious battery power of mobile nodes for route information exchange and route computation. However, when mobiles are outside of the maximum communication range of a BS or an AP, a distributed ad hoc routing protocol, such as Dynamic Source Routing (DSR) [John96] or Ad hoc On-Demand Distance Vector Routing (AODV) [Perk99], is desirable.

In MCNs, a hybrid routing approach is commonly used. Route control is shared by the BS and mobile nodes. For example, in Cellular Base Routing (CBR) [Li02] and Cellular Based Source Routing (CBSR) [Li03], mobile nodes collect neighbourhood information and send it to the BS for route computation. This helps reduce the route computation overhead at relaying nodes.

In MCNs, not only a source node can initiate a relaying request, a relaying node (or forwarding agent) can also take the initiative by advertising their free channels (available capacity) for relaying [Agge01, Li02, Wu00]. Hence, routing overhead can be shared among source nodes and relaying nodes.

2.2.4.2. Path selection metric

Different MCN routings or relaying proposals have different path selection metrics including BS reachability, hop count, path loss, link quality, signal strength, bit error rate (BER), carrier-to-interference ratio, delay-sensitivity, throughput, power, battery level, mobile speed and energy consumption.

In some proposals, relaying nodes provide the reachability information of BS to neighbouring nodes such that mobile nodes can select the best next hop relaying node to reach the BS. Imposing a hop count limit helps to avoid excessive packet delay, but reduces the chance of distant mobile nodes getting relaying paths to the BS. Choosing paths based on the smallest number of hops also raises fairness and energy efficiency issues. The hop count issue was discussed in the Base-Centric Routing (BCR) [Hsu02] of MCN-p [Lin00]. Some routing strategies are based on the link quality which can be expressed as a function of path loss, signal strength, BER and carrier-to-interference ratio. Delay and throughput are commonly used metrics because they directly reflect the network performance. Minimum power routing is important in CDMA-based MCNs to reduce interference to achieve high cell capacity. Battery level, speed, and energy consumption of the mobile nodes are useful metrics for assuring the reliability of relaying paths. In A-Cell Adapting Routing (ACAR) [Tam05], relaying nodes are chosen based on these metrics.

2.2.5. Load balancing

In cellular networks, a major problem is hot spot (congested area). While mobile users in hot spots tend to experience high call blocking, neighbouring cells of hot spots may still have channels available. Obviously, the radio resource is not fully utilized. Load balancing in cellular networks helps alleviate the hot spot problem by relaying traffic load from congested cells to other less-congested cells. This in turn helps reduce call blocking, utilize the radio resources, and increase the system throughput.

iCAR [De02] and PARCelS [Zhou02] are two load balancing schemes for balancing the load among cells through relaying. In iCAR, low cost limited mobility ARSs are placed in hot spots for relaying traffic out of the hot spots. This strategy is still costly and not flexible enough to handle the highly dynamic load situation in 3G networks. PARCelS uses mobile nodes for relaying. When a BS is congested, mobile nodes search best routes to other non-congested cells. Route information is forwarded to BSs for selection. This strategy requires considerable routing overhead in mobile nodes and does not take advantage of the presence of powerful BSs. In addition, both schemes do not take into account the load balancing among relaying nodes which could greatly affect the load balancing performance.

Balancing load among MTs is important to avoid power over-consumption of some relaying nodes, which affects route availability and connectivity. Although this issue is more related to routing, balancing load among cells and MTs is important to achieve good network performance. For example, during a load balancing process among cells, which source nodes in a highly loaded cell should be chosen for re-routing their traffic to other slightly loaded cells? Of course, this may depend on location of the source nodes and the availability and reliability of the relaying nodes for relaying the traffic. A-Cell load balancing (ALBA) [Tam05] scheme addresses this issue.

ALBA [Tam05] is a dynamic load balancing scheme for CDMA MCNs. The basic idea of the scheme is to shift traffic load as much of the required amount as possible from a highly loaded cell to slightly loaded cells. Note that relaying routes for load migration may not exist especially in a highly dynamic-load network, such as 3G systems.

2.2.6. Other issues

Power Control

Power control adjusts the transmission power of the BS and mobile nodes. It also helps maintain connectivity and minimize power consumption of mobile nodes. In 3G, or CDMA systems, power control is even more important to reduce signal interference and avoid the near-far problem [Holm04]. The near-far problem occurs when a mobile node which is closer to the BS outshouts the other mobile nodes which are far away from the BS. That is, the signals of the mobile node which is closer to the BS dominate the reception of the BS. Reducing interference helps maximize the cell capacity. In MCNs, this issue is more complicated due to the complexity of the transmission patterns. In [Radw06a], the cell capacity gain in a CDMA MCNs system based on perfect power control is quantified. In [Radw06b], the impact of power control on energy consumption of mobile nodes is demonstrated. Power control also helps stabilize the cell capacity, which is important for the call admission control mechanism, which decides whether a call should be accepted or not, to perform admission control effectively.

Security

Similar to ad hoc networks, the security issues in MCNs are important, including secure routing, authentication of users, and security in charging and rewarding schemes for packet forwarding. Unlike ad hoc networks, however, MCNs have a centralized authority for the registration and auditing process. This gives MCNs better ability in preventing and detecting security attacks. In [Xie06], a secure macro/micro-mobility protocol based on multi-hop cellular Internet Protocol (IP) was proposed to prevent various security threats, such as forged BS, unauthorized network access, registration attacks (registration poisoning, bogus registration, and registering replay attack), multi-hop paging/routing cache poisoning, and multi-hop routing attacks (anti-integrity, impersonations, anti-confidentiality, and duplications). Macro-mobility refers to the mobility across local domains whereas micro-mobility refers to the mobility within a local domain. The idea is based on registration and certificate-based authentication. In general, protocols are designed based on light-weight cryptographic techniques, such as symmetric key systems instead of heavy-weight public key systems to avoid high computational overhead. This is especially important in a MCN where mobile devices are resource constrained.

Incentives for relaying

As co-operation in packet forwarding is important for these networks to be successful, various charging and rewarding schemes [Avoi05, Jako03, Luo07, Sale03, Weyl04] have been proposed to encourage users to co-operate in packet forwarding. The issue of selfish nodes which refuse to pay and/or cheat to obtain rewards or free packet delivery services and the issue of false accusation of honest nodes of misbehaviour are also discussed.

2.3. Cell Size

Wireless technology affects the cell size (the coverage of a BS) which in turn affects the cell or system capacity. A small cell has high cell or system capacity and a large cell has low capacity. In MCNs, cell coverage not only depends on cell size, but also on network topology and/or node density. The coverage and capacity of a cell affects the total demand that can be served and, thus affects the cell throughput. The relation makes the cell size an important design factor for MCNs.

2.3.1. Cell size in wireless networks

Determining optimal cell sizes of different cells dynamically in a cellular network is a challenging problem especially in a 3G system where a cell's coverage is inversely related to its capacity. Cell size of a wireless network also brings other issues, such as cost, user mobility, and handoff. These issues are discussed below.

2.3.1.1. Cell size versus cost

In wireless networks, cell size depends on the wireless technology that is chosen which in turn affects the cost. In cellular networks, since frequency bands are licensed, the BS is allowed to transmit with high power. Hence, a large cell size can be achieved. However, the cost of frequency spectrum licensing and equipment is also high. A cheaper alternative is to use WLAN technology which uses ISM bands that require no frequency licensing fee. Also, an AP (the BS of a WLAN) is inexpensive. Furthermore, a high data-rate can be achieved. A limitation of WLAN technology is that the cell size can only be small.

2.3.1.2. Cell size versus coverage

In wireless networks, the coverage of a cell is basically the cell size or the communication range of a BS or an AP. In cellular networks, a large cell size is allowed. Cellular communication service providers have the option of using a large cell size or a small cell size to suit the capacity and/or the coverage needs. But, in WLANs, the cell size is restricted to be small to avoid significant interference among ISM band users. Small cell size could increase the chance of handoff or disconnection of an ongoing call if the mobile users have high mobility. In fact, WLAN is designed for indoor and low mobility usage.

2.3.1.3. Cell size versus cell capacity

Cell size of a wireless network not only affects the cell coverage, but also the cell or system capacity. The cell capacity is usually measured as number of admitted calls or total available data rates. A large cell has low capacity and a small cell has high capacity.

In a bandwidth-limited system, such as 2G TDMA or WiMAX orthogonal frequency division multiple access (OFDMA), a small cell size allows higher frequency reuse among cells and, hence, a higher system capacity. However, the coverage (network reachability) is reduced.

In an interference-limited system, such as 2G or 3G CDMA system, mobile users send signals on the same frequency, but use different orthogonal codes. In this case, the signals of a user become the noise or interference signals to the others. Thus, a portion of the signal power of a mobile terminal is required to overcome the noise whereas a part of the signal power is used to overcome the path-loss [Rapp02] (propagation loss or signal attenuation) along the signal path. The higher the number of mobile users or the data rates, the higher the signal power margin (interference margin [Holm04]) is required to overcome the noise. Thus, a lower signal power margin remains for the propagation loss. That is, the communication range or cell size needs to be smaller. Thus, a small cell has a high capacity whereas a large cell has small capacity. This is one main motivation behind the ODMA [Anti99] which was proposed for the 3G system.

2.3.1.4. Cell size versus user mobility and handoff

Cell size computation should account for user mobility too. The decision of cell size of a cellular network is influenced by user mobility. The combined effect of cell size and user mobility affects the frequency of handoff of ongoing calls of mobile users moving from one cell to an adjacent cell. The handoff requires a channel change when a mobile user moves from one cell to another. If the handoff involves changing over two different frequency bands, the mobile user may feel the disturbance of the call.

In an environment where user mobility is high, using a large cell size reduces the frequency of handoff because the chance of mobile users move outside of the cell is lower. This also helps assure the quality of the connections.

2.3.1.5. Static cell size

The cell size of a cellular network is usually decided during a planning stage. A service provider may use a high BS transmission power to achieve a large coverage or a low BS transmission power (small cell size) to achieve high capacity depending on the characteristics of the service areas. For example, in a city center, high capacity is preferred because the density of mobile subscribers is high. On the contrary, on a highway, a large coverage may be needed because of high speed of users. In addition, a static approach cannot cope with the dynamic nature of users in a 3G network in which mobile users are provided with wide range of data-rates service. The capacity demand does not only depend on the mobile user density of the service areas, but also the services that the mobile users use.

2.3.2. Cell size in MCNs

The problem of determining the optimal cell size in a MCN is more involved than in a cellular network because the coverage in a MCN not only depends on the cell size, but also on the availability of mobile nodes for setting up relaying paths which in turn depends on the network topology and node density.

2.3.2.1. Cell size versus coverage and radio resource utilization

To achieve high cell capacity without deteriorating the coverage or network reachability, some MCN proposals, such as MCN-p [Lin00] and A-Cell [Safw03], assume a small cell size and a dense network. A small cell size is used to achieve a high cell capacity whereas a dense network provides sufficient relaying nodes for relaying signals. However, dense networks may not always occur in practice. When the network is sparse, the performance of MCNs could be greatly degraded because distant mobile nodes may not find relaying paths to relay their signals to the BS to use the available or abundant capacity.

MCN proposals such as iCAR [De02] and PARCelS [Zhou02] do not specify the cell size, leaving an implication that a general large or medium cell size is used. However, the cell capacity with these sizes may not be able to meet the demands in a cell. Having a large cell size reduces the ability to increase the cell capacity to meet high demand. Thus, a static cell size strategy cannot adapt the network topology and traffic pattern to maximize the radio resource utilization and the system throughput.

2.3.2.2. Cell size versus routing and packet delay

In MCNs, the cell size not only affects the cell capacity and the coverage, but also the routing efficiency and the packet delay.

In a MCN, BSs usually take part in route computation for the relaying paths because they are more powerful and have unlimited energy supply. To compute the route, BSs need to collect information about mobile nodes such as their locations and traffic load. After routes are computed, the route information needs to be sent back to the mobile nodes for execution. To reduce the latency of communicating routing information between BSs and mobile nodes, relaying paths having fewer hops are preferred. A larger cell size requires a smaller number of hops of a relay path for signals of distant mobile nodes to reach the BSs. Thus, for routing purposes, a larger cell size is preferred because it helps increase the speed of route discovery and maintenance and, hence, the efficiency of routing. In addition, a smaller number of hops also reduces packet delay, which is especially important for delay-sensitive applications.

Although a large cell size is preferable for routing or delay-sensitive applications, a large cell size results in low cell or system capacity. In a dense network, a small cell size is preferable for achieving high capacity to meet high demands as the network reachability is not an issue. ACAR [Tam05] was designed based on this idea. For ACAR, a large cell size is used for routing discovery and maintenance. After a route is set-up or updated, a small cell size allowing a higher capacity through multi-hop short-range relaying path is used for actual data communications.

Summary

In a cellular network, cell size affects the cost, cell coverage, cell capacity, mobility and handoff. The combined effect of cell size and cell capacity in turn affects the total demands that can be served and, hence, the radio resource utilization and system throughput. A static cell size strategy can not adapt the dynamic nature of 3G cellular networks.

In a MCN, cell size affects the coverage, radio resource utilization, routing efficiency and packet delay. Given a network topology, network density, and traffic patterns in a MCN, how to achieve a good balance between cell size and cell capacity to achieve maximum throughput is an important question. The cell size issue has not been addressed until we recently introduced an optimal cell size concept in [Tam06b]. The details of the concept are discussed in Chapter 5.

2.4. Channel Assignment

The ultimate goal of channel assignment is to maximize the radio resource utilization. However, the issues related to channel assignment may be quite different in different contexts. In the following sections, we discuss such issues along with some existing channel assignment schemes in cellular networks and MCNs.

2.4.1. Channels in cellular systems

Before discussing channel assignment, we briefly explain what a channel is in our context. As mentioned in Chapter 1, a channel can be a frequency, a time-slot, a code, or a time-slot code pair depending on which wireless technology is used.

In 1G wireless systems, FDMA is used. An allocated frequency spectrum for the systems is sub-divided into frequency bands. A mobile user is assigned a pair of frequencies for a full duplex communication with the BS. In this case, a channel is a frequency band.

In a 2G wireless system, in which TDMA is used, each frequency band in the FDMA system is divided into a number of time-slots. Each user is assigned one or more time-slots for communication. This access method not only increases the utilization of the bandwidth, but also the number of mobile users that can be served and the flexibility of providing different data-rates for different services.

In 2G narrowband CDMA or 3G W-CDMA systems, a user or a connection is distinguished from other users or connections by using a different code. The information bit of a data unit of a mobile user is spread over a transmission medium by a spreading code [Holm04]. All the users can transmit their signals to their BSs at the same time on the same frequency such that they can be distinguished by the BSs. In this way, the frequency spectrum can be fully reused in every cell in the systems. The trade-off is that the signals of one user become the noise of the other users. The noise needs to be minimized. In these systems, a channel is represented by a code. If the TDD mode is used, a channel is represented by a combination of a time-slot and a code.

2.4.2. Channel assignment in cellular networks

In cellular networks, channel assignment may be viewed as a channel reuse problem or as a switching point problem.

2.4.2.1. Channel reuse problem

The goal of channel reuse is to assign a minimum number of channels to every requested call in the cells such that interference constraints are satisfied. Channels or frequencies can be reused among cells as long as the distance among those cells is sufficiently large such that the signal interference level among the cells does not exceed the required level. The problem is NP-Complete. [Fu06] provides a genetic algorithm to solve this problem.

2.4.2.2. Switching point problem

In reference [Noor04], the channel assignment problem is presented as a switching point problem. A switching point is a position in a TDD transmission frame at which the direction of signal transmission is reversed. For example, an uplink (from mobile node to BS) transmission is changed to a downlink (from BS to mobile node) transmission and vice versa. Improper switching point assignment reduces the utilization of the channels. In a TDD CDMA system, a channel is a time-slot code pair. Three channel allocation schemes, Fixed Channel Allocation (FCA), Dynamic Channel Allocation (DCA), and DCA with Adaptive Switching Point (DCA-ASP), are discussed [Noor04]. These schemes are designed for a 3G TDD CDMA system. For both FCA and DCA, only a single switching point in a transmission frame is considered (see Figure 2.4). The switching point is set and fixed when the system is initialized. Thus, the number of uplink slots and downlink slots and their relative positions in a transmission frame is fixed. We briefly discuss these schemes as follows.

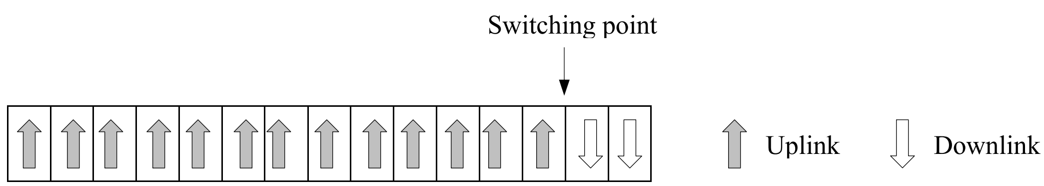


Figure 2.4: A switching point in a TDD CDMA transmission frame

For FCA, channels are assigned without considering the channel quality of time-slots. FCA assigns channels (time-slots) randomly to each call (connection) based on the direction (uplink or downlink) of the traffic flow of the call on a first-come, first-serve fashion. Thus, a call request may be assigned a time-slot having bad channel quality, e.g., high BER. In this case, the number of good packets that can be successfully delivered is small.

For DCA, a channel or times-slot is assigned to a traffic flow based on the quality of the channel. This helps avoid assigning a bad channel to a traffic flow such that the packet is corrupted. Still, the position of the switching point is predetermined when the system is initialized. Therefore, DCA cannot adapt dynamic traffic patterns which could be symmetric or asymmetric in terms of uplink and downlink transmission. In a highly dynamic traffic environment, without a proper channel assignment scheme, the radio resource could be greatly under-utilized.

To address the limitations of DCA, DCA-ASP was proposed [Noor04]. DCA-ASP supports the movement of multiple switching points to dynamically adjust the bandwidth to suit the traffic for uplink and downlink.

2.4.3. Channel assignment in MCNs

In MCNs, channel assignment may involve assigning channels among cells, which we call *inter-cell channel assignment*. Channels assigned in adjacent cells should be different to avoid signal interference. Channel assignment may also involve assigning channels inside a cell, which we call *intra-cell channel assignment*. Furthermore, channel assignment may involve assigning channels dedicated for relaying, which are called *dedicated relaying channel assignment*, and assigning channels among mobile nodes, including source nodes and relaying nodes, which we call *nodal-channel assignment*. Intra-cell channel assignment may involve dedicated relaying channel assignment and/or nodal channel assignment. These are basically channel reuse problems rather than switching point problems.

The choice of wireless technology or medium access strategy influences the decision of whether an inter-cell and/or a nodal channel assignment are needed. If a contention-based medium access scheme, such as WLAN Distributed Coordination Function (DCF), is used, only inter-cell channel assignment is required. MCN-p [Lin00] and MCN-b [Lin00] are designed based on WLAN technology. If a multiple access scheme, such as 2G TDMA or 3G W-CDMA, is used, inter-cell channel assignment and/or intra-cell channel assignment are required.

Inter-cell channel assignment

As mentioned in Section 2.4.2.1, channel assignment in a cellular network usually deals with assigning channels (frequencies) to cells of the network to maximize channel reuse. Each cell is assigned a number of channels which are different from those assigned to its adjacent cells to avoid interference. Each mobile node communicates with its own BS. As a cellular network is merely a special case of MCN, channel assignment among cells or inter-cell channel assignment is exercised in MCNs.

Dedicated relaying channel assignment

Some MCN proposals, such as MADF [Wu00], set aside some dedicated channels for relaying. As the relaying channels are different from the cellular channels for BS communication, this helps avoid the signals in the relaying channels interfering the signals on the cellular channels. However, the question of how many channels should be set aside is raised. Improper channel assignment increases the chance of channel idling of forwarding channels and, hence, wasting of radio resource. In other words, dedicated forwarding channel assignment may not maximize the channel reuse.

Nodal channel assignment

In MCNs, source nodes may communicate with the BS directly or indirectly through other intermediate relaying nodes. Thus, channels are needed to be assigned to source nodes and relaying nodes.

Some existing nodal channel assignment schemes are the A-Cell Channel Assignment (ACA) [Safw04], Random Slot Assignment (RSA) [Alri05], Delay-Sensitive Slot Assignment (DSSA) [Alri05], and Fixed Channel Assignment [Li06].

ACA, RSA, and DSSA are designed for A-Cell [Safw03] which is based on TDD W-CDMA technology. A-Cell also uses directional antennas to increase spatial reuse and reduce interference and power consumption.

ACA is formulated as an Integer Programming problem to optimize the channel reuse in an A-Cell. Given a network topology and traffic patterns in an A-Cell, the task of ACA is to find the minimum number of channels required to satisfy the call requests without causing signal collisions.

In reference [Li06], the Fixed Channel Assignment scheme and the clustered multihop cellular networks (cMCNs) were proposed. The channel assignment scheme is designed for the cMCN which consists of mobile nodes as cluster heads surrounding a cell. The clustered heads communicate with the BS to obtain the information of channel availability and assign channels to the relaying nodes. The channel assignment scheme neither provides an optimal solution in channel reuse nor does address of the delay issue.

Packet delay

In MCNs, as multi-hop relaying is involved, packet delay increases. Improper channel assignment would greatly increase packet delay. ACA does not address the packet delay issue.

In reference [Alri05], the delay issue with respect to channel assignment in MCNs is addressed and two channel assignment schemes, the RSA and DSSA, were proposed. RSA assigns channels or time-slots randomly. Obviously, RSA cannot guarantee minimum delay because the randomness of the slot assignment does not minimize packet delay. To address the delay issue, DSSA [Alri05], a heuristic channel assignment scheme, is designed. Simulation results show that DSSA outperforms RSA in terms of throughput and packet delay. Although DSSA lowers packet delay, it does not guarantee minimum packet delay.

Directional antennas

Directional antennas [Gyod00, Chen04, Dimo08] help reduce interference and power consumption, increase channel reuse and decouple multi-path routes. They were proposed in A-Cell [Safw03] to increase channel reuse. Our MCN proposal, AMC [Tam06b], is also based on the use of directional antennas with the optimal cell size model in a single-cell case. Although directional antennas have their merits, the interference patterns of the networks using directional antennas are quite different from those using omni-directional antennas. When designing a channel assignment scheme for the network in which directional antennas are used, the interference patterns need to be considered.

2.5. Summary

In this chapter, we discussed the network architecture, the design factors, and related issues of MCNs. The design factors are wireless technology, cell size, relaying device, wireless interface, communication mode, and supporting technology. The related issues are cost, capacity, coverage, routing, channel assignment, load balancing, and security. A classification of existing MCN proposals based on these design factors is presented. An in-depth discussion on cell size and channel assignment is also provided.

We showed that the wireless technology is the most important design factor for MCNs because it raises a number of issues, such as cost, capacity, coverage, QoS, channel assignment and power control. If cellular technology is used, the cost is high, the data-rate is medium, and channel assignment is required. But, the cell size is more controllable and the QoS can be assured. If WLAN technology is used, the cost is low and high data rates can be achieved. However, the cell size is limited to small and QoS assurance is difficult to achieve. In addition, small cell size requires high node density to ensure network connectivity for distant nodes. But, high node density may not be the case in practice. Thus, most MCN proposals are designed based on cellular technology or a combination of cellular and WLAN technology.

For cell size, we discussed its related issues in cellular networks and MCNs. In cellular networks, cell size affects the cost, the coverage, capacity, mobility, and handoff. For example, in 3G systems, the cell size is basically the coverage and is inversely related to the cell capacity. We also mentioned that static cell size strategy cannot cope with the dynamic of traffic load in a 3G network. In MCNs, cell size not only affects cost, capacity and coverage, but also the radio resource utilization, throughput, routing efficiency, and packet delay. The cell size and node density together affects the coverage or network reachability which in turn affects the total demand that can be served and, hence, the radio resource utilization and system throughput. Cell size affects the routing efficiency because the number of hops of a relaying path is reduced as cell size increases. Little research work has been done on the cell size issue.

For channel assignment, although its ultimate goal is to maximize radio resource utilization, the issues related to channel assignment could be quite different in different contexts. In cellular networks, channel assignment can be viewed as a channel reuse problem among cells or as a switching point problem within a cell. In MCNs, in addition to channel reuse among cells, channel assignment also involves channel reuse among mobile nodes. Furthermore, channel assignment in MCNs could greatly affect the packet delay. In fact, no existing channel assignment schemes are designed to minimize the packet delay in a TDD W-CDMA MCN.

Chapter 3

System Model

In this chapter, we provide a system model for a MCN environment with TDD W-CDMA. We discuss the duplex schemes and multiple access schemes of cellular networks. We discuss the system components and the interactions among them. We present a network model for TDD W-CDMA MCN. The model can be applied to omni-directional or directional antennas in a MCN environment. The assumptions of the model are also provided.

3.1. System Features

Duplex schemes

A two-way communication can be facilitated by a half-duplex or full-duplex communication method. For a half duplex communication, a channel is shared by two users who communicate with each other. When one user uses the channel to speak, the other user cannot use the channel to speak and can only listen. In other words, the two users need to take turns to speak. A full duplex communication method allows simultaneous traffic sending in both directions by using two channels. Each user uses one channel to speak and the other one to listen. In this way, both of them can speak at the same time.

In cellular systems, there are two types of full duplex communication schemes: FDD and TDD [Rapp02]. For FDD, two frequencies are used: one frequency for uplink (from a mobile node to the BS) transmission and the other frequency for downlink (from the BS to a mobile node) transmission. For TDD, at least two time-slots are required: one for uplink communications and the other one for downlink communications.

Figure 3.1a shows the operation of FDD. In the figure, frequencies *A* and *B* are respectively used for uplink and downlink communications. In Figure 3.1b, time-slots 1 and 3 are used for uplink communications whereas the time-slots 2 and 4 are used for downlink communications. For a MCN, TDD is preferred to FDD because TDD does not need to synchronize the two frequencies for each mobile node whereas FDD does.

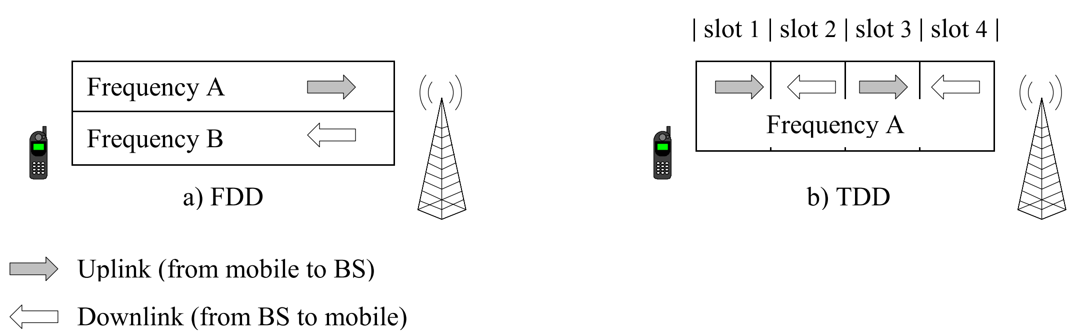


Figure 3.1: Operation of FDD and TDD

Multiple access schemes

In cellular networks, there are three major types of multiple access schemes for mobile nodes to access the radio resources of the systems. The schemes are FDMA, TDMA and CDMA.

For FDMA, the licensed frequency band is divided into sub-frequencies. Each sub-frequency is allocated to a user as a channel for communication. To achieve full duplex communications, a user is allocated a pair of frequencies, one for uplink and one for downlink, for communicating with the BS. FDMA is used in 1G cellular systems which are mainly designed for voice communications. Obviously, using FDMA does not fully utilize the capacity of the channel because users at both ends of the communication link seldom speak at the same time. In other words, the channel is idle when one of the users is listening instead of speaking. Figure 3.2a illustrates the FDMA design. In the figure, there are three frequencies; each represents one channel. To increase the utilization of the radio resources, TDMA is introduced.

TDMA can be considered as an overlay scheme on FDMA. For TDMA, each frequency or channel of FDMA is divided into time-slots. Each user is allocated one or more time-slots. This not only avoids the wastage of the channel capacity due to the intermittent idling of the channel during the user conservations, but also increases the number of users that can be served. Furthermore, this provides flexibility in providing different data-rates for the users for data-services, such as short messaging and internet browsing. TDMA is a mainstream multiple access scheme that is used in 2G cellular systems. Figure 3.2b illustrates the concept of TDMA. In the figure, the duration of a frequency (channel) is divided into 5 time-slots; each time-slot represents one channel.

Although TDMA increases the utilization of the radio resources, like FDMA, frequency reuse among cells is still not maximized as neighbouring cells cannot use the same frequency bands. The higher the frequency reuse factor, the higher the utilization of the radio resources. To increase the frequency reuse, CDMA technology is used. In CDMA, each user is assigned a code (a spreading code) which is orthogonal to other codes in terms of coding and decoding. In this way, signals from or for different users can be distinguished. By using CDMA, all the cells in the system can reuse the whole frequency band. Thus, the frequency reuse is maximized. However, since all the users are transmitting on the same frequency band, though by using different codes, the signals of one user become the noise of the other users. The interference level should be controlled to a minimum to avoid affecting the quality of the signals of each user. This interference issue is handled by a power control mechanism in CDMA systems. Narrowband CDMA is used in some 2G cellular systems. Figure 3.2c illustrates the CDMA method. In the figure, all four codes are transmitted on one frequency band and the duration is not sub-divided into time-slots.

TDD W-CDMA

W-CDMA is the air-interface technology for 3G wireless systems. It is basically a CDMA technology in which a wider band than that of 2G is used. A wider band allows a higher capacity in terms of number of users that can be served and/or data-rates that can be provided. In a 3G system, the data-rate can be up to 2Mbps. TDD W-CDMA is one of the full duplex modes of 3G systems. This mode has a better radio resource utilization especially for asymmetric traffic as more time-slots can be allocated for either downlink or uplink traffic. In addition, in a MCN environment, TDD makes the design of transceiver simpler than FDD. Using TDD avoids the complications of coordinating the frequencies being used for uplink and downlink transmissions. Note that 2G TDMA and 2G TDD CDMA can be considered as special cases of TDD W-CDMA in terms of channel assignment. Thus, in our system model, a TDD W-CDMA access scheme is used. In fact, there is a proposal using TDD CDMA for future generation wireless networks [Esma03].

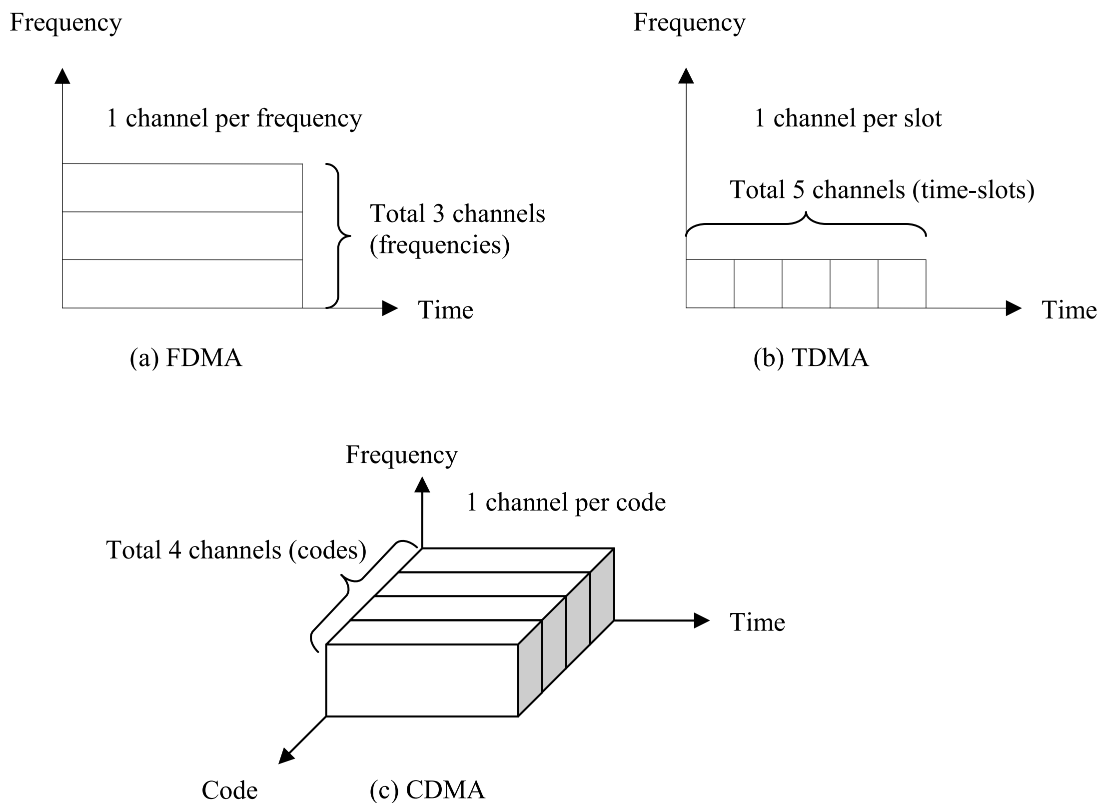


Figure 3.2: Three multiple access techniques: a) FDMA, b) TDMA and c) CDMA

According to the 3G Universal Mobile Telecommunication Systems (UMTS) standard [Holm04], the TDD W-CDMA mode has 15 time-slots, each slot can have up to 16 codes depending on the spreading factor that is used. Figure 3.3 illustrates an example of a transmission frame of the 3G UMTS TDD mode. In this example, 13 slots are allocated for uplink transmission and 2 slots are allocated for downlink transmission.

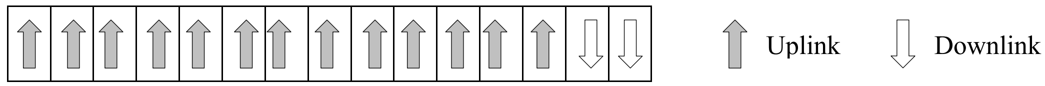


Figure 3.3: A TDD transmission frame of 3G UMTS

Power Control

As mentioned in Section 2.2.6, power control helps maintain network connectivity, minimize power consumption of mobile nodes, reduce signal interference to maximize the cell capacity, and address the near-far problem [Holm04]. In 3G networks, there are open-loop power control and closed-loop power control mechanisms. In open-loop power control, the BS is not required to communicate with mobile nodes to adjust their transmission power. Thus, their transmission power may not be as accurate as required. However, open-loop power control helps provide a rough estimate of initial power setting of a mobile node at the beginning of a connection. In closed-loop power control, the BSs regularly monitor the interference conditions and command the mobile nodes to adjust their power if necessary. In MCNs, the interference pattern is more complex as mobile nodes not only transmit signals to the BSs but also to their relaying nodes.

Directional antennas

Directional antennas help focus the transmission power in a particular direction with a specified beam-angle [Gyod00, Chen04, Dimo08]. In this way, the signal interferences among mobile nodes can be reduced. This also increases channel reuse. Power consumption of mobile nodes can also be reduced as the transmission power of mobile nodes is directed in a specified direction and the beam is set at a certain angle. The interference pattern is quite different from the case in which omni-directional antennas are used. For example, neighbouring nodes of a transmitting node are not interfered by the signals of the transmitting node if they do not fall in the transmission zone of the transmitting node. Thus, channel assignment scheme needs to be developed to adapt to this type of environment.

3.2. System Components

The basic components for the system are admission control, routing, channel assignment and cell size determination. Figure 3.4 shows the interactions between these components are explained.

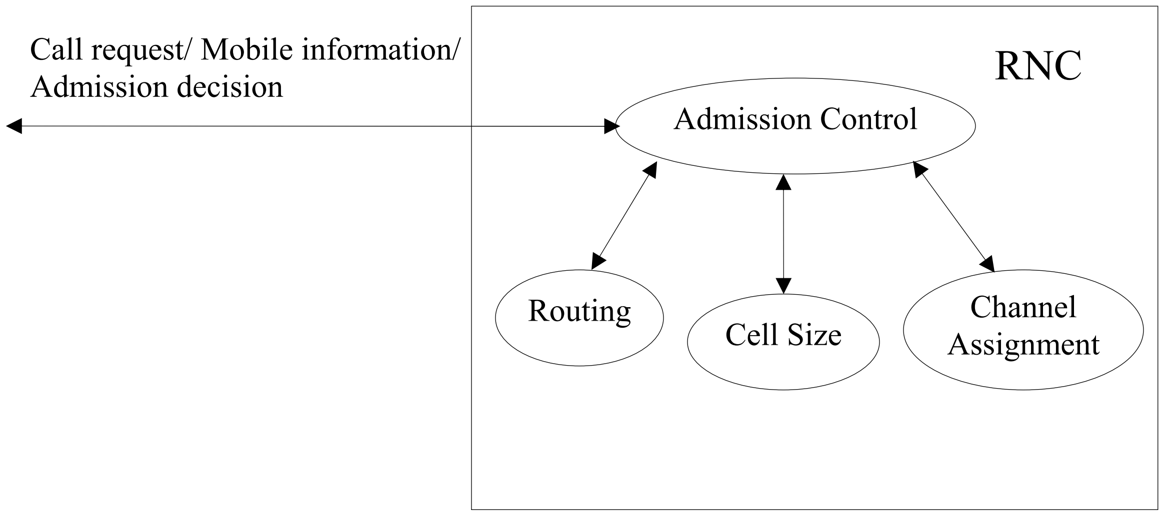


Figure 3.4: Interactions between the cell size and channel assignment components

When a mobile node initiates a call request, the request is sent to the BS, which passes the request to the call admission control algorithm at the RNC. The RNC calls the routing algorithm to find the relaying paths for the source node. If a centralized routing scheme, such as ACAR [Tam05], is used, information such as location information of mobile nodes provided by GPS can be collected through control channels, such as the Common and Dedicated Control Channels [Holm04], of the BS. The information is passed to the routing algorithm at the RNC to compute the paths. Note that other information or parameters may be collected depending on the routing protocol (see Section 2.2.4.2). Possible paths from the source node to the BS are computed and passed to the cell size algorithm for selection based on the cell sizes and the capacities of the cells. If there is a valid path from the source node and the BS has enough capacity, the path information is passed to the channel assignment algorithm for assigning channels to each node on the path; otherwise, the admission control function sends a signal to the source node to deny the call request (block the call). If channel assignment succeeds, the routes, cell sizes, and channel assignment information are passed to the BS which in turn sends the information to the mobile nodes through the Common and Dedicated Control Channels [Holm04]. BSs also adjust the cell sizes based on the computed cell sizes. Note that for the OCA and OCS schemes, all the existing paths are required for the computation, whereas for the MSWF, SCSF, and HTCSF schemes only the path pertaining to the source node is required.

The RNC continues to monitor the network conditions. If a route is broken or the link quality of a route is below the required level, new route discovery is triggered. If a new route is found, the cell size and channel assignment functions are activated to compute the new cell sizes and channel assignment for the source nodes.

3.3. Network Model

Figure 3.5 illustrates a typical TDD W-CDMA MCN environment in which directional antennas are used. The triangular regions represent the transmission zones of directional antennas. In the figure, there are source nodes, source points, relaying nodes, relaying points, BSs, and relaying paths. When a source node makes a call request, a relaying path is set up to relay signals for the call. For example, H-I-G-b_1 is a relaying path in which node H is a source node and nodes I and G are relaying nodes. A call can be considered as a connection, traffic flow or session. We represent the initiation point of a connection

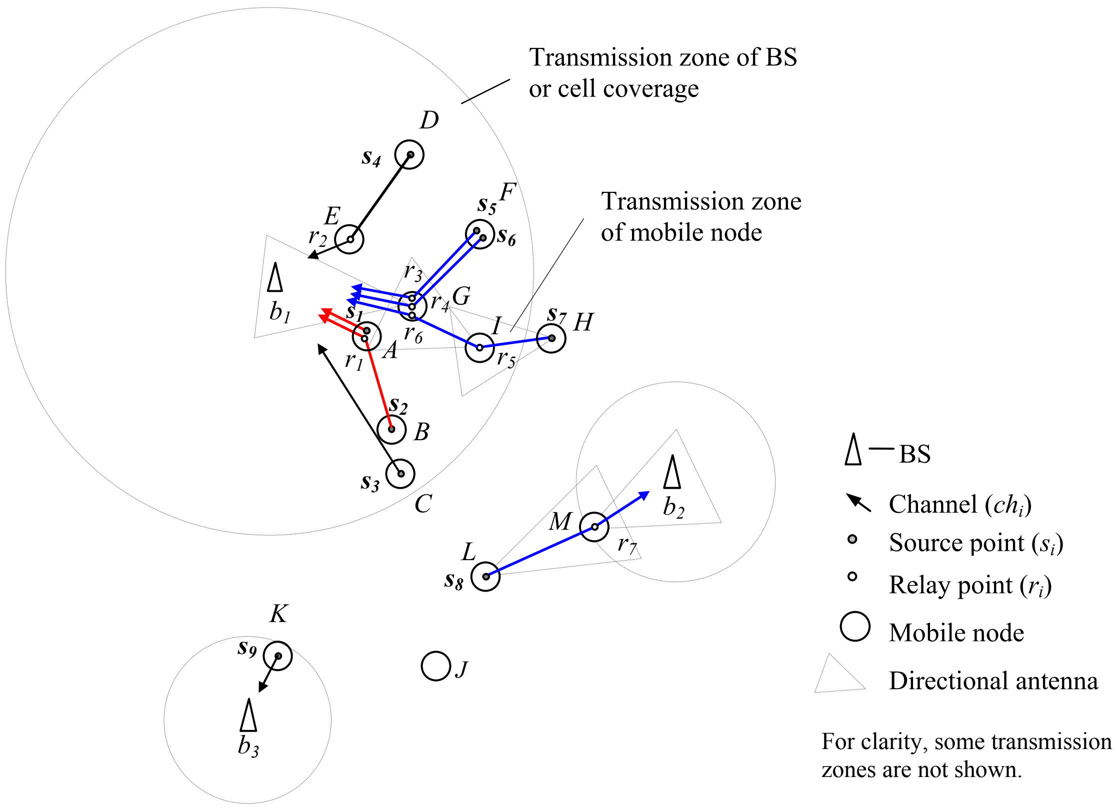


Figure 3.5: Topology in a MCN with directional antennas

in a source node by a virtual point called source point s_i. Each source point s_i is assigned exactly one channel with a demand of traffic $\mu(s_i)$. As TDD W-CDMA is used, a channel is a time-slot t and code c pair. If TDMA is used, a channel is represented by a time-slot. Each source node may have several source points. For example, in Figure 3.5, source node F has two source point s_5 and s_6. Each source point supports a different connection for a different service. Each connection is relayed by a relaying point r_j in each relaying node on a relaying path. A relaying node may have several relaying points; each relays a different connection. A relaying node can be a source node itself. For example, source node A is also a relaying node. A mobile node may choose to relay or not to relay signals for other nodes. For example, node B chooses not to relay signals for node C. Thus, node

C has to send its signals to the BS directly. As the transmission range of the BS could be larger than the relaying range of mobile nodes, single-hop long range communications and multi-hop short range communications may co-exist inside a cell. This model is applicable to omni-directional antennas cases by changing the angular transmission zones of the mobile nodes to circular transmission zones.

Note that the BS only needs to communicate with the *last-hop node* on a relaying path for a source point s_i. For example, nodes *A*, *C*, *E*, *G*, *M* and *K* are last-hop nodes. A source node can be a relaying node and a last-hop node itself (see node *A*).

3.4. Assumptions

Relaying paths are input to the cell size and channel assignment schemes and are handled by a routing protocol. In this work, we assume a routing protocol exists and can be utilized in our system. Thus, in this model, we assume relaying paths are computed and provided by a MCN routing protocol, such as ACAR [Tam05]. Other existing MCN routing protocols such as CBSR [Li03] and CBR [Li02] can also be considered.

The model can be applied to mobile users with high mobility such as users travelling on highways. In this environment, frequent broken relaying paths may occur. In this case, an effective routing scheme is needed to repair the broken relaying paths or rediscover and re-establish new paths for the ongoing calls. After the relaying paths are repaired or re-established, channel assignment and cell size computation can be performed. Although mobility of users can affect the system performance, it does not affect the optimality of the channel assignment and cell size schemes.

We also assume a power control mechanism is used. Power control in a cellular system helps minimize the interference and, hence, maintain the cell capacity and coverage. In a MCN environment, power control is a challenging task as many mobile nodes may transmit simultaneously to their relaying nodes and/or the BSs. The signal transmission pattern is quite different from that of a cellular network. Power control in MCNs is beyond the scope of this work.

All mobile nodes are assumed to use a fixed short transmission range except the last-hop nodes that are within the communication range of the BS. The last-hop nodes can communicate at a flexible range as large as the BS range. Using a fixed transmission range for general mobile nodes helps simplify the relay architecture whereas using short transmission range helps reduce transmission power and, hence, the interference.

3.5. Summary

In this chapter, we presented a system model for a TDD W-CDMA MCN. This model can be applied to any TDD MCN as they are merely special cases of a TDD W-CDMA MCN. We discussed different duplex schemes and multiple access schemes for the model. TDD duplex scheme is chosen instead of FDD for a MCN because it is not easy to synchronize the uplink and downlink frequencies of FDD mode of mobile nodes in a MCN environment. TDD W-CDMA is chosen because it is a current 3G technology. In addition, other cellular technologies, such as TDMA and TDD CDMA, can be considered as special cases of TDD W-CDMA in terms of channel assignment. We also discussed the advantages of using directional antennas and the channel assignment issues raised because of using directional antennas. We provide a network model for a TDD W-CDMA MCN environment in which directional antennas are used.

Chapter 4

Channel Assignment in Multi-hop Cellular Networks

In this chapter, we propose an optimal channel assignment scheme, called OCA, and a heuristic channel assignment scheme, called MSWF, to minimize the packet delay in a TDD W-CDMA MCN. OCA provides an optimal channel assignment which guarantees minimum delay in the networks and can be used as an unbiased benchmark tool to study the performance of different network conditions and networking schemes. However, like most optimal schemes, OCA is computationally expensive and inefficient for a large real-time problem. In this case, MSWF, an efficient and effective heuristic scheme, is a better alternative. Simulation results show that MSWF achieves on average 95% of the delay performance of the optimal solution obtained by using OCA and is effective at different cell sizes, achieving high throughout and low packet delay.

This chapter is organized as follows. In the next section, we present a channel assignment model for a TDD W-CDMA MCN. We examine the channel conflicts and delay components in the model. In Sections 4.2 and 4.3, we present the OCA and MSWF schemes, respectively. In Section 4.4, the complexity analysis of the schemes is provided. In Section 4.5, a simulation model for studying the performance of MSWF with respect to OCA and a single-hop case is presented. Simulation results are also discussed.

4.1. Channel Assignment Model

In this section, we present a channel assignment model to model the channel assignment in a TDD W-CDMA MCN environment. We begin with a discussion of the channel assignment and channel conflicts in these networks. Then, we describe the main delay component, packet relaying delay of these networks. Note that this model is applicable to a TDD CDMA or TDD TDMA MCN as they are special cases of a TDD W-CDMA MCN. Figure 4.1 illustrates a typical scenario of channel assignment in a TDD W-CDMA MCN environment in which directional antennas are used.

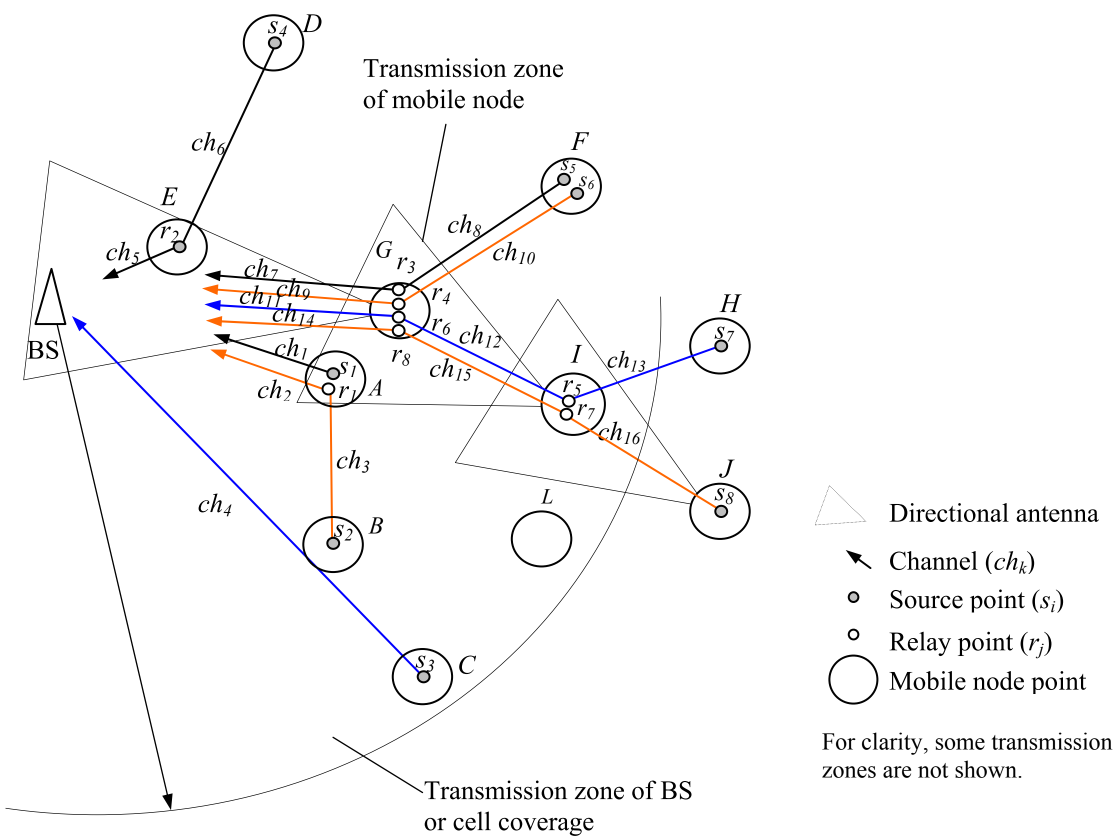


Figure 4.1: Topology and channel assignment in a MCN with directional antennas

4.1.1. Channel assignment

As mentioned in Section 3.2, when a source node makes a call request, a relaying path is set up to relay signals for the call. A call is considered as a connection which is represented by a virtual point, called source point s_i. A source node may have several source points; each point represents a different connection. Each connection is relayed by a relay point r_j on each relaying node on a relaying path. A relaying node may have several relaying points; each relaying point relays a connection. A relaying node can be a source node itself. Each source point or relaying point on the path is assigned exactly one channel for the connection. In a TDD CDMA cellular network or MCN, a channel ch_k is represented by the pair (time-slot t, code c).

4.1.2. Channel conflicts

When assigning a channel to a connection of a mobile node, channel conflicts need to be avoided. We define two types of channel conflicts in this network environment: co-channel and co-time-slot conflicts.

- *Co-channel conflict* -

Case 1: When a relaying node receives signals from more than one transmitting node, the channels of the transmitting nodes must be different; otherwise, signal collisions occur. For example, in Figure 4.1, node *G* is receiving signals from node *F* and *I*. Then, channels (ch_8 and ch_{10}) that are assigned to node *F* have to be different from channels (ch_{12} and ch_{15}) that are assigned to node *I*.

Case 2: When a receiving node is within the transmission zone of another transmitting node, the receiving channels of the receiving node have to be different from the transmitting channels of the transmitting node; otherwise, signal collisions occur. For example, node *A* is in the transmission zone of node *I*. Then, channel ch_3 that is assigned to node *B* for receiving node *A* and channels (ch_{12} and ch_{15}) that are assigned to node *I* must be different.

Case 3: A source node or relay node may serve several connections simultaneously. Each connection must be assigned a different channel. For example, channels (ch_8 and ch_{10}) that are assigned to node *F* have to be different from each other. Also, channels (ch_7, ch_9, ch_{11} and ch_{14}) of node *G* must be different.

- *Co-time-slot conflict* -

A node cannot physically receive and transmit data on the same time-slot using the same frequency. Assigning a channel to a current node (a mobile node which is being assigned a channel) with the same time-slot as that of the channel of the next-hop node (a mobile node which is next to the current node on a relaying path towards the BS) would cause packet loss because the next-hop node is simply not able to receive the packet. For example, in Figure 4.1, if node *I* is the current node, then node *G* is the next-hop node on the relaying path *J*-*I*-*G*-BS. The time-slots of the channels (ch_7, ch_9, ch_{11} and ch_{14}) that are assigned to node *G* have to be different from the time-slots of the channels (ch_8, ch_{10}, ch_{12}, and ch_{15}) that are assigned to node *F* and node *I*, respectively; otherwise, node *G* cannot receive the signals from node *F* and *I*.

4.1.3. Relaying delay

One fundamental issue in MCNs is packet delay because multi-hop relaying is involved which increases the delay. In a TDD MCN environment, the packet delay consists of four components: packet delay and time-slot waiting time at the source node, packet transmission time, packet propagation time, and time-slot waiting time at relaying nodes (see Figure 4.2). Among them, the time-slot waiting time of a packet at the relaying nodes, which we call *relaying delay*, significantly affects the packet delay. When a packet arrives at a relaying node, it has to wait until its allocated time-slot for sending out. The relaying delay at each hop of a relaying path is accumulated along the path. Thus, improper time-slot assignment significantly increases the relaying delay (or the packet delay) especially when the number of hops of the relaying path is large. Thus, an effective channel assignment scheme is important to reduce or minimize the relaying delay and, hence, the packet delay.

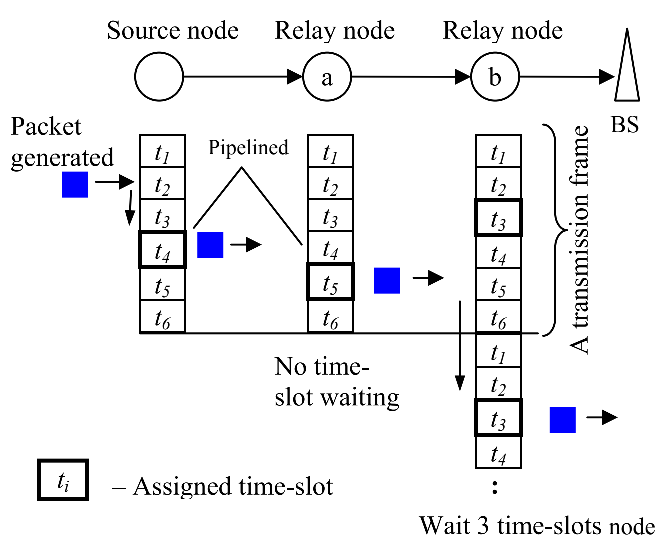


Figure 4.2: Time-slot waiting time (relaying delay) at relaying nodes on relaying path

4.1.4. Delay bounds

Different network topologies, network densities, hop count limits, number of available channels, and routing schemes give different channel interference patterns in terms of co-channel and co-time-slot conflicts. These conflicts in turn affect the resulting packet delay of a channel assignment scheme. Thus, it is difficult to find a delay performance function for a channel assignment scheme with respect to the problem size or the number of network nodes. Instead, a lower and an upper bound *per-hop packet* (*relaying*) *delay* are relatively easy to obtain. The per-hop packet delay is equal to the transmission time at the current node plus the time-slot waiting time at the next-hop relaying node. Note that the time-slot waiting time of a packet at a source node, which varies and depends on the arrival time of the generated packet rather than on the channel assignment scheme, is not included in the per-hop packet delay.

The *per-hop lower bound packet delay* can be achieved when there is no or little channel conflict such that a perfectly pipelined condition is achieved. In this case, the time-slot waiting time for an arrived packet at a relaying node is zero, i.e., a packet arrived at a relaying node can be sent out immediately. Assume the total time including transmission time, propagation time, and processing time of a packet is within one time-slot, the per-hop lower bound packet delay is 1 time-slot.

The *per-hop upper bound packet delay* is one time-slot for packet transmission, propagation, and processing plus the longest time-slot waiting time for an arrived packet waiting at a relaying node. Thus, the delay is equal to the number of time-slots of a transmission frame minus 1. For example, in Figure 4.2, if time-slot t_4 instead of t_3 is assigned to relaying node *b* for sending the packet which arrives at node *b*, then the packet relaying delay is 5, i.e., the number of time-slots per a transmission frame, which is 6, minus 1. Note that the per-hop upper bound packet delay is valid only when the arrival rate of the packets is less than or equal to the departure rate of the packets. That is, there is no backlog of packets in the queues of the time-slots at the relaying nodes.

4.2. Optimal Channel Assignment (OCA)

In this section, we present our OCA scheme for a TDD W-CDMA MCN. The scheme resides and is executed in the network controller, such as the RNC in the 3G UMTS [Holm4]. Each network controller connects a number of BSs. We assume the RNC has global information of the position, data-rate, route, and channel assignment of all mobile nodes involved in communications.

Problem definition

Given a set of relaying paths, the task is to find a channel assignment to minimize the total packet relaying delay and to ensure that no signal collision, channel conflict, or co-time-slot conflict occurs.

4.2.1. OCA formulation

To solve the problem, we formulate the OCA scheme as an Integer Programming problem as follows. We start with the set of relaying paths from source nodes to the BS found by the routing algorithm deployed in the system. Let V be the set of (virtual) points that determine the relaying paths such that no two paths intersect except at the BS. Let S be the set of source points, and R the set of relaying points such that $S \cup R = V$, and let $|S|$

= n and | R | = m. Several points from set V may correspond to the same physical node. For example, in Figure 4.1, node A contains the source point s_1 and the relaying point r_1. Figure 4.3 illustrates the relaying paths and virtual points corresponding to the traffic flows in the topology in Figure 4.1.

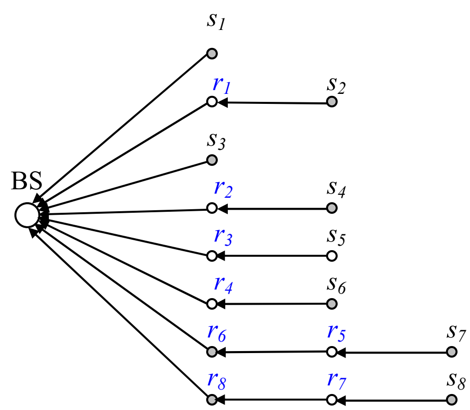


Figure 4.3: Virtual points for the relaying paths corresponding to the topology in Figure 4.1

We consider the time slot t and code c of a channel (t, c) to be positive integers. Let T and C be the maximum number of time slots and codes that can be used by a network node (BS or a mobile node) in the system. For example, in the TDD mode of UMTS (WCDMA) [Holm04], T=15 and maximum C=16. To describe a channel assignment for the points in V, we define two binary {0, 1} variables:

$$x(u,t) = \begin{cases} 1, & if\ u \in V\ is\ assigned\ time-slot\ \ t \\ 0, & Otherwise \end{cases} \tag{4.1}$$

$$y(u,c) = \begin{cases} 1, & if\ u \in V\ is\ assigned\ \ code\ \ c \\ 0, & Otherwise \end{cases} \tag{4.2}$$

To model the co-channel conflict phenomenon, we define the collision graph $G = (V, E)$ whose vertex set is the set of all source points and relaying points of mobile nodes that a network controller manages. An edge in this graph exists between two vertices u and v if and only if assigning the same channel for transmission to both u and v leads to a signal collision at some node a in the network (e.g. node G in Figure 4.1) or to co-channel conflict if u and v are virtual points of the same node (e.g. node F in Figure 4.1). Figure 4.4 is a collision graph corresponding to the scenario in Figure 4.1. In Figure 4.4, an edge exists between s_5 and r_5, s_5 and r_7, s_6 and r_5, s_6 and r_7, s_5 and s_6, r_5 and r_7, because of co-channel conflicts.

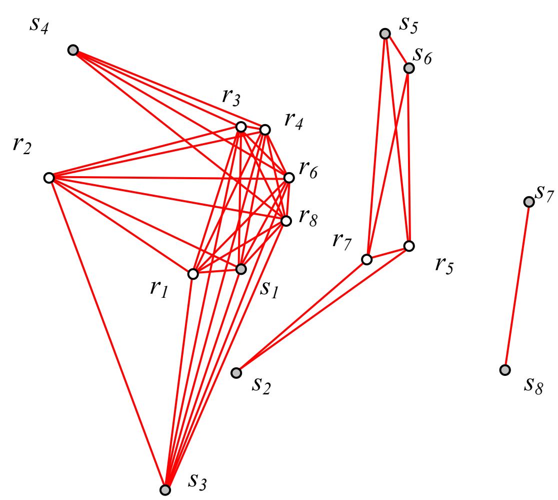


Figure 4.4: Collision graph corresponding to the co-channel conflicting points in Figure 4.1

Similarly, we can model the co-time-slot conflict by using a consecutive graph $G_c = (V, E_c)$. The receiving time-slots of a mobile node (i.e., the transmitting time-slots of a sender node of this receiving node) have to be different from its transmitting time-slots. An edge in this graph exists between two vertices which are in two consecutive nodes respectively along a path (e.g. node *F* and *G* in Figure 4.1). Figure 4.5 is the consecutive graph for the network topology in Figure 4.1. In Figure 4.5, an edge exists between s_5 and r_3, s_5 and r_4, s_5 and r_6, s_5 and r_8. This is because the time-slot of the transmitting channel ch_8 of source point s_5 of node *F* has to be different from the time-slots of the transmitting channels ch_7, ch_9, ch_{11}, and ch_{14} of relaying point r_3, r_4, r_6, and r_8, respectively, of node *G* (see Figure 4.1).

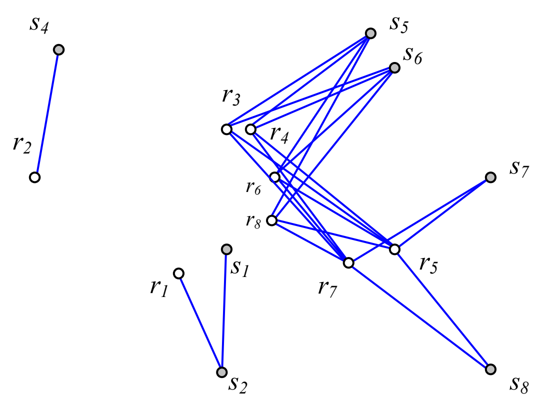


Figure 4.5: Consecutive graph corresponding to the co-time-slot conflicting points in Figure 4.1

We can now describe the linear constraints on variables $x(u, t)$ and $y(u, c)$ used in our Integer Programming formulation.

First, we enforce that exactly one channel (t, c) is assigned to every source point or relaying point of a mobile node represented by a vertex *u* in *V*:

$$\sum_{1 \le t \le T} x(u,t) = 1, \ \sum_{1 \le c \le C} y(u,c) = 1, \forall u \in V \tag{4.3}$$

Given an edge $(u, v) \in E$ from the collision graph, we have a valid channel assignment only if the vertices u and v are not assigned the same channel. For all possible channels, we can write,

$$x(u, t) + x(v, t) + y(u, c) + y(v, c) \leq 3, \forall (u, v) \in E, 1 \leq t \leq T, 1 \leq c \leq C. \quad (4.4)$$

Given an edge $(u, v) \in E_c$ from the consecutive graph, we have a valid channel assignment only if the vertices u and v are not assigned the same time-slot. For all possible time-slots t, we can write,

$$x(u, t) + x(v, t) \leq 1, \ \forall (u, v) \in E_c, 1 \leq t \leq T. \quad (4.5)$$

Any assignment of {0, 1} values to variables x and y that satisfies constraints (4.3)-(4.5) defines a valid channel assignment. However, we are interested in a channel assignment that minimizes total packet relaying delay. The objective function that models this delay is more difficult to express. For two consecutive vertices u and v on a relaying path P, the delay incurred if distinct time slots t_u and t_v are assigned, is $(t_v - t_u) \bmod T$. For example, in Figure 4.2, the total packet relaying delay = $(t_5 - t_4) \bmod 6 + (t_3 - t_5) \bmod 6 = 1 \bmod 6 + (-2) \bmod 6 = 1 + 4 = 5$. Therefore, the packet relaying delay $\delta(u,v)$ from point u to v can be written as a quadratic expression,

$$\delta(u,v) = \sum_{\substack{1 \leq t_u, t_v \leq T \\ t_u \neq t_v}} x(u,t_u) \cdot x(v,t_v) \cdot [(t_v - t_u) \bmod T]. \quad (4.6)$$

The objective function is $$F = \min \sum_{P \in W} \sum_{u,v \in P} \delta(u,v), \quad (4.7)$$

where P is an input relaying path and W is the set of input relaying paths. Function F is quadratic, but it can be linearized at the expense of increasing the problem size to make the formulation suitable for linear programming solvers. We define a new set of non-negative variables for every pair of consecutive vertices u, v on a path and every pair of time slots t_u and t_v. Our goal is to make $z\ (u, v, t_u, t_v)$ equal to the product of variables $x(u, t_u) \cdot x(v, t_v)$. We can achieve this goal with the following three constraints,

$$z(u, v, t_u, t_v) \leq x(u, t_u), \text{ and } z(u, v, t_u, t_v) \leq x(v, t_v) \tag{4.8}$$

$$z(u, v, t_u, t_v) \geq x(u, t_u) + x(v, t_v) - 1 \tag{4.9}$$

For example, if either one of $x(u, t_u)$ or $x(v, t_v)$ is zero, the non-negative variable $z(u, v, t_u, t_v)$ has to be zero, and when both $x(v, t_u)$ and $x(u, t_v)$ are one, $z(u, v, t_u, t_v)$ is one because of constraint (4.9). OCA finds the optimal channel assignment for the systems such that the delay function is minimized.

$$F = \min \sum_{P \in W} \sum_{u,v \in P} \sum_{\substack{1 \leq t_u, t_v \leq T \\ t_u \neq t_v}} z(u,v,t_u,t_v) \cdot [(t_v - t_u) \operatorname{mod} T], \tag{4.10}$$

$$= \min \sum_{\substack{u,v \in V \\ 1 \leq t_u, t_v \leq T,\ t_u \neq t_v}} z(u,v,t_u,t_v) \cdot [(t_v - t_u) \operatorname{mod} T], \tag{4.11}$$

subject to constraints (4.3), (4.4), (4.5), (4.8) and (4.9) and $x(u, t_u)$, $y(u, c_u)$, $x(v, t_v)$, $y(v, c_v) \in \{0, 1\}$, $\forall\, u, v \in V$, $1 \leq t_u \leq T$, $1 \leq c_u \leq C$, $1 \leq t_v \leq T$, $1 \leq c_v \leq C$, $z(u, v, t_u, t_v) \geq 0$.

Although OCA guarantees optimal solution in terms of minimum delay and can be used as an unbiased benchmark tool for performance evaluation of different network conditions and network schemes, like most optimization program, OCA is computationally expensive and inefficient for large real-time channel assignment problems. A heuristic scheme is more suitable in this case.

4.3. Minimum Slot Waiting First (MSWF)

We, herein, propose a heuristic channel assignment scheme, called MSWF, for a TDD W-CDMA MCN. MSWF provides a good result compared to the optimal solution offered by OCA. Unlike OCA, which solves an off-line (non-real time) global optimization problem of minimizing the packet relaying delay of the system, MSWF is a greedy scheme providing a locally optimal solution, focusing on minimizing delay for the packets on the path of a new call. The information about other nodes, which do not interact or interfere with the nodes on the new path, does not need to be processed. The channel assignments of existing paths are not affected.

The design of MSWF is based on two principles:

- Eliminate conflicting channels
- Select channels which contribute minimum relaying delay (time-slot waiting time).

4.3.1. The MSWF Scheme

MSWF mainly consists of two phases: the Proposing phase and the Checking phase. When assigning a channel to a node (virtual point) on a relaying path, a channel which contributes minimum relaying delay (time-slot waiting time) is proposed for the node. We call the node the *current node* and the selected channel the *proposed channel*. The channel is checked for channel conflicts based on four rules: *a*, *b*, *c*, and *d*. Rules *a* and *b* are used for checking co-time-slot conflicts whereas Rules *c* and *d* are used for checking co-channel channels. If no rules are violated, the proposed channel is accepted (selected); otherwise, the channel is eliminated. The following are the two phases.

Proposing phase:

- A channel that contributes the lowest relaying delay is proposed to the current node on the path.

Checking phase:

Rule a. The current node itself is not receiving on the time-slot of the proposed channel.

Rule b. The *next-h*op node is not transmitting on or temporary assigned with the time-slot of the proposed channel.

Rule c. Nodes on the other routes, having their *transmission zones* in which the next-hop node falls, are not transmitting on the proposed channel.

Rule d. Nodes that are in the transmitting zone of the current node are not receiving on the proposed channel.

A *next-hop node* is a relaying node that is one hop closer to the BS than the current node on the relaying path. For example, in Figure 4.1, if the current node is *I*, the next-hop node is *G*. A *transmission zone* is the region covered by the antenna.

Steps of the algorithm

Figure 4.6 illustrates six major steps of MSWF. Channels are proposed (assigned) starting from the *last-hop node* of the path all the way back to the source node. A last-hop node is the node closest to the BS on a path.

In Step 1, an input path is supplied by a routing algorithm in response to a new call request.

In Step 2, the last-hop node becomes the current node (Nd_{curr}). For example, in Figure 4.1, BS←*G*←*I*←*J* (or BS←r_8←r_7←s_8 in terms of virtual points) is a path where node *G* is the last-hop node which is also the current node.

In Step 3, a channel is proposed to the current node, i.e., the last-hop node. An available channel (a time-slot *t* and code *c* pair) in which the time-slot has the largest index is proposed to the current node. The proposed channel is checked by using Rules *a* and *d*. If Rule *a* is violated, all the channels in time-slot *t* are eliminated. If only Rule *d* is violated, the channel is eliminated. In either case, the channel assignment fails. MSWF continues to search for a not-tried available channel, in which the time-slot has the largest index, for the current node. If the channel satisfies the rules, the channel is temporarily assigned to the current node and the channel assignment is a success and the channel assignment process continues with Step 4; otherwise, the channel assignment is a failure and the channel assignment process is terminated.

In Step 4, the current node becomes the next-hop node whereas the successor of the current node becomes the current node. A successor of a current node is a node which is farther away from the BS than the current node on the relaying path. For example, node *I* becomes the current node whereas node *G* becomes the next-hop node.

Input: a relaying path for a connection.
Output: a channel assignment for each node on the path that minimizes the packet relaying delay.

1. Input a path
2. The last-hop node to the BS becomes the current node (Nd_{curr})
3. Assign an arbitrary channel (*ch*) to Nd_{curr}
   Do *PROPOSE* a not-tried available c*h* (*t*, *c*) where index of *t* is largest
      If Nd_{curr} not receiving on *t*, (Rule a)
         If *Nds* in Tx_{zone} of Nd_{curr} not receiving on this c*h*, (Rule *d*)
            c*h* is accepted (selected) for Nd_{curr} .
         Else
            c*h* is eliminated and ch assignment fails.
      Else
         All *chs* in *t* are eliminated and *ch* assignment fails.
      Endif.
   While (c*h* assignment fails and not-tried available c*hs* exist).
   If *ch* assignment fails, return. Else continue to Step 4.

4. Nd_{curr} and successor of Nd_{curr} becomes Nd_{next} and Nd_{curr}, respectively.

5. Assign a *ch* to Nd_{next}
   Do *PROPOSE* a not-tried available *ch* with *t* closest to that of *ch* of Nd_{next}
      If Nd_{curr} is not receiving on *t*, (Rule *a*)
      and Nd_{next} is not transmitting on or temporary assigned with *t*, (Rule *b*)
         If *Nds* on other routes, having their Tx_{zone}
            in which Nd_{next} falls, not transmitting on this *ch*, (Rule *c*)
         and *Nds* in Tx_{zone} of Nd_{curr} not receiving on this *ch*, (Rule *d*)
            c*h* is accepted (selected) for Nd_{curr} .
         Else
            c*h* is eliminated and *ch* assignment fails.
         Endif.
      Else
         All *chs* in *t* are eliminated and *ch* assignment fails.
      Endif.
   While (*ch* assignment fails and not-tried available *chs* exist).
   If *ch* assignment fails, return. Else repeat Steps 4 and 5 until the source node is reached.

6. If *ch* assignment is a success, update BS_{tbl} *and* Nd_{tbl}.

Definitions:

- **Current Node** (Nd_{curr}) is a node (virtual point) to be proposed a channel.
- **Next-Hop Node** (Nd_{next}) is a node next to current node towards BS.
- **Last-Hop Node** is a node nearest to the BS on the path.
- **Node** (*Nd*) is a source node or relaying node.
- **Time-slot** (*t*) is a time slot of a transmission frame.
- **Code** (*c*) is a spreading code [Holm04] in CDMA systems.
- **Channel** (c*h*) is a time-slot code pair (*t, c*).
- **BS Channel Table** (BS_{tbl}) is a channel information table in BS.
- **Nodal Channel Table** (Nd_{tbl}) is a channel information table in a node.
- **Transmission Zone** (Tx_{zone}) is the coverage of antenna.

Figure 4.6: The MSWF Algorithm

In Step 5, MSWF starts to assign a channel to the current node. An available channel, having a time-slot (index) which is successive and the closest to the time-slot (index) of the channel assigned to the next-hop node, is proposed to the current node. The aim is to minimize the relaying delay. The proposed channel is then tested by using Rules *a*, *b*, *c*, and *d*. If no rules are violated, the channel is temporarily assigned to the current node. If Rule *a* or *b* is violated, all the channels in the time-slot of the proposed channel are eliminated. If only Rule *c* or *d* is violated, the proposed channel is eliminated. In both cases, the channel assignment fails, MSWF continues to search a not-tried available channel for the current node. If all the not-tried available channels do not satisfy the rules and the channel assignment still fails, the channel assignment process is finished. If channel assignment is a success, Steps 4 and 5 are repeated for channel assignment for the next current node. This iterative process continues until the source node (source point) is assigned a channel or a channel assignment fails.

In Step 6, if channel assignment is a success, i.e., each node on the path is temporarily assigned a channel, the call (connection) is accepted and the channel information table of the nodes and the BS are updated; otherwise, the temporary assignments are removed.

Note that the reason of choosing the last-hop node as the starting point of channel assignment is that the node is closest to the BS which has the largest number of incident (receiving) channels, i.e., a large number of co-channel conflicts. As a relaying node on the path is farther away from the BS, the number of incident channels at the relaying node is smaller. Thus, there are fewer co-channel conflicts and more available channels for selection. This helps increase the chance of a successful channel assignment. The Dsatur algorithm [Brél79] is a well known heuristic for coloring of a graph. The algorithm performs color assignment starting with a vertex with the highest degree (the largest number of adjacent vertices) in a graph. That is, the vertex is also the most constrained one. The idea behind MSWF is similar to that of Dsatur.

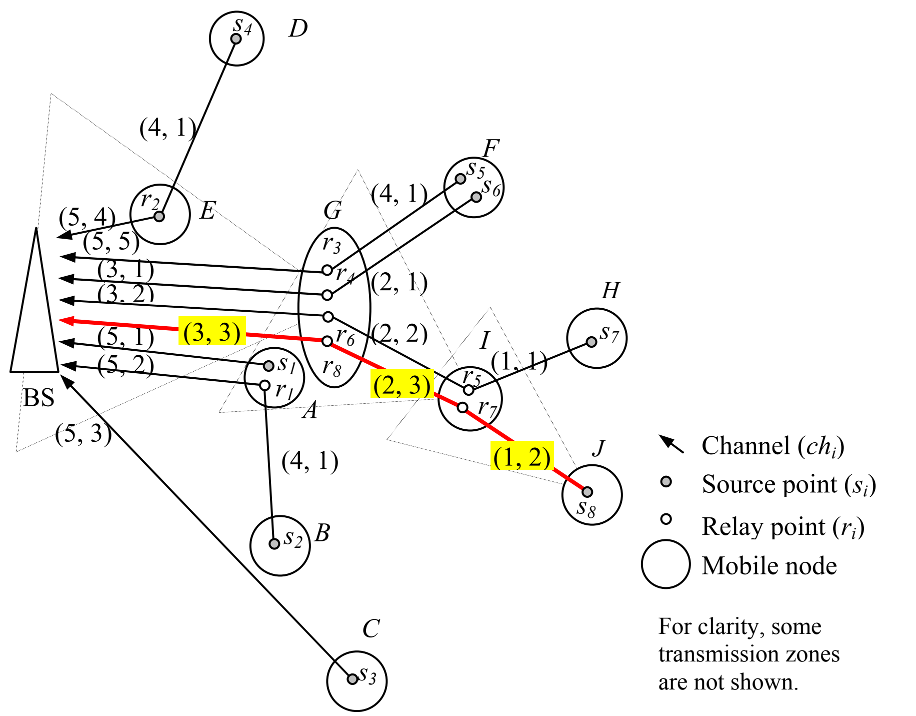


Figure 4.7: Channel assignment for the virtual points corresponding to the scenario in Figure 4.1

4.3.2. Illustration of MSWF algorithm

To further illustrate the MSWF algorithm, a channel assignment setting based on the scenario in Figure 4.1 is used (see Figure 4.7). Assume there are five time-slots per a transmission frame and five codes per each slot for BS uplink transmission. Assume all the connections or paths are already set up with channels assigned by MSWF except the connection along the path BS←*G*←*I*←*J* (or BS←r_8←r_7←s_8 in terms of virtual points).

MSWF starts assigning a channel at the last-hop node *G* (the current node) for the virtual point r_8 on the path. Channel (4, 1), which is an available channel having time-slot 4 which is the largest index among the time-slot indexes of the available channels in the BS, is proposed for r_8 of node *G*. Since node *G* is receiving on time-slot 4 using channel (4, 1) for the source point s_5, Rule *a* is violated. Thus, all the available channels in time-slot 4 are eliminated. The next not-tried available channel with the largest time-slot (index) 3 is (3, 3). Since neither node *G* is receiving on time-slot 3 nor the nodes inside the transmission zone of node *G* are receiving on this channel, Rules *a* and *d* are satisfied. The channel is temporary assigned to r_8 and channel assignment is a success. Note that there is no need to check Rules *b* and *c* since the next hop node is the BS instead of a relaying node.

MSWF continues to assign a channel for r_7 of node *I*, which is a successive node of node *G* on the path. Node *I* becomes the current node whereas node *G* becomes the next-hop node. Channel (2, 3) is an available channel of node *I* with time-slot 2 which is the closest time-slot that precedes time-slot 3 of the channel (3, 3) assigned to r_8 of the next-hop node *G*. This channel contributes minimum delay for packets that have been sent from node *I* and that have arrived at node *G* and that have also left node *G*. This channel also satisfies Rules *a*, *b*, *c*, and *d*. Thus, this channel is accepted and temporary assigned to node *I*.

Channel assignment continues for s_8 of node *J*. Node *J* becomes the current node whereas node *I* becomes the next-hop node. Channel (1, 2) is chosen for s_8 of node *J* for minimizing the relaying delay of this hop. The channel satisfies all the rules. The channel is accepted and temporary assigned to node *J*. Since node *J* is the source node, no more nodes need to be assigned a channel. All the nodes on the relaying path are temporarily assigned a channel such that the delay on the path is also minimized. The call (connection) is accepted and the channel information tables on the nodes and the BS are updated with the results of the channel assignment. Note that in this example, the channel assignment solution is also optimal as the time-slot waiting time for a packet arrived at a relaying node on any of the relaying paths is zero. The packet delay of the system is minimized.

When the call is finished or dropped, the channel of each node on the path for the call is de-allocated.

4.3.3. Worse case of MSWF algorithm

Considering the same example in Figure 4.7, assuming node *B* is assigned all the codes in time-slot 5 and node *J* is assigned all the codes in time-slot 2 except channel (2, 2), then relaying point of r_7 can only be assigned a channel with time-slot 4. If we further assume that node *H* is assigned all the codes in time-slot 1 and node *J* receives on time-slot 3 for other connection, then source point s_8 can only be assigned a channel in time-slot 5. In this case, the time-slots of the channel assigned to the path BS$\leftarrow r_8 \leftarrow r_7 \leftarrow s_8$ are 3, 4, and 5 respectively. The relaying delay for this path is the highest (worst) and is equal to [(3 - 4)mod 5 + (4 - 5)mod 5] = 4 + 4 = 8 time-slots whereas, in previous example, the delaying delay is [(3 - 2)mod 5+ (2 - 1)mod 5] = 1 + 1 = 2 time-slots.

4.4. Complexity Analysis

The channel assignment problem is similar to the graph-coloring problem [Corm01], which is to determine the minimum number of colors needed to color a given graph such that adjacent vertices must have different colors. In addition to the coloring constraints, the channel assignment problem has additional constraints on delay. As the graph-coloring problem is NP-hard, we conjecture that the channel assignment problem is also NP-hard.

Note that we formulated the channel assignment problem as an Integer Programming Problem (referred to as the OCA scheme), which is also NP-hard [Corm01]. The running time for solving an Integer Programming Problem using the best known algorithm is exponential. Thus, when the problem size is large, the time required to obtain the solution is prohibitive. Note that OCA considers all the existing paths and channel assignments for computation. By contrast, a heuristic scheme, such as the MSWF scheme, is more efficient. For MSWF, only a path for the new call needs to be considered. For each mobile node on a relaying path, a channel is proposed. In this case, we only need to check if the channels assigned to existing active mobile nodes have conflicts with the proposed channel. In the worst case, we need to check all the mobile nodes for each node on the relaying path. Assume the total number of mobile nodes is n and the number of hops of the relaying path is small or is a constant. Then, the time complexity of MSWF is $O(n)$ where n is the number of mobile nodes. Thus, MSWF is more efficient in terms of running time.

4.5. Performance Evaluation

In this section, we study the performance of MSWF with respect to OCA and the single-hop case for different conditions in terms of number of time-slots, number of hop counts, nodal densities, and cell sizes.

4.5.1. Simulation model and parameters

Our simulation model is a single-cell (see Figure 4.8). The number of relaying nodes varies from 0 to 160 in increments of 40. We choose 160 as the maximum number of relaying nodes because, at this setting, the network is dense enough such that most of the source nodes can reach the BS through multi-hopping. Further increase in the number of relaying nodes has little effect on meeting further demand at the BS. We separate the role of source node and relaying node so that the case in which no mobile nodes are willing to relay signals can be captured. The different number of relaying nodes is used to model different network densities and traffic patterns. All the nodes are uniformly distributed in a circular region with a radius of 1100 m centered at the BS.

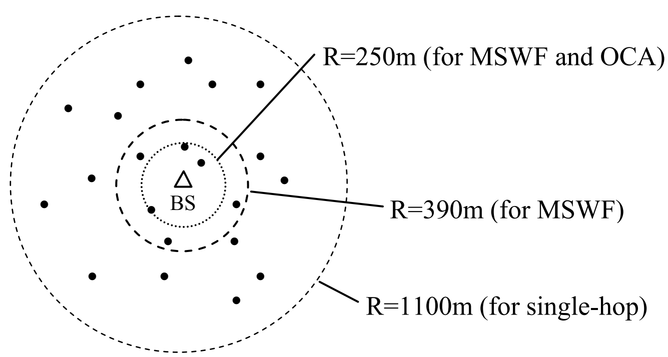


Figure 4.8: A single-cell model

In this simulation, two scenarios are provided. Scenario 1 is used to study the performance of MSWF with respect to the optimal solution provided by OCA. In this scenario, a smaller number of channels and a smaller hop count compared to Scenario 2 are used to reduce the computation time for OCA. Scenario 2 is used to study the performance of MSWF with respect to the single-hop case for different cell sizes. As MSWF is not computationally expensive, the problem size can be made larger.

Scenario 1: (*small problem: 25 channels and 45 source nodes*)

In this scenario, the transmission range of the BS and mobile nodes is 250m with cell capacity of 1035 kbps. Each TDD data transmission frame is 3.33 milliseconds long and has 5 time-slots. This frame size is 1/3 of the standard data transmission frame size, which is 10 milliseconds and 15 time-slots, in the WCDMA standard [Holm04]. All 5 time-slots are assigned for the BS uplink transmission. Each time-slot can be assigned maximum 5 codes. Thus, there are 25 channels (time-slot code pairs). Each code corresponds to a data rate of 41.4 kbps, which is 3 times 13.8 kbps (the data-rate of one code with spreading factor 16 in the WCDMA standard [Holm04]). Each call uses one code at a constant bit rate. The maximum number of hops is set to 4 to avoid excessive delay. The number of source nodes is 45 which is larger than 25. Each source generates call requests at an average rate of 0.5 calls per minute following a Poisson distribution. The average duration of each call is 1 minute with an exponential distribution. Each mobile node is equipped with a directional antenna with a 45^{o} beam angle. Note that the number of source nodes is larger than the number of available channels. This is used to create to a high channel competition situation for studying the effectiveness of the channel assignment schemes. Table 4.1 shows the simulation parameters. The BS or mobile capacities corresponding to the ranges are taken from the Table 5.2 (Section 5.5).

Scenario 2: (large problem: 65 channels and 70 source nodes)

In this scenario, for the MSWF case, the transmission ranges (cell size) of the BS are 250m and 390m with cell capacity of 1035 kbps and 828 kbps respectively whereas the transmission range of mobile nodes is 250m. For the single-hop case, the transmission range of the BS is 1100m with cell capacity of 207 kbps. Each TDD data transmission frame is 10 milliseconds long and has 15 time-slots according to the WCDMA standard [5]. The number of uplink time-slots and the number of downlink time-slots are 13 and 2 respectively. For MSWF-R250m case, each time-slot can be assigned 5 codes. Thus, there are 65 uplink channels. For MSWF-R390m case, each time-slot can be assigned with 4 codes and there are 52 channels. Each code corresponds to a data rate of 13.8kbps [Holm04]. For the single-hop case, each time-slot can be assigned with maximum 1 code. In both cases, each call uses one code at a constant bit rate. The maximum number of hops is set to 7 to avoid excessive delay. The number of source nodes is 70, which is slightly larger than 65, the number of available channels in the BS. This ensures that when the network is dense enough, the reachable demand from the source nodes is higher than the cell capacity such that a congested condition for channel assignment can be established to test the effectiveness of MSWF. The call request rate, call holding time, beam angle of directional antenna, and simulation time are the same as that of Scenario 1. This scenario is mainly used for studying the performance of MSWF in using different cell sizes, radii of 250m and 390m, and as compared to the single hop case.

Table 4.1: SIMULATIONS PARAMETERS

	Scenario 1	Scenario 2		
	OCA/MSWF- s5 h4	MSWF- s15 h7	MSWF- s15 h7	Single-hop
BS or last-hop nodal range	250m	250m	390m	1100m
Nodal range	250m	250m	250m	1100m
BS or last-hop nodal capacity	1035 kbps	1035 kbps	828 kbps	207 kbps
Nodal capacity	1035 kbps	1035 kbps	1035 kbps	207 kbps
Number of time-slots/frame	5	15	15	15
Number of time-slots (uplink)	5	13	13	13
Number of codes/ time-slot	5	5	4	1
Max. hop count	4	7	7	1
Data rate per code	41.4 kbps	13.8kbps		
Number of source nodes	45	70		
Call request rates	0.5 calls/min.			
Call holding time	1 min.			
Antenna	directional antenna with beam angle 45^o			
Simulation duration	15 mins.			

The simulation is implemented using OPNET Modeler 10.0A [Opne08]. The optimization package is MOSEK version 4 [Mose08]. A 90% confidence level with 10% confidence intervals is used in the simulation [Kinn96].

For MSWF and the single-hop case, OPNET Modeler is used for the whole simulation process. For OCA, the OPNET Modeler is used to generate the network topology. Then, Euclidean shortest paths are used as relaying paths and the collision and consecutive graphs are computed. The graphs are translated into an input file for the MOSEK solver to compute an optimal channel assignment for each source point and relaying point. The resulting channel assignment is transferred to OPNET where the simulation resumes and the throughput and packet delay are obtained.

As mentioned in Chapter 3, we assume there is a routing protocol to provide relaying paths. We assume nodes to be static (or with limited mobility) to simulate a pedestrian environment. In this case, the effectiveness of routing protocol does not affect the performance of the schemes. We assume there is a power control mechanism to control the transmission power of the BSs and the mobile nodes. We assume a perfect physical medium and sufficient battery capacity of mobile nodes for relaying signals.

4.5.2. Performance metrics

We use cell throughput and packet delay as the performance metrics for studying the performance of MSWF with respect to OCA, different network topologies, densities, and cell sizes.

Cell throughput – the number of packets that the BS receives per second. High throughput means high channel reuse and more nodes being served.

Packet delay - the time required for a packet sent from the source node to reach the BS.

Per-hop packet delay - the time required for a packet to be transmitted from one node and to wait in the next-hop relaying node before it was transmitted again.

Low packet delay or per-hop packet delay indicates the effectiveness of the channel assignment. Note that our results will always yield minimum packet delay globally for OCA and locally for MSWF.

Coefficient of variance (C.O.V.) of per-hop packet delay – This is the standard deviation of per-hop packet delay divided by the average per-hop packet delay. This measures the variation in packet delay among different relaying nodes or paths. High C.O.V represents some nodes or paths having high delay as compared to the average delay.

4.5.3. Simulation results

In this section, we first discuss the simulation results of MSWF on per-hop packet delay and C.O.V. of per-hop packet delay with respect to that of OCA and various parameters. Then, we study the performance of MSWF with respect to the single-hop case and different cell sizes in terms of throughput and packet delay. In the figures, the notation sN_1 hN_2 RN_3 represents N_1 time-slots per frame, N_2 hop count limits, and cell size of radius N_3, e.g., for s15 h7 R250m, there are 15 time-slots per frame, the hop count limit is 7, and the communication range of the BS is 250m in radius.

Considering Figure 4.9, the per-hop packet delay of OCA is equal to 1 time-slot, which is the same as the lower bound of per-hop packet delay mentioned in Section 4.1.4. This shows the optimality characteristic of OCA. For MSWF cases, the number of relaying nodes has more influence on the per-hop packet delay. When the network is sparse to medium dense, most source nodes cannot reach the BS using the available channels. More channels are available for selection and fewer paths are set for relaying signals. Thus, fewer co-channel conflicts occur and it is relatively easier to resolve the channel conflicts to achieve perfectly pipelined condition of time-slot assignment along the input paths. Thus, the per-hop packet delay is equal to or close to the lower bound, one. Among the MSWF cases, the larger the number of time-slots per frame and the number of hops, the higher the per-hop packet delay.

Although the per-hop packet delay increases as the number of relaying nodes increases, the delay is still close to the per-hop packet delay lower bound (1 time-slot) and far away from the per-hop packet delay upper bounds, 4 time-slots for the 5 time-slots (MSWF-s5 h4 R250m) case and 14 time-slots for the 15 time-slots (MSWF-s15 h7 R250m) case.

On average, the per-hop packet delays of MSWF and OCA are 1.057 time-slots and 1 time-slot, respectively. Thus, MSWF achieves approximately 95% of the delay performance of OCA. We define the delay performance as 1 / per-hop packet delay. If only high node density, i.e., when the number of relaying nodes is 160, is considered, the average per-hop packet delay of MSWF is 1.143 time-slots which is still close to the lower bound. Similar relative results are observed for other parameters. Thus, MSWF provides a good result compared to the optimal solution provided by OCA and with respect to the delay bounds.

In Figure 4.10, the coefficient of variance (C.O.V) of per-hop packet delay of MSWF increases when the number of relaying nodes increases. Among the MSWF cases, the MSWF-s5 h4 case has the lowest C.O.V. because it has a smaller hop count limit than that of the MSWF-s15 h7 cases. The hop count limits the range of variation of packet delay. The trade-off is that the network reachability of the MSWF-s5-h4 case is lower. This means that fewer distant nodes can reach the BS.

In Figures 4.11 and 4.12, we can generally observe that when the number of relaying nodes is zero, the network condition is reduced to a single hop case. Thus, the average delay of the two MSWF cases and the single-hop case are the same. The delay reflects the average time-slot waiting time of a packet in a source node. The throughput of the MSWF cases is lower than that of the single-hop case because MSWF has a smaller cell size than that of the single-hop case. In fact, the single-hop case has the largest coverage (network reachability) such that all the source nodes can reach the BS and all the available channels (although the number of channels is small) are used.

For MSWF cases, when the number of relaying nodes is small, i.e., when the nodal density is low, many source nodes still cannot reach the BS due to the lack of relaying paths. Thus, the increase in throughput and delay of MSWF is little. The MSWF case with cell radius of 390m has a higher throughput because the coverage is larger.

When the number of relaying nodes increases, the throughput of MSWF increases because more source nodes, especially the distant source nodes, have paths to reach the BS. The delay also increases as the number of hops on each relaying path increases.

When the number of relaying nodes further increases, the throughput increases until all the available channels are used. At this stage, the throughput value is the highest. The MSWF case with cell radius of 250m has a higher throughput than that of the MSWF case with cell radius of 390m because the former has a larger capacity (more channels) and more demands can be served. There is no further increase in delay because the number of hops of most of the relaying paths reaches the hop count limit.

Note that the results of the MSWF-s5 h4 R250m case of Scenario 1 are not included in Figures 4.11 and 4.12 because the settings are quite different from those of Scenario 2.

Although MSWF has a higher delay than that of the single-hop case, with the consideration of the fact that MSWF has a larger hop count limit, MSWF still achieves a good performance in terms of minimizing delay. In fact, the delay difference between the MSWF cases and the single-hop case is approximately equal to the average number of hops per path times the per-hop packet delay lower bound. This demonstrates the effectiveness of MSWF in minimizing delay.

Figure 4.13 illustrates the delay-throughput characteristics of MSWF. The MSWF case with a larger cell size of R=390m has slightly better delay throughput performance than the MSWF case with small cell size of R=250m. Nevertheless, the observed low delay at relatively high throughput is achievable due to the fact that MSWF minimizes the packet relaying delay.

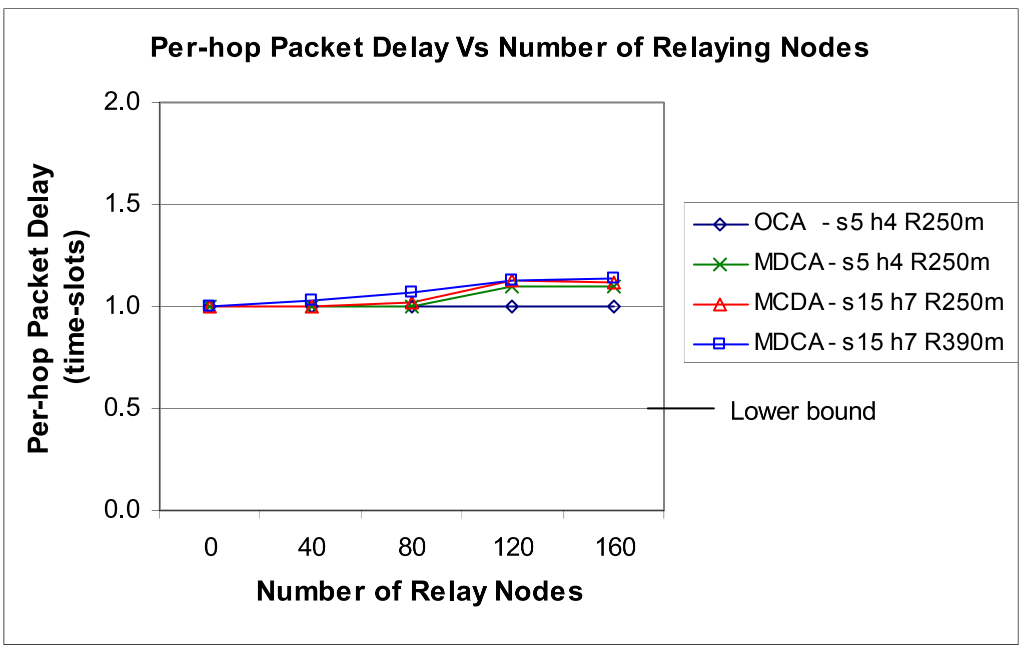


Figure 4.9: Per-hop Packet Delay (Scenarios 1 and 2)

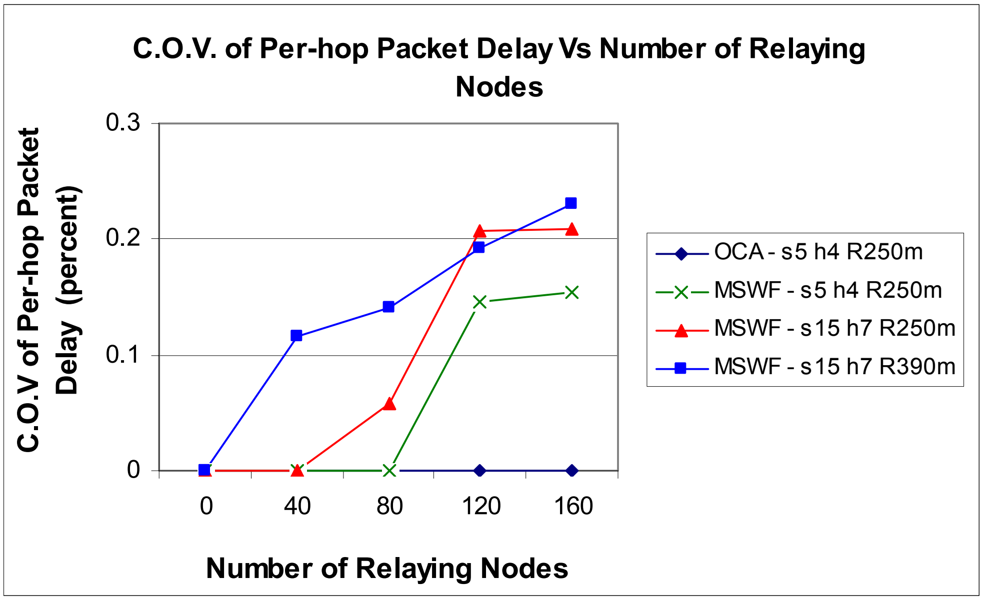


Figure 4.10: C.O.V. of Per-hop Packet Delay (Scenarios 1 and 2)

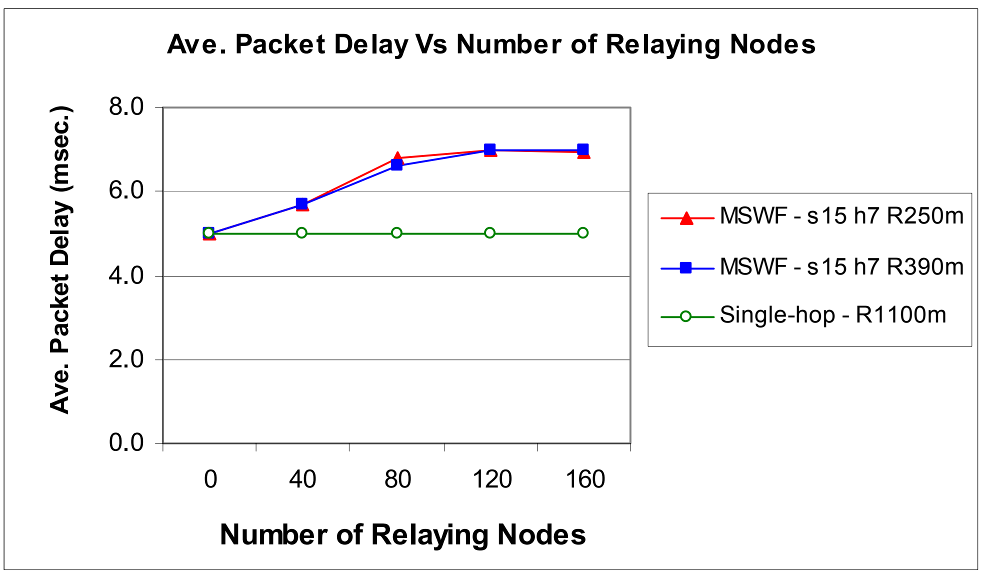


Figure 4.11: Packet Delay (Scenario 2)

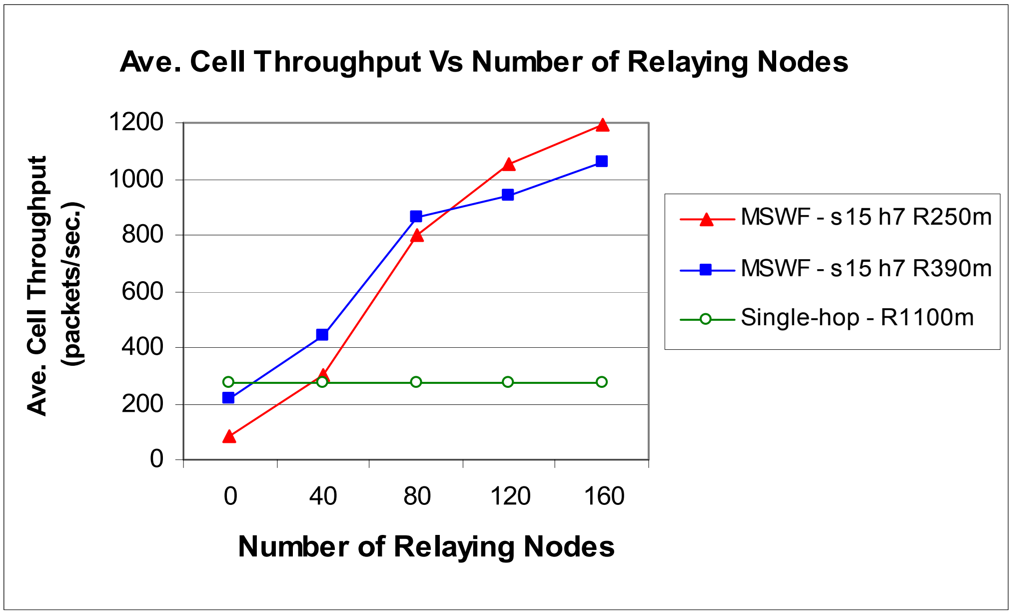


Figure 4.12: Cell Throughput (Scenario 2)

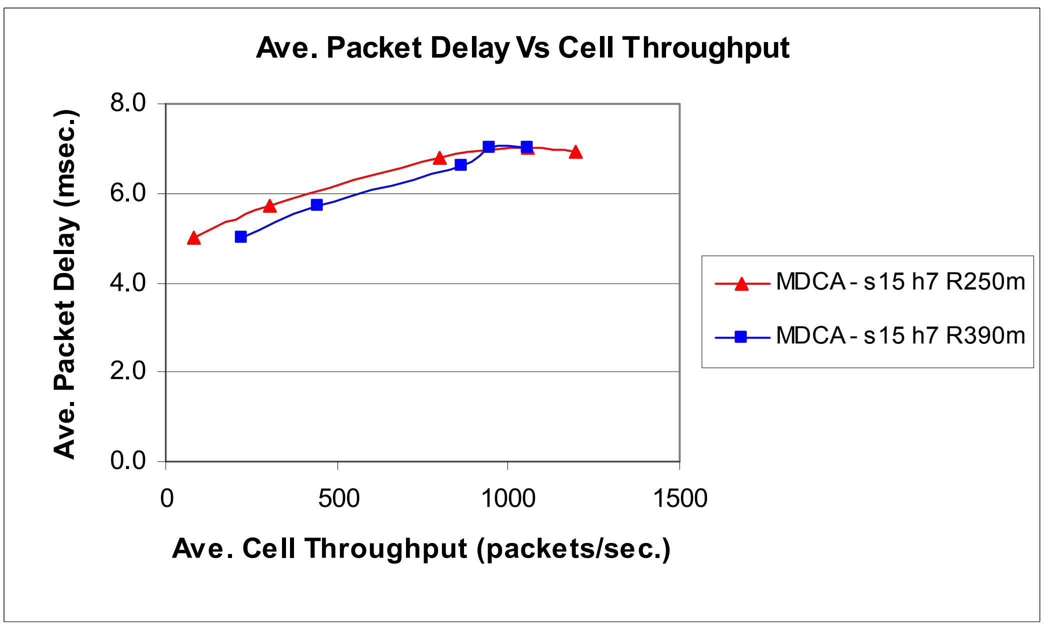


Figure 4.13: Packet Delay versus Cell Throughput (Scenario 2)

4.6. Limitations

As mentioned in Section 4.2, OCA is computationally expensive and inefficient for large real-time channel assignment problems. When the problem size is large, OCA is computationally prohibitive. MSWF is more efficient for large real-time problems. However, MSWF does not guarantee minimum delay.

Both OCA and MWSF require overhead for computing the channel assignment, communicating the channel assignment results between the BS and the mobile nodes, and storing the channel assignment information at the BS, RNC, and the mobile nodes. In MCNs, coding and decoding of the signals at each relaying node may be involved. This requires additional overhead for signal processing, increases battery power consumption of mobile nodes, and increases the complexity of the system. When the user mobility is high or the link quality is low, frequent broken route may occur. This triggers new route discovery for the broken route. New channel assignments are required for the new routes.

4.7. Summary

In this Chapter, we proposed the OCA scheme and a heuristic channel assignment scheme, called MSWF, to minimize the delay in a TDD W-CDMA MCN or any TDD MCN. OCA provides an optimal channel assignment solution in terms of minimum packet relaying delay and can also be used as an unbiased benchmark tool for performance comparison of different network conditions and networking schemes. However, OCA is computationally expensive and inefficient for large real-time channel assignment problems. MSWF provide a good result compared to the optimal solution provided by OCA in terms delay and processing time. Simulation results show that

MSWF achieves on average 95% of the delay performance of OCA. MSWF has a low per-hop packet delay which is close to the per-hop lower bound packet delay. It achieves high throughput and low delay as compared to the single-hop case and is applicable to different cell sizes.

Chapter 5

Cell Size in Multi-hop Cellular Networks

In this chapter, we propose the OCS scheme and two heuristic cell size schemes, called HTCSF and SCSF in a TDD W-CDMA MCN environment. OCS finds the optimal cell sizes that maximize the system throughput of MCNs. OCS can also be considered as a network planning aid in cellular networks and MCNs. Although OCS provides an optimal solution, like most optimization program, OCS is computationally expensive and not suitable for large real-time problems. A heuristic is preferable in this case. Thus, in this chapter, we also propose two heuristics, HTCSF and SCSF. Simulation results show that they provide good results compared to the optimal solution provided by OCS. HTCSF achieves higher throughput than that of SCSF.

This chapter is organized as follows. In the next section, we discuss the importance of cell size and the optimal cell size in a single-cell and a multi-cell MCN environment. In Section 5.2, we introduce the OCS for finding the optimal cell sizes in a MCN environment. In Section 5.3, we propose SCSF whereas in Section 5.4, we propose HTCSF. In Section 5.5, we discuss the complexity of the schemes. In Section 5.6, we present the simulation model and discuss the results.

5.1. Cell Size

In this section, we describe the importance of cell size and the concept of optimal cell size in a single-cell and a multi-cell MCN environment.

5.1.1. The importance of cell size

Cell size affects the network reachability. As mentioned in the network model in Chapter 3, the BS only needs to communicate with the last-hop node on a relaying path for a source point s_i. The demand (data-rates) of source nodes can reach the BS as long as their last-hop nodes are within the communication range (cell size) of the BS. In other words, their demands cannot be served by the BS if their last-hop nodes are outside the range of the BS. For example, in Figure 5.1, nodes *A*, *B*, and *D* are the last-hop nodes. Cell size also affects the cell capacity. The smaller the cell size, the higher the cell capacity can be achieved in a CDMA system. Thus, the cell size is an important factor that affects the network reachability and cell capacity which in turn affect the total demands that can be served which gives the system throughput.

5.1.2. Optimal cell size - single-cell case (Capacity-demand model)

To explain the concept of optimal cell size, we introduce a Capacity-demand model. The model models the relationship between the capacity function of a cell and the demands of the mobile users with respect to the cell size. The *capacity* represents the cell capacity of the CDMA technology and is a decreasing function over the cell size. The *demand* represents the total data rate requested by the source nodes and is an increasing step function over the cell size. Figure 5.1 shows a typical single-cell TDD W-CDMA MCN environment and a graph for illustrating the optimal cell size in a single-cell case. In the MCN environment, mobile nodes use directional antennas to transmit signals. But, for clarity, the transmission zones of directional antennas are not shown.

In the graph in Figure 5.1, there are three demand values (μ_1, μ_2, and μ_3) plotted as three line segments; each value represents the cumulated demand of the source nodes having their last-hop nodes within the corresponding communication range (cell size) of the BS. For example, $\mu_1 = \mu(s_1)$ *and* $\mu_2 = \mu(s_1) + \mu(s_2) + \mu(s_3)$. Two capacity functions, U and U', are used to illustrate two typical cases.

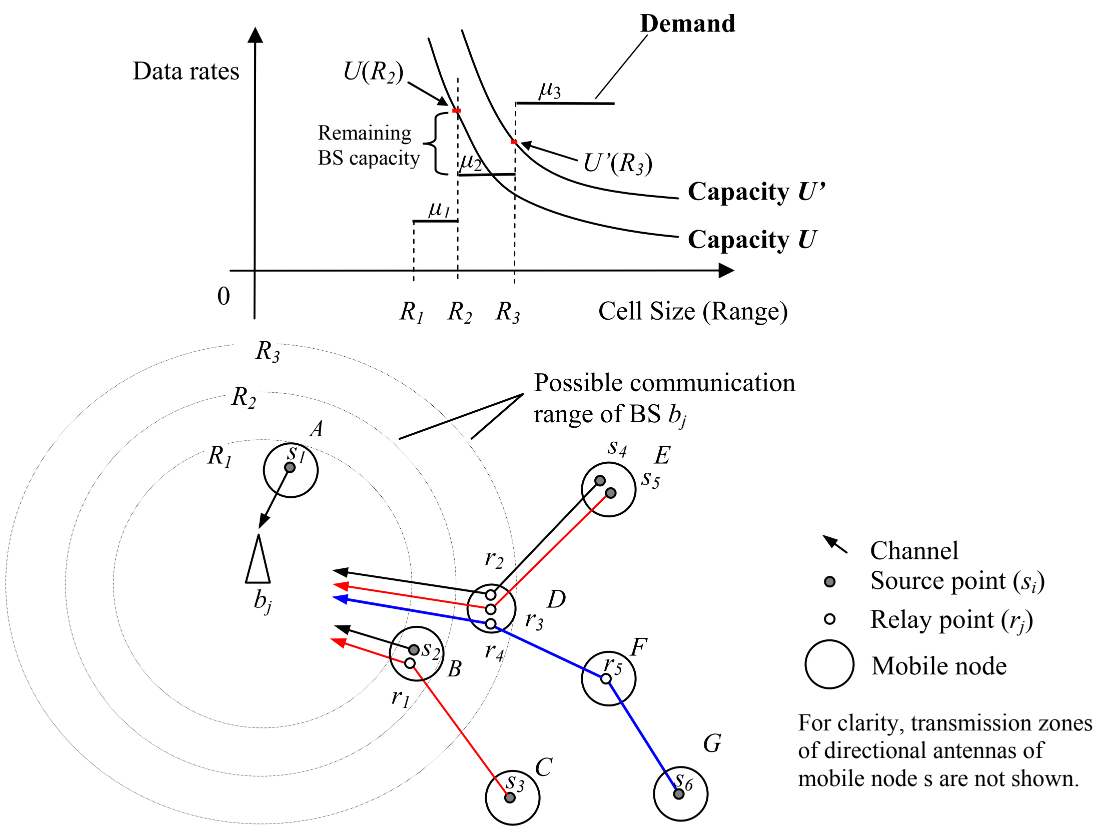


Figure 5.1: Capacity-demand model

In the first case, the curve of the capacity function *U* intersects with the demand segment μ_2, which is the total demand originated from source nodes *A*, *B* and *C*. The demand of node *C* is relayed through node *B* to the BS. The optimal cell size is at range R_2 because, at this range, the demand that can be served is maximized and equal to μ_2 and the remaining capacity ($U(R_2) - \mu_2$) for future calls is also maximized.

In the second case, the curve of the capacity function *U'* is in-between the demand segments μ_2 and μ_3. In this case, if the total demands that can be served by the capacity corresponding to the range R_3 is more than μ_2, then R_3 is the optimal cell size; otherwise, the optimal cell size is at R_2. This problem is a special case of knapsack problem [Levi03] which involves selecting a number of objects (each of them has a size and a value) into a bag which has a capacity such that the total value is maximized. In this case, the objects are the demands. Each demand has a value which is the data rate requested by a corresponding source point. The cell capacity corresponds to the bag's capacity. The capacity function may be obtained through on-line computation, prediction or experiments.

5.1.3. Optimal cell size - multi-cell case

In a multi-cell environment, the task of finding the optimal cell sizes not only requires the computation of the cell sizes, but also involves the selection of a BS among several possible BSs for a connection (source point) of a source node. The decision is dependent on the availability of relaying paths (each path ends at a different neighbouring BS), the locations of the last-hop nodes of the paths, the cell sizes, and the corresponding capacities of the cell sizes. For example, in Figure 5.2, connection s_7 requested by node *I* can be assigned to BS b_2 or b_3. Assigning the connection to b_3 gives a lower overall system capacity compared to that of assigning it to b_2 because the cell size of b_3 needs to be larger to cover the last-hop node *J* of the relaying path for the connection in node *I*. Assume that a source point has one path per each BS. The distance between the last-hop node of the path and the BS is an important input parameter for determining the optimal cell size.

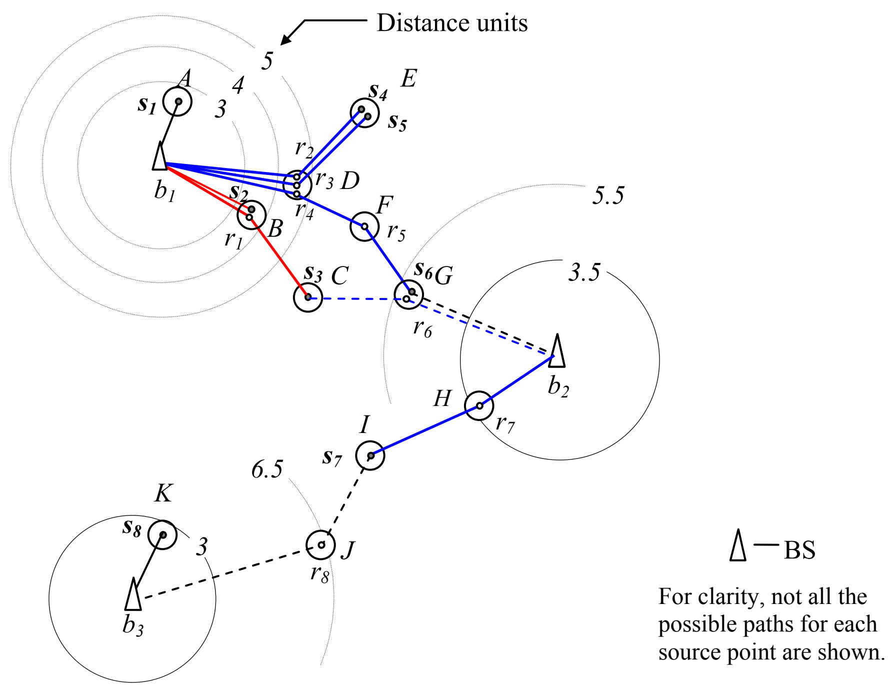


Figure 5.2: Cell size in a multi-cell environment

We denote by *V* the set of (virtual) points that determine the relaying paths, which do not intersect with one another except at the BSs. Several points from set *V* may be in the same physical node. Let *S* be the set of source points $\{s_1, s_2, \ldots, s_m\}$ and *L* be the set of relaying points $\{r_1, r_2, \ldots, r_k\}$, and *B* be the set of BSs $\{b_1, b_2, \ldots, b_n\}$ such that $S \cup L \cup B = V$. Each source point may have several relaying paths; each path has a different nearby BS as target. In other words, for each BS, there is a set of relaying paths; each path relays the signals of a unique source point to the BS. This setting is similar to the setting in the channel assignment model in Section 4.1 except that a set of BSs is included.

To determine whether or not a source point can reach the BS, we just need to know the distance of the last-hop node of the relaying path for the source point from the BS. Assume a routing protocol provides a set of relaying paths for each source point for each BS, the topology in Figure 5.2 can be translated into three sets of relaying paths; each set of paths ends at a different BS. From these paths we can construct a bipartite distance graph G (V_h, E) for each BS where V_h is the set of vertices $\{v(s_1), v(s_2),\ldots, v(s_m)\}$ representing the last-hop relaying points towards to the BSs and E is the set of edges connecting a vertex in V_h to a BS. Each edge is weighted by the distance $d(v(s_i), b_j)$ of the last-hop relaying node having the last-hop relaying point $v(s_i)$ on the path reaching the BS b_j. If no relaying path is found for a source point for a BS, the distance of that source point to that BS is set to infinite (∞). Figure 5.3 shows the distance graph for the BS b_1 for the scenario in Figure 5.2. For example, for s_3 in Figure 5.3, $v(s_3) = r_1$ and $d(v(s_3), b_1) = d(r_1, b_1) = d_3 = 4$.

Let $R(b_j)$ be the set of communication ranges $\{R_1, R_2, \ldots, R_n\}$ of BS b_j in non-decreasing order where R_n represents the distance between the BS and the last-hop node of the relaying path for a source point s_i to the BS. The ranges can be obtained based on the distance graph G. For example, in Figure 5.3, the set of the ranges for BS b_1 is {3, 4, 5} where the number represents the units of distance. Note that the set of ranges can also be a set of predetermined values depending on the standard and/or configuration of the system in practice. For example, the set of ranges for BS b_1 could be {1, 2, 3, 4, 5, 6, 7}.

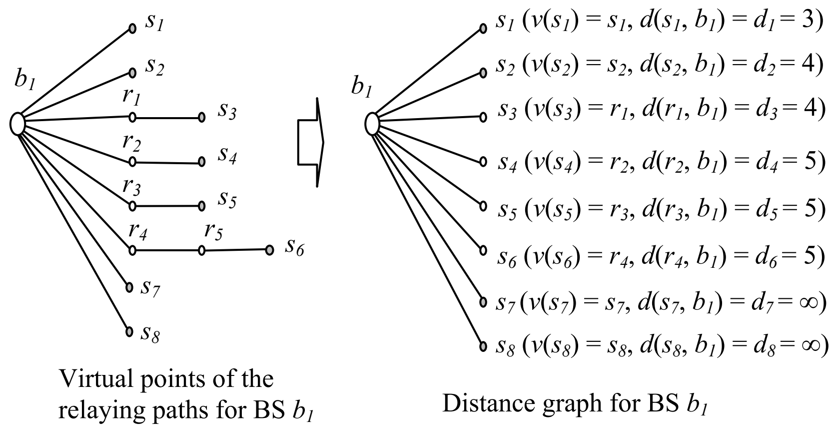


Figure 5.3: Distance graph *G* for base station b_1 in Figure 5.2

5.2. Optimal Cell Size (OCS)

In this section, we propose the OCS scheme. OCS is an Integer Program used to find the optimal cell sizes that maximize the system throughput in a TDD W-CDMA MCN and a cellular network. Note that a cellular network is merely a special case of MCN. In the followings, the problem definition and detailed formulation of OCS are given.

Problem definition

Given a set of BSs and a set of relaying paths of all source nodes to the BSs, each source node has one or more traffic demands, each path goes to a different BS, and a cell capacity function, the task is to find the optimal cell sizes (communication ranges of the BSs) that maximize the system throughput (servable demand).

5.2.1. OCS formulation

Based on the Capacity-demand model in Section 5.1.2, $\mu(s_i)$ represents the demand of the source point s_i. We seek to maximize the total demand that can be served. More specifically, we need to decide (i) what communication range (cell size) to use for each BS and (ii) which BS to assign to a connection request. To solve for optimal cell sizes, we formulate the problem as an Integer Programming problem.

To model the assignment of the connection of a source point s_i of a mobile node to a BS b_j, we use a connection assignment variable $x(b_j, s_i)$. For example, in Figure 5.2, if s_4 is assigned to b_1, then $x(b_1, s_4)$ is equal to 1.

$$x(b_j, s_i) = \begin{cases} 1, & \text{if } s_i \text{ is assigned to } b_j \\ 0, & \text{Otherwise} \end{cases} \tag{5.1}$$

To ensure that a source point is assigned to no more than one BS, the total sum of the values of the variables x is constrained to be less or equal to 1. For example, the sum of the values of $x(b_1, s_1)$, $x(b_2, s_1)$, and $x(b_3, s_1)$ must be less than or equal to 1. The sum is zero if s_1 cannot be assigned to any of the BSs (b_1, b_2, b_3) due to insufficient capacities of the BSs or a lack of a relaying path to the BSs.

$$\sum_{b_j \in B} x(b_j, s_i) \le 1, \quad \forall b_j \in B, \forall s_i \in S \tag{5.2}$$

To model a communication range of a BS b_j, we define a range variable $y(b_j, R_k)$. The value of $y(b_j, R_k)$ is equal to 1 if R_k is less than or equal to the chosen communication range R_c of the BS; otherwise, $y(b_j, R_k)$ is equal to 0. For example, in Figure 5.3, the set of possible ranges is {3, 4, 5}. If the R_c is 4, then $y(b_1, 3)$ and $y(b_1, 4)$ are equal to 1 whereas $y(b_1, 5)$ and $y(b_1, \infty)$ are equal to 0.

$$y(b_j, R_k) = \begin{cases} 1, & \text{if } R_k \le R_c \\ 0, & \text{Otherwise} \end{cases} \tag{5.3}$$

The relationship between the ranges of a BS b_j can be expressed in a set of k-1 inequalities in a compact form as follows. For example, if R_c is 4, the value of $y(b_1, 3)$ is restricted to 1 because $y(b_1, 3) \geq y(b_1, 4)$ where $y(b_1, 4) = 1$.

$$y(b_j, R_1) \geq y(b_j, R_2) \geq \cdots \geq y(b_j, R_c) \geq \cdots \geq y(b_j, R_n), \forall b_j \in B, \qquad (5.4)$$

where $R_1 < R_2 < \ldots < R_c < \ldots < R_n$.

To relate the range variable y to the connection assignment variable x, we formulate a third constraint as follows. A source point s_i can be served by a BS b_j only if the BS has chosen a communication range which is able to reach the last-hop node of the relaying path for the source point. For example, if the distance $d(v(s_i), b_j)$ between the last-hop node having the relaying point $v(s_i)$ for the source point s_i and the BS b_j is less than or equal to a chosen range R_c, then $x(b_j, s_i)$ can be 1 or 0 depending on whether or not it is admitted by BS b_j. If $d(v(s_i), b_j)$ is larger than R_c, then $x(b_j, s_i)$ must be equal to 0.

$$x(b_j, s_i) \leq y(b_j, R_k), \forall b_j \in B, s_i \in S(b_j, R_k), \qquad (5.5)$$

where $S(b_j, R_k)$ is a set of source points s_i with $R_{k-1} < d(v(s_i), b_j) \leq R_k$. Sets $S\{b_j, R_k\}$ are disjoint for different values of R_k and thus there is at most one constraint (5.5) for every pair of BS and source point. For example, in Figure 5.3, for R_k or $d = 3$, $S(b_1,3)=\{s_1 \mid 0 < d_1 \leq 3\}$ and $x(b_1, s_1) \leq y(b_1, 3)$. For R_k or $d = 4$, $S(b_1,4)=\{s_2, s_3 \mid 3 < d_2 \leq 4, 3 < d_3 \leq 4\}$ and $x(b_1, s_2) \leq y(b_1, 4)$; $x(b_1, s_3) \leq y(b_1, 4)$. If the chosen range R_c is 4, then $y(b_1, 4)$ is equal to 1. Since $y(b_1, 3) \geq y(b_1, 4)$ in constraint (5.4), $y(b_1, 3)$ is also equal to 1. In this case, $x(b_1, s_1)$ can be 1 or 0 depending on whether or not s_1 is assigned to b_1. The same rationale applies to $x(b_1, s_2)$ and $x(b_1, s_3)$.

To ensure that the total demand of the source points served by a BS is not larger than the capacity of that BS for the chosen range, knapsack inequalities are used. Let $U(b_j, R_k)$ be the capacity decrease of BS b_j when the communication range of BS b_j is increased from range R_{k-1} to range R_k. In particular, we denote by $U(b_j, 0)$ the maximum capacity of the BS corresponding to the smallest possible communication range. By convention, we choose $U(b_j, R_k)$ to be positive. The summation of the product between the value of each variable $x(b_j, s_i)$ and its demand $\mu(s_i)$ (data rate) has to be less than or equal to the capacity at that chosen range of b_j. For example, assume $U(b_1,0)=25$, $U(b_1,R_k)=5 \;\forall\; k$, and $R_c=4$, then the capacity is $25-5\cdot y(b_1,1)-5\cdot y(b_1,2)-5\cdot y(b_1,3)-5\cdot y(b_1,4)-5\cdot y(b_1,5)=25-5(1)-5(1)-5(1)-5(1)-5(0)=5$. The total demand that can be served by the BS is

$$\sum_{s_i \in S} \mu(s_i) \cdot x(b_j, s_i) \le U(b_j, 0) - \sum_{1 \le k \le R(b_j)} U(b_j, R_k) \cdot y(b_j, R_k), \forall b_j \in B \qquad (5.6)$$

The objective is to maximize the total demand that can be served. The total servable demand is dependent on the number of reachable requests, the data-rate of each request, and the cell capacity. Thus, the objective function is to maximize the value of the summation of the products of each variable x and its corresponding demand $\mu(s_i)$ (data rate) for all BSs. OCS finds the optimal cell sizes of the system such that the total demand is maximized.

$$D = \max \sum_{b_j \in B, s_i \in S} \mu(s_i) \cdot x(b_j, s_i), \qquad (5.7)$$

subject to constraints (5.2), (5.4), (5.5), and (5.6), and $x(b_j, s_i), y(b_j, R_k) \in \{0,1\}$, $\forall\, s_i, b_j \in V$, and $R_k \in R$.

In the optimal solution, a variable $x(b_a, s_b)$ having value of 1 represents the source point s_b assigned to BS b_a. For example, if $x(b_1, s_1)$=1, $x(b_1, s_2)$=0, and $x(b_1, s_3)$=1, then s_1 and s_3 are assigned to BS b_1 whereas s_2 is not. Range variable $y(b_c, R_d)$ having a value of 1 in the optimal solution where R_d is the largest determines the optimal range of BS b_c. For example, if $y(b_1, 3)$=1, $y(b_1, 4)$=1, and $y(b_1, 5) = 0$, then the optimal range (cell size) of BS b_1 is 4.

5.3. Small Cell Size First (SCSF)

Although OCS provides an optimal solution for cell size, it is computationally expensive and may not suitable for large real-time problems. In this case, a heuristic which provides good results compared to the optimal solution provided by OCS is more suitable. We, herein, propose our first heuristic, called SCSF, which provides good results compared to the optimal solutions provided by OCS.

5.3.1. The SCSF scheme

The idea of SCSF is to maximize the cell or system capacity of a MCN by choosing a relaying path of a source point such that the required communication range (cell size) of the BS is the smallest. In other words, the distance of the last-hop node of the chosen path from the BS is the shortest. Also, the BS should have enough remaining cell capacity to meet the demand of the source point. Figure 5.4 illustrates the SCSF algorithm.

Steps of the SCSF algorithm

In Step 1, when a new call arrives, input the demand (data-rate) μ of a new call request.

In Step 2, a set of n relaying paths of the call, each path goes to a different BS, is assumed to be provided by a routing algorithm.

In Step 3, the distance of the last-hop node of each path from its corresponding BS is computed and stored with its BS as an entry on a list. We call the distance *last-hop distance*.

In Step 4, the list is sorted in an ascending order according to the values of the last-hop distances. Thus, the last-hop distance of the entry at the head of the list is the shortest. The BS of the entry requires the shortest communication range (smallest cell size) to cover the last-hop node of that call, i.e., the shortest communication range for the call (source node) to reach. The BS is a potential target BS (cell) for the call.

In Step 5, the entry starting from the head of the list, i.e., the first entry, is checked to see if the BS can be the target BS or if the BS has enough capacity to meet the demand.

In Step 6, the last-hop distance of the entry is checked. If it is less than or equal to the current communication range of the corresponding BS, continue with Step 7; otherwise, go to Step 8.

In Step 7, the remaining capacity of the BS is checked. If the remaining capacity is enough to meet the demand μ of the call, the current cell range of the BS is the target cell range and the BS is the target BS for the call. Iteration of the loop is finished; otherwise, go to Step 10, i.e., repeat Step 6 for the next entry on the list.

Input: the demand of a new call, the relaying paths for the call to the BSs, and the capacity function of the BSs.
Output: the target BS for the call and the range of the BS.

1. Input the demand μ of a new call.
2. Input n relaying paths for the call.
3. Compute and store the *last-hop distance* of each relaying path on a list.
4. Sort the list based on the last-hop distances in an ascending order.

5. Loop through the list,
6. If last-hop distance <= current cell range of the corresponding BS,
7. If the remaining capacity of the BS is enough to meet the demand μ,
   The current cell range is the *target cell range* for the call.
   The BS is the *target BS* for the call.
   Done and break the loop.
   Endif.
   Else
8. Find a larger but minimum cell range that covers the last-hop node.
9. If the remaining capacity of the range is enough for the demand μ,
   The larger cell range is the target cell range for the call.
   The BS is the target BS for the call.
   Done and break the loop.
   Endif.
   Endif.
10. End Loop (i.e., Repeat Step 6 for the next entry on the list)
11. If not Done,
   The call is blocked (i.e., not enough capacity for the call)
   Endif.

Definitions:

- **Last-hop node** is a node nearest to the BS on a relaying path.
- **Last-hop distance** is the distance of the last-hop node on a relaying path from its corresponding BS.
- **Target cell range** is the communication range of the target BS.
- **Target BS** is the BS to which a new call is proposed to connect.

Figure 5.4: The SCSF Algorithm

In Step 8, as the last-hop distance of the entry is greater than the current communication range (cell size) of the BS, a larger cell range to cover the last-hop node of the relaying path needs to be used.

In Step 9, if a larger but minimum cell range to cover the last-hop node is found and the remaining cell capacity of the BS at that range is enough to meet the demand μ of call, the range is the target cell range and the BS is the target BS. Iteration of the loop is finished.

In Step 10, if the assignment is not finished (i.e., not Done) and the iteration of the list is not exhausted, repeat Step 6 to try the next entry on the sorted list.

In Step 11, after the iteration is finished, if no Target BS is found (i.e., not Done), the call is blocked.

5.3.2. Illustration of SCSF algorithm

We use the topology and traffic pattern in Figure 5.5 to illustrate the SCSF algorithm. In the figure, assume that the connections for node *A*, *B*, *E*, *G*, *I*, and *K* are established, a new call request represented by a source point s_3 is placed at node *C* with a demand μ, and three relaying paths for the call to the three BSs (b_1, b_2, b_3) are C-B-b_1, C-G-b_2, and C-b_3 respectively. The last-hop distances of the paths to their corresponding BSs are $s_3(r_1, d(r_1, b_1)) = 4$, $s_3(r_6, d(r_6, b_2)) = 5.5$, and $s_3(s_3, d(s_3, b_3))= \infty$. Note that, as mentioned in Section 3.2, all mobile nodes are assumed to use a short range to communicate with other mobile nodes. Thus, in Figure 5.5, node *K* will not relay signals for node *C* as it is far away from node *C*. The last-hop distances of the paths with their BSs are stored on a list which is then sorted in ascending order based on the value of the distances. The sorted list becomes $\{(4, s_3, b_1), (5.5, s_3, b_2), (\infty, s_3, b_3)\}$.

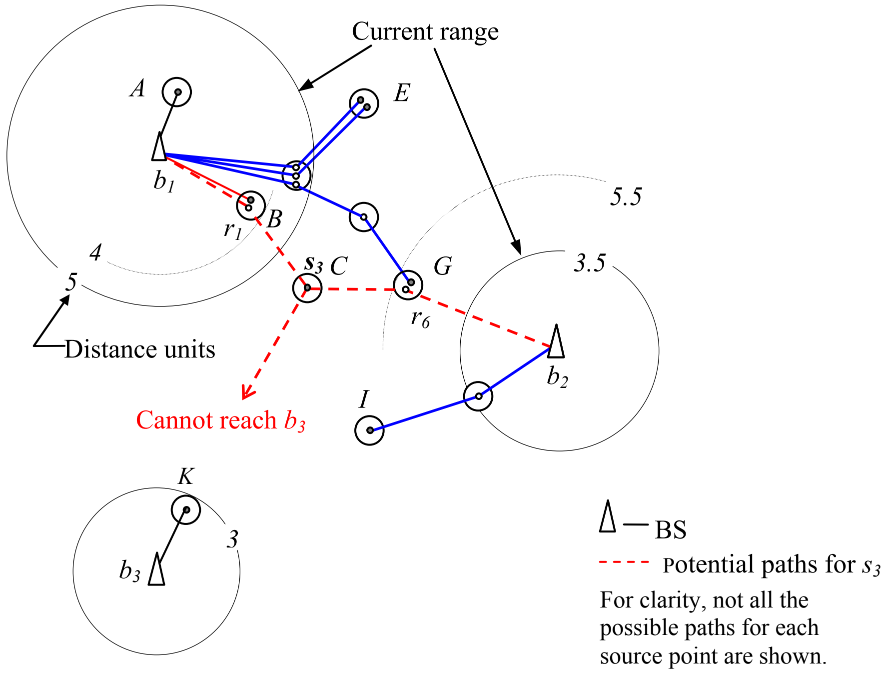


Figure 5.5: An example to illustrate the SCSF algorithm

The first potential target BS for the call is b_1 because b_1 requires the smallest cell range (4 units of distance) to cover the last-hop node of the path for the call. Since the last-hop distance of the path is 4 units which is smaller than the current cell range (5 units of distance) of b_1, the current cell range is a potential target cell range for the call. If the remaining capacity of b_1 of the current range is enough to meet the demand of the call, then b_1 is chosen as the target BS and the current cell range of b_1 is the target cell range; otherwise, the BS with the next shortest last-hop distance on the list will be the potential target BS, and, in this case, b_2 is the next potential target BS. Since the current cell range of b_2 is smaller than the last-hop distance of the last hop node of the path for the call to b_2, a larger cell range (5.5 units of distance) is required to cover the last-hop node of the path. If the remaining capacity (the capacity at range of 5.5 units of distance minus the used capacity for existing connections) is larger than the demand of the call, then b_2 is the target BS and the target cell range is 5.5 units of distance. If not, b_3 which is in the next entry on the list will be the potential target BS. Since the last-hop distance of the path for the call to b_3 is ∞, i.e., unreachable, there is no target BS in this case.

5.4. Highest Throughput Cell Size First (HTCSF)

Although SCSF is more adaptive than a static cell size strategy and may yield high throughput in some situations, it may not achieve high system throughput in some cases. For example, if the current cell range of a BS is large, i.e., there exists a last-hop node of a relaying path of a source node which is far away from the BS, the cell capacity is low. The cell capacity could be used up by only a small number of users. A new call could have a higher chance to be blocked even though its relaying path has short last-hop distance because the cell capacity is not enough. If the existing call continues for a long time, the cell capacity and the system throughput as well will be kept low during the period. In other words, the number of calls or the total data-rates or the system throughput is kept low. To avoid the domination of the source point which requires long last-hop distance, it may be better to sacrifice some of these existing calls that require a long or the longest last-hop distance for the BS. In this case, the cell range can be reduced for higher cell capacity for accepting not only the new call, but also more future calls. Based on this idea, we herein, propose a throughput-based cell size scheme, called HTCSF.

5.4.1. The HTCSF scheme

The basic idea of HTCSF is to adjust the cell size to achieve the highest throughput (i.e., the highest demand that can be served) when a new call arrives. In this case, when a new call arrives, the existing call(s) requiring long last-hop distance may need to be dropped so that a smaller cell size can be used to achieve higher cell capacity for the new call and future calls. Figure 5.6 shows the HTCSF algorithm which is the same as the SCSF algorithm except Step 7 in which an ELSE statement for checking if there is alternative connection assignment to achieve a better system throughput is added.

Steps of the HTCSF algorithm

All the steps are the same as the steps in SCSF except Step 7. In Step 7, if the remaining capacity of the BS is enough to meet the demand μ of the call, the current cell range of the BS is the target cell range and the BS is the target BS for the call; otherwise, find a smaller cell range such that the total demand that can be served is greater than or equal to the total current demand that can be served. If the smaller cell range is found, the range is the target cell range and the BS is the target BS.

By using the strategy in Step 7, the total demand that can be served is the highest and the capacity and/or remaining (unused) capacity increases. Higher remaining capacity also increases the chance of accepting future calls. The limitation of HTCSF is that existing calls having last-hop nodes which are far away from the BS, i.e., near to the edge of the current cell range, are likely to be dropped for a new call, which has a last-hop node closer to the BS when the cell size decreases. This may affect QoS provisioning and requires a QoS strategy to handle the situation. The QoS strategy could be a subject of future work.

Input: the demand of a new call, the relaying paths for the call to the BSs, and the capacity function of the BSs.
Output: the target BS for the call and the range of the BS.

1. Input the demand μ of a new call.
2. Input n relaying paths for the call.
3. Compute and store the *last-hop distance* of each relaying path on a list.
4. Sort the list based on the last-hop distances in an ascending order.

5. Loop through the list,
6. If last-hop distance <= current cell range of the corresponding BS,
7. If the remaining capacity of the BS is enough to meet the demand μ,
    The current cell range is the *target cell range* for the call.
    The BS is the *target BS* for the call.
    Done.

* Else
    If the total demand that a smaller cell range can meet > = the total demand that the current cell range can meet,
        The smaller cell range is the target cell range.
        The BS is the target BS for the call.
        Done.
    Endif.
  Endif.
Else

8. Find a larger but minimum cell range that covers the last-hop node.
9. If the remaining capacity of the range is enough for the demand μ,
    The larger cell range is the target cell range for the call.
    The BS is the target BS for the call.
    Done.
  Endif.
Endif.
10. End Loop (i.e., Repeat Step 6 for the next entry on the list)
11. If not Done,
    The call is blocked (i.e., not enough capacity for the call)
  Endif.

Definitions:
- **Last-hop node** is a node nearest to the BS on a relaying path.
- **Last-hop distance** is the distance of the last-hop node on a relaying path from its corresponding BS.
- **Target cell range** is the communication range of the target BS.
- **Target BS** is the BS to which a new call is proposed to connect.

Figure 5.6: The HTCSF Algorithm

5.4.2. Illustration of HTCSF algorithm

Similar to the illustration of SCSF, we use the example in Figure 5.5 to illustrate the HTCSF algorithm. Again, assume the connections for nodes *A*, *B*, *E*, *G*, *I*, and *K* are established and a new call request s_3 arrives. The relaying paths for the call for each of the BSs are *C*-*B*-b_1, *C*-*G*-b_2, and *C*-b_3. The set of last-hop distances of the paths, each goes to its corresponding BS, with their BSs is $\{(4, s_3, b_1), (5.5, s_3, b_2), (\infty, s_3, b_3)\}$.

BS b_1 is the first potential target BS for the call because the last-hop distance (4 units of distance) for the call is less than the current cell range (5 units of distance) of b_1. Then, if the remaining capacity of b_1 is enough to meet the demand of the call, b_1 is chosen as the target BS and the current cell range of b_1 remains the target cell range; otherwise, a smaller cell range of b_1 that covers the largest total demand giving the highest throughput is preferred. In this case, if the demand of the call is greater than or equal to the total demands of nodes *E* and *G*, a smaller cell range (4 units of distance) that covers the last-hop node of the path for the new call is used. Note that even if the demand of the call is equal to the total demand of nodes *E* and *G*, it may be worthy of sacrificing source nodes *E* and *G* for a smaller cell size to achieve a higher cell capacity as well as remaining (available) capacity. This increases the available capacity for future calls. If there is no smaller cell range that achieves at least the current demand, then go to Step 10 to check the next entry on the last-hop distance list for the potential target BS. In this case, b_2 is the next potential target BS. Steps 6 to 9 are repeated and so on.

5.5. Complexity Analysis

The cell size problem is a generalization of the Knapsack Problem which is an NP-hard problem [Corm01]. The idea of the Knapsack Problem is that given a knapsack of a certain capacity, fill the knapsack with objects which have various sizes and values such that the total value is maximized. For the cell size problem, a connection can be considered as an object which has a value of 1. The demand (data-rate) of the connection can be considered as the size of the object. In this case, the cell size problem is more complex than the Knapsack Problem because there is more than one BS and each BS may have different capacities. In other words, there is more than one knapsack and the capacity of each knapsack may be different. As the Knapsack Problem is NP-hard, we conjecture that our cell size problem is also NP-hard.

As the cell size problem is formulated as an Integer Programming Problem which is also NP-hard [Corm01], the running time for solving the cell size problem using the best known algorithm for integer programming is exponential. Thus, when the problem size is large, the time required to obtain the solution is prohibitive. By contrast, the heuristic cell size schemes, SCSF and HTCSF, are more efficient. For the SCSF algorithm, when a new call request of a source node is placed, several possible relaying paths for the source node are provided by a routing protocol. Each of the paths goes to a different BS. Only the last-hop distances of the paths are needed for the cell size computation. The last-hop distances are sorted in an ascending order and the capacity of the BS starting from the head of the list is checked. Sorting requires O(m log m) time where m is the number of BSs and checking requires O(m) time. Thus, the time complexity is O(m log m) plus O(m) which is O(m log m). The HTSCF algorithm requires an additional step of checking if the cell size can be reduced to sacrifice some existing calls which require a large cell size for the same or more servable demand. Thus, it requires to sort the last-hop distances of the mobile nodes which have ongoing calls. Sorting requires O(n log n) time where n is the number of mobile nodes which have ongoing calls. Thus, the time complexity of HTCSF is O(m log m) plus O(n log n). Since the number of mobile nodes n is much larger than the number of BSs m, the time complexity of HTCSF is O(n log n).

5.6. Performance Evaluation

We quantify the performance gain of HTCSF with respect to SCSF, OCS, a multi-hop small cell size (SCS) case and a general single-hop large cell size (LCS) case. The settings of SCS are the same as that of OCS/ SCSF/ HTCSF except that the cell size of SCS is small and fixed where as the cell size of OCS/ SCSF/ HTCSF is adjustable.

5.6.1. Simulation model and parameters

Our simulation model is a 3-Cell model (see Figure 5.7). Each cell has 25 source nodes and a number of relaying nodes varying from 0 to 160 in increments of 40. We choose 160 as the maximum number of relaying nodes because, at this setting, the network is dense enough such that most of the source nodes can reach the BS through multi-hopping. Further increase in the number of relaying nodes has little effect on meeting demand at the BS. The source nodes and relaying nodes are uniformly distributed over a circular area with a radius of 1.1 km centered at each BS. The roles of source node and relaying node are separated so that the case of no mobile nodes willing to relay signals can be captured.

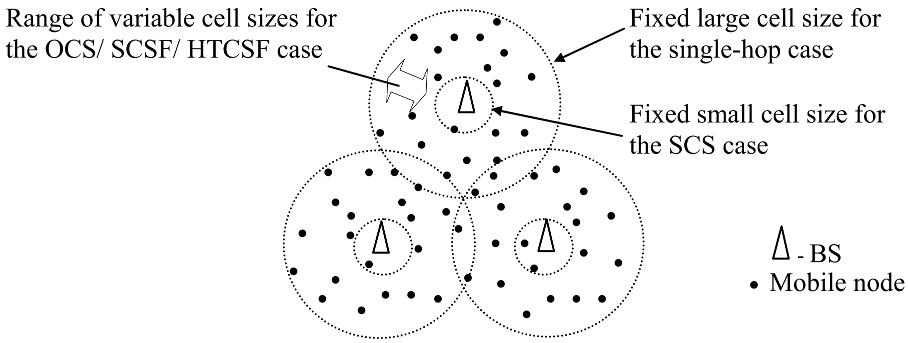


Figure 5.7: 3-Cell model

In this model, we vary the number of relaying nodes to model different nodal densities, traffic patterns and network topologies. Table 5.1 shows the simulation parameters. The cell ranges for SCS and the single-hop LCS case are 250m and 1.1 km, respectively. For OCS/ SCSF/ HTCSF, the cell size ranges from 250m to 1.1 km. The capacities corresponding to ranges are shown in Table 5.2. The values in the table are obtained by using the capacity function *U*(*R*) that we approximated by using curve fitting based on the sample data of uplink ranges with respect to the uplink capacity (data rates) in a suburban area for W-CDMA networks in [Holm04]. The sample data is shown in Table 5.3. The capacity function $U(R)= 245.56R^4-2717.13R^3+11245.43R^2-20736.4R+14511.99$, where *U*(*R*) is the cell capacity in terms of date-rate (kbps) and *R* is the transmission range of the BS. As the cell range *R* increases, the capacity *U*(*R*) of the cell in terms of data-rate in kbps decreases. The curve fitting software is available in MATLAB [Math08]. Note that we scale down the capacity values by a factor of 9.5 to reduce the simulation time. Since our interest is in the relative throughput gain of HTCSF as compared to that of OCS, SCSF and SCS, the absolute value is less significant in this case. For all the cases (OCS, SCSF, HTCSF and SCS), the communication range of the mobile nodes is fixed at 250m with a capacity of 1035 kbps unless they are the last-hop nodes. Each data transmission frame is 10ms long and consists of 15 time-slots according to the TDD WCDMA standard [Holm04]. Each time-slot can be assigned at most 5 codes and each code corresponds to a data rate of 13.8 kbps [Holm04]. Each call uses three slots and one code per slot at a constant bit rate. The duration of each call is 15 minutes. The maximum number of hops is set to 7 to avoid excessive delay. Each mobile node is equipped with a directional antenna with a 45^{o} beam angle to increase the spatial reuse. In all cases, we assume to use our heuristic channel assignment scheme, MSWF, which is described in previous chapter. A channel is represented by a time-slot and code pair. The duration for the simulation is 15 minutes. The simulation is modeled with OPNET Modeler 10.0 A [Opne08]. The optimization package used is MOSEK version 5 [Mose08]. A 90% confidence level with confident intervals [$\bar{x}$ - 10% $\bar{x}$, $\bar{x}$ +10% $\bar{x}$] is used in the simulation [Kinn96].

Table 5.1: SIMULATIONS PARAMETERS

	OCS/ SCSF/ HTCSF	SCS	Single-hop LCS
BS or last-hop nodal range	250 ~ 1100 m	250 m	1100 m
Nodal range	250 m		1100 m
Number of time-slots/ frame	15		
Max. hop count	7		
Data rate per code	13.8 kbps		
Number of source nodes/ cell	25		
Call request rates	0.5 calls/min.		
Call holding time	15 min.		
Antenna	directional antenna with beam angle 45^{o}		
Channel assignment scheme	MSWF		
Simulation duration	15 min.		

Table 5.2: CAPACITY CORRESPONDING TO THE CELL RANGE

Range (m)	250	390	560	780	1100
Capacity (kbps)	1035	828	621	414	207
Capacity (codes/ time-slot)	5	4	3	2	1

Table 5.3: UPLINK RANGE VERSUS CAPACITY

Range (km)	1.1	1.4	1.75	2.25	2.7	3.1
Capacity (kbps)	2048	1024	384	144	64	32

The OPNET Modeler is used to generate the network topology. The distance graph is computed using the Euclidean shortest paths as relaying paths. For the OCS case, the distance graph is input to MOSEK to compute the optimal cell sizes and the connection assignment of each source point. The optimal cell sizes and connection assignments are transferred back to OPNET for the simulation to obtain the throughput and the other performance metrics.

In this simulation, our focus is to quantify the relative throughput among HTCSF, SCSF, OCS, SCS and the single-hop LCS case. We assume the mobile nodes to be static (or with limited mobility) so that the results will not be affected by the effectiveness of the routing protocol. High mobility may cause frequent disconnections and may require updating of the paths by the routing protocol and the cell sizes by HTCSF, SCSF and OCS. This increases the control overhead, but does not affect the optimality of OCS or the effectiveness of HTCSF and SCSF. We assume perfect power control so that the cell capacity function remains unchanged during the simulation. We also assume a perfect physical medium, and sufficient battery capacity for relaying signals.

5.6.2. Performance metrics

We use the following metrics to evaluate the performance of HTCSF, SCSF, OCS, SCS, and the single-hop LCS case.

Cell throughput – the number of packets received at BSs receives per second. High throughput represents a good choice of cell size that provides a good balance between the cell capacity and the network reachability that maximizes the demands being served.

Call acceptance ratio (*AR*) – the ratio between the number of accepted calls and the total number of calls. High call acceptance ratio represents a good cell size that provides a good combination of cell capacity and network reachability to achieve high demands being served.

Cell size – the communication range of the BS. Large cell size increases the network reachability, but reduces the cell capacity and vice versa.

Packet delay - the time required for a packet sent from the source node to reach the BS. Low packet delay represents the effectiveness of the channel assignment scheme.

5.6.3. Simulation results

Considering Figures 5.8, 5.9, and 5.10, we observe that when the number of relaying nodes is zero, the HTSCF, SCSF, OCS and SCS cases are reduced to the single-hop case. For the SCS case, many source nodes cannot reach the BS to use the available cell capacity because the cell size is small (low network reachability). Therefore, the cell (or BS) throughput and the call acceptance ratio (*AR*) are low. For the large cell size single-hop case, although the cell size is large (high network reachability), the cell capacity is too small to meet the demand. Thus, the throughput and the *AR* are also low. For the OCS case, the average cell size used for the BSs is 731m in radius to maximize the demands that can be served. This provides an optimal balance between the cell capacity and the network reachability to maximize the demands that can be served. Therefore, the throughput and *AR* are the highest. The single-hop case has higher throughput and *AR* than that of SCS because it has a better combination of coverage and capacity than that of SCS. The throughput of both SCSF and HTCSF cases is higher than that of the SCS and the single-hop LCS cases because they can adjust the cell sizes to better values to accommodate more demand.

As the number of relaying nodes increases, more source nodes can reach the BSs through multi-hop relaying which increase the reachable demand. Thus, the throughput and *AR* of SCS increase. In the HTSCF, SCSF and OCS cases, the cell sizes are adjusted to a smaller value to achieve a higher cell capacity to meet the increased reachable demand. Thus, the throughput of all the three cases increases whereas the OCS case has the highest throughput.

When the number of relaying nodes reaches 120 and beyond, the network reachability is no longer an issue, but the capacity is because most source nodes can reach the BSs through relaying. OCS uses the same (small) cell size as that of SCS to achieve maximum capacity to serve the demands. Thus, OCS and SCS achieve the highest throughput. OCS has a slightly higher throughput than that of SCS because OCS makes use of the abundant relaying paths and assigns the source nodes (points) optimally among the three BSs whereas SCS does not have this feature. The throughput performance of HTCSF and SCSF also increases. HTCSF and SCSF has a lower throughput than that of the SCS case because the cell size of HTCSF and SCSF could be dominated by some source nodes which require a large cell size (small cell capacity) to cover them. Among the SCSF and HTCSF cases, when the network is very dense, HTCSF has higher throughput than SCSF because HTCSF has the ability to sacrifice some existing calls of some source nodes which requires large cell size in favour of a smaller cell size with higher cell capacity for a new call and future calls. SCSF is less flexible compared to HTCSF in this case. Although SCS is better than HTCSF and SCSF in terms of throughput in a high dense network, HTCSF and SCSF are better in term of flexibility to accommodate the calls especially when the traffic pattern is unknown in prior and/or the density of the network is not uniformly distributed.

In Figure 5.11, although the packet delay of HTSCF, SCSF, OCS and SCS cases is higher than that of a single-hop LCS case because multi-hopping is involved, the delay is still considered to be low in a multi-hop environment. This is due to the effectiveness of our channel assignment scheme, MSWF.

Figure 5.12 shows the delay-throughput characteristic of HTSCF, SCSF, OCS and SCS which follows a general trend in multi-hopping situation. We observe that OCS always achieves the highest throughput and *AR* because OCS always gives the optimal cell sizes for different nodal densities and traffic patterns and assigns the source nodes (points) optimally among the three BSs. Such characteristic makes OCS a useful planning aid in network planning process. In general, OCS has on average 67% throughput higher than that of SCS when the network density is sparse to medium (see Figure 5.8). HTCSF and SCSF on average achieve 78% and 71%, respectively, of the performance of OCS in terms of throughput. Hence, we claim HTCSF and SCSF provides good results compared to the optimal solution provided by OCS. Both HTCSF and SCSF outperform SCS when the network is sparse and the single-hop LCS case regardless of the network density.

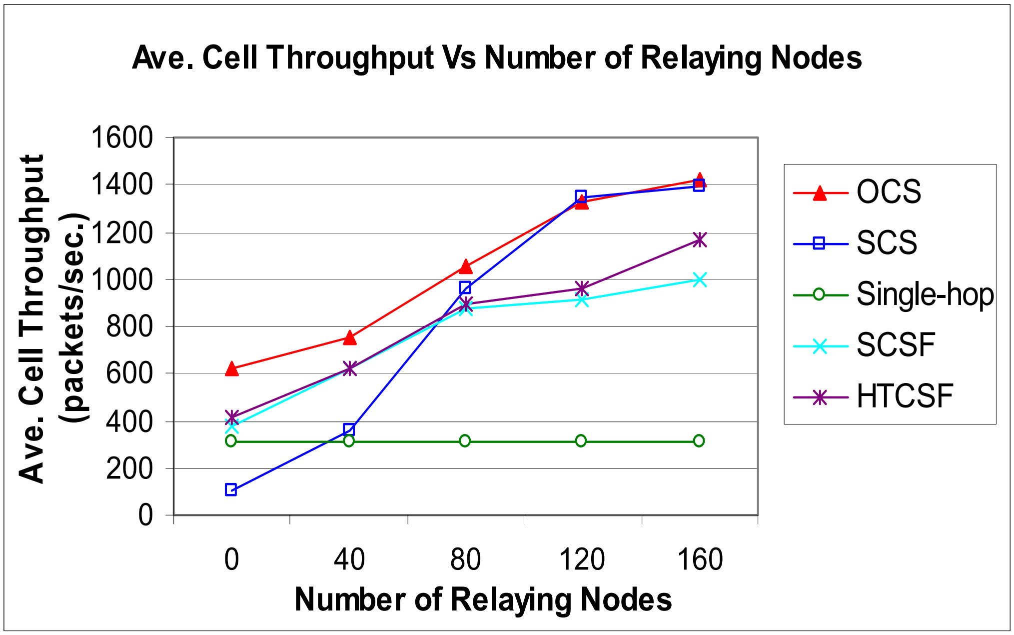


Figure 5.8: Cell Throughput

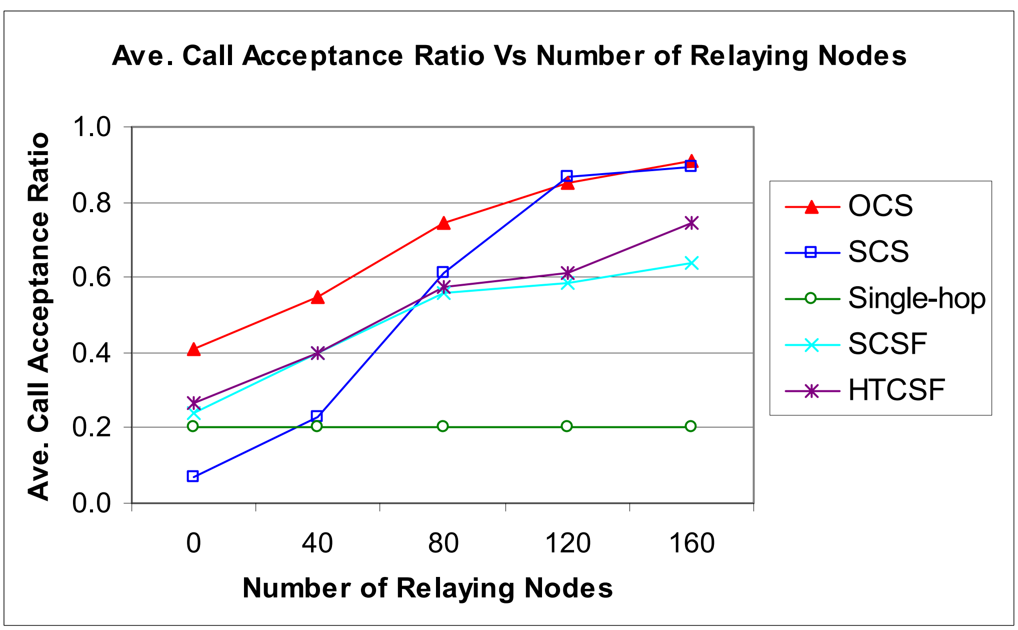


Figure 5.9: Call Acceptance Ratio

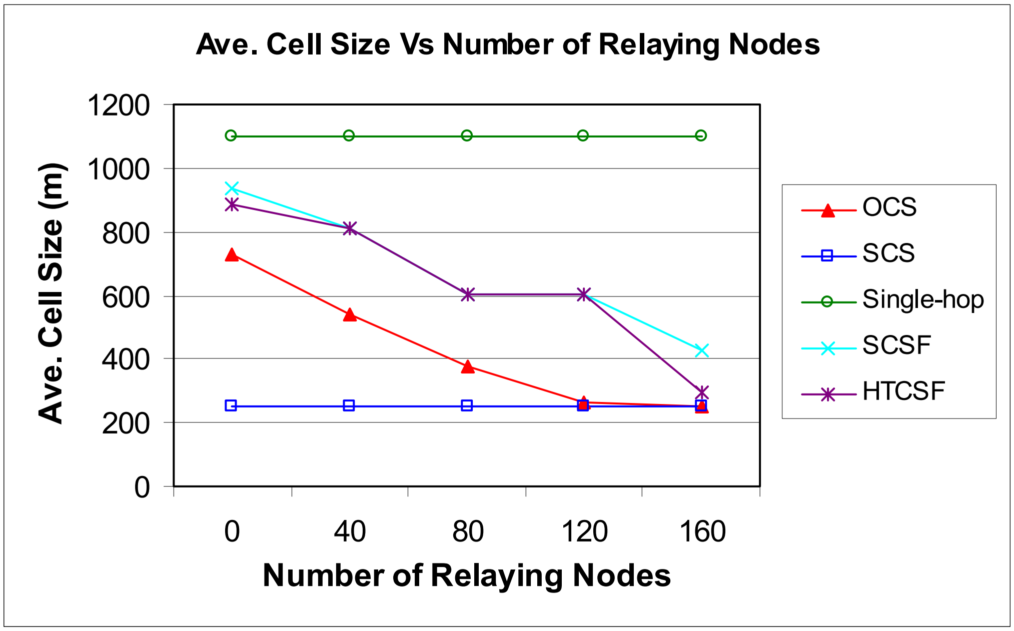


Figure 5.10: Cell Size

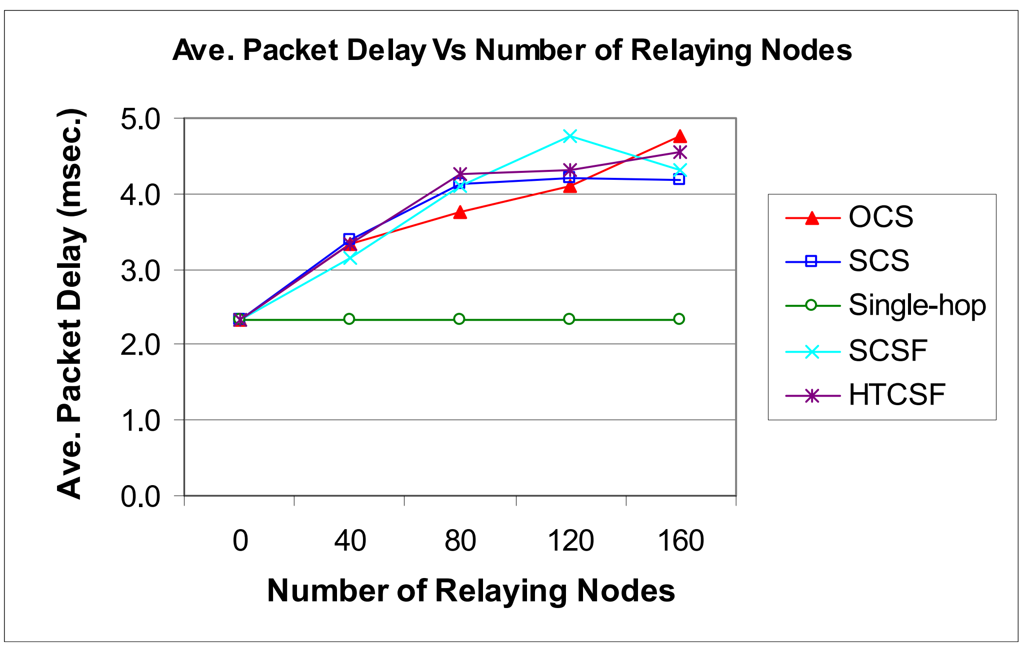


Figure 5.11: Packet Delay

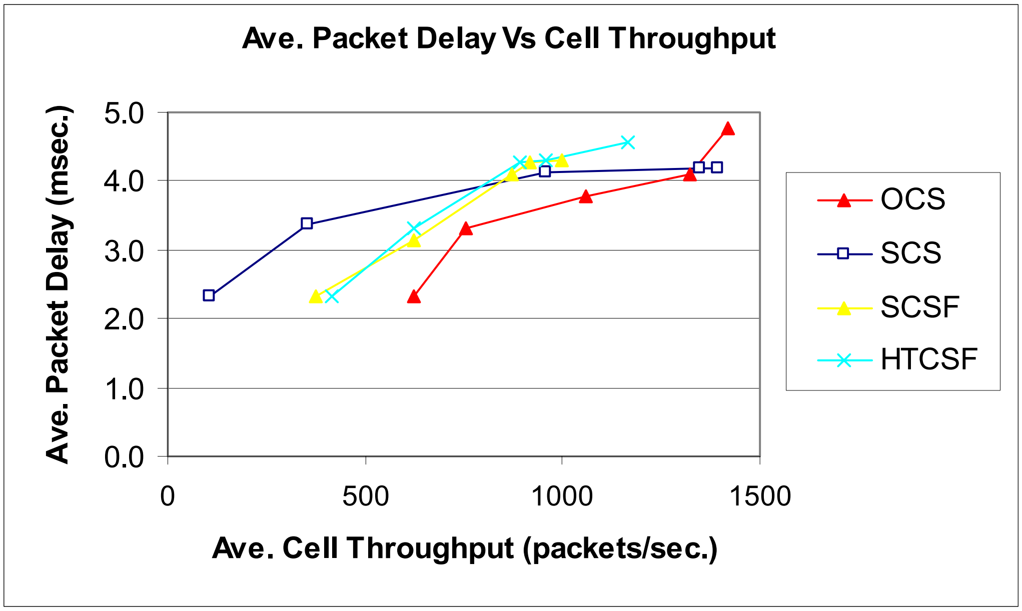


Figure 5.12: Packet Delay versus Cell Throughput

5.7. Limitations

As mentioned at the beginning of this chapter, OCS is computationally expensive and inefficient for large real-time cell size problems. When the problem size is large, OCS is computationally prohibitive. SCSF and HTCSF are more efficient for large real-time problems. However, SCSF and HTCSF cannot guarantee maximum throughput.

OCS, SCSF and HTCSF require overhead for collecting information, such as the location of the last hop nodes and the demand of the source nodes need to compute the cell size. Control overhead is required to communicate the cell size results between the BS and the RNC. This also increases the complexity of the system.

The cell size of SCSF may be dominated by source nodes, which require a large cell size to cover. Thus, throughput cannot be maximized. HTCSF is better than SCSF in terms of throughput. However, it is required to sacrifice (drop) some existing calls which may not be acceptable to the users. A QoS strategy may need to be designed to handle this situation.

5.8. Summary

In this chapter, we proposed the OCS scheme and two heuristics, called SCSF and HTCSF, for a TDD W-CDMA MCNs. OCS computes optimal cell sizes which provide an optimal balance between cellular capacity and coverage that maximizes the system throughput. OCS can be considered as a network planning aid for cellular systems and MCN. As OCS is computationally expensive, the heuristics, SCSF and HTCSF, are more suitable for solving large real-time cell size decision problems. Simulation results show that OCS has on average 67% throughput higher than that of the fixed SCS multi-hop case when the network is sparse and medium dense. Simulation results also show that HTCSF and SCSF on average achieve 78% and 71%, respectively, throughput performance of OCS. HTCSF and SCSF provide good results compared to the optimal solutions provided by OCS, outperforms SCS when the network is sparse and the single-hop LCS case regardless of the network density.

Chapter 6

Conclusions and Future Work

For multi-hop cellular networks (MCNs), cell size and channel assignment are two important design factors. Cell size is related to the cost, capacity, routing efficiency, and packet delay in these networks. A combination of cell size and network density also affects the coverage or network reachability which in turn affects the total demand that can be served and, hence, the radio resource utilization and system throughput. Channel assignment affects the channel reuse and packet delay in these networks. In a time division duplex (TDD) wideband code division multiple access (W-CDMA) MCN where directional antennas are used, to reuse the channels while minimizing packet delay and avoiding signal collisions is a non-trivial task. Improper channel assignment would greatly affect the packet delay and the system throughput. In this thesis, we study the channel assignment problem and cell size problem in TDD W-CDMA MCN. We proposed schemes to solve these problems.

For the channel assignment problem, we proposed the Optimal Channel Assignment (OCA) scheme to minimize the packet delay in these networks and any TDD MCN. OCA can also be used as an unbiased benchmark tool for performance comparison of different network conditions and networking schemes. As OCA is computationally expensive, we also proposed a heuristic, called Minimum Slot Waiting First (MSWF), for solving large real-time problems. Simulation results show that MSWF provides a good result compared to the optimal solution provided by OCA in terms of packet delay. MSWF has a low per-hop packet delay which is close to the per-hop packet delay lower bound. It achieves high throughput as compared to the single-hop case and is applicable for different cell sizes. It also achieves low delay in MCNs.

For the cell size problem, we proposed the Optimal Cell Size (OCS) scheme and two heuristics, called Small Cell Size First (SCSF) and Highest Throughput Cell Size First (HTCSF), for these networks. OCS computes optimal cell sizes providing an optimal balance between cellular capacity and coverage that maximizes the system throughput. OCS can also be considered as a network planning aid for cellular systems and MCN. As OCS is computationally expensive, the heuristics, SCSF and HTCSF, are more suitable for solving large real-time cell size problems. Simulation results show that OCS outperforms the fixed small cell size (SCS) multi-hop case when the network is sparse and medium dense. Simulation results also show that HTCSF and SCSF provide good results compared to the optimal solution provided by OCS in terms of throughput. HTCSF and SCSF outperform SCS when the network is sparse and the single-hop large cell size (LCS) case regardless of the network density.

Future work

Channel assignment and cell size are only two aspects in MCNs. To realize the concept of MCNs, other aspects such as mobility, handoff, routing, power control, and QoS need to be considered.

In an environment with high user mobility, the chance of broken relay paths increases. Frequent disconnections and/or handoffs of ongoing calls may occur. The frequency of route re-discovery and re-establishment also increases, which in turn increases the control overhead of the system. Therefore, an effective routing protocol and an effective handoff scheme need to be developed.

In a CDMA network, power control is important to keep the interference level low and to maintain the capacity and cell size, which greatly affects the performance of the networks. In 3G systems, there are open loop power control and close-loop power control mechanisms to control the power of mobiles and BSs. The former does not require mobile nodes to give feedback of their power information to the BS whereas the latter one does. In a CDMA MCN environment, the interference pattern is more complex as there are many simultaneous signal transmissions for communication among mobile nodes and between mobile nodes and the BSs. An effective power control scheme needs to be designed.

Quality of service (QoS) is a fundamental issue in wireless networks. In 3G or future generation wireless networks, mobile users are provided with different services, such as on-line video gaming, streaming video, and video conferencing, which have different QoS level requirements in terms of delay, response time, and data-rates. In MCNs, multi-hopping increases packet delay and mobility and limited battery power of relaying nodes cause link instability or failure. These factors make QoS provisioning in MCNs more difficult. Other performance metrics such as fairness, delay jitter, aggregate throughput and packet loss ratio may also need to be considered. A comprehensive QoS framework may be needed to address these issues.

BIBLIOGRAPHY

[Agge01] G. N. Aggelou and R. Tafazolli, "On the Relaying Capability of Next-Generation GSM Cellular Networks", *IEEE Personal Communications*, vol. 8, no. 1, Feb. 2001, pp. 40-47.

[Alri05] M. Al-Riyami, A. M. Safwat and H. S. Hassanein, "Channel Assignment in Multi-hop TDD W-CDMA Cellular Networks," *IEEE International Conference on Communications* (*ICC*), vol. 3, May 2005, pp. 1428–1432.

[Anti99] 3rd Generation Partnership Project (3GPP); Technical Specification Group Radio Access Network; "Opportunity Driven Multiple Access," Sophia Antipolis, Valbbonne, 1999. (3G TR 25.924 version 1.0.0)

[Avoi05] G. Avoine, "Fraud Within Asymmetric Multi-hop Cellular Networks," *Financial Cryptography and Data Security*, vol. 3570 of LNCS, Feb. 2005, pp.1-15.

[Brél79] D. Brélaz, "New Methods to Color the Vertices of a Graph," *Communications of the ACM*, vol. 22, no. 4, April 1979, pp. 251–256.

[Bhar04] B. Bhargava, X. Wu, Y. Lu and W. Wang, "Integrating Heterogeneous Wireless Technologies: a Cellular Aided Mobile Ad hoc Network (CAMA)," *ACM Mobile Networks and Applications*, vol. 9, no. 4, Aug. 2004, pp.393–408.

[Chan03] R. Chang, W. Chen and Y. Wen, "Hybrid Wireless Network Protocols," *IEEE Transaction s Vehicular Technology*, vol. 52, no. 4, July 2003, pp. 1099-1109.

[Chen03] C.T. Chen, S. Tekinay and S. Papavassiliou, "Geocasting in Cellular Ad Hoc Augmented Networks," *IEEE Vehicular Technology Conference (VTC)*, vol. 3, Oct. 2003, pp. 1858-1862.

[Chen04] S. Chen, A. Hirata, T. Ohira and N. C. Karmakar, "Fast Beamforming of Electronically Steerable Parasitic Array Radiator Antennas: theory and experiment," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 7, July 2004, pp. 1819-1832.

[Corm01] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, "Introduction to Algorithms," 2nd ed., MIT Press, 2001.

[De02] S. De, O. Tonguz, H. Wu and C. Qiao, "Integrated Cellular and Ad hoc Relay (iCAR) Systems: Pushing the Performance Limits of Conventional Wireless Networks," *IEEE Annual Hawaii International Conference on System Sciences (HICSS),* vol. 9, Jan. 2002, pp. 3899-3906.

[Dimo08] T.D. Dimousios, C. I. Tsitouri, S. C. Panagiotou and C. N. Capsalis, "Design and Optimization of a Multipurpose Tri-band Electronically Steerable Passive Array Radiator (ESPAR) Antenna with Steerable-beam-pattern for Maximum Directionality at the Frequencies of 1.8, 1.9 and 2.4 GHz with the Aid of Genetic Algorithms," *IEEE Antennas and Propagation Conference*, Mar. 2008, pp. 253-256.

[Esma03] R. Esmailzadeh, M. Nakagawa and A. Jones, "TDD-CDMA for the 4th Generation of Wireless Communications," *IEEE Wireless Communications*, vol. 10, no.4, Aug. 2003, pp. 8-15.

[Fu06] X. Fu, A. G. Bourgeois, P. Fan and Y. Pan, "Using a Genetic Algorithm Approach to Solve the Dynamic Channel-Assignment Problem," *International Journal of Mobile Communications*, vol. 4, no. 3, 2006, pp. 333–353.

[Gyod00] K. Gyoda and T. Ohira, "Design of Electronically Steerable Passive Array Radiator (ESPAR) Antennas," *IEEE Antennas and Propagation Society International Symposium*, vol. 2, July 2000, pp.922-925.

[Holm04] H. Holma and A. Toskala, "*WCDMA for UMTS, Radio Access for Third Generation Mobile Communications*," 3rd ed., John Wiley & Sons, 2004.

[Hsu02] Y. C. Hsu and Y. D. Lin, "Base-Centric Routing Protocol for Multi-hop Cellular Networks," *IEEE GLOBECOM*, vol. 1, Nov. 2002. pp. 158 – 162.

[Jako03] M. Jakobsson, J.-P. Hubaux and L. Buttyán, "A Micro-Payment Scheme Encouraging Collaboration in Multi-hop Cellular Networks," *Financial Cryptography*, vol. 2742 of LNCS, Jan. 2003, pp.15-33.

[John96] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad hoc Wireless Networks," *Mobile Computing*, Kluwer Academic Publishers, vol. 353, 1996, pp. 153-181.

[Kinn97] J. J. Kinney, "*Probability: An Introduction with Statistical Applications*," 1st ed., John Wiley & Sons, 1997.

[Ko00] Y.-B. Ko and N. H. Vaidya, “Location-Aided Routing (LAR) in Mobile Ad hoc Networks,” *Wireless Networks*, vol. 6, no.4, July 2000, pp.307-321.

[Kwon02] Y. H. Kwon and D. C. Lee, “An Uplink Packet Relay Protocol for CDMA Cellular-like systems,” *MILCOM 2002*, vol. 2, Oct. 2002, pp. 940–945.

[Kudo05] E. Kudoh and F. Adachi, “Power and Frequency Efficient Wireless Multihop Virtual Cellular Concept,” *IEICE Transactions on Communications*, vol. E88-B, no. 4, April 2005, pp.1613–1621.

[Luo07] H. Luo, X. Meng, R. Ramjee, P. Sinha and Li Li, “The Design and Evaluation of Unified Cellular and Ad-Hoc Networks,” *IEEE Transactions on Mobile Computing*, vol. 6, no. 9, Sept. 2007, pp. 1060–1074.

[Levi03] Anany Levitin, “*Introduction to the Design and Analysis of Algorithms*,” Addison-Wesley, 2003.

[Li02] H. Li, M. Lott, M. Weckerle, W. Zirwas and E. Schulz, “Multihop Communications in Future Mobile Radio Networks,” *IEEE Personal, Indoor and Mobile Radio Communications,* vol. 1, Sept. 2002, pp. 54–58.

[Li03] H. Li, D. Yu, and H. Chen, “New Approach to Multihop – Cellular Based Multihop Network,” *IEEE Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 2, Sept. 2003, pp. 1629–1633.

[Li06] X.J. Li, P. Han and P. H. J. Chong, “A Fixed Channel Assignment Scheme for Multihop Cellular Network,” *IEEE GLOBECOM*, Nov 2006, pp. 1–5.

[Li08] X. J. Li, B. Seet and P. H. J. Chong, “Multihop Cellular Networks: Technology and Economics,” *Computer Network: The International Journal of Computer and Telecommunications Networking,,* vol. 52, no. 9, June 2008, pp. 1825-1837.

[Lin00] Y. D. Lin and Y.C. Hsu, “Multihop Cellular: A New Architecture for Wireless Communications,” *IEEE INFOCOM, IEEE Computer and Communications Societies.* vol. 3, Mar. 2000, pp.1273–1282.

[Math08] MATLAB, www.mathworks.com, Dec. 2008

[Opne08] The MOSEK Optimization Software, www.mosek.com, Dec. 2008.

[Mose08] OPNET Technologies Inc., www.opnet.com, Dec. 2008.

[Norr04] L. Norr, and A. Anpalagan, "Dynamic Channel Allocation in TDD-CDMA systems," *Telecommunications, IEEE Canadian Review*, summer 2004, pp.9-13.

[Perk99] C. Perkins and E. Royer, "Ad hoc On-demand Distance Vector Routing," *IEEE Workshop on Mobile Computing Systems and Applications*, Feb. 1999, pp. 90-100.

[Radw06] A. Radwan and H. Hassanein, "Capacity Enhancement in CDMA Cellular Networks using Multi-hop Communication," *IEEE ISCC*, June 2006, pp. 832–837.

[Rapp02] T. S. Rappaport, "*Wireless Communications, Principles and Practice*," 2nd ed., Prentice Hall, 2002.

[Safw03] A. Safwat, "A-Cell: A Novel Multi-hop Architecture for 4G and 4G+ Wireless Networks," *IEEE VTC*, vol. 5, Oct. 2003, pp.2931-2935.

[Safw04] A. Safwat, "ACA: Channel Assignment in Ad hoc, 4G and Beyond Wireless Networks with Directional Antennas," *IEEE International Conference on Communications (ICC),* vol. 6, June 2004, pp.3143 – 3147.

[Sale03] N. B. Salem, L. Buttyán, J.-P. Hubaux and M. Jakobsson, "A Charging and Rewarding Scheme for Packet Forwarding in Multi-hop Cellular Networks," *ACM Mobile Ad hoc Networking and Computing (MobiHoc),* June 2003, pp. 13-24.

[Tam05] Y. H. Tam, A. M. Safwat and H. S. Hassanein, "A Load Balancing and Relaying Framework for TDD W-CDMA Multi-hop Cellular Networks," *IFIP-TC6 Networking Conference*, May 2005, pp. 1267-1280.

[Tam06a] Y. H. Tam, H. S. Hassanein and S. G. Akl, "Effective Channel Assignment in Multi-hop W-CDMA Cellular Networks," *ACM International Wireless Communications and Mobile Computing Conference (IWCMC)*, July 2006, pp. 569-574.

[Tam06b] Y. H. Tam, H. S. Hassanein, S. G. Akl, and R. Benkoczi. "Optimal Multi-hop Cellular Architecture for Wireless Communications," *IEEE P2MNet*, Nov. 2006, pp. 738-745.

[Tam07] Y. H. Tam, R. Benkoczi, H. S. Hassanein, and S. G. Akl. "Optimal Channel Assignment in Multi-hop Cellular Networks," *IEEE GLOBECOM*, Nov. 2007, pp. 731-735.

[Tam08] Y. H. Tam, R. Benkoczi, H. S. Hassanein, and S. G. Akl. "Optimal Cell Size in Multi-hop Cellular Networks," *IEEE GLOBECOM*, Nov. 2008, pp. 1-5.

[Tane03] A. S. Tanenbaum, "*Computer Networks*," 4th edition, Prentice Hall, 2003.

[Tao07] Z. Tao, A. Li, K. H. Teo and J. Zhang, "Frame Structure Design for IEEE 802.16j Mobile Multihop Relay (MMR) Networks," *IEEE GLOBECOM*, Nov. 2007, pp. 4301-4306

[Toh01] C. K. Toh, "*Ad hoc Mobile Wireless Network, Protocols and Systems*," 1st ed., Prentice Hall, 2001.

[Weyl04] A. Weyland and T. Braun, "Cooperation and Accounting Strategy for Multi-hop Cellular Networks," *IEEE LANMAN*, April 2004, pp.193-198.

[Wu00] X. Wu, S. H. G. Chan, and B. Mukherjee, "MADF: A Novel Approach to Add an Ad hoc Overlay on a Fixed Cellular Infrastructure", *IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 2, Sept. 2000, pp. 549-554.

[Xie06] B. Xie, A. Kumar, D. P. Agrawal and S. Srinivasan, "Secured Macro/micro-mobility Protocol for Multi-hop Cellular IP," *Journal of Pervasive and Mobile Computing*, vol. 2, no. 2, April 2006, pp. 111-136.

[Yama02] Y. Yamao, T. Ostu, A. Fujiwara and S. Yoshida, "Multi-hop Radio Access Cellular Concept for Fourth Generation Mobile Communications System", *IEEE Personal, Indoor and Mobile Radio Communications*, vol. 1, Sept. 2002, pp. 59–63.

[Zade02] A.N. Zadeh, B. Jabbari, R. Pickholtz and B. Vojcic, "Self-organizing packet radio ad hoc networks with overlay (SOPRANO)," *IEEE Communications Magazine*, vol. 40, no. 6, June 2002, pp.149–157.

[Zhou02] J. Zhou and Y. Yang, "PARCelS: Pervasive Ad hoc Relaying for Cellular Systems," *Med-Hoc-Net,* Sep. 2002.