Linköping Studies in Science and Technology
Dissertation No. 1029

Applications of Resource Optimization in Wireless Networks

Patrik Björklund

Linköpings universitet
INSTITUTE OF TECHNOLOGY

Department of Science and Technology
Linköpings Universitet
SE-601 74 Norrköping, Sweden

Norrköping 2006

Applications of
Resource Optimization
in Wireless Networks

Patrik Björklund

patbj@itn.liu.se
http://www.liu.se
Department of Science and Technology

ISBN 91-85523-41-0 ISSN 0345-7524

Printed by LiU-Tryck, Linköping, Sweden, 2006.

Abstract

The demand for wireless communications is increasing every year, but the available resources are not increasing at the same rate. It is very important that the radio resources are used in an efficient way allowing the networks to support as many users as possible. The three types of networks studied in this thesis are frequency hopping GSM networks, ad hoc multi-hop networks and WCDMA networks.

One type of network with a promising future is ad hoc multi-hop networks. The users in this kind of networks communicate with each other without base stations. Instead the signal can be sent directly between two users, or relayed over one or several other users before the final destination is reached. Resources are shared by letting the users transmit in time slots. The problem studied is to minimize the number of time slots used, when the users broadcast. Two different optimization models are developed for assigning time slots to the users. A reduction of the number of time slots means a shorter delay for a user to transmit next time.

The rapid growth of the number of subscribers in cellular networks requires efficient cell planning methods. The trend of smaller cell sizes in urban areas for higher capacity raise the need for more efficient spectrum usage. Since the infrastructure of a second generation cellular system, such as GSM, already exists, and the available bandwidth of an operator is limited, frequency planning methods are of utmost importance. Because of the limited bandwidth in a GSM network, the frequencies must be reused. When planning a GSM network the frequencies can not be reused too tightly due to interference. The frequency planning problem in a GSM network is a very complex task. In this thesis an optimization model for frequency assignment in a frequency hopping GSM network is developed. The problem is to assign frequencies to the cells in the network, while keeping the interference to a minimum. Different meta heuristic methods such as tabu search and simulated annealing are used to solve the problem. The results show that the interference levels can be reduced to allow a capacity increase.

The demand for sending more information over the wireless communication systems requires more bandwidth. Voice communication was handled well by the second generation cellular systems. The third generation of mobile telecommunication systems will handle data transmissions in a greater extent. The last type of network considered in the thesis is a WCDMA cellular network. The aim is to schedule the transmission of packet data from the base station to the users. Scheduling models that maximize the utility are developed for both the downlink shared channel and the high speed downlink shared channel.

Acknowledgments

I am very grateful to all the people that have supported me during the work with this thesis. First of all, I would like to thank my supervisors Peter Värbrand and Di Yuan for the fantastic support and advice during the work with this thesis. For the ad hoc part I wish to thank the research group at the Department of Communication Systems, Swedish Defence Research Agency (FOI), for the technical discussions and the test data. I am also very grateful to Joachim Samuelsson who was very helpful with the GSM part of this thesis. For the WCDMA part I am very grateful for the input from Niclas Wiberg, Eva Englund and Gunnar Bark form Ericsson Reseach in Linköping. Without their help the last part of the thesis would never be finished. My colleagues at the Department of Science and Technology, deserve special thanks for their support.

And, I am very grateful to my wife Ingela for her support and encouragement at all times. And at last to my lovely children Adrian and Oliver who has made my life so rich.

Norrköping, July 2006

Patrik Björklund

Contents

Abbreviations

3G	Third Generation
3GPP	3rd Generation Partnership Project
AD	Add and Drop
AMC	Adaptive Modulation and Coding
AMPS	Advanced Mobile Telephone System
AUC	Authentication Center
BAP	Bandwidth Allocation Problem
BCCH	Broadcast Control Channel
BSC	Base Station Controller
BSS	Base Station Subsystem
BTS	Base Transceiver Station
CDMA	Code Division Multiple Access
CN	Core Network
CQI	Channel Quality Index
DCA	Dynamic Channel Assignment
DCH	Dedicated Channel
DSCH	Downlink Shared Channel
EDGE	Enhanced Data Rates for GSM Evolution
FAP	Frequency Assignment Problem
FAPH	Frequency Hopping Assignment Problem
FCA	Fixed Channel Allocation
FDMA	Frequency Domain Multiple Access
FH	Frequency Hopping
GGSN	Gateway GPRS Support Node
GMSC	Gateway Mobile Switching Center
GPRS	General Packet Radio System
GR	Greedy
GSM	Global System for Mobile Communication
HARQ	Hybrid Automatic Repeat Request
HLR	Home Location Register
HSBAP	High Speed Downlink Packet Access Bandwidth Allocation Problem
HSDPA	High Speed Downlink Packet Access
HS-DPCCH	Uplink High Speed Dedicated Physical Control Channel
HS-DSCH	High Speed Downlink Shared Channel
HSN	Hopping Sequence Number
HS-SCCH	High Speed Shared Control Channel
HSUPA	High Speed Uplink Packet Access
IP	Integer Programming
LP	Linear Programming
LSCF	Set Covering for Link Slot Formulation
LSF	Link Slot Formulation

MAIO	Mobile Allocation Index Offset
MAL	Mobile Allocation List
ML	Modify Length
MLP	Minimum Length Scheduling Link Oriented Problem
MNP	Minimum Length Scheduling Node Oriented Problem
MS	Mobile Station
MSC	Mobile Switching Center
NMT	Nordic Mobile Telephone
NSCF	Set Covering Node Slot Formulation
NSF	Node Slot Formulation
NSS	Network and Switching Subsystem
OL	Optimize Length
OVSF	Orthogonal Variable Spreading Factor
PF	Proportional Fair
PLMN	Public Land Mobile Network
PSTN	Public Switched Telephone Network
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RNC	Radio Network Controller
RR	Round Robin
SF	Spreading Factor
SGSN	Serving GPRS Support Node
SIR	Signal to Interference Ratio
SNR	Signal to Noise Ratio
STDMA	Spatial Time Division Multiple Access
STRX	Super Transceiver
SZ	Size of the restricted neighborhood in tabu search
TACS	Total Access Communication System
TBL	Tabu List Length
TCH	Traffic Channel
TDMA	Time Division Multiple Access
TRX	Transceiver
TTI	Time Transmission Interval
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
UTRA	Universal Terrestrial Radio Access
UTRAN	UMTS Terrestrial Radio Access Network
VLR	Visitor Location Register
WCDMA	Wideband Code Division Multiple Access

Chapter 1

Introduction

The development of wireless communications started in the 1850's with the insight of J.C. Maxwell that energy could be transported without wires. The first to explore the ideas was Heinrich Hertz who transmitted a spark a few meters to a receiver. The electromagnetic waves were known as "Hertzian waves". But the first who used these electromagnetic waves for communication was an Italian engineer Guglielmo Marconi. In 1901 he was the first to transmit a message over the Atlantic Ocean from Cornwall to Massachusetts. Mobile communication is one of the fastest growing technologies today and the number of cellular phones is increasing rapidly every year. There is also a growing interest for wireless communication via laptops and personal digital assistants. One of the problems in mobile communications is that the resources available for communication are limited. The telecommunication operators are forced to invent new methods to increase the capacity of their existing networks. This is very important, because the grade of service for the customers must be maintained.

In the beginning of wireless communications the users were multiplexed by given different frequencies. This method is known as Frequency Domain Multiple Access (FDMA). When the number of applications and users grew the radio spectrum needed a classification to avoid unnecessary interference. The classification was done by the International Telecommunication Union which works for an efficient use of the radio spectrum. The radio spectrum is divided into different frequency bands depending on the applications. The frequencies are a limited resource in wireless communication systems and must be used with care. To multiplex more than one user to a frequency methods such as Time Division Multiple Access (TDMA) was used. In TDMA, a frequency is divided into several time slots (TDMA frame), with each time slot being equivalent to a transmitting channel. A user can transmit during one or several of these time slots. The TDMA frame should not be too long since it introduces delays. Another way to multiplex the users is Code Division Multiple Access (CDMA). In CDMA the users share one frequency but are multiplexed by different codes that are unique for each user. This method is used by some of the third generation (3G) cellular systems. The number of codes assigned to a frequency is limited and should be used as efficient as possible.

One limited resource is the radio spectrum that must be used in a very efficient way. Time slots or codes that are multiplexed to the frequencies must also be used very efficient. In this thesis methods are developed to use these resources effectively. The different network types studied in this thesis are multi-hop ad hoc networks, frequency hopping Global System for Mobile Communications (GSM) networks and 3G Wideband Code Division Multiple Access (WCDMA) networks.

1.1 Ad Hoc Networks

An ad hoc network is formed without any fixed central administration units. Instead the network consists of several mobile nodes or radio units that communicate with each other over a wireless interface. The radio units within the reach of each other can establish a direct communication link, if the signal-to-noise ratio is strong enough. Radio units that are not in the reach of each other must relay the signal over other nodes. Since a signal can be relayed, the networks are also referred to ad hoc multi-hop networks. Mobile ad hoc multi-hop networks have mainly been considered for military applications, since a centralized network would not be beneficial. However, there has also been a growing interest in ad hoc networks in the commercial sector in recent years. The main reason for this is that there is an increasing need for laptops, personal digital assistants etc. being able to communicate with each other.

One widely spread access method for ad hoc multi-hop networks is Spatial Time Division Multiple Access (STDMA). STDMA is an access method that takes into account that users are usually spread out geographically. Therefore, two users with a sufficient spatial separation can use the same time slot for transmission. Since users share the time slots, the frame size of STDMA can often be smaller than the frame size of TDMA.

Scheduling the users in ad-hoc networks is important and different approaches exist. Example of scheduling problems in ad hoc networks are assigning transmission rights to nodes or links with the objective to minimize the length of the STDMA frame or maximizing the throughput. The former problem is investigated in this thesis. Minimizing the STDMA frame is of interest since the schedule repeats itself from one frame to the next. Reusing the time slots achieve efficient resource utilization.

1.2 Evolution of Cellular Systems

The evolution of mobile cellular systems can be divided into three generations. The first generation of cellular systems included only analog systems. This generation used FDMA to handle several users simultaneously. In Europe different standards were introduced, for example the Nordic Mobile Telephone (NMT) in Scandinavia and Total Access Communication System (TACS) in United Kingdom, Spain, Austria and Ireland. In the United States the Advanced Mobile Telephone System (AMPS) was introduced. In Japan the JTACS/NTACS system were used. All these systems were introduced at the beginning of the 1980s. AMPS had the largest penetration of the analog systems in the world.

The research in the digital field soon made it possible to consider digital cellular systems. The performance of a cellular radio system depends upon the co-channel interference, and voice quality could be increased by the use of digital transmission. Since digital systems can tolerate a much higher level of co-channel interference, the frequency reuse could be tighter and a capacity increase could be possible. In Europe GSM was the first digital cellular system and was launched in 1992. GSM has over 2 billion subscribers worldwide and over 600 million subscribers in Europe [1]. Originally, GSM operated only in the 900 MHz band (GSM-900), but was soon extended to operate in the 1800 MHz (GSM 1800) and 1900 MHz (GSM 1900) bands. GSM 1800/1900 is particularly used in suburban areas with a very dense traffic situation. The first phase of GSM only supported a subset of all the services in the standard. GSM phase 2 was an extension of phase 1 with supplementary services. The latest phase of the GSM system phase 2+ includes packet data services such as General Packet Radio Service (GPRS) and Enhanced Data Rates for GSM Evolution (EDGE). In the United States two different standards were introduced, D-AMPS and IS-95.

The demand for sending more information over the wireless communication systems requires more bandwidth. Voice communication was handled well by the second generation cellular systems, such as GSM. The third generation of mobile telecommunication systems will also handle data transmissions in a greater extent. Earlier generations have been circuit switched which means that a call occupies the channel during the duration of the call. The 3G systems have packet data switching instead, which means that user data are sent in small packets and this results in that the same channel can be used by others. The 3G systems are defined in the IMT-2000 specification (International Mobile Telecommunications-2000) that includes a global standard. Some of the systems includeed in the IMT-2000 specification are Universal Mobile Telecommunications Service (UMTS), in which WCDMA is included, CDMA-2000 and TD-SCDMA.

1.3 Overview of GSM

The GSM system uses a multiple TDMA and FDMA scheme, where every frequency is divided into eight time slots. For GSM-900 two spectra are used, 890-915 MHz for uplink connection and 935-960 MHz for downlink connection. These bands are shared between one or several telecom operators. The carrier spacing in GSM is 200 kHz, and for GSM-900 a total of 124 carriers are available. Many books and reports cover the topic of cellular systems, e.g. [10, 12, 42, 43, 52, 65, 131].

1.3.1 Network Elements in GSM

The structure of the GSM system can be divided into three main parts, the Mobile Station (MS), the Base Station Subsystem (BSS) and the Network and Switching Subsystem (NSS), see Figure 1.1. The BSS has the physical equipment that provides radio coverage for the cell, Base Transceiver Station (BTS) and Base Station Controller (BSC). The BTS controls the radio interface to the MS and includes transmitting/receiving devices and antennas that serve each cell bounded to the BTS. The BSC manages all the radio related functions of GSM such as MS handover, radio channel assignment and the collection of cell configuration data. A BSC can control several BTSs. The NSS deals with the connections between the BSC and internal and external networks. The NSS includes the main switching functions of GSM and databases for the subscribers. The main role of NSS is to manage communications between GSM and other networks. The NSS includes the Mobile Switching Center (MSC), Gateway Mobile Switching Center (GMSC), Home Location Register (HLR) and Visitor Location Register (VLR). The MSC performs the telephony switching functions for the GSM network. It controls calls from other telephony and data networks, such as the Public Switched Telephone Network (PSTN) and Public Land Mobile Networks (PLMN). The VLR database stores information about all the mobile subscribers who are visiting an MSC service area. There is one VLR for each MSC in the network. The HLR is a centralized network database that contains all mobile subscribers belonging to a specific telecom operator, dependending on the number of subscribers there can be more than one HLR in a network. The NSS contains also architecture for GPRS, i.e. Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN).

1.3.2 Frequency Planning in GSM

One way to use the available radio resources efficiently in GSM is to reuse the frequencies among the cells in a network. Reusing the frequencies in a network does introduce some issues that are not encountered otherwise. Two users who

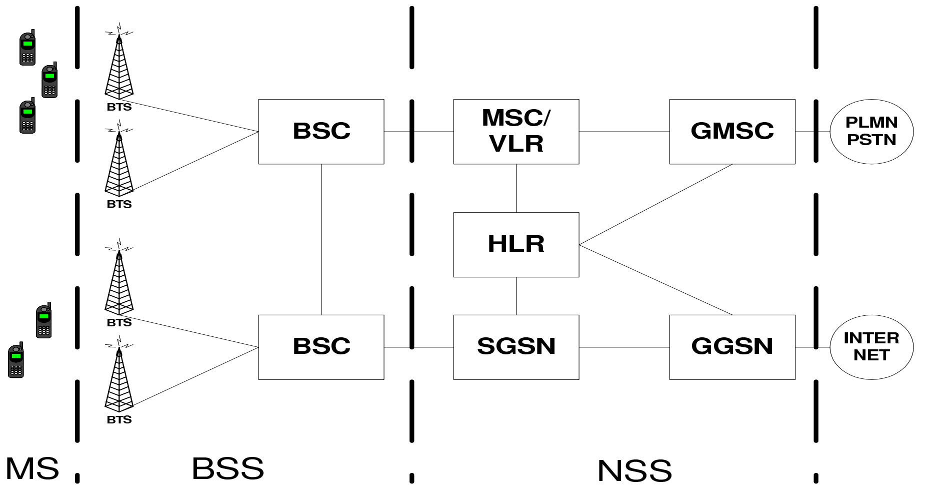

Figure 1.1: GSM system architecture.

are using the same frequency in two different locations in the network may interfere with each other. Therefore care must be taken when allocating the frequencies in a wireless network. If the same frequency is assigned to two users in cells that are close to each other, the interference may cause a dropped call or lost data packet. Since the number of mobile communication subscribers in the networks is growing rapidly, the telecom operators must replan their existing networks to increase capacity. The planning process is very complex and techniques from the discipline of *operations research* or *optimization* can be used. The telecom operators can make large cost savings if they take advantage of operations research. One example is that if the interference in a network can be reduced, more users can be connected with the same grade of service.

A cell is characterized by its area of coverage and the frequencies assigned to it. The area of coverage of the cell is highly dependent on the type of antenna, placement of the antenna and transmitted power [28, 43]. To cover a larger area, several cells are needed, and since an operator has a limited spectrum, the frequencies must be reused. The frequency reuse in the network is limited by co-channel interference. On the other hand, increasing the reuse distance gives lower capacity in the network. From a network planning perspective, it is an interesting trade off to analyze - capacity vs quality. The co-channel interference is the most limiting factor for frequency reuse. Assuming a hexagonal cell pattern, the possible reuse K are given by $K = a^2 + ab + b^2$ where a and b are natural numbers. The reuse factor can obtain the values $K = 1, 3, 4, 7, 9, 12, 13, ...$ and $K = 1$ means that all channels are used in all cells [12]. The reuse factor is often denoted x/K, x is the reuse factor for base stations and K is the reuse factor for the cells. A common reuse pattern for the TCH frequencies are 3/9, which means that a reuse factor $K = 9$ is used and every channel group is covered with 3 ($x = 3$) three-sector sites [42, 72, 122, 143, 150]. In Figure 1.2 a 1/3 reuse

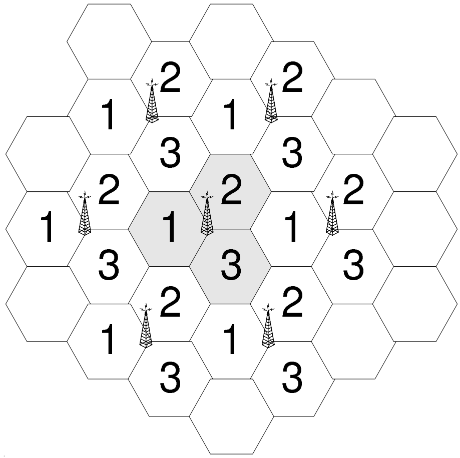

Figure 1.2: A cell structure with a 1/3 reuse.

pattern is shown.

One way to increase the capacity is by splitting the cells into smaller cells when the traffic grows. This method is very expensive since new base stations are required. Often a small part of the cell is a high density area (hotspot) with a high traffic load. A hierarchical cell structure with an overlaid macrocellular layer and an underlying microcellular layer is often used. The macrocells are usually described as a cell with a radius of a couple of hundred meters to several tens of kilometers. The micro cells are small and has a radius of less than 500 meters. The micro cells cover only the hotspots and since the average distance between these spots is large will the resulting interference between microcells be small. Methods for sharing the spectrum between micro and macro cells can be found in [139, 147]. Several capacity enhancement methods are presented in [40, 47, 89, 100, 121, 143].

One way to reduce the interference in a GSM network is to change the frequency after each TDMA frame. This feature is called frequency hopping. In the thesis the problem of assigning frequency lists to the cells in a frequency hopping GSM network with the objective to minimize the interference is investigated.

1.4 Overview of WCDMA

The UMTS standardization started as early as 1992. The first release of standards was announced 1999. The standardization of UMTS is handled by the 3rd Generation Partnership Project (3GPP) which is a forum with members such as manufactures, telecom operators and standardization institutes. The initial release, Release '99 described the radio access technologies UTRA-TDD and UTRA-FDD and standardized the GSM/GPRS network as a core. UTRA-TDD uses different time slots to separate uplink and downlink communication and UTRA-FDD uses different frequencies. WCDMA supports frequency division duplex and a telecom operator is assigned several pairs of carriers of 2x5 MHz. In WCDMA the spectrum for the uplink is assigned to 1710-1785 MHz and for the downlink to 2110-2170 MHz. The mobile telecom operators in a country need to share the available spectrum. Next release was planned for 2000 but 3GPP decided to split the release into two parts, Release 4 and Release 5. Release 4 introduced quality of service in the fixed network. Release 5 specifies a different core network and a high speed downlink packet access channel with significant higher data rate. Release 6 comprises the use of multiple input multiple output antennas, enhanced multi media services, security enhancements and many more management features. Several books which treat WCDMA have been written, e.g. [67, 79, 130].

1.4.1 WCDMA Architecture

WCDMA systems use the same network architecture used by the second generation mobile telecommunications systems such as GSM. The elements in the network that handles radio related functionality are grouped into a radio access network UMTS Terrestrial Radio Access Network (UTRAN). The network elements that are responsible for switching and routing calls and data to external networks are called the Core Network (CN). The last part of the architecture is the User Equipment (UE) that contains mobile terminals. The architecture of a WCDMA network is shown in Figure 1.3. UTRAN contains two elements, Node B and a Radio Network Controller (RNC). Node B is exactly the same as a base station. Several Node Bs are grouped and connected to a RNC. The RNC has a similar functionality as BSC in GSM, i.e control the radio resources. The RNC also handles power and handover control. UTRAN supports soft handover, which occurs between Node B's controlled by different RNCs. The CN has similar architecture as a GSM network, see Section 1.3.1. The elements in the CN are VLR, HLR, MSC, GMSC and SGSN/GGSN.

1.4.2 Code Division Multiple Access

The access method used in WCDMA is CDMA. In CDMA user data is assigned a unique code sequence, i.e. channelization code. The channelization code is used for encoding the transmitted signal. The receiver knows the code sequences for the user and can decode and recover the original user data. This can be done since cross correlations between the channelization code of a desired user data and the codes of other data streams are small. The encoding process spreads

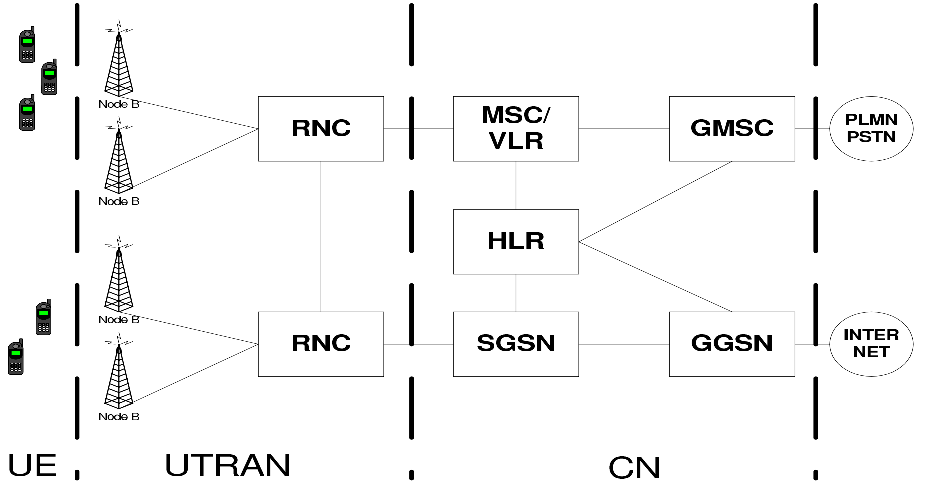

Figure 1.3: WCDMA network architecture.

the spectrum of the information signal over a larger bandwidth. The spectral spreading of the transmitted signal gives CDMA its multiple access capability. The information bits are spread over the larger bandwidth by multiplying the user data with quasi-random bits (chips) derived from the spreading codes. In WCDMA a constant chip rate of 3.84 Mcps is used, which gives a bandwidth of 5 MHz.

The channelization codes used are Walsh-Hadamard codes [123] and also referred to Orthogonal Variable Spreading Factor (OVSF) codes and can be represented by a coding tree [67]. Orthogonality between channelization codes in the tree is guaranteed if one code has not been generated based on another. The different levels of the code tree are assigned a Spreading Factor (SF). The SF determines the number of chips per bit or symbol for a channelization code. Since the chip rate is constant, a constant SF generates a constant number of chips per bit or symbol. The purpose of the OVSF is to change the bit rate by changing the number of chips per bit or symbol. Increasing the OVSF generates a lower bit rate since each bit contains more chips.

In WCDMA the limited resources in a cell are transmission power and number of channelization codes. The problem encountered in the thesis is how to allocate the channelization codes without exceeding the power limitations to the users in the downlink of a cell.

1.5 Contributions

In the context of resource optimization in wireless networks we have developed several models and algorithms to improve resource utilization.

Part I

- Models for assigning transmission rights to nodes or links in a STDMA ad hoc network are presented. There is also a presentation of a column generation solution technique together with primal heuristics.

The content of Part I has been published in [19, 20, 22].

Part II

- A frequency assignment model for a frequency hopping GSM network is presented. The model is a generalization of the classical frequency assignment problem. Instead of assigning frequencies we assign lists of frequencies. Heuristic solution strategies for frequency hopping GSM networks are also presented.

The content of Part II has been published in [19, 23].

Part III

- A linear integer optimization model is presented for the Downlink Shared Channel in WCDMA. The optimization model is solved by dynamic programming strategy suited for online implementation.
- A scheduling model for the High Speed Downlink Packet Access and two greedy algorithms are presented.

The scheduling of Downlink Shared Channel with dynamic programming is published in [21]. The scheduling of the High Speed Downlink Packet Access is not yet published.

1.6 Outline of the Thesis

This thesis consists of three parts, all related to resource allocation in wireless networks. The first part handles ad hoc networks. Part II handles second generation cellular system, GSM. The last part handles a third generation cellular system, WCDMA, which can be seen as an extention of GSM.

Part I: Resource Optimization in Ad Hoc Networks

The first part of the thesis deals with ad hoc networks. The optimization approach is to minimize the length of the transmission frame. Chapter 2 contains a general description of ad hoc networks and the most important parameters that affect the communication quality. Two different assignment strategies are discussed, assigning nodes and assigning communication links. In Chapter 3, two different optimization models and solution strategies are discussed. The solution strategies used are a direct solution approach with commercial mathematical programming software and a column generation approach. Chapter 4 contains computational results for both optimization models.

Part II: Frequency Assignment in Frequency Hopping GSM Networks

The second part of the thesis deals with the frequency assignment problem in a frequency hopping GSM network. Chapter 5 explains some basic properties of frequency hopping and frequency assignment models for frequency hopping and non hopping GSM networks. Different solution strategies are explained in Chapter 6. The techniques are meta heuristics, such as tabu search and simulated annealing. Chapter 7 contains the computational results for the heuristics.

Part III: Resource Optimization in WCDMA Networks

The third part of the thesis deals with the problem of assigning channelization codes for the downlink shared packet data channels in WCDMA. In Chapter 8 the Downlink Shared Channel is investigated. A linear integer optimization model is developed for assigning channelization codes to users in a cell and a dynamic programming solution strategy is also presented. The last chapter in Part III, Chapter 9, deals with the new data packet access concept High Speed Data Packet Access. For assigning channelization codes and bit rates to the users a linear integer optimization model is developed. Two greedy algorithms are used to find feasible solutions to the model.

The contents of the three parts in the thesis are based on some published articles. The last chapter of the thesis, Chapter 10, presents some research that have been done by different authors since these articles. There is also a discussion of future reseach.

Part I

Resource Optimization in Ad Hoc Networks

Chapter 2

Spatial Time Division Multiple Access

2.1 Network Model

An ad hoc network can be characterized by a directed graph $G = (N, A)$, where the node set N represents the radio units, and the arc set A represents the communication links. Typically, the cardinality of A is much less than $|N|(|N|-1)$, meaning that the network is sparsely connected. A sample network of 20 nodes is shown in Figure 2.1. A directed link (i, j) belongs to A if its signal-to-noise ratio (SNR) is greater or equal to a given threshold, that is, if

$$SNR(i,j) = \frac{P_i}{L_b(i,j)N_r} \geq \gamma_0, \tag{2.1}$$

where P_i is the transmitting power of i, $L_b(i, j)$ is the path-loss between i and j, N_r is the effect of the thermal noise, and γ_0 is the threshold.

Several assumptions are commonly made in STDMA scheduling. First, a node cannot transmit and receive simultaneously. Secondly, a node can receive data from at most one other node at any time. Finally, we assume that a link is error-free if and only if the signal-to-interference ratio (SIR) is above a threshold γ_1 (possibly equals γ_0). For link (i, j), the SIR-criterion is formulated as

$$SIR(i,j) = \frac{P_i}{L_b(i,j)(N_r + \sum_{k\in K, k\neq i} \frac{P_k}{L_b(k,j)})} \geq \gamma_1. \tag{2.2}$$

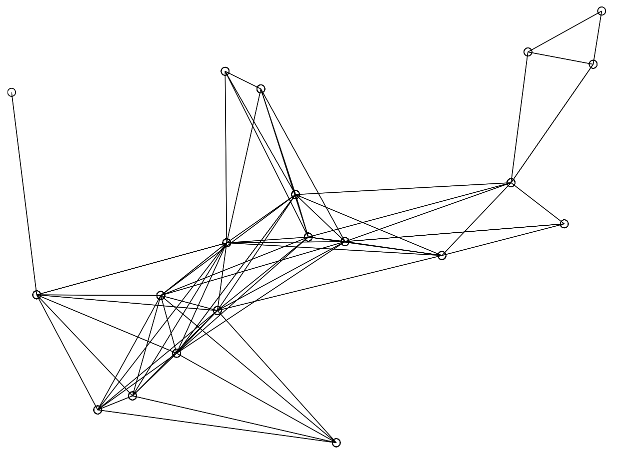

Figure 2.1: An ad hoc network of 20 nodes.

In (2.2), K is the set of nodes that are in simultaneous transmission. The term $\sum_{k \in K, k \neq i} \frac{P_k}{L_b(k,j)}$ is thus the accumulated interference with respect to link (i, j). The effect of Rayleigh fading is not taken into consideration which results in the assumption that no part of the transmission power of node i is treated as interference. We also assume that the transmitting power for the nodes are set to a constant values, e.g. every node transmits at its maximum power when sending data to another node. It is often assumed, e.g. [57] that $P_i = P, \forall i \in N$, and $L_b(i,j) = L_b(j,i), \forall (i,j) \in A$. These assumptions are, however, not necessary for our mathematical models or solution methods.

2.2 Assignment Strategies

There are two possibilities to assign the time slots: node-oriented assignment and link-oriented assignment. In the former strategy, a node is assigned one or several time slots in the schedule. In each of these slots, the node may use any of its (outgoing) links for transmitting data to another node. This assignment strategy is well-suited for broadcast traffic. In link-oriented assignment, a link is assigned one or several time slots for point-to-point communication between a specific pair of nodes. Empirically, link-oriented assignment achieves a higher spatial reuse than that of node-oriented assignment [58].

Note that the SIR-criterion (2.2) leads to different constraints for the two assignment strategies. For node-oriented assignment, a time slot can be assigned to node i only if all the outgoing links of i satisfy (2.2). If a time slot is assigned to link (i, j) using link-oriented assignment, then it is required that (2.2) is satisfied for this particular link.

A variety of heuristic algorithms for STDMA scheduling can be found in the literature. Example of algorithms that minimize the frame length for node oriented scheduling can be found in [36, 37, 41, 49, 95] and for link oriented scheduling in [35, 64, 118, 119]. In [136] both link and node assignment strategies are evaluated. If the SIR requirement is relaxed the problem can be seen as a variation of the graph coloring problem, see [96]. In the classical graph coloring problem, see e.g. [24, 29], a graph is given with undirected arcs between the nodes. The problem is to assign one color to every node in the graph, with requirements that two adjacent nodes can not be assigned the same color. The objective is to minimize the number of colors used in the graph. An extension to the problem of minimizing the frame length is to take traffic variations into account, traffic senstive algorithms are presented in [56, 57, 58, 132].

2.3 Problem Definition

If traffic distribution is not taken into consideration, then the length of the STDMA schedule determines the efficiency of the spatial reuse of the time slots. We define two optimization problems, denoted by MNP and MLP, for minimum-length scheduling for node-oriented and link-oriented assignments, respectively.

Given the set of nodes N, the path-loss between every pair of nodes (i.e. $L_b(i,j)$, $\forall i,j \in N : i \neq j$), the transmitting power of each node (i.e. $P_i, \forall i \in N$), the noise effect N_r, and the two thresholds γ_0 and γ_1, the objective of MNP is to find a minimum-length schedule, such that every node receives at least one time slot, and such that the following constraints are satisfied.

- Two end nodes of a link must be assigned different time slots. (This is because a node cannot both transmit and receive in a time slot.)
- Two nodes, both having directed links to a third node, must be assigned different time slots. (This is because a node cannot receive from more than one node in a time slot.)
- A time slot can be assigned to a node only if all the outgoing links of the node satisfy the SIR-constraint (2.2).

For link-oriented assignment, the corresponding problem MLP amounts to finding a minimum-length schedule, such that every link receives at least one time slot, and such that the following constraints are satisfied.

- Two links that share a common node, irrespective of the link directions, must be assigned different time slots. (Note that this constraint comprises the first two constraints of MNP.)

- A time slot can be assigned to a link only if the SIR-constraint (2.2) for the link is satisfied.

Chapter 3

Scheduling Models for STDMA

We study MNP and MLP using mathematical programming formulations. We first present two linear integer formulations: a node-slot formulation for MNP, and a link-slot formulation for MLP. We then formulate the two problems using set covering formulations, for which we will derive the column generation method.

3.1 A Node-Slot Formulation

Let $T = \{1, ..., |T|\}$ be a set of time slots. To ensure feasibility of MNP, it is sufficient to have $|T| = |N|$. We introduce the following binary variables.

$$x_{it} = \begin{cases} 1 & \text{if time slot } t \text{ is assigned to node } i \\ 0 & \text{otherwise.} \end{cases}$$

$$y_t = \begin{cases} 1 & \text{if time slot } t \text{ is used} \\ 0 & \text{otherwise.} \end{cases}$$

MNP can be formulated using the following node-slot formulation (NSF).

$$\text{[NSF]} \quad z_N^* = \quad \min \sum_{t \in T} y_t \tag{3.1}$$

$$\text{s.t.} \quad \sum_{t \in T} x_{it} \geq 1, \forall i \in N \tag{3.2}$$

$$x_{it} \leq y_t, \forall i \in N, \forall t \in T \tag{3.3}$$

$$x_{it} + \sum_{j:(j,i) \in A} x_{jt} \leq 1, \ \forall i \in N, \forall t \in T \tag{3.4}$$

$$\frac{P_i/N_r}{L_b(i,j)} x_{it} + \gamma_1(1 + M_{ij})(1 - x_{it}) \geq \gamma_1(1 + \sum_{k \in N: k \neq i,j} \frac{P_k/N_r}{L_b(k,j)} x_{kt}), \forall (i,j) \in A, \forall t \in T \tag{3.5}$$

$$x_{it} = 0/1, \ \forall i \in N, \forall t \in T \tag{3.6}$$

$$y_t = 0/1, \ \forall t \in T. \tag{3.7}$$

The objective function (3.1) minimizes the total number of time slots. Constraint (3.2) ensure that every node is assigned at least one slot. Constraint (3.3) state that a slot is used (i.e. $y_t = 1$) if it is assigned to any node. Constraint (3.4) ensure that different time slots are assigned to two nodes if they are the two end nodes of a link, or if both have links to a third node. The SIR-criterion is defined in (3.5). If slot t is not assigned to node i (i.e. $x_{it} = 0$) and M_{ij} is sufficiently large, then (3.5) is redundant. If $x_{it} = 1$, the constraint reads $\frac{P_i/N_r}{L_b(i,j)} \geq \gamma_1(1 + \sum_{k \in N: k \neq i,j} \frac{P_k/N_r}{L_b(k,j)} x_{kt})$, which corresponds to the SIR-criterion (2.2).

To ensure that (3.5) is redundant when $x_{it} = 0$, M_{ij} can be set to $M_{ij} = \sum_{k \in N: k \neq i,j} (P_k/N_r)/L_b(k,j)$, i.e. the sum of the potential interference from all nodes other than i and j. However, not all nodes in the set $\{k \in N : k \neq i,j\}$ will transmit simultaneously, because these nodes must also satisfy constraints (3.4) and (3.5). It is therefore possible to compute a smaller value of M_{ij}, which improves the linear programming relaxation (LP-relaxation) of NSF, see Appendix B.

The formulation NSF contains two types of symmetry. First, there are many solutions that correspond to the same assignment, but with different time slots allocated. To break this type of symmetry, we can enforce that slot t can be used only if slot $t-1$ is used, by adding the constraints $y_t \leq y_{t-1}, t = 2, ..., |T|$. The second type of symmetry is related to the fact that swapping the nodes of any two slots does not affect the objective function value. Such symmetry can be partially eliminated by requiring that node i must be assigned a slot with an index less or equal to i, i.e. $x_{it} = 0, \forall i, t : i < t$.

3.2 A Link-Slot Formulation

Problem MLP can be formulated using a link-slot formulation (LSF), which uses the following variables.

$$x_{ijt} = \begin{cases} 1 & \text{if time slot } t \text{ is assigned to link } (i,j) \\ 0 & \text{otherwise.} \end{cases}$$

$$y_t = \begin{cases} 1 & \text{if time slot } t \text{ is used} \\ 0 & \text{otherwise.} \end{cases}$$

$$v_{it} = \begin{cases} 1 & \text{if node } i \text{ is transmitting in time slot } t \\ 0 & \text{otherwise.} \end{cases}$$

Formulation LSF is stated below.

$$
\begin{aligned}
\text{[LSF]} \quad z_L^* = \quad & \min \sum_{t \in T} y_t \\
\text{s.t.} \quad & \sum_{t \in T} x_{ijt} \geq 1, \ \forall (i,j) \in A && (3.8) \\
& x_{ijt} \leq y_t, \ \forall (i,j) \in A, \forall t \in T && (3.9) \\
& \sum_{j:(i,j) \in A} x_{ijt} + \sum_{j:(j,i) \in A} x_{jit} \leq 1, \forall i \in N, \forall t \in T && (3.10) \\
& x_{ijt} \leq v_{it}, \forall (i,j) \in A, \forall t \in T && (3.11) \\
& \frac{P_i/N_r}{L_b(i,j)} x_{ijt} + \gamma_1 (1 + M_{ij})(1 - x_{ijt}) \geq \gamma_1 (1 + \\
& \sum_{k \in N : k \neq i,j} \frac{P_k/N_r}{L_b(k,j)} v_{kt}), \forall (i,j) \in A, \forall t \in T && (3.12) \\
& x_{ijt} = 0/1, \ \forall (i,j) \in A, \forall t \in T && (3.13) \\
& v_{it} = 0/1, \ \forall i \in N, \forall t \in T && (3.14) \\
& y_t = 0/1, \ \forall t \in T. && (3.15)
\end{aligned}
$$

In LSF, the cardinality of T can be set to $|A|$ in order to guarantee feasibility. Constraint (3.8) and (3.9) correspond to (3.2) and (3.3), respectively. Constraint (3.10) state that two adjacent links must be assigned different time slots. The two sets of variables, x and v, are linked to each other in constraint (3.11) (note that v_{it} has the same meaning as x_{it} in NSF), and are used to derive the SIR-constraint (3.12). Similar to the case of node-oriented assignment, M_{ij} in the SIR-constraint (3.12) can be set to $M_{ij} = \sum_{k \in N : k \neq i,j} (P_k/N_r)/L_b(k,j)$, although it is possible to compute a smaller value for this coefficient. To break the symmetry in LSF, we can add the constraints $y_t \leq y_{t-1}, t = 2, ..., |T|$, and $x_{ijt} = 0, \forall (i,j), t : B_{ij} < t$, where B_{ij} denotes the index of link (i,j).

Formulations NSF and LSF are straightforward linear integer models. From a computational point of view, the two formulations are not suitable to use. In particular, the numbers of variables and constraints grow rapidly with respect to the network size.

3.3 Set Covering Formulations

In this section, we reformulate MNP and MLP using set covering formulations. The LP-relaxations of the set covering formulations can be efficiently solved using a column generation method. The optimal solutions of the LP-relaxations also enable optimal or near-optimal integer solutions.

An instance of a set covering problem is characterized by a finite set S, and a collection of sets C. Each member in C comprises a subset of the elements in S. A feasible solution is a set cover of S, that is, a subset $C' \subseteq C$ such that every element in S belongs to at least one member of C'.

The set covering formulations of MNP and MLP are based on the concept of transmission groups. A transmission group is a group of nodes or a group of links that can be in simultaneous transmission, and can therefore share the same time slot. Let L_N and L_A be the sets of transmission groups of nodes and links, respectively. We define one binary variable for each transmission group.

$$x_l = \begin{cases} 1 & \text{if transmission group } l \text{ is assigned a time slot} \\ 0 & \text{otherwise.} \end{cases}$$

The set covering formulation of MNP is stated below.

$$\begin{aligned} \text{[NSCF]} \quad z_N^* = \quad & \min \sum_{l \in L_N} x_l && (3.16) \\ \text{s.t.} \quad & \sum_{l \in L_N} s_{il} x_l \geq 1, \forall i \in N && (3.17) \\ & x_l = 0/1, \forall l \in L_N. && (3.18) \end{aligned}$$

In NSCF, s_{il} is an indication parameter that is one if group l contains node i, and zero otherwise. The objective function (3.16) minimizes the total number of assigned time slots. Constraint (3.17) ensure that every node belongs to at least one group that is assigned a time slot. It is apparent that MNP is a set covering problem in which $S = N$ and $C = L_N$.

Problem MLP can be formulated using the following set covering formulation.

$$\begin{aligned} \text{[LSCF]} \quad z_L^* = \quad & \min \sum_{l \in L_A} x_l && (3.19) \\ \text{s.t.} \quad & \sum_{l \in L_A} s_{ijl} x_l \geq 1, \forall (i,j) \in A && (3.20) \\ & x_l = 0/1, \forall l \in L_A. && (3.21) \end{aligned}$$

In LSCF, parameter s_{ijl} indicates whether link (i, j) belongs to group l (i.e. $s_{ijl} = 1$ if group l contains (i, j), and zero otherwise). MLP is a set covering problem in which $S = A$ and $C = L_A$.

The two set covering formulations have a very simple constraint structure. The complexity lies mainly in the cardinality of the two sets L_N and L_A. For networks of realistic size, there are huge numbers of transmission groups. However, this difficulty can be overcome using a column generation approach that effectively exploits the structure of the two formulations.

3.4 A Column Generation Solution Method

Originally presented in [53, 54], column generation is a decomposition technique for solving a structured linear program (LP) with few rows but many columns (variables). Column generation decomposes the LP into a master problem and a subproblem. The master problem contains a subset of the columns. The subproblem, which is a separation problem for the dual LP, is solved to identify whether the master problem should be enlarged with additional columns or not. Column generation alternates between the master problem and the subproblem, until the former contains all the columns that are necessary for finding an optimal solution of the original LP.

Column generation is especially attractive for problems that can be formulated using set covering formulations, which typically contain a huge number of columns, although very few of them are used in the optimal solution. We note that this method has been proposed in [103] for solving the graph coloring problem, which has a similar structure to our STDMA scheduling problems.

3.4.1 Node-oriented Assignment

To apply column generation to MNP, we consider the LP-relaxation of NSCF.

$$z_N^{LP} = \min \sum_{l \in L_N} x_l \tag{3.22}$$

$$\text{s.t.} \quad \sum_{l \in L_N} s_{il} x_l \geq 1, \ \forall i \in N \tag{3.23}$$

$$0 \leq x_l \leq 1, \ \forall l \in L_N. \tag{3.24}$$

The column generation master problem is the same as the above LP-relaxation, except that the set of transmission groups, L_N, is replaced by a subset $L_N^0 \subseteq L_N$.

To ensure feasibility of the master problem, the set L_N^0 must satisfy $\sum_{l \in L_N^0} s_{il} \geq 1, \forall i \in N$. One particular choice of L_N^0 is the node set N (i.e. the set of transmission groups derived by TDMA).

When the master problem is solved, we need to identify whether it can be improved by adding new columns (transmission groups) to L_N^0. In LP terms, this amounts to examining whether there exists any transmission group $l \in L_N$, for which the corresponding variable x_l has a strictly negative reduced cost. Using LP-duality, the reduced cost $\bar{c}_l$ of variable x_l is

$$\bar{c}_l = 1 - \sum_{i \in N} \bar{\beta}_i s_{il}, \tag{3.25}$$

where $\bar{\beta}_i, \forall i \in N$, are the (optimal) dual variables to (3.23). Clearly, there exists at least one variable with negative reduced cost if and only if the minimum of (3.25) is negative. We are thus interested in the following optimization problem.

$$\min_{l \in L_N} \bar{c}_l = 1 - \max_{l \in L_N} \sum_{i \in N} \bar{\beta}_i s_{il}. \tag{3.26}$$

The column generation subproblem is equivalent to (3.26), but formulated differently. We use the following variables in the subproblem.

$$s_i = \begin{cases} 1 & \text{if node } i \text{ is included in the transmission group} \\ 0 & \text{otherwise.} \end{cases}$$

The subproblem can be formulated as follows.

$$\begin{aligned}
\max \quad & \sum_{i \in N} \bar{\beta}_i s_i && (3.27) \\
\text{s.t.} \quad & s_i + \sum_{j:(j,i) \in A} s_j \leq 1, \ \forall i \in N && (3.28) \\
& \frac{P_i/N_r}{L_b(i,j)} s_i + \gamma_1 (1 + M_{ij})(1 - s_i) \geq \gamma_1 (1 + \\
& \sum_{k \in N: k \neq i,j} \frac{P_k/N_r}{L_b(k,j)} s_k), \forall (i,j) \in A && (3.29) \\
& s_i = 0/1, \ \forall i \in N. && (3.30)
\end{aligned}$$

Note the similarity between the constraints in the subproblem and those in the node-slot formulation NSF. The main difference is that the subproblem only considers one transmission group for one time slot, but in NSF transmission groups for all time slots are to be determined simultaneously.

If the optimal solution to the subproblem results in a strictly negative reduced cost, the corresponding new transmission group is added to the master problem, which is then reoptimized, and the column generation method proceeds to the next iteration. Otherwise, the LP-relaxation of NSCF has been solved to optimality, and the optimum of the master problem equals z_N^{LP}.

3.4.2 Link-oriented Assignment

The column generation method for link-oriented assignment is very similar to that for node-oriented assignment. The LP-relaxation of LSCF reads

$$
\begin{aligned}
z_L^{LP} = \quad & \min \sum_{l \in L_A} x_l && (3.31)\\
\text{s.t.} \quad & \sum_{l \in L_A} s_{ijl} x_l \geq 1, \ \forall (i,j) \in A && (3.32)\\
& 0 \leq x_l \leq 1, \ \forall l \in L_A. && (3.33)
\end{aligned}
$$

The corresponding master problem is the above LP-relaxation defined for a subset of transmission groups $L_A^0 \subseteq L_A$. Given the optimal solution of the master problem, finding the transmission group with minimum reduced cost is equivalent to the following optimization problem, where $\bar{\beta}_{ij}, \forall (i,j) \in A$, are the (optimal) dual variables to (3.32).

$$\min_{l \in L_A} \bar{c}_l = 1 - \max_{l \in L_A} \sum_{(i,j) \in A} \bar{\beta}_{ij} s_{ijl}. \tag{3.34}$$

We use two sets of variables in the formulation of the subproblem.

$$s_{ij} = \begin{cases} 1 & \text{if link } (i,j) \text{ is included in the transmission group,} \\ 0 & \text{otherwise,} \end{cases}$$

$$v_i = \begin{cases} 1 & \text{if node } i \text{ is transmitting,} \\ 0 & \text{otherwise.} \end{cases}$$

Using these variables, the subproblem, which is an alternative way of formulating (3.34), can be stated as follows.

$$\max \sum_{(i,j)\in A} \bar{\beta}_{ij} s_{ij} \tag{3.35}$$

$$\sum_{j:(i,j)\in A} s_{ij} + \sum_{j:(j,i)\in A} s_{ji} \leq 1, \forall i \in N \tag{3.36}$$

$$s_{ij} \leq v_i, \ \forall (i,j) \in A \tag{3.37}$$

$$\frac{P_i/N_r}{L_b(i,j)} s_{ij} + \gamma_1(1+M_{ij})(1-s_{ij}) \geq \gamma_1(1+$$

$$\sum_{k\in N:k\neq i,j} \frac{P_k/N_r}{L_b(k,j)} v_k), \forall (i,j) \in A \tag{3.38}$$

$$s_{ij} = 0/1, \ \forall (i,j) \in A \tag{3.39}$$

$$v_i = 0/1, \ \forall i \in N. \tag{3.40}$$

As for the case of node-oriented assignment, the column generation method alternates between the master problem and the subproblem, until the value of (3.34) is non-negative.

3.4.3 Enhancements

The performance of the column generation method depends on the computing effort of one iteration (in particular for solving the subproblem), as well as the total number of iterations before reaching optimality. Both factors become crucial if we wish to solve large-scale network instances within reasonable computing time.

We propose two enhancements for accelerating the convergence of the method. Solving the subproblems, which are integer programs, may require excessive computing time, making large-scale networks out of reach of the column generation method. Indeed, a straightforward implementation of the method failed to solve the LP-relaxation of LSCF for network instances with more than 40 nodes. To overcome this difficulty, we propose the following modification to the method. Instead of solving the subproblems to optimality, a threshold (less than zero) is used for termination control. In particular, we halt the solution process of the subproblem after a time limit. If the best solution found so far yields a reduced cost that is less or equal to the threshold, we terminate the solution process, and add the corresponding column (which is the transmission group with the best known reduced cost so far) to the master problem. Otherwise, the threshold is divided by a factor of two, and the solution process is resumed for another limited amount of time, after which the (new) threshold is used for termination control. In addition, we impose an upper bound of the threshold (i.e. the threshold is not increased if it becomes greater than or equal to this bound). In our implementation, the time limit is set to 10 seconds, the initial value of the threshold is -4, and the upper bound is set to -0.1.

Note that the above enhancement does not compromise the solution optimality. In particular, the upper bound of the threshold ensures that optimality is reached within a finite number of iterations.

The second enhancement concerns the generation of maximum feasible groups, We call a transmission group maximum feasible, if the addition of any new node (or link) will make the group infeasible. Note that, for both NSCF and LSCF, there exists at least one optimal solution in which all the transmission groups are maximum feasible. By adding maximum feasible groups to the master problem, we attempt to minimize the number of transmission groups that the method needs to generate before reaching optimality.

To find a maximal feasible group, we incorporate an additional step after solving the subproblem. Let $\bar{s}_i, \forall i \in N$ be a solution (not necessarily optimal) to the subproblem for node-oriented assignment. The solution can be made maximum feasible by considering the following problem, obtained by doing two modifications to the subproblem. First, we replace the objective function (3.27) by $\max \sum_{i \in N} s_i$, which maximizes the total number of nodes. In addition, we add the constraints $s_i = 1, \forall i \in N : \bar{s}_i = 1$ to the subproblem. It can be easily realized that solving this modified problem yields a maximum feasible group, for which the reduced cost is less than or equal to that of the original subproblem solution. Similarly, given a subproblem solution $\bar{s}_{ij}, \forall (i,j) \in A$ for link-oriented assignment, the corresponding modified subproblem amounts to maximizing $\sum_{(i,j) \in A} s_{ij}$, with the additional constraints $s_{ij} = 1, \forall (i,j) \in A : \bar{s}_{ij} = 1$.

The above step for finding a maximum feasible group may take long computing time (in particular for link-oriented assignment). We have therefore chosen to set a time limit (10 seconds in our implementation) in this step. The best solution found within the time limit is the transmission group added to the master problem.

3.5 Two Heuristic Procedures

The column generation method solves the LP-relaxations of NSCF and LSCF. If some variables are fractional-valued in the LP-optimum, the solution does not represent a feasible schedule. To obtain integer solutions, enumeration schemes (such as the branch-and-price technique in [103]) or heuristics are necessary. We consider two heuristic procedures for generating integer solutions.

The first procedure is a straightforward post-processing step of the column generation method. Specifically, we consider the optimal integer solution for the transmission groups that have been generated in the column generation method. We do this by imposing the integrality constraints to all the variables in the master problems. The resulting integer problems (which, in fact, are restricted versions of NSCF and LSCF) are then solved to optimality using a linear integer solver. We use z_N^{IP} and z_L^{IP} to denote, respectively, the numbers of time slots found for node-oriented and link-oriented assignments using this solution procedure.

The second procedure for generating feasible schedules is an iterative greedy algorithm. In one iteration, the algorithm constructs a feasible transmission group, which is assigned a time slot. The algorithm for node-oriented assignment is as follows. Initially, all the nodes are stored in a list. In our implementation, the nodes are stored following the order of their indices. The first node in the list is added to the transmission group, and removed from the list. Next, the algorithms considers the second node in the list. The node is added to the group and removed from the list if all the constraints of MNP are satisfied. Continuing in this fashion, the algorithm scans through the list and adds as many nodes as possible to the group. A time slot is then assigned to the group, and the algorithm proceeds to the next iteration. We note that the time for checking the feasibility of a transmission group is polynomial in $|N|$. Consequently, the algorithm has a polynomial time complexity. Below we provide a formal description of the algorithm, where S_t is the group of nodes that are assigned time slot t, and Q is the list.

1. Initialization.
   (a) Set $S_t = \emptyset, t = 1, \dots, |N|$.
   (b) Set $t = 1$.
   (c) Store the nodes in list Q.
2. Repeat until Q is empty:
   (a) For $i = 1, \dots, |Q|$:
      i. Let n_i be the node at position i of Q.
      ii. If $S_t \cup \{n_i\}$ is a feasible transmission group, set $S_t = S_t \cup \{n_i\}$.
   (b) Set $Q = Q \setminus S_t$.
   (c) Set $t = t + 1$.

The greedy algorithm for link-oriented assignment is identical to the above algorithm, except that all the entities are defined for links instead of nodes. We therefore leave out the explicit description of the algorithm for the case of link-oriented assignment. We use z_N^G and z_L^G to denote the numbers of time slots generated by the greedy algorithm for node-oriented and link-oriented assignments, respectively.

By applying the column generation method and the two heuristic procedures to the same network instance, we are able to obtain lower and upper bounds to the optimal schedule lengths. Specifically, the inequalities $z_N^{LP} \leq z_N^* \leq \min\{z_N^{IP}, z_N^G\}$ and $z_L^{LP} \leq z_L^* \leq \min\{z_L^{IP}, z_L^G\}$ hold.

Chapter 4

Numerical Results

We have used six different test networks of various sizes in our numerical experiments. These networks are provided by the Swedish Defence Research Agency (FOI). The numbers of the nodes and links range from 10 to 60, and from 26 to 396, respectively. For each of the test networks, the following computations have been carried out. First, we use the general linear integer solver CPLEX (version 7.0) [71] to solve the NSF and the LSF. Solving these two formulations requires excessive computing time for large networks, we have therefore set a time limit of 10 hours. We then apply our column generation method to solve the LP-relaxations of the set covering formulations, NSCF and LSCF. The column generation method is implemented using AMPL [46] and CPLEX. The latter is used to solve both the master problem and the subproblem. Finally, we apply the heuristics described in the previous section to compute feasible schedules. We have conducted our experiments on a Sun UltraSparc station with a 400 MHz CPU and 1 GB RAM.

4.1 Node-oriented Assignment

Computational results for node-oriented assignment are summarized in Table 4.1. The second column in the table shows the number of network nodes, which equals the number of time slots in the TDMA schedule. For the formulation NSF, the table displays the number of time slots of the best integer solution found within the time limit, the lower bound provided by the LP-relaxation, and the computing time. Note that, if the computing time is less than the limit (10 hours), then the best integer solution has been proven to be optimal; otherwise CPLEX has either not found the optimal solution, or did not verify optimality. For the formulation NSCF, which is solved using the column generation method,

	$\|N\|$	NSF (CPLEX)			NSCF (Column Generation)				Heuristic
		Slots	LP	Time	Slots (z_N^{IP})	LP (z_N^{LP})	Iter.	Time	Slots (z_N^G)
N10	10	10	10	0.1s	10	10	1	0.1s	10
N20	20	16	16	1s	16	16	10	3s	16
N30	30	21	16	10h*	21	21	12	7s	21
N40	40	15	10	10h*	15	14	43	32s	16
N50	50	28	16	10h*	23	23	46	1m19s	26
N60	60	31	17	10h*	26	26	60	4m31s	30

Table 4.1: Numerical results for node-oriented assignment. *) The maximum computational time for CPLEX is set to 10h.

we show the number of slots of the integer solution (i.e. z_N^{IP}), the LP-bound (i.e. z_N^{LP}), the number of column generation iterations, and the computing time. The last column in the table shows the number of time slots of the feasible schedule found by the greedy algorithm (i.e. z_N^G). The results of the greedy algorithm are obtained with very little computing effort (less than a couple of seconds) for all the test networks, we have therefore not included the solution time of this algorithm in the table.

Based on the results in Table 4.1, we make the following observations. Formulation NSF can be used to solve small network instances to optimality. However, for large networks, this formulation is clearly not computationally efficient. In particular, CPLEX did not manage to solve the problem to optimality for any of the networks with more than 20 nodes. Moreover, the LP-relaxation of NSF is very weak, when compared to the solutions found by the other methods.

The set covering formulation NSCF is more computationally efficient than NSF. The LP-relaxation of this formulation can be solved efficiently using the column generation method. We observe that the LP-relaxation provides very tight lower bounds. In addition, the transmission groups generated in the column generation procedure lead to a feasible schedule that is optimal or near-optimal. For five of the six networks, the number of time slots of the feasible schedule, z_N^{IP}, is equal to the lower bound z_N^{LP}, and is therefore optimal. For network N40, the two values differ by one slot.

Our results show that the maximum possible spatial reuse in STDMA, which can be measured as the ratio between $|N|$ and z_N^{LP}, increases by network size. For the networks used in our experiments, this ratio ranges from 1.0 to 2.86.

We note that the greedy algorithm performs well for small networks. In particular, for networks with 30 nodes or less, the schedules found by the greedy algorithm are optimal. For the other networks the relative difference between z_N^G and z_N^{LP} is up to 15.4%.

	$\lvert A\rvert$	LSF (CPLEX)			LSCF (Column Generation)				Heuristic
		Slots	LP	Time	Slots (z_L^{IP})	LP (z_L^{LP})	Iter.	Time	Slots (z_L^G)
N10	26	17	11	7h30m	17	17	20	6s	19
N20	134	–	24	10h*	70	70	175	4m22s	77
N30	176	–	28	10h*	94	93	111	4m23s	103
N40	184	–	21	10h*	45	43	360	15m47s	44
N50	296	–	31	10h*	85	84	445	1h32m35s	108
N60	396	–	–	10h*	115	114	874	2h53m16s	131

Table 4.2: Numerical results for link-oriented assignment. *) The maximum computational time for CPLEX is set to 10h.

4.2 Link-oriented Assignment

Table 4.2 summarizes the computational results for link-oriented assignment. Here, the number of time slots of the TDMA schedule equals the number of links $|A|$, which is shown in the second column of the table. The other columns have the same meaning as those in Table 4.1. In addition, we use '–' to denote that no solution of LSF is obtained within the time limit.

From an optimization point of view, scheduling of link-oriented assignment is much more challenging than that of node-oriented assignment, because the former involves a much larger solution space. We observe that, using formulation LSF, CPLEX could only solve the problem for the network with 10 nodes. For other networks, no feasible schedule could be found within the time limit. The column generation method solved the LP-relaxation of LSCF for all networks, although the solution process took considerably more time than the case of node-oriented assignment. The LP-bound is very close to the integer optimum. For two of the six networks, optimal schedules were found using the transmission groups generated in the column generation method. For the other cases, the difference between z_L^{IP} and z_L^{LP} is one or two time slots.

We observe that spatial reuse is achieved for all the networks. In particular, the ratio between $|A|$ and z_L^{LP} varies between 1.52 and 4.09. It can also be noted that link-oriented assignment achieves higher spatial reuse than node-oriented assignment.

The performance of the greedy heuristic varies by network instance. For network N40, it found the best know solution (which may be optimal). For the other networks the relative difference between z_L^{LP} and z_L^G is up to 28.5%.

4.3 Conclusions

Resource optimization is a crucial issue for ad hoc networks. A particular optimization problem concerns finding a STDMA schedule with minimum length. We have studied this optimization problem for node-oriented and link-oriented assignment strategies. Using set covering formulations, we are able to derive a column generation method which efficiently solves the LP-relaxations. We have also evaluated two approaches for finding feasible schedules. The first approach applies integer programming to the transmission groups generated in the column generation method, and the second approach is a greedy algorithm.

Several conclusions can be drawn from our computational study. First of all, the LP-relaxations of the set covering formulations yield very tight bounds. These bounds are very useful for benchmarking the performance of heuristic algorithms. Secondly, applying integer programming to the transmission groups generated in the column generation procedure often leads to optimal or near-optimal solutions. Moreover, the greedy algorithm performed well for some of the cases in our experiments (particularly for node-oriented assignment). For other cases, the solutions found by the algorithm are up to 28.5% from optimality. The main advantage of this algorithm is its simplicity, which makes it an interesting candidate for distributed implementations.

Part II

Frequency Assignment in Frequency Hopping GSM Networks

Chapter 5

Frequency Assignment in Frequency Hopping Networks

This chapter contains a description of how frequency hopping works in a GSM network. Frequency hopping is an important feature in a GSM network since the interference can be decreased. This chapter includes also a discussion of the classical frequency assignment problem and a frequency assignment model for frequency hopping GSM networks.

5.1 Frequency Hopping in GSM Networks

One way to increase the performance and the capacity in a GSM network is to adopt slow Frequency Hopping (FH). FH improves the performance when the received signal suffers from multipath fading, and FH reduces the required signal to interference ratio [129, 145]. The infrastructure of a GSM network comprises of BTS (see Section 1.3.1) located at a number of sites. A BTS site consists of one or several cells. A transmission facility is often referred to as a transceiver (TRX). Typically, a cell has several TRXs, and one frequency is allocated to each TRX. The capacity of a cell can therefore be measured in the number of TRXs. The capacity of a frequency is divided into eight time slots using TDMA. A time slot that carries user traffic is known as Traffic Channel (TCH). For administrative purposes, e.g. providing mobile terminals with control information, there is also one Broadcast Control Channel (BCCH) in each cell. This channel consumes one time slot, and is operated by one TRX in the cell. FH means that a TRX shifts frequency at every new TDMA frame, the hopping is performed approximately 217 times per second. For a specific user, a new frequency is used from one time slot to the next. The advantages of FH are twofold, frequency diversity and interference diversity. The drawback with FH is that the network components become more complex and expensive. Performance evaluations for FH GSM networks can be found in [38, 77, 81, 104, 124, 142].

5.1.1 Frequency Diversity Gain

In wireless communications, the transmitted signal is subject to scattering. The signal from the source is often reflected against objects and the signal that enters the receiver is a combination of many different reflected components. These multipath propagation components have traveled different ways and will have different phases at the receiver antenna. In some cases, the phase difference will lead to a weak (attenuated) summarized signal at the receiver, a so called fading dip. A fading dip may cause severe signal degradation and information loss at the receiver and should be avoided. This multipath fading which is dependent on phase differences between signal components is very rapid, and this phenomenon is also called fast fading or Rayleigh fading. The phase difference at the receiver depends on the frequency of the wave components, since the path length is different for different frequencies, which results in the fading being different at different locations. If the receiver is moving, the user will experience fading dips frequently, and the distance between the fading dips depends on the frequency used. Changing frequency continuously is one way to reduce the influence of fading dips and the probability of good link quality is increased. The advantage of using FH is that the channel will not suffer severe fading dips under longer time periods. The frequency interference gain is significant when the mobile terminal stands still or moves slowly, and decreases when the speed is higher [12, 43, 151]. There are several factors that affect the frequency diversity gain. The number of hopping frequencies is an important factor that affects the frequency diversity gain. Higher gain is achieved with more hopping frequencies, but more than 8 hopping frequencies give less improvement [127, 141].

5.1.2 Interference Diversity Gain

Besides frequency diversity gain, FH also introduces interference diversity gain. Without FH, neighboring cells with adjacent or co-channel frequencies to the carrier frequency will interfere continuously. The interference can be severe for certain frequencies and may lead to information loss. FH means that different frequencies will interfere with the carrier at different time slots. Instead of having a constant level of interference between two cells, the interference is spread among several frequencies. For a specific carrier, only some time slots will suffer from high interference. From a system point of view, strong interferers are shared between users and so called interference averaging is achieved. A GSM network can be planned for average interference instead of worst case interference. The interference diversity gain is dependent on the number of hopping frequencies and interference levels from the interfering cells. A significant interference diversity gain is obtained with as little as 3 hopping frequencies [31, 110].

The interference diversity gain consists of two parts, both related to the network load. When pseudo-random FH is introduced, the probability of being hit by another MS is averaged out between all frequencies in the available spectrum. The first part leads to a gain linear to the network load, e.g. a halving of the network load gives a gain of 3 dB. No loading gain is achieved in a non hopping network since no averaging of interference of the different frequencies occurs. A non hopping network must be considered for a worst case scenario. The other part corresponds to the coding scheme used, and the coding becomes much more efficient with lower loads, leading to a fractional loading gain [141, 143, 149].

5.1.3 Frequency Hopping Strategies

FH can be performed in two different ways, cyclic mode or random mode. In cyclic hopping, all interfering sectors use the same hopping sequence, and a channel will be affected by the same interferers all the time. No interference diversity is obtained with cyclic hopping [110]. Random hopping uses sequences that are pseudo-randomly generated and interference diversity is obtained. Every cell which performs FH must be assigned with a Hopping Sequence Number (HSN) which may take 64 different integer values (0,1,2,..,63). HSN=0 implies cyclic FH, and the remaining numbers are different pseudo-random sequences.

The physical ways to perform FH in a cell are baseband hopping and synthesized hopping. In baseband hopping, every TRX is assigned to a fix frequency, while the channel is shifted between the TRXs. The number of frequency carriers must be equal to the number of TRXs. The other scheme is synthesized hopping, where a channel stays at a certain TRX, while the TRX shifts among the available frequencies. In the latter method, the number of carriers is greater than or equal to the number of TRXs. The procedure is to allocate different Mobile Allocation Lists (MAL) of frequencies to all TRXs. The MAL contains all the hopping frequencies for a specific TRX. Since the number of frequencies can be more than the number of transceivers in a cell when synthesized hopping is used, the reuse factor can be changed while the number of TRXs is fixed.

MAIO management is a method for avoiding interference in multi-sector sites. The Mobile Allocation Index Offset (MAIO) is a parameter which takes as many values as the number of frequencies. Carriers using the same HSN but different MAIO, never use the same frequency within the same burst. Usually channels in the same cell use the same HSN but different MAIOs. By using synthesized hopping, it is possible to allocate the same frequencies to all the sectors of a site, since the MAIO management technique will prevent the frequencies being used simultaneously [140, 143].

5.1.4 Common and dedicated spectrum bands for BCCH and TCH frequencies

There are two strategies for using the frequency band for the TCH and BCCH carriers, common band and dedicated bands. In the common band strategy both TCH and BCCH carriers use the entire band, and a TCH can receive interference from both TCH and BCCH carriers. In the dedicated bands strategy the frequency band is split into two parts, one for the TCH frequencies and one for the BCCH frequencies. In the second strategy a TCH carrier can only be interfered by another TCH. FH is not performed at the TRXs that operates the BCCH. Simulations done show that dedicated spectrum bands give less interference than common band strategy which suffered from severe disturbance in the downlink [89]. It is easier to put extra TCH-TRX in an existing cell when dedicated bands are used since the BCCH frequency plan does not have to be changed.

5.2 Classical Frequency Assignment

One very well-known problem in operations research is the classical frequency assignment problem (FAP). The FAP belongs to the class of NP-complete problems, which means that the problem probably can not be solved in polynomial time. There exist two major approaches to deal with the FAP in wireless networks: Fixed Channel Assignment (FCA) and Dynamic Channel Assignment (DCA). In FCA, a channel is assigned to a connection beforehand and can not be changed on-line. In DCA the channels are changed on-line whenever the radio connection suffers from interference and the quality requirements are not fulfilled [12, 86]. In the following we concentrate on FCA methods.

5.2.1 Different Versions of the Frequency Assignment Problem

There exist several versions of FAP [86], e.g. minimum span, minimum order, minimum blocking and minimum interference FAP. In the minimum span FAP, certain soft and/or hard restrictions are given for the quality of the network. These restrictions can be interference and separation requirements. The soft restrictions can be violated at some penalty cost, but the hard restrictions must be met. The objective is to minimize the difference between the minimum and maximum frequencies assigned, the span. A number of heuristics have been developed for solving the minimum span problem [27, 70, 134]. Exact methods such as lower bounding procedures are presented in [7, 8, 16, 17, 50, 69, 135]. Most of the algorithms used are related to graph coloring techniques, and excellent surveys in the field of frequency assignment are [73, 86] for example. In [138] different algorithmic approaches are investigated. A FAP related to the minimum span problem is minimum order FAP. Instead of minimizing the span of frequencies, the number of frequencies is minimized. The minimum blocking and minimum interference FAP use a fixed span of frequencies. In the minimum blocking FAP, the goal is to assign frequencies in such a way that the overall blocking probability of the network is minimized. The FAP that is in focus of this thesis, is the minimum interference problem. All the frequencies in a telecom operators spectra are assumed to be used in the frequency assignment. The objective is to minimize the total sum of interference in the telecom operators network. In the rest of this chapter the discussion focuses on the minimum interference FAP.

5.2.2 Minimum Interference Frequency Assignment Problem

In the minimum interference FAP, the total sum of the interference levels is minimized. From the available number of frequencies, exactly one frequency is assigned to one TRX. To avoid intolerable interference, a minimum carrier spacing (separation) is introduced between different TRXs within the same cell and BTS. Significant interference may occur between two TRXs using the same or adjacent frequencies. The co-channel, adjacent-channel interference and separation are always given between pairs of TRXs. These parameters are specified in three square matrices. In FAP, the BCCH frequencies and TCH frequencies can be planned separately when different spectra are used. Hence, the FAP can be considered as two problems, one for the TCH and one for the BCCH.

The minimum interference FAP can be stated as follows:

Given a list of TRXs and a frequency spectrum, a list of feasible frequencies for each TRX, co-channel and adjacent channel interference matrices and separation requirements for each pair of TRXs. The problem is to assign exactly one frequency to each TRX such that no requirements are violated and the total cumulative interference is minimized.

The mathematical expression of the minimum interference FAP is:

$$
\begin{aligned}
&[FAP] \\
&\min \quad \sum_p \sum_q c_{pq}^{co} \left(\sum_{i \in F_p} \sum_{j \in F_q} a_{ij}^{co} x_{ip} x_{jq} \right) + \\
&\qquad\quad \sum_p \sum_q c_{pq}^{adj} \left(\sum_{i \in F_p} \sum_{j \in F_q} a_{ij}^{adj} x_{ip} x_{jq} \right) \qquad (5.1)
\end{aligned}
$$

$$
\begin{aligned}
\text{s.t.} \quad & \sum_{i \in F_p} x_{ip} && = && 1, \forall p && (5.2) \\
& x_{ip} + x_{jq} && \leq && 1, \forall i \in F_p, j \in F_q : f_{ij} < d_{pq}, \forall p, q && (5.3) \\
& x_{ip} && = && 0/1, \forall i \in F_p, \forall p && (5.4)
\end{aligned}
$$

The variables in the FAP model are:

$$
x_{ip} = \begin{cases} 1 & \text{if frequency } i \text{ is assigned to TRX } p \\ 0 & \text{otherwise.} \end{cases}
$$

The parameters used in the FAP model are

c_{pq}^{co}	=	co-channel interference between interfering TRX p and serving TRX q.
c_{pq}^{adj}	=	adjacent channel interference between interfering TRX p and serving TRX q.
f_{ij}	=	the distance between the frequencies i and j.
d_{pq}	=	separation requirement between TRX p and TRX q.
F_p	=	the set of feasible frequencies at TRX p.

The parameters a_{ij}^{co} and a_{ij}^{adj} in the objective function are assignment parameters and can be expressed as follows

$$
a_{ij}^{co} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are the same frequency} \\ 0 & \text{otherwise.} \end{cases}
$$

and

$$
a_{ij}^{adj} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are adjacent frequencies} \\ 0 & \text{otherwise.} \end{cases}
$$

The co-channel and adjacent-channel interference form two matrices. The co-channel interference matrix and the adjacent-channel interference matrix may not be symmetric, i.e. usually $c_{pq}^{co} \neq c_{qp}^{co}$ and $c_{pq}^{adj} \neq c_{qp}^{adj}$. The separation requirement d_{pq} between TRXs is stored in a matrix which is symmetric. The separation constraint (5.3) states that no two frequencies with a distance f_{ij} less than separation requirement d_{pq} can be assigned to TRX p and TRX q. When i and j are equal, the distance is $f_{ij} = 0$. If there are no separation requirements between a pair of TRXs, the corresponding d_{pq} should be less than zero. The constraint (5.2) ensures that only one of the feasible frequencies in the set F_p can be assigned to TRX p.

In the literature, several heuristics have successfully generated good feasible solutions. In [30] a tabu search heuristic is used for the problem of maximizing the sum of traffic loads offered by different regions. Other approaches with a tabu search heuristic are presented in [32]. Several heuristic algorithms are tested in [25, 117]. Graph theoretical planning techniques are presented in [51]. In [26] an entirely new integer programming formulation for the fixed span problem is developed and a heuristic method to solve the model is presented. In [92] an evolutionary optimization method is used to the fixed span problem.

5.3 Frequency Assignment in Frequency Hopping Networks

Most of the research in frequency assignment deals with non hopping networks. The corresponding planning problem becomes much more complex as soon as FH is introduced. The FAP in FH networks (FAPH) is a generalization of classical FAP. From an optimization point of view, FAPH is in general considerably more difficult than FAP. The main reason for this complexity lies in the larger number of possible lists of frequencies.

5.3.1 Previous Research

The frequency hopping FAP has not been investigated to the same extent as the classic FAP. In [128] a model using Markov chains is developed. The model measures a test subscriber in a dynamic traffic case. In the model, two types of cost functions are used, the long term bit erasure rate and the bad quality cost function. The cost functions have different complexity levels and the solutions are obtained by a dynamic channel assignment policy in an FCA network.

Some important contributions are the work from a research group at Aalborg University, CPK, Denmark, which will be covered in more detail. The contributions are from [141, 143]. The model is developed to study the combined effect from all the serving channels in the resource allocation process. An allocation algorithm has been developed, which takes into consideration the impact of frequency and interference diversity as well as fractional loading gain. The algorithm is a heuristic algorithm, a combination of a search tree and a dynamic constraint setting. The idea is to minimize the interference cost function in a given network with existing base stations and frequencies. The model in [141, 143] includes both co-channel and adjacent-channel interference, and the use of dedicated spectrum bands for the TCH and BCCH carriers gives two separate cost functions, one for the TCH layer and one for the BCCH layer. The input data are two input files, one describing the interference between cells (interference matrix) and the other describing the requirements for each cell, i.e. how many and which frequencies are allocated (carrier database). A relationship between the interference reduction factor (frequency diversity gain) and the MS speed is presented. The relationship has been developed through simulations, where both random and sequential FH have been investigated. Hopping over as much as 8 frequencies gives a frequency diversity gain of 39 % compared to a non-hopping network. A relationship for the interference diversity gain as a function of network load is developed. Most of the gain is achieved when 3 frequencies are used for hopping. In the interference cost function presented in [141, 143], the adjacent-channel interference is the co-channel interference weighted by a factor 0.015. The initial solution is obtained by randomly choosing frequencies for each cell. Constraints such as separation requirements are not violated which give a feasible initial solution. The algorithm continues to change the frequencies with the highest cost function coefficients. To obtain good solutions the algorithm includes constraints to cut off the worst interferers.

5.3.2 Problem Definition

The frequency assignment problem includes both TCH and BCCH frequencies, but with separate spectra. Therefore the assignment problem can be divided into two separate and similar planning problems. The interference from the BCCH spectrum to the TCH spectrum is negligible compared to the interference between frequencies in each band. Since it is necessary to have one BCCH in each cell, only one of the TRXs in a cell is used for the BCCH. The planning of the BCCH frequencies can be considered as a classic FAP. The remaining TRXs in a cell will be assigned with TCHs. The TCH-TRXs in a cell are grouped into a Super Transceiver (STRX). Instead of allocating frequencies to single TCH-TRXs, a MAL is allocated to every STRX. The MAL must contain at least as many frequencies as the number of TRXs in the STRX. Since synthesized FH is used, it is possible to allocate a number of frequencies to an STRX, which is larger than the number or TRXs. The case in which there is only one TRX in the STRX and the MAL length is one is identical to the classical FAP.

The minimum interference FAPH can be stated as follows:

Given a set of STRXs and a set of frequencies, a set of feasible frequencies for each STRX, the minimum and maximum lengths of the MAL for each STRX, the co-channel and the adjacent channel interference for every pair of STRXs, the frequency diversity gain and the interference diversity gain, the separation requirement for each STRX, as well as the separation requirements between some pairs of STRXs. The problem is to assign a MAL to each STRX, such that no separation requirement is violated, and that the total cumulative co-channel and adjacent-channel interference is minimized.

5.3.3 Mathematical Formulation

To formulate FAPH mathematically, we introduce the following parameters.

c_{pq}^{co}	$=$	the co-channel interference between STRX p and STRX q, where p is the interfering STRX.
c_{pq}^{adj}	$=$	the adjacent channel interference between STRX p and STRX q, where p is the interfering STRX.
f_{ij}	$=$	the distance between MALs i and j in the frequency spectrum.
d_{pq}	$=$	the minimum distance between the MALs at STRX p and STRX q.
R_p	$=$	the set of feasible frequencies at STRX p.
N_p	$=$	the number of TRXs in STRX p.

The variables in the FAPH model are:

$$x_{ip} = \begin{cases} 1 & \text{if MAL } i \text{ is assigned to STRX } p \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned} &[FAPH] \\ &\min \quad \sum_p \sum_q (c_{pq}^{co} N_p) \left(\sum_{i \in R_p} \sum_{j \in R_q} a_{ij}^{co} x_{ip} x_{jq} \right) + \\ &\qquad\quad \sum_p \sum_q (c_{pq}^{adj} N_p) \left(\sum_{i \in R_p} \sum_{j \in R_q} a_{ij}^{adj} x_{ip} x_{jq} \right) \end{aligned} \tag{5.5}$$

$$\text{s.t.} \quad \sum_{i \in F_p} x_{ip} = 1, \forall p \tag{5.6}$$

$$x_{ip} + x_{jq} \leq 1, \forall i \in R_p, j \in R_q : f_{ij} < d_{pq}, \forall p, q \tag{5.7}$$

$$x_{ip} = 0/1, \forall i \in R_p, \forall p \tag{5.8}$$

The scaling factors a_{ij}^{co} and a_{ij}^{adj} for the FAPH are

$$a_{ij}^{co} = \frac{|M_i \cap M_j|}{|M_i||M_j|} g(|M_i|, |M_j|) \tag{5.9}$$

$$a_{ij}^{adj} = \frac{|(f : f \in M_i, f+1 \in M_j)| + |(f : f \in M_i, f-1 \in M_j)|}{|M_i||M_j|} g(|M_i|, |M_j|). \tag{5.10}$$

M_i and M_j are the MALs for STRX p and STRX q respectively. The factor $g(|M_i|, |M_j|)$ is a function which includes the gain from FH. The gain factor is explained in more depth in Section 7.1. For the computation of a_{ij}^{co} and a_{ij}^{adj} see Appendix A. The classical FAP is a special case of FAPH where a_{ij}^{co} and a_{ij}^{adj} are either zero or one. The set R_p contains all the feasible MALs for an STRX p with respect to separation requirements and minimum and maximum lengths of the MAL. In the classic FAP, the set R_p can only contain MALs of length 1, i.e. $|R_p| = 1, \forall p$. Constraint (5.6) ensures that every STRX is assigned to exactly one MAL. Constraint (5.7) takes care of the separation requirement between pairs of STRXs.

Chapter 6

Solution Methods for Frequency Hopping Networks

One way to obtain good solutions is to apply different heuristic methods. The meaning of the word *heuristic* is find or discover and comes originally from the Greek word *heuriskein*. A heuristic does not necessarily solve the problem to optimality, but usually delivers good quality feasible solutions.

Local search is used in many heuristic methods. Suppose we have a minimization problem over a set of feasible solutions. Local search considers only a subset of the solution space, because very often it is not practical to consider the entire space. In continuous optimization, the subset or neighborhood is usually defined as all solutions within a given distance from the present solution. In the discrete case we have to replace the distance by some other measure, e.g. all solutions that can be reached by changing the value of one binary variable. The local search method searches the neighborhood of the current solution to find an improvement. The steepest descent method selects the neighbor which gives the largest improvement compared to the present solution. The solution space is explored by repeatedly moving from the current to neighboring feasible solutions. If no solution in the neighborhood results in a lower cost function value, the current solution is taken as the final solution, which is a local optimum. If the neighborhood is large, there is a greater probability of finding good solutions. On the other hand, a large neighborhood results in increased computational time. Since the solution space can contain several local optima, methods are needed to continue the search for other local optima. Such heuristics are sometimes called meta heuristics, and two which are well known in this class are Tabu Search and

Simulated Annealing.

6.1 Construction of an Initial Solution

To obtain an initial solution for the meta heuristic algorithms, a feasible solution such that the MAL length for STRX p equals the minimum requirement is considered. Such an assignment is a feasible solution to the corresponding classical FAP. A constructive heuristic is used to create an initial solution. The constructive heuristic considers one STRX at a time and assigns the minimum number of frequencies required. The frequencies for a certain STRX can be chosen in different ways. One way is to randomly choose the minimum number of feasible frequencies to an STRX, and another is to choose the frequencies in a greedy manner. In the greedy procedure, the frequency that yields the least additional interference is chosen until the minimum number of frequencies are assigned to the MAL. Computational experiments show that choosing frequencies randomly gives very poor initial solutions. The solutions of a tabu search heuristic from a very poor initial solution also tends to be poor [25], therefore the initial solutions will be constructed in a greedy manner. Another concern in constructing an initial solution is the order in which the STRXs are considered. The STRXs can be taken randomly or by their indices. Another strategy is to first assign MALs to STRXs that have a significant performance impact on the rest of the network, for example by sorting the STRXs in a decreasing order according to the sum of the interferences to other STRXs.

6.2 Mobile Allocation List Generation

To use solution strategies such as local search or meta heuristics, it is important to define the neighborhood structure of the specific problem, in our case the FAPH. Usually, the neighborhood consists of all neighboring solutions that can be reached from the present solution by some predefined exchange operation, which we call *a move*.

In our case, we have as many neighbors as there are STRXs in the problem. The exchange operation for STRX p, is defined as replacing the current MAL of STRX p with the optimal MAL, while all other MALs are being fixed (see Figure 6.1).

Definition: *The neighborhood is defined by all solutions that can be obtained by replacing the MAL for a single STRX by the optimal one, with all other MALs being fixed.*

In order to evaluate a neigbor, we need to find the optimal MAL for an STRX.

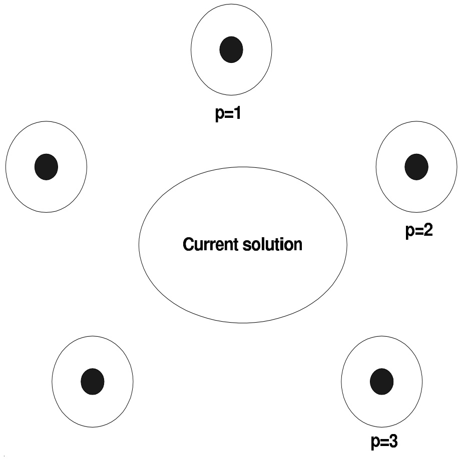

Figure 6.1: Illustration of the neighborhood structure.

This can be formulated as an integer programming problem, see Section 6.2.1.

6.2.1 Integer Programming Model for MAL Generation

The MAL generation problem can be formulated as an integer programming problem. The procedure for the integer programming approach is to generate and assign a MAL to an STRX p. When an STRX is chosen the rest of the solution is fixed. Since the rest of the solution is known it is only the MAL of STRX p that will be changed. The MAL that will be assigned to STRX p has length l. The MAL assigned to STRX p must at least contain N_p frequencies, otherwise one or several TRXs in STRX p are not assigned a frequency. The problem of finding the optimal MAL for STRX p can be approached by solving the fixed length problem with all feasible length values, $m_p \leq l \leq n_p$.

m_p	$=$	the minimum length of the MAL assigned to STRX p.
n_p	$=$	the maximum length of the MAL assigned to STRX p.

The index p is kept fixed in the following mathematical descriptions, i.e. for a specific STRX p, the optimal MAL of length l is constructed. The MALs assigned to the other STRXs do not change during the MAL construction for STRX p.

The values in the co-channel and adjacent-channel interference matrices must be scaled to the current situation by the scaling factors a_{ij}^{co} and a_{ij}^{adj}. Therefore it is important to know which frequencies are common in STRX p and STRX q and which are adjacent. The length of the MAL for STRX p is known to be l, and the co-channel scaling factor between STRX p and STRX q is determined by the common frequencies. No co-channel interference will occur if MAL p and MAL q are disjoint sets. The integer programming model uses the following variables

$$z_{vp} = \begin{cases} 1 & \text{if frequency } v \text{ is choosen in MAL for STRX } p \\ 0 & \text{otherwise.} \end{cases}$$

y^{co}_{pq} = the number of common frequencies between STRX p and STRX q.

y^{adj}_{pq} = the number of frequencies between STRX p and STRX q such that the spacing is 1.

The following parameters are used in the integer programming model

$$e^{co}_{vq} = \begin{cases} 1 & \text{if common frequency } v \text{ is present in MAL for STRX } q \\ 0 & \text{otherwise.} \end{cases}$$

$$e^{adj}_{vq} = \begin{cases} 1 & \text{if adjacent frequency } v \text{ is present in MAL for STRX } q \\ 0 & \text{otherwise.} \end{cases}$$

F_p = the set of feasible frequencies at STRX p.

d_p = the minimum distance between the frequencies in the MAL assigned to STRX p.

d_{pq} = the minimum distance between the MALs at STRX p and STRX q.

f_{vw} = the distance between MALs v and w in the frequency spectrum.

The parameter d_p is the frequency separation requirement for the MAL assigned to STRX p. For two frequencies to be both present in the MAL at STRX p, they must have a distance of at least d_p in the frequency spectrum.

Assume that the MALs for STRX p and STRX q are M_i and M_j. The number of common frequencies for the STRX p and STRX q can be expressed as

$$y^{co}_{pq} = |M_i \cap M_j|. \tag{6.1}$$

The number of adjacent frequencies can be expressed as

$$y^{adj}_{pq} = |(f : f \in M_i, f+1 \in M_j)| + |(f : f \in M_i, f-1 \in M_j)|. \tag{6.2}$$

Scaled co-channel and adjacent-channel interference parameters $\overline{c}^{co}_{pq}$ and $\overline{c}^{adj}_{pq}$, are

$$\overline{c}^{co}_{pq} = \frac{c^{co}_{pq} N_p g(l, |M_j|)}{l|M_j|} \tag{6.3}$$

$$\overline{c}_{pq}^{adj} \quad = \quad \frac{c_{pq}^{adj} N_p g(l, |M_j|)}{l|M_j|}. \tag{6.4}$$

A model formulation for optimizing a MAL for STRX p while the rest of the network is fixed can be expressed as

$$\min \sum_{q:q\neq p} (\overline{c}_{pq}^{co} + \overline{c}_{qp}^{co}) \cdot y_{pq}^{co} + \sum_{q:q\neq p} (\overline{c}_{pq}^{adj} + \overline{c}_{qp}^{adj}) \cdot y_{pq}^{adj} \tag{6.5}$$

$$\begin{array}{llcll}
\text{s.t.} & \displaystyle\sum_{v\in F_p} z_{vp} & = & l & (6.6) \\
& \displaystyle\sum_{v\in F_p} e_{vq}^{co} z_{vp} & = & y_{pq}^{co} \quad \forall p \neq q & (6.7) \\
& \displaystyle\sum_{v\in F_p} \sum_{w\in F_q : f_{vw}=1} e_{wq}^{adj} z_{vp} & = & y_{pq}^{adj} \quad \forall p \neq q & (6.8) \\
& z_{vp} + z_{wp} & \leq & 1 \quad \forall v, w \in F_p : f_{vw} < d_p & (6.9) \\
& z_{vp} & = & 0 \quad \forall v \in F_p, \exists w \in M_j : & \\
& & & \quad f_{vw} < d_{pq} & (6.10) \\
& z_{vp} & = & 0/1 \quad \forall v \in F_p & (6.11) \\
& y_{pq}^{co}, y_{pq}^{adj} & \in & Z^+ \quad \forall p \neq q. & (6.12)
\end{array}$$

The length of the MAL for STRX p is l and constraint (6.6) ensures that the number of feasible frequencies chosen is l. Constraint (6.7) summarizes the number of common frequencies between STRX p and STRX q. In a similar way, constraint (6.8) deals with the sum of adjacent frequencies between STRX p and STRX q. The separation requirement in the MAL for STRX p must be valid. Constraint (6.9) ensures that two frequencies v and w can not be assigned to the same MAL, unless the internal separation requirement d_p is fulfilled. Finally constraint (6.10) ensures that the separation requirement between two STRXs is valid. If frequency w is assigned to MAL M_j, the frequency v can not be assigned to the MAL for STRX p if the distance is smaller than d_{pq}.

The model can be expressed in a more compact form. The constraints (6.7) and (6.8) can be substituted into the cost function (6.5). The cost function can be rewritten as

$$\begin{aligned}
& \sum_{q:q\neq p} (\overline{c}_{pq}^{co} + \overline{c}_{qp}^{co}) y_{pq}^{co} + \sum_{q:q\neq p} (\overline{c}_{pq}^{adj} + \overline{c}_{qp}^{adj}) y_{pq}^{adj} \\
= & \sum_{q:q\neq p} (\overline{c}_{pq}^{co} + \overline{c}_{qp}^{co}) \sum_{v\in F_p} e_{vq}^{co} z_{vp} + \sum_{q:q\neq p} (\overline{c}_{pq}^{adj} + \overline{c}_{qp}^{adj}) \sum_{v\in F_p} \sum_{w\in F_q : f_{vw}=1} e_{wq}^{adj} z_{vp} \\
= & \sum_{v\in F_p} \sum_{q:q\neq p} (\overline{c}_{pq}^{co} + \overline{c}_{qp}^{co}) e_{vq}^{co} z_{vp} + \sum_{v\in F_p} \sum_{w\in F_q : f_{vw}=1} \sum_{q:q\neq p} (\overline{c}_{pq}^{adj} + \overline{c}_{qp}^{adj}) e_{wq}^{adj} z_{vp}
\end{aligned}$$

$$= \sum_{v \in F_p} b^{co}_{vp} z_{vp} + \sum_{v \in F_p} b^{adj}_{vp} z_{vp}$$

The parameter $b_{vp} = b^{co}_{vp} + b^{adj}_{vp}$ is a sum of contributions from co-channel and adjacent channel interference and does not depend on the variable z. b^{co}_{vp} and b^{adj}_{vp} have following expressions:

$$b^{co}_{vp} = \sum_{q:q \neq p} (\overline{c}^{co}_{pq} + \overline{c}^{co}_{qp}) e_{vq} \tag{6.13}$$

$$b^{adj}_{vp} = \sum_{w \in F_q : f_{vw}=1} \sum_{q:q \neq p} (\overline{c}^{adj}_{pq} + \overline{c}^{adj}_{qp}) e_{wq} \tag{6.14}$$

The mathematical formulation for finding the optimal MAL length l can be expressed in the compact form as

$$\min \sum_{v \in F_p} b_{vp} z_{vp} \tag{6.15}$$

$$\text{s.t.} \sum_{v \in F_p} z_{vp} = l \tag{6.16}$$

$$z_{vp} + z_{wp} \leq 1 \quad \forall v, w : f_{vw} \leq d_p \tag{6.17}$$

$$z_{vp} = 0 \quad \forall v \in F_p, \exists w \in M_j : f_{vw} < d_{pq} \tag{6.18}$$

$$z_{vp} = 0/1 \quad \forall v \in F_p \tag{6.19}$$

6.2.2 A Greedy Algorithm

One way of solving the MAL generation problem is to use commercial optimization software. In this case we find the optimal solution with Integer Programming (IP). Another possibility is to use a heuristic, e.g. the following greedy heuristic:

Algorithm: Greedy (GR)

1. Start with an empty MAL, $S^0 = \emptyset$ and $t = 0$.
2. Choose a feasible frequency f with respect to separation requirements, whose additional cost is minimum. If no such frequency exists, stop.
3. Set t=t+1, $S^t = S^{t-1} \cup \{f\}$ and update separation requirements.
4. If $|S^t| = l$, stop otherwise go to step 2.

In this case we have a slightly modified neighborhood, since we no longer can guarantee an optimal MAL. The MALs generated using the GR approach tends to be near-optimal. Usually in local search, all neighbors are different from the current solution. However, in our case it may happen that the current MAL is already optimal, and this can be handled in two ways; either we do not allow this, i.e. we force the model to find the best MAL - but not the same as the current one, or we simply allow the same solution to be generated. These two strategies are controlled by two parameters, Modify Length (ML) and Optimize Length (OL). These parameters can be used one at a time or weighted together. The ML approach always chooses the best MAL with different length compared to the one that is already assigned to the STRX. The OL approach always selects the best MAL no matter what length the previous assigned MAL had. When both ML and OL are used simultaneously, different weight combinations can be used. The parameter weights can take integer values from 0 to 10. For example when OL=1 and ML=1 the probability of choosing one of the options is the same. If one of the weights is set to zero, the corresponding parameter is not active.

Another possibility of a neighbor structure is to base the exchange strategy on classical add and/or drop operations (AD). In this case we will not regenerate the same MAL. Preliminary computational results showed however, that this strategy is not as efficient as the previous one.

6.3 The Tabu Search Algorithm

The tabu search heuristic, see [55], uses a local search method, and by applying different attributes in the solution space, certain solutions can attain tabu status. To avoid unnecessary repetition of moves around a local optimum, i.e. cycling, the solutions which attain tabu status can not be chosen for a predefined number of iterations or a certain time span, called the *tabu tenure*. Often it is too expensive in terms of computational effort to store all the recent solutions, so instead certain key attributes are stored. In tabu search, the attributes and tabu tenure to a solution which has tabu status are stored in a tabu list. On some occasions, the tabu status is too strict and may block good moves. Aspiration criteria are introduced into tabu search when a tabu status for a solution can be over-ridden. One simple example of the aspiration criteria is to allow a move to be made when the new solution has a better solution than the currently best found. The neighborhood in tabu search is the original neighborhood with solutions from the tabu list removed.

The FAPH model described in Section 5.3.3 can be solved by a tabu search heuristic, a meta heuristic that has been successful in solving many difficult combinatorial optimization problems. The tabu search heuristic is a combined MAL generation and assignment procedure. Starting from an initial feasible solution, the tabu search attempts to change the current MAL solution, following the neighborhood definition in Section 6.2, and search for an improved objective function value. Feasibility is kept in every iteration and the best solution found after termination gives the final solution.

Evaluating all neighbors in every iteration will lead to very long computation times, and makes it difficult to escape from local optima. Therefore, a restricted neighborhood is chosen. This restricted neighborhood contains only a subset of all the STRXs in the network.

MAL modification of a certain STRX in the restricted neighborhood is an important part of the tabu search algorithm. The solution procedure for the tabu search heuristic, from the current solution is to pick one of the STRXs and keep the MALs of the other STRXs fixed. Generating an optimal MAL for an STRX while keeping the others fixed can be formulated as a linear integer problem as described in Section 6.2.1.

When the MAL is chosen for one of the STRXs in the restricted neighborhood, the new MAL and interference value are stored, and the original MAL is assigned. The procedure continues by choosing the next STRX in the restricted neighborhood and repeating the MAL modification part. When the solution procedure has gone through all STRXs in the restricted neighborhood, the MAL which yields the lowest interference value is assigned to the STRX. The STRX with a modified MAL is added to the tabu list. A new restricted neighborhood is sampled, and the solution procedure is repeated. The STRXs in the tabu list cannot be chosen during the length of the tabu tenure, except when an aspiration criteria is valid. If a better solution than the best one is found, the aspiration criteria overrides the tabu status of the STRX.

Algorithm: Tabu Search

1. Given a feasible initial solution.
2. Select the size of the restricted neighborhood (SZ), length of the tabu tenure (TBL) and the number of iterations (M). Select IP or GR as solution strategy, and the values of OL and ML.
3. For the number of iterations $i = 1, .., M$.
   (a) Sample a restricted neighborhood in the network
   (b) For the STRXs in the restricted neighborhood
      i. Calculate the MAL for STRX p.
      ii. Store the calculated MAL and interference value and assign the original MAL to STRX p.
   (c) Assign the MAL to the STRX in the restricted neighborhood that results in the smallest interference value, taking tabu list and aspiration criteria into account.

6.4 The Simulated Annealing Algorithm

Simulated annealing as a meta heuristic for optimization problems has been around for more than 20 years. The origin of the technique of simulated annealing comes from cooling heated materials. In the beginning of the 1980s the idea was born that this could be used in the search of feasible solutions of optimization problems, see [82]. The simulated annealing technique can be seen as an extension of the local search method. In simulated annealing, moves which deteriorate the solution are allowed, e.g. in a minimization problem uphill moves are accepted. The acceptance of an inferior solution is described by a probability function. The probability function depends on a control parameter t (temperature) and the magnitude of the difference δ. δ is the difference between the new and the current cost function values. The acceptance probability for an inferior solution can be expressed as

$$p(t, \delta) = e^{-\frac{\delta}{t}}. \tag{6.20}$$

The temperature parameter is not constant, and after one or several iterations, t is decreased. When t becomes smaller, the acceptance probability decreases. To examine whether an inferior solution will be accepted or not, a number r between 0 and 1 is randomly generated. If r is smaller than the acceptance probability, the inferior solution is accepted. It is easier to accept poor solutions in the beginning of the simulation.

The temperature is reduced after a number of iterations by a reduction function $\alpha(t)$. The simplest reduction function used is $\alpha(t) = at$, where $a < 1$. Another commonly used cooling function is $\alpha(t) = \frac{t}{1+\beta t}$ where β is a small value. For other reduction functions see [9, 120]. The initial temperature must be "hot" enough to allow the final solution to be independent of the initial solution. If the temperature parameter is not sufficiently high at the start, the algorithm will only explore a small fraction of the solution space. There can be a lot of modifications to the simulated annealing algorithm, e.g. using different acceptance probability functions.

The initial solution is constructed in a greedy manner as explained in Section 6.1. In the simulated annealing algorithm, an STRX is randomly selected from the total number of STRXs. The optimal MAL is calculated in exact the same way as in the tabu search algorithm, which is described in Section 6.3. Calculate the deviation δ, which is the difference between the newly found cost function value and the current cost function value. If the new solution has a lower cost function value compared to the present solution, δ is negative. The new solution is assigned and a new STRX is again randomly selected. If the deviation δ is positive, the new solution is accepted according to a probability function. To evaluate whether the solution is accepted or not, a number r, $r \in (0, 1)$, is randomly generated. If r is smaller than the acceptance probability, the new solution is assigned.

Algorithm: Simulated Annealing

1. Given an initial solution.

2. Select IP or GR, weights for the ML and OL parameters, and the number of iterations M.

3. For $i = 1, ..., M$.

   (a) Randomly select an STRX.

   (b) Calculate MAL for the randomly selected STRX.

   (c) Calculate the deviation δ.

   (d) If $\delta < 0$, assign the MAL to the randomly selected STRX.

   (e) If $\delta \geq 0$ generate r uniformly in the range [0,1]. If $r < e^{-\frac{\delta}{t}}$ assign the MAL to the randomly selected STRX otherwise keep the present solution.

   (f) Reduce the temperature according to the cooling function.

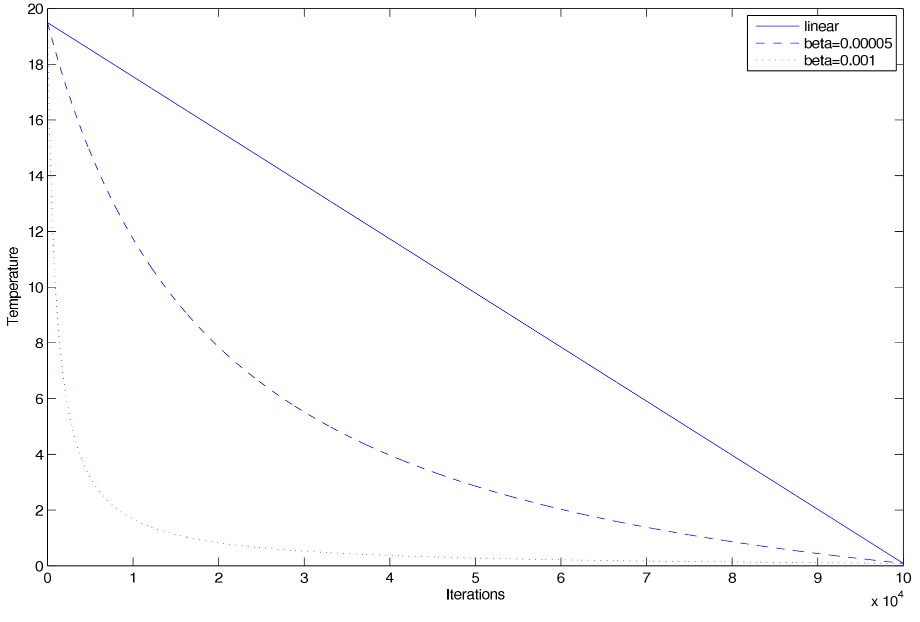

Figure 6.2: The cooling function with different beta values and M=100 000 iterations.

Instead of using temperatures at the start and at the end of the procedure, acceptance probabilities are given. The probability P_0 corresponds to the probability that a solution that is one unit worse is accepted in the first iteration of the simulation. The probability P_1 corresponds to the probability at the end of the simulation in a similar manner. From the acceptance probabilities, the beginning and final temperatures are calculated. The cooling function for the temperature which handles both the linear and non linear case is

$$t(i) = t_0 + \frac{t_1 - t_0}{\frac{M-1}{1+\beta(M-1)}} \frac{i-1}{1+\beta(i-1)}. \tag{6.21}$$

The parameters t_0 and t_1 are the temperatures in the beginning and at the end of the simulation, M is the total number of iterations and i is the iteration number, starting with $i = 1$ and ending with $i = M$. The cooling function for different beta values are compared in Figure 6.2. $\beta = 0$ corresponds to a linear cooling function and increasing the β value results in a decreased probability for accepting a worse solution.

Chapter 7

Computational Results for Frequency Hopping

7.1 Input data

The data needed to solve the FAPH are a *carrier database*, *interference matrices* and *separation requirements*. The carrier database consists of the number of sites, number of cells at each site and the number of carriers in each cell. The number of carriers/cell is equivalent to the number of TRXs/cell. Instead of considering each TRX separately, all TRXs in a cell are grouped into STRXs, see Figure 7.1. The interference matrices are the co-channel interference matrix and the adjacent-channel interference matrix. The interference matrices are given in the form of STRX to STRX interference. The interference matrices are $N \cdot N$ matrices, with N equals the number STRXs in the network. The co-channel interference matrix describes for each STRX the amount of traffic (in Erlang) that it fails to serve if all TRXs suffer from co-channel interference continuously. Similarly, the adjacent-channel matrix describes the amount of traffic (in Erlang) an STRX fails to serve when all its TRXs are disturbed by adjacent-channel interference all the time. In the input data, the interference-matrix values are scaled with a voice activity factor equal to 0.5 due to the use of discontinuous transmission. The use of power control will decrease the amount of interference in the network and a factor of magnitude 0.5 is used to compensate for power control.

Data from a real life GSM network (URBAN) from a larger city is used. The GSM network contains 309 BTS with totally 421 cells and 785 TRXs. The network has 421 BCCH carriers and 364 TCH carriers. The spectrum for this GSM network consists of 53 frequencies which are split into two separate spectra,

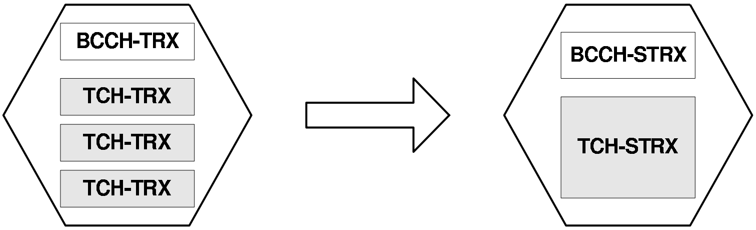

Figure 7.1: Example of how several TRXs in a cell can be transformed into STRXs.

Instance	Cells	TRXs	BCCH TRXs	TCH TRXs	TCH STRXs	TCH freqs	BCCH freqs
URBAN	421	785	421	364	305	29	24
NORR25	65	181	65	116	53	28	22
NORR85	187	482	187	295	134	26	15
NORR400	892	2592	892	1700	802	27	10

Table 7.1: Size and characteristics of different network instances.

one for the BCCH which consists of 24 frequencies and the other for the TCH which consists of 29 frequencies. We will only consider the planning problem for the TCH part since our main focus is on planning FH networks. The TCH TRXs are grouped into 305 STRXs. Several synthesized networks are also used: NORR25, NORR85 and NORR400. Table 7.1 shows the sizes and characteristics of the instances.

The interference matrices have values of very different magnitude and also contain many elements which are zero. The co-channel interference matrix of the URBAN instance, has a largest value of 1013 and smallest $6.25 \cdot 10^{-6}$ and only approximately 16 % of the elements are nonzero. The adjacent-channel interference matrix for the URBAN instance has a span from $6.25 \cdot 10^{-6}$ to 423 and only 7 % of the elements are nonzero.

Separation requirements are needed to avoid severe interference. Three different types of separation requirements are needed.

- *Intracell separation* which is the separation requirement between two frequencies assigned in the same cell. Separation is usually 3 which means at least two frequencies in between are required.
- *Intrasite separation* which is the separation requirement between frequencies assigned to different STRXs in the same site. Separation is usually 2, which means no co-channel or adjacent-channel frequencies are allowed.
- *Intercell separation* which is the separation between two cells not in the

Scenario	Network load (%)	Mobile terminal speed (m/s)
1	30	1
2	60	6

Table 7.2: Network scenarios.

	Number of frequencies								
Scenario	1	2	3	4	5	6	7	8	≥ 9
1	0	-2	-3.2	-4.2	-5.0	-5.6	-5.8	-6.0	-6.0
2	0	-1	-1.6	-2.1	-2.5	-2.8	-2.9	-3.0	-3.0

Table 7.3: Frequency diversity gain in dB.

same site. This type of separation is sometimes present for some STRXs and is usually at most 2.

The separation requirements are described by the separation matrix, which has the form of $N \cdot N$. The values in the diagonal represent the intracell separation while the others are intercell and intrasite separation.

When using FH, two types of gain are introduced, frequency diversity gain and interference diversity gain. In the model these types of gain are summarized in a function $g(|M_i|, |M_j|)$ which takes a positive value less or equal to 1. The gain function $g(|M_i|, |M_j|)$ depends on the number of frequencies in the MALs assigned to both the serving and interfering STRX. The frequency diversity gain is dependent on the speed of the MS and the number of frequencies $|M_j|$ used in the serving cell. The interference diversity gain is dependent on the network load and the number of frequencies in both the serving and the interfering cell. The interference diversity gain comprises of two parts, one directly related to the network load and the other linked to the fractional loading gain. The fractional loading gain is a function of the product of the MAL lengths for the serving cell and interfering cell. The two scenarios we use in the optimization model are shown in Table 7.2 and originates from [143]. Scenario 1 has 30 % load and the MS moves with a speed of 1 m/s while in scenario 2 the load is doubled to 60 % and the speed of the MS is increased to 6 m/s. Frequency diversity gain and interference diversity gain for both scenarios originate from practical experience [127] and simulation results [141, 143, 149]. Table 7.3 and Table 7.4 show the gain factors for the two scenarios. The interference diversity gain factor is linearly interpolated from Table 7.4.

The frequency hopping method used is synthesized hopping. The frequencies are not divided into clusters such as for example 3/9 and 4/12. The TRXs are assumed to have a bandwidth large enough to handle the feasible frequencies in the spectrum, and all TRXs can handle all frequencies as long as the con-

	Product of two MAL lengths				
Scenario	1	4	9	16	≥ 16
1	0	-6.1288	-6.3788	-6.48	-6.48
2	0	-2.72	-2.82	-2.92	-2.92

Table 7.4: Interference diversity gain in dB. (The gain factors for other values are linearly interpolated from the values in the table.)

Solution procedure	SZ (%)	TBL	ML/OL	Iterations	Interference
GR	3	100	7/1	100 000	293.92
IP	3	100	7/1	100 000	293.36
GR	15	3	7/1	1000	399.85
IP	15	3	7/1	1000	396.23
GR	50	7	7/1	1000	420.21
IP	50	7	7/1	1000	416.71

Table 7.5: Comparing simulation results from IP and GR approaches in URBAN instance, scenario 2.

straints to the FAPH are satisfied. The initial solutions were constructed in a greedy manner for the different instances, see Section 6.1. For the real life network URBAN the initial solutions have interference values 364.25 and 527.76 for scenario 1 and scenario 2, respectively. The results obtained from the different solution strategies for the FAPH are compared to the non hopping FAP and 1/1 planning.

7.2 Tabu Search Heuristic Experiments

To evaluate the FAPH model for a FH GSM network, a tabu search heuristic is used. To solve the single STRX linear integer problem which arises in the implementation, an add/drop procedure (AD), greedy procedure (GR) or a commercial optimization software tool is used (IP). The optimization software used is CPLEX 6.6 and the GR and AD algorithms are implemented in Matlab 6.0. Each of the solution procedures can be used stand-alone or be weighted together. To evaluate the influence of the solution procedure parameters, GR and IP approaches are compared, also weights of AD and GR combinations are compared. In Table 7.5 solutions from GR and IP are compared for the URBAN instance.

The IP procedure gives a slight improvement over the GR approach. The two simulations on the top of Table 7.5 with a restricted neighborhood of STRXs

(SZ) of 3 % and tabu list length (TBL) of 100 have been computed with 100 000 iterations for the URBAN instance with scenario 2. The IP is only slightly better, but the simulation time is much longer. The IP procedure needed 37.5 hours and the GR procedure needed 17.8 hours for 100 000 iterations. The other results in Table 7.5 are for 1000 iterations. When SZ is 15 % and TBL is 3, the solution time for the GR procedure is 47 minutes and 105 minutes for the IP procedure. Even if the parameter settings are changed, the differences in interference values are small when comparing the GR and IP solution procedures. In the forthcoming simulations, the greedy solution procedure will be used instead of IP since the computational time of the latter is almost two times longer and the improvement is very moderate. In Figure 7.2, a comparison is made between GR and IP parameter setting with ML/OL=7/1, SZ=15 %, TBL=3 for the URBAN instance and scenario 2. The interference value as a function of the simulation time, shows that GR obtains a good solution faster than the IP procedure. Similar results are obtained also for the other instances.

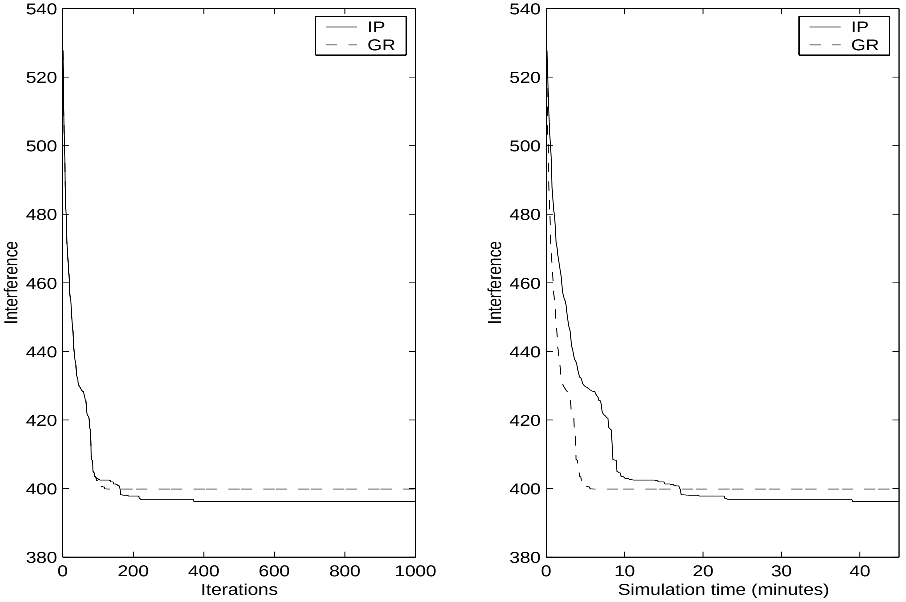

Figure 7.2: Comparision between *GR* and *IP* solution procedure with ML/OL=7/1, SZ=15 %, TBL=3 for the URBAN instance scenario 2.

A comparison between AD and GR solution strategies has also been made. Comparing AD and GR shows that GR leads to better solutions independent of the other parameter values used. When using different weights of the AD and GR parameters, simulations showed that GR and AD gave less interference than weighting the two strategies together. Other simulations made for both scenarios indicate that a GR solution strategy is preferable.

7.2.1 Modify Length or Optimize Length

Using a GR solution procedure the influence of the OL and ML parameters have a significant impact on the solution found. When the OL parameter was large compared to the ML parameter, the simulation often got caught in a local optimum very fast. Even when the OL parameter was of the same size as the ML parameter, it did not force the algorithm out of a local optimum. Experiments indicate that a parameter setting of ML=7 and OL=1 gave the best performance in both scenarios encountered.

7.2.2 Size of the Restricted Neighborhood and Tabu Length

In the restricted neighborhood, the STRXs are chosen randomly. Simulations have been made for different sizes SZ of the restricted neighborhood, from 3% to 100%, (100% corresponds to all STRXs being taken into consideration). Simulations have also been made when the SZ is randomized in every iteration. Different tests show that the interference calculated is high and no good solutions were found. When SZ becomes larger than 40 %, a local optimum is found rather quickly and the algorithm is unable to force the solutions out of the local optimum.

The combination of using a small SZ and a long TBL, results in such behavior that the algorithm does not get caught in a local minimum. Simulating over 100 000 iterations with SZ of 3 % and 5 % and TBL between 50 and 100 decreases the interference values significantly. Interference values under 300 are obtained for the URBAN instance scenario 2 when SZ is as small as 3% and TBL is 100, see Table 7.7. Similar results are obtained for scenario 1, where a SZ of 3 % and TBL of 100 results in an interference of 150.32, see Table 7.6. The initial solution has a MAL of minimum length, the number of frequencies assigned in the MAL equals the number of TRXs in the STRX. The average MAL length ratio for the initial solution is 1. The tabu search simulations assign more frequencies to the MALs and obtain lower interference in the network. The average MAL length ratios have been increased to a value of between 1.170 and 1.225 for scenario 2 and between 1.431 and 1.552 for scenario 1, see Table 7.7 and Table 7.6. The MAL lengths have increased by 17.0 % to 22.5 % for scenario 2 and 43.1 % to 55.2 % for scenario 1. Scenario 1 for which the network load is only 30 % and with a relatively slow MS speed of 1 m/s can use more frequencies in the MALs due to a higher FH gain.

SZ (%)	TBL	Interference	Increase in MAL length	Solution Time (s)
5	50	177.98	1.495	99729
5	65	174.54	1.478	102646
5	75	167.16	1.500	100290
5	100	153.03	1.459	102083
3	50	160.26	1.431	62120
3	65	153.45	1.519	61545
3	75	152.02	1.508	61887
3	100	150.32	1.552	61758

Table 7.6: Interference with a small neighborhood and long tabu list for URBAN instance and Scenario 1. ML=7 and OL=1. The number of iterations is 100 000.

SZ (%)	TBL	Interference	Increase in MAL length	Solution Time (s)
5	50	358.98	1.170	107713
5	65	350.26	1.132	106335
5	75	341.60	1.173	106801
5	100	323.45	1.165	107525
3	50	346.66	1.203	64339
3	65	323.65	1.181	64054
3	75	311.01	1.165	63996
3	100	293.92	1.225	64122

Table 7.7: Interference with a small neighborhood and long tabu list for URBAN instance and Scenario 2. ML=7 and OL=1. The number of iterations is 100 000.

Instance	Scenario	Interference FAPH	Increase in MAL length	Solution Time (s)	Interference FAP
URBAN	1	150.32	1.552	61758	517.61
NORR25	1	12.98	1.888	7317	111.56
NORR85	1	37.41	2.210	21303	393.85
NORR400	1	213.25	2.039	152812	2168.05
URBAN	2	293.92	1.225	60122	517.61
NORR25	2	47.42	1.629	7503	111.56
NORR85	2	144.80	1.881	22060	393.85
NORR400	2	792.42	1.741	157908	2168.05

Table 7.8: Computational results when using tabu search for different networks. The parameter settings are SZ=3%, TBL=100, ML=7 and OL=1.

7.2.3 Tabu Search Results

The best solution found for the non hopping URBAN network has an interference value of 517.61. In this case, the gain from FH is set to zero and the MALs for the STRXs are at minimum length. Performing FH over the MALs of minimum length decreases the interference to 291.32 and 380.35 for scenario 1 and scenario 2 respectively. In Table 7.8, results from FAPH simulations are compared to the non hopping case. The best results for the instances NORR25, NORR85 and NORR400 are obtained when SZ is small (3 %) and TBL is long (100). The lowest interference for the instances NORR25, NORR85 and NORR400 are obtained when the ML parameter has a much higher weight than the OL parameter. The gain of FH is less in scenario 2 since the speed of the MS is faster and the network load is higher.

7.3 Simulated Annealing Experiments

In the tabu search heuristic, a small neighborhood and a long tabu list gave the lowest interference values for both scenarios. This indicates that a simulated annealing approach would be efficient for obtaining low interference values. In the simulated annealing algorithm, the acceptance probabilities P_0 and P_1 are used. P_0 is the probability that a solution, one unit worse than the current, is accepted in the beginning of the simulation. Similarly in the same way, the probability P_1 corresponds to the end of the simulation.

The values of the acceptance probabilities are very important for obtaining low interference values in simulated annealing. In the tabu search heuristic approach, it was found that in order to avoid to get caught in a local optimum, the ML influence was more important than OL. In the simulated annealing simulations, various parameter settings are investigated to see how the different parameters influence the outcome of the final result. Different test problem instances have been used for the computational simulations. The MAL generation/assignment in the simulated annealing algorithm is solved by GR and IP. The difference after 100 000 iterations shows that IP results in only a slightly better interference value than to GR. The IP results take approximately two times longer to obtain than for GR. The IP solution approach is only 3 to 5 percent better compared to the GR solution technique. In the following calculations GR will be used.

7.3.1 Influence of P_0 and P_1 Probabilities

The parameters P_0 and P_1 correspond to temperatures at the beginning and at the end of the simulation. If the cooling is too fast, the likeliness of the solution to move away from a local optimum decreases since the probability of accepting a worse solution decreases too rapidly. The cooling function used, see equation (6.21), has a β parameter to set the rate of the cooling. When $\beta = 0$ a linear cooling function is used, otherwise the cooling function is non linear, see Section 6.4.

The influence of the β parameter during the simulations is if β is too large the simulated annealing algorithm will only explore a tiny fraction of the solution space. If the cooling function is linear the algorithm does not explore the local optimum thoroughly. A β value of 0.00005 is compared to a linear cooling function for the URBAN instance scenario 1 in Figure 7.3. Which β value to choose depends on the instance, scenario and other parameter settings. In Figure 7.4 the cumulative interference is shown for different values of the β parameter. It is clearly shown that when the β parameter is increased, the solution values become less fluctuative. When a linear cooling function is used, the probability that solutions which give a worse cost function value will be accepted is large, this is indicated by the fluctuative cost function values. When the β parameter is large, a local search behavior dominates and a worse interference value is obtained. The best interference value after 100 000 iterations is plotted in Figure 7.4, which corresponds to $\beta = 0.0005$. The β parameter value which yields the best results is different for different ML, OL, P_0 and P_1 parameter combinations.

7.3.2 The Modify Length and Optimize Length Parameters

The other parameters of interest are ML and OL. The simulated annealing approach is less sensitive to changes in these parameters compared to the tabu search solution procedure. Different combinations of ML/OL have been tested and these parameters do not have a strong influence as they do in the tabu

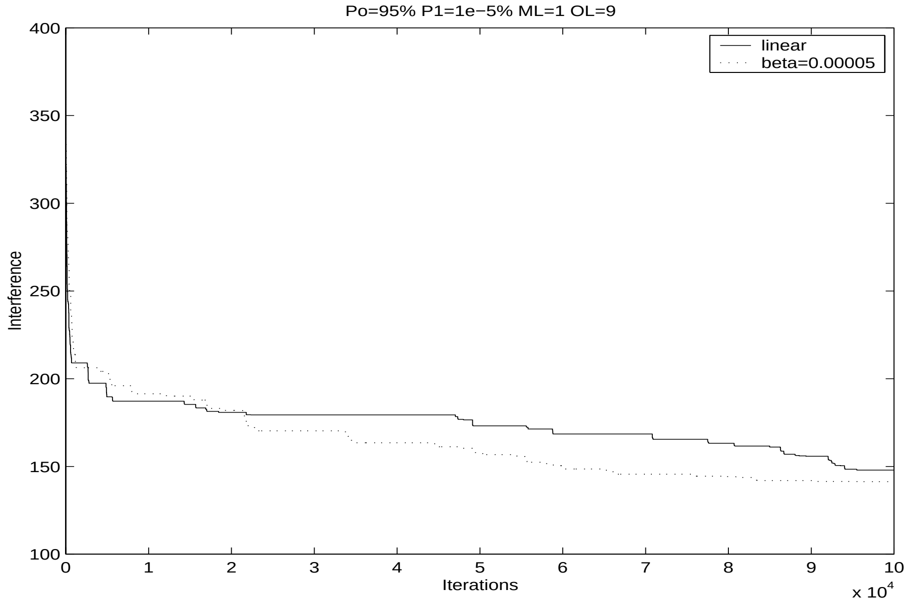

Figure 7.3: The gain when using $\beta = 0.00005$ instead of a linear cooling function for scenario 1. The parameter settings are $P_0 = 95$ %, $P_1 = 10^{-5}$ %, ML=1 and OL=9.

search strategy. The ML parameter is used to force the simulation out of local optimum. In absence of ML (ML/OL=0/1), the algorithm behaves as a local search procedure. When an OL procedure is used, the best MAL is always chosen even if the acceptance probability is high. The solution gets caught in a local optimum, since a solution that is worse is never chosen in the optimization algorithm. An ML setting always chooses a different MAL length than the optimal (the second best). Simulations for 100 000 iterations indicate that the ML/OL settings have only a slight impact on the final solution. It is though important that both parameters are present if a good solution is to be obtained. The ML procedure does not result in good solutions when simulating over too few iterations, since the solution space is not explored enough.

7.3.3 Simulated Annealing Results

The simulations made for the frequency hopping network are compared with these for a non hopping network. To obtain solutions for the FAP, the interference diversity and frequency diversity gains are set to zero. When the ML parameter has a non zero value, the algorithm forces the solution to assign more than the minimum numbers of frequencies in the MALs. During the simulation the solution is infeasible from a non hopping FAP point of view. After the pre-

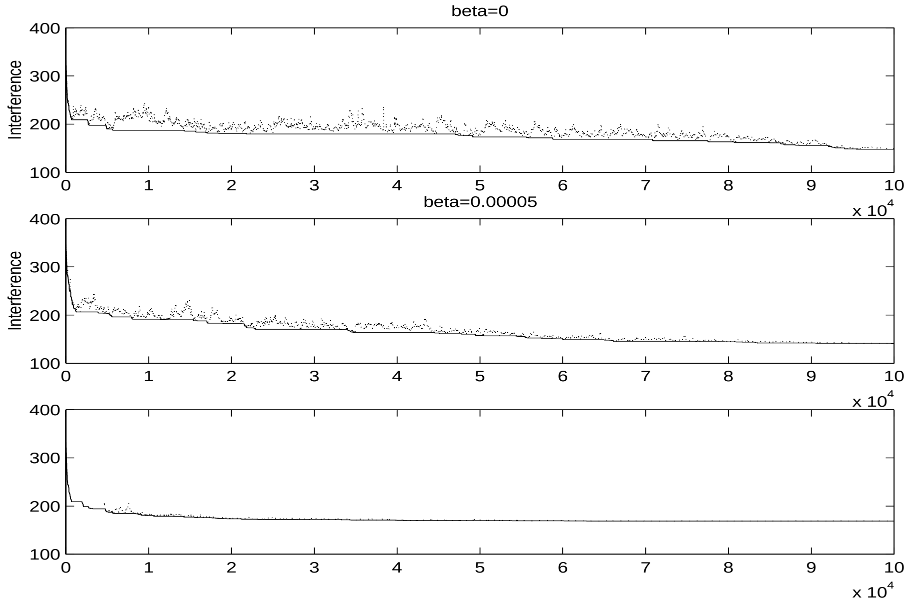

Figure 7.4: The total interference in scenario 1 for $\beta = 0$, $\beta = 0.00005$ and $\beta = 0.001$ for 100 000 iterations. The parameter settings are $ML = 1, OL = 9, P_0 = 95$ % and $P_1 = 10^{-5}$ %.

defined number of iterations the ML parameter is set to zero and a local search is performed. Since the interference and frequency diversity gains are set to zero, the final solution has a MAL with the minimum number of frequencies, a feasible solution to FAP.

In Table 7.9 the solutions found for the different instances with a fixed parameter setting are compared with those found for the non hopping network. The computational experiments for the different instance networks are done over 100 000 iterations. The parameter setting used for the computations are $P_0 = 95\%$, $P_1 = 10^{-3}\%$ and $ML/OL = 1/5$. The cooling function used for the URBAN instance is linear for both the FAPH and FAP. The interference values for the FAPH are much lower than to FAP, because of frequency diversity gain and interference diversity gain. The simulated annealing results for the URBAN instance indicate that interference can be reduced up to approximately 68 % and 40.5 % for scenario 1 and scenario 2 when FH is implemented. The URBAN instance has 55.8 % and 10.2 % more frequencies for scenario 1 and scenario 2 compared to the classic FAP. The reason why the algorithm assigns more frequencies in the MAL for scenario 1 is that the MS speed is higher and the network load is higher for scenario 2, which decreases the gain of frequency hopping. The difference in simulation time between different parameter settings for a specific instance is small.

Scenario 1					
Instance	Interference FAPH	Increase in MAL length	Solution Time (s)	Interference FAP	1/1 planning
URBAN	148.64	1.558	7052	472.44	788.70
NORR25	12.50	1.784	4537	111.28	14.82
NORR85	36.81	2.207	5265	395.07	39.24
NORR400	202.52	2.091	9727	2092.96	207.25
Scenario 2					
Instance	Interference FAPH	Increase in MAL length	Solution Time (s)	Interference FAP	1/1 planning
URBAN	280.96	1.102	7663	472.44	3572.02
NORR25	44.62	1.431	4929	111.28	67.11
NORR85	143.39	1.824	5548	395.07	177.72
NORR400	775.02	1.802	10227	2092.96	938.62

Table 7.9: Computational results with a fixed parameter setting $P_0 = 95\%$, $P_1 = 10^{-3}\%$ and ML/OL=1/5.

Many computational experiments have been made for the real life network URBAN. The best known FAPH results for the URBAN instance after 100 000 iterations is 143.11 for scenario 1 and 267.39 for scenario 2. The increase in the MAL lengths is 1.511 for scenario 1 and 1.124 for scenario 2. The best result is 369.71 if classical FAP is used. Several simulations have been made for the URBAN instance with 1 000 000 iterations, the best known values see Figure 7.5. The computational time is approximately ten times longer. The best interference values for scenario 1 is 128.07 and uses 48.4 % more frequencies in average to the MALs compared to the non hopping FAP. For scenario 2 the best interference found is 237.26 and with MALs with an average increase of 11.3 % in the length. A linear cooling function was used in the computations for 100 000 and 1 000 000 iterations.

7.3.4 1/1-planning and FAP Solutions

An interesting question is how much the interference would decrease when FH is applied to the FAP solution. The best FAP value obtained for the URBAN instance after 100 000 iterations is 369.71. If FH is performed for the MALs calculated in the FAP solution, the interference can be reduced. In this case the MALs are of minimum length, i.e. the number of frequencies in the MAL equals the number of TRXs in the STRX. Performing FH with MALs of minimum length results in a total interference of 208.31 for scenario 1 and 269.14 for scenario 2. Comparing the results of FH with MALs of minimum length, and the solution from the FAPH algorithm, the largest relative decrease is for scenario 1.

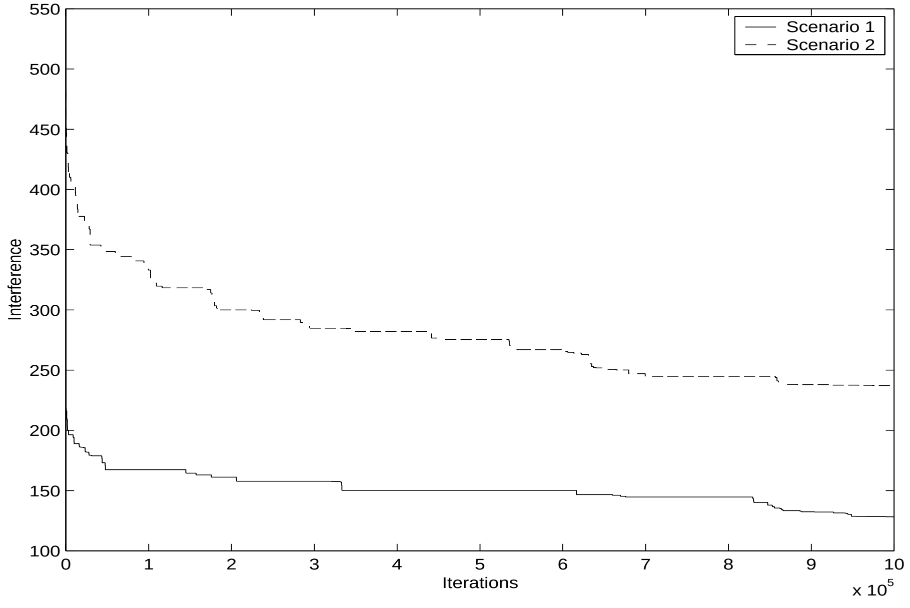

Figure 7.5: Best known interference value for scenario 1 and scenario 2 after 1 000 000 iterations.

For scenario 2, the difference between the FAPH solution and FAP with hopping is small. The benefit of using the FAPH solution is that the increase in MAL length can be used as a capacity expansion, see Section 7.3.5.

1/1-planning is a cell planning strategy with a reuse pattern of 1/1 which means that all frequencies are assigned to all STRXs. Frequency hopping is performed over all frequencies in the cell. The obvious advantage of the strategy is its convenience since it is easy to add more TRXs in a cell when more capacity is desired. No replanning of the frequencies is necessary since the present plan is valid. All the MALs contain the maximum number of frequencies and now a probability exists that two adjacent cells simultaneously transmit on the same frequency. The interference would decrease when the length of the MALs increases. The URBAN instance has 29 TCH frequencies and the maximum MAL length for the 1/1 planning strategy is 29. The total interference value is very high. A comparison between 1/1-planning and FAPH is shown in Table 7.9.

7.3.5 Capacity Expansion

When using the FAPH algorithm several MALs for the STRXs have a relative increase greater than 1. This means that the number of frequencies assigned in the MAL is larger than the number of TRXs in the STRX. This fact can be used to increase the capacity in the network since additional TRXs can be added while keeping the solution feasible. The increase in capacity is not free since the cost of additional TRXs and network modifications must be taken into account. When more TRXs are added to the network, the interference will increase. To evaluate the capacity expansion, the best results found for 100 000 iterations are used. The capacity is increased by setting the total number of TRXs in a cell equal to the MAL length. The capacity expansion is only computed for the URBAN instance. In scenario 1, 193 additional TRXs are added to the 364 existing ones, which gives a capacity expansion of 53 %. The interference increased from 143.11 to 340.40. For scenario 2, 53 TRXs are added. The increased capacity is approximately 14% with the interference increasing from 267.39 to 310.34.

7.4 Comparing Tabu Search and Simulated Annealing Results

To compare the solutions between tabu search and simulated annealing, the URBAN instance is used. In Figure 7.6 the best known tabu search results from scenario 1 and scenario 2 (SZ=3%, TBL=100, ML/OL=7/1) are compared to simulated annealing results ($P_0 = 95\%$, $P_1 = 10^{-3}\%$ and ML/OL=1/5). The total solution time for 100 000 iterations for simulated annealing is 1.96 h and 2.13 h for scenario 1 and scenario 2 respectively. The tabu search solution is truncated to exactly match the solution times for simulated annealing. In scenario 1, simulated annealing has a lower interference value after approximately 0.3 h. For scenario 2, simulated annealing take approximately 0.6 h to obtain lower interference. Comparing several tabu search and simulated annealing computations, the same conclusion can be drawn. For a fixed solution time simulated annealing has lower interference values than tabu search.

7.5 Conclusion

The main issue of this part of the thesis is how frequency hopping can lower the interference in the network. By admitting a fully flexible frequency assignment method with no frequency groups, large improvements in interference can be achieved using frequency hopping. For the frequency hopping assignment problem, an integer model is developed. Instead of assigning different frequencies to each TRX, a list of frequencies is assigned to the cell, a MAL. The TRXs in each cell are grouped into a STRX. The goal is now to assign MALs to the STRXs in the network with the purpose of minimizing the co-channel and adjacent-channel interference.

Different heuristic solution strategies are investigated such as tabu search and

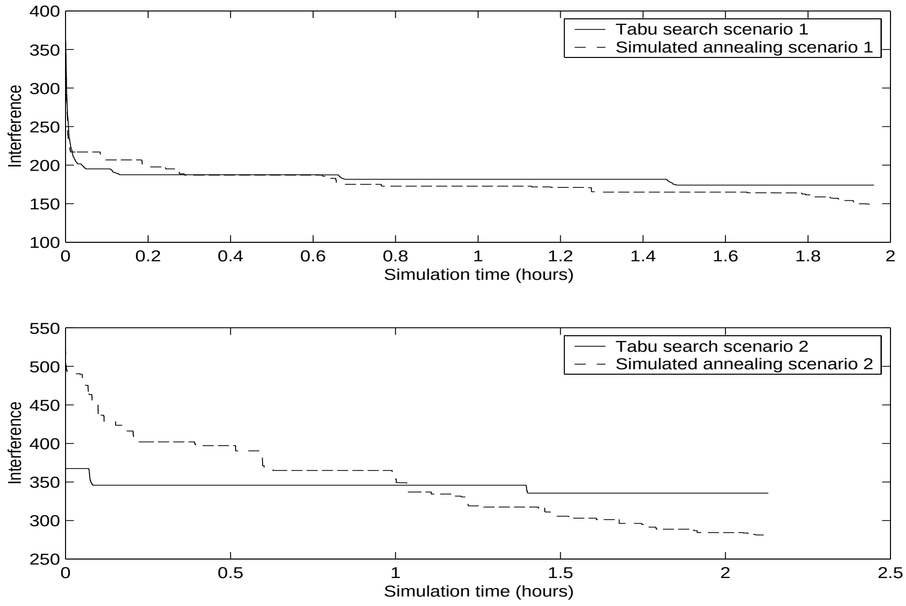

Figure 7.6: Comparison between tabu search and simulated annealing for the URBAN instance. The tabu search has the parameter values ML/OL=7/1 and SZ=3 % and TBL=100. The simulated annealing has the parameter values ML/OL=1/5, $P_0 = 95\%$, $P_1 = 10^{-3}\%$ for scenario 1 and scenario 2. The cooling function of the simulated annealing algorithm is linear.

simulated annealing. The simulations obtained from these two different heuristic solution strategies show that frequency hopping reduces the interference in the network. Comparing tabu search and simulated annealing strategies it is shown that a good solution can be obtained more rapidly with the use of the latter method. In the tabu search heuristic a small neighborhood size and long tabu list gives the best results. In the simulated annealing approach the cooling function had the largest impact on how good the solution would be. Which cooling function to use depends on the different instances, parameter settings and scenarios. The results from the FAPH are compared with the 1/1-planning strategy. 1/1-planning which has reuse one, imply to assign all available frequencies to all STRXs. The interference is very high for the URBAN instance when 1/1-planning is used even if the hopping is performed over all frequencies. Despite the interference averaging over all frequencies, severe interference occurs.

Part III

Resource Optimization in WCDMA Networks

Chapter 8

Scheduling the Downlink Shared Channel

In WCDMA there are two main categories of transport channels, common channels and dedicated channels. The dedicated transport channels, are what the name reveals, channels directed or dedicated for a specific mobile terminal. The common transport channels are used by several mobile terminals simultaneously. The channel used for data transfer and speech in the downlink is the downlink Dedicated Channel (DCH). The DCH does not offer flexible fast data transfer so an additional channel is the Downlink Shared Channel (DSCH). DSCH is a channel intended for packet data in the downlink. A high speed data channel in the downlink also exists, High Speed Downlink Packet Access (HSDPA) which will be discussed in Chapter 9.

8.1 Overview of DSCH

The DSCH is intended to carry dedicated user data. The DSCH is always associated with the downlink DCH. The reason that the DSCH needs to be associated to a DCH is to ensure the availability of power control [67]. The DCH also provides information to the mobile terminal when it has to decode the DSCH and which spreading code to decode it with. Switching the bit rate in the downlink is a little bit more complex compared to the uplink. In the uplink the Node B can detect which SF a mobile terminal is using and change SF without informing the Node B beforehand. The mobile terminal does not have that feature, which results in that the Node B cannot change the SF in the downlink without informing the mobile terminal. The mobile terminal must prepare so the correct SF is applied in the mobile terminal before the data arrives.

The DSCH uses OVSF, which can be changed in powers of two between 4 and 256. The DSCH capacity can be shared among several users on a frame-by-frame basis. Each frame has a duration of 10 ms. The transmission of high peak rate and low duty cycle data in a DSCH channel does not lead to low utilization as with the use of DCH channels. The reason is that the DSCH has a short Time Transmission Interval (TTI) of 10ms. That allows several users to be active simultaneously using lower data rates or a few users to be active with high data rates. The DSCH can handle as many as three channelization codes with top SF 4 to provide 2.8 Mbps on the physical channel. A good overview of the DSCH can be found in [67] and for more detailed analysis see [2, 3].

Bandwidth allocation and related issues in WCDMA systems have received much attention during recent years. A simple dynamic channel allocation algorithm is proposed and compared to a fixed channel allocation strategy in [44, 88]. In [61], the authors present a dynamic resource scheduling mechanism that adopts closed loop power control, and minimizes the power needed to satisfy QoS-constraints. An extension of this mechanism is discussed in [62]. The authors of [93] examine, from a game theory perspective, bandwidth allocation with continuous utility functions. The authors of [11, 94] consider bandwidth allocation models similar to the one discussed in this chapter of the thesis. Their models, however, do not explicitly address the power control mechanism in WCDMA networks. Assignment of OVSF codes in the downlink is investigated in [45]. In [63] a first approach with downlink bandwidth allocation and dynamic programming is presented. The model presented does not consider a limited amount of codes. Although the focus in this chapter is downlink bandwidth allocation, similar issues exist for the uplink direction. Resource allocation and power control for the uplink have been studied in, for example, [83, 106, 108, 109].

8.2 Signal to Interference Ratio in the Downlink

The power of the Node B is a limited resource and is shared by all users in the cell. If a mobile terminal downloads a session which requires large bandwidth, a large part of the total power of the Node B is used. This means that only a small number of users are allowed to use high bit rates. In the formulation we use $M = \{1, \ldots, |M|\}$ to denote the set of users connected to the serving cell. The transmission rates are characterized by a discrete set. Let the index set of the physical channel bit rates be $R = \{1, \ldots, |R|\}$, and R_j the bit rate of $j, j \in R$. The SIR for a user i in a cell with Node B n can be formulated as

$$SIR_{in} = \frac{W}{R_j} \frac{p_{in} g_{in}}{\sum\limits_{m \neq n} (P_m g_{im}) + \alpha_i g_{in} P_n + N_0}, \tag{8.1}$$

where the power p_{in} is the power from Node B n to mobile terminal i. The path loss between mobile terminal i and Node B n is described by g_{in}. P_m is the power from interfering Node B m and α_i is the orthogonality of the channel for user i. Since WCDMA uses orthogonal channelization codes in the downlink to separate users, the channels should be orthogonal at the receiver. But multipath propagation yields that the different channels may not be orthogonal. The orthogonality parameter α_i varies between 0 and 1, and a value of 1 means that all of Node B n power is treated as interference. W is the chip rate and N_0 is the thermal noise power.

To support rate j of user i, the closed loop power control mechanism assigns an amount of power to user i, such that the SIR defined in equation (8.1) meets a threshold γ_i, that is,

$$\frac{W}{R_j}\frac{p_{in}g_{in}}{\sum\limits_{m\neq n}(P_m g_{im}) + \alpha_i g_{in} P_n + N_0} \geq \gamma_i. \tag{8.2}$$

The threshold γ_i is dependent on the type of service, multipath profile and terminal speed [67]. The SIR requirement for data services is in the range of 2 dB to 5 dB [62, 84]. In other words, to allocate rate j to user i, the tranmission power, denoted by p_{ij}, is given by the following equation:

$$\begin{aligned}
p_{ij} &= \frac{\gamma_i\left(\sum\limits_{m\neq n}(P_m g_{im}) + \alpha_i g_{in} P_n + N_0\right)}{g_{in}\dfrac{W}{R_j}} && (8.3)\\
&= \frac{\gamma_i\alpha_i}{\dfrac{W}{R_j}}P_n + \frac{\gamma_i\left(\sum\limits_{m\neq n}(P_m g_{im}) + N_0\right)}{g_{in}\dfrac{W}{R_j}} && (8.4)\\
&= \beta_{ij}^1 P_n + \beta_{ij}^2, && (8.5)
\end{aligned}$$

where

$$\beta_{ij}^1 = \frac{\gamma_i\alpha_i}{\dfrac{W}{R_j}} \tag{8.6}$$

and

$$\beta_{ij}^2 = \frac{\gamma_i \left(\sum_{m \neq n} (P_m g_{im}) + N_0 \right)}{g_{in} \frac{W}{R_j}}. \tag{8.7}$$

8.3 Optimization Model for the DSCH

The goal of bandwidth allocation is to assign each user a transmission rate, with the objective of maximizing network performance in terms of the service to the users. It is possible for the system to deny any downlink bandwidth to a user.

In the subsequent text, the term utility is used for the degree of service offered by the network. For user i, the utility of assigning transmission rate j is denoted by u_{ij}. One simple example of the utility function is the transmission rate (bandwidth) itself, i.e. $u_{ij} = R_j, \forall i, \forall j$. This utility function models the efficiency in utilizing the power and code resources, or, equivalently, the throughput. Other utility functions can be used to, for example, model both resource utilization and user fairness [152]. The bandwidth allocation problem amounts to maximizing the total system utility, subject to the constraints that every user is assigned at most one of the transmission rates, and that the total power and channelization codes used in every cell does not exceed the limitations. In the following we do not use the index n for the different Node Bs since we consider only one cell for scheduling and treat the other cells as interferers. The cell that will be treated in the optimization model is denoted *serving cell* and will have a power limit for the DSCH of P^{max} W.

The optimization model has the following variables

$$x_{ij} = \begin{cases} 1 & \text{if bit rate } j \text{ is assigned to user } i \\ 0 & \text{otherwise.} \end{cases}$$

$$p_{ij} = \text{power needed if bit rate } j \text{ is assigned to user } i.$$

The model uses the following parameters

$$u_{ij} = \text{utility for user } i \text{ if bit rate } j \text{ is assigned.}$$

The bandwidth allocation problem (BAP) can be formulated as follows

$$
\begin{array}{llcll}
[BAP] & & & & \\
\max & \sum_{i\in M}\sum_{j\in R} u_{ij}x_{ij} & & & (8.8)\\
\text{s.t.} & \sum_{i\in M}\sum_{j\in R} p_{ij} & \leq & P^{max} & (8.9)\\
& \sum_{i\in M}\sum_{j\in R} 2^{1-j}x_{ij} & \leq & C^{max} & (8.10)\\
& \sum_{j\in R} x_{ij} & \leq & 1, \quad \forall i & (8.11)\\
& p_{ij} & = & (\beta^1_{ij}P+\beta^2_{ij})x_{ij}, \quad \forall i,j & (8.12)\\
& x_{ij} & = & 0/1, \quad \forall i,j & (8.13)\\
& p_{ij} & \in & R^+, \quad \forall i,j & (8.14)
\end{array}
$$

Constraint (8.9) states that the total available DSCH power P^{max} for the serving cell is not allowed to be exceeded. The constraint (8.10) handles that the codes assigned maintain their orthogonality. C^{max} is the number of code trees with a top node corresponding to SF=4 (minimum SF for the DSCH). Constraint (8.11) states that a user is assigned at most one bit rate level. Constraint (8.12) refers to the the SIR equation (8.1) and describes how the power assigned to the mobile terminals depend on the bit rate. The total Node B power P for the serving cell is set to a constant value.

Combining constraints (8.9) and (8.12) gives

$$\sum_{i\in M}\sum_{j\in R}(\beta^1_{ij}P+\beta^2_{ij})x_{ij} \quad \leq \quad P^{max}. \tag{8.15}$$

Letting $a_{ij} = \beta^1_{ij}P + \beta^2_{ij}$ and $b_j = 2^{1-j}$ the bandwidth allocation problem for the serving cell can be formulated using the following model.

$$
\begin{array}{llcll}
\max & \sum_{i\in M}\sum_{j\in R} u_{ij}x_{ij} & & & (8.16)\\
\text{s.t.} & \sum_{i\in M}\sum_{j\in R} a_{ij}x_{ij} & \leq & P^{max} & (8.17)\\
& \sum_{i\in M}\sum_{j\in R} b_j x_{ij} & \leq & C^{max} & (8.18)\\
& \sum_{j\in R} x_{ij} & \leq & 1, \quad \forall i & (8.19)\\
& x_{ij} & = & 0/1, \quad \forall i,j. & (8.20)
\end{array}
$$

The reformulated BAP model is a multiple choice knapsack problem [102, 116].

8.4 Dynamic Programming

Knapsack problems can be solved using general integer programming methods (e.g., linear programming relaxation and branch-and-bound enumeration). For these methods, however, it is very hard to predict the computational complexity a priori.

Dynamic programming is a well-known type of algorithm for many combinatorial problems in computer science. This solution technique is applicable to problems with the property that parts of an optimal solution are themselves optimal. Finding the optimal solution for such a problem can be performed recursively by computing the optimum to a sequence of similar problems. Dynamic programming offers an effective approach for knapsack problems when the data size is restricted. We refer to [102, 116] and the references therein for a detailed treatment of dynamic programming applied to knapsack problems.

The constraint (8.18) has a discrete right-hand-side C^{max} which can take the integer values of 1,2 or 3. The number of codes available is also integral and the constraint can be easily converted to fit the dynamic programming solution structure. We scale both sides with 64, since the largest SF which can be used is 256. The coefficient b_j for SF=256 is 1/64. To apply dynamic programming all coefficients in the constraint need to be integer values. Scaling the coefficients by 64 gives the coefficient for SF=4 equal to 64, SF=8 equal to 32, ..., SF=256 equal to one. The coefficient now becomes $\bar{b}_j = 64b_j$ and right-hand side becomes $\bar{C}^{max} = 64C^{max}$ which gives the constraint

$$\sum_{i \in M} \sum_{j \in R} \bar{b}_j x_{ij} \leq \bar{C}^{max}. \tag{8.21}$$

But still dynamic programming cannot be applied directly, because the left-hand-side coefficients of (8.17), a_{ij}, are fractional-valued. We therefore consider an approximation of the problem, in which these coefficients are quantized to values of a discrete set. With a sufficiently large number of quantization steps, the approximation becomes equivalent to the problem with the original coefficients. Let Q denote the number of quantization steps. In the quantization process, the inequality (8.17) is multiplied by Q at both sides, and thus becomes

$$\sum_{i \in M} \sum_{j \in R} Q a_{ij} x_{ij} \leq Q P^{max}. \tag{8.22}$$

The coefficients of this inequality are then rounded to integer numbers. To ensure that any feasible solution to the quantized problem is also feasible to the original problem, QP^{max} is rounded downwards, whereas all Qa_{ij} are rounded upwards. In other words, (8.17) is replaced by the following inequality after quantization,

$$\sum_{i \in M} \sum_{j \in R} \bar{a}_{ij} x_{ij} \leq \bar{P}^{max}, \tag{8.23}$$

where $\bar{a}_{ij} = \lceil Q a_{ij} \rceil$, and $\bar{P}^{max} = \lfloor Q P^{max} \rfloor$.

To apply dynamic programming to the quantized problem, think of the right-hand side of (8.21) and (8.23) as two different sets of states $\{0, 1 \ldots, \bar{C}^{max}\}$ and $\{0, 1 \ldots, \bar{P}^{max}\}$. The subset of users $\{x_1, \ldots, x_i\}$ defining a sequence of stages for $i = 1, \ldots, |M|$. To ease the presentation, we introduce an artifical user with index 0, for which $\bar{a}_{ij} = 0$ and $\bar{b}_j = 0, \forall j$.

Let $U_i(\lambda_1, \lambda_2)$ denote the optimal utility of allocating bandwidth to the first i users using a maximum of λ_1 energy units and λ_2 code units. The problem can be formulated as

$$\begin{aligned}
U_i(\lambda_1, \lambda_2) = \quad & \max \sum_{m=1}^{i} \sum_{j \in R} u_{mj} x_{mj} && && (8.24) \\
\text{s.t.} \quad & \sum_{m=1}^{i} \sum_{j \in R} \bar{a}_{mj} x_{mj} \leq \lambda_1 && && (8.25) \\
& \sum_{m=1}^{i} \sum_{j \in R} \bar{b}_j x_{mj} \leq \lambda_2 && && (8.26) \\
& \sum_{j \in R} x_{mj} \leq 1, && m \in 1, \ldots, i && (8.27) \\
& x_{mj} = 0/1, && m \in 1, \ldots, i, \forall j \in R. && (8.28)
\end{aligned}$$

We define $U_0(\lambda_1, \lambda_2) = 0, \lambda_1 = 1, \ldots, \bar{P}^{max}$ and $\lambda_2 = 1, \ldots, \bar{C}^{max}$. It is clear that $U_{|M|}(\bar{P}^{max}, \bar{C}^{max})$ gives us the optimum to the quantized problem. We need however a recursion to define $U_{|M|}(\bar{P}^{max}, \bar{C}^{max})$ using values of $U_i(\lambda_1, \lambda_2)$ for $i < |M|$ and $\lambda_1 < \bar{P}^{max}$, $\lambda_2 < \bar{C}^{max}$.

Consider the problem of finding $U_i(\lambda_1, \lambda_2)$ with $i \geq 1$. In the optimal solution to this problem, user i is either allocated a rate of zero, or one of the positive rates. In the former case, $U_i(\lambda_1, \lambda_2) = U_{i-1}(\lambda_1, \lambda_2)$, whereas in the latter case $U_i(\lambda_1, \lambda_2) = \max_{j \in R}(u_{ij} + U_{i-1}(\lambda_1 - \bar{a}_{ij}, \lambda_2 - \bar{b}_j))$. This leads us to the following recursion:

$$U_i(\lambda_1,\lambda_2) = \max\{U_{i-1}(\lambda_1,\lambda_2), \max_{j\in R:\lambda_1-\bar{a}_{ij}\geq 0,\lambda_2-\bar{b}_j\geq 0}(u_{ij}+U_{j-1}(\lambda_1-\bar{a}_{ij},\lambda_2-\bar{b}_j))\} \tag{8.29}$$

Starting from $U_0(\lambda_1,\lambda_2) = 0$ for $0 \leq \lambda_1 \leq \bar{P}^{max}$, $0 \leq \lambda_2 \leq \bar{C}^{max}$ we can use the above recursion to calculate $U_1(\lambda_1,\lambda_2),\ldots,U_{|M|}(\lambda_1,\lambda_2)$ for all possible values of λ_1 and λ_2. This process can be visualized in the following table, in which the rows are the users, the columns are the power levels, and the entries are the corresponding utility values. Since we have two constraints to consider in the dynamic programming procedure we have $\bar{C}^{max}$ tables.

	0	$\ldots$	$\lambda_1-\bar{a}_{ij}$	$\ldots$	λ_1	$\ldots$	$\bar{P}^{max}$
0	0	$\ldots$	0	$\ldots$	0	$\ldots$	0
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
$i-1$	0	$\ldots$	$U_{i-1}(\lambda_1-\bar{a}_{ij},\lambda_2)$	$\ldots$	$U_{i-1}(\lambda_1,\lambda_2)$	$\ldots$	$U_{i-1}(\bar{P}^{max},\lambda_2)$
i	0	$\ldots$	$U_i(\lambda_1-\bar{a}_{ij},\lambda_2)$	$\ldots$	$U_i(\lambda_1,\lambda_2)$	$\ldots$	$U_i(\bar{P}^{max},\lambda_2)$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
$\lvert M\rvert$	0	$\ldots$	$U_{\lvert M\rvert}(\lambda_1-\bar{a}_{ij},\lambda_2)$	$\ldots$	$U_{\lvert M\rvert}(\lambda_1,\lambda_2)$	$\ldots$	$U_{\lvert M\rvert}(\bar{P}^{max},\lambda_2)$

In the table above, an entry in row i can be calculated using appropriate entries in row $i-1$ together with the values of $\bar{a}_{ij}$, u_{ij}, $j \in R$ and λ_2. Note that it is not necessary to compute all the entries of the above table, as some power levels do not match any feasible solution. The overall optimal utility is found by taking the maximum of the last row of the table when $\lambda_2 = \bar{C}^{max}$. To trace the optimal solution itself, it is necessary to store the optimal decisions (transmission rates) used to generate the utility values in a second table. Starting from row $|M|$ and the maximum-utility entry, the rate generating this entry is the optimal rate for user $|M|$, and the optimal utility of the previous row can be found. Continuing in this fashion back to user one gives the optimal bandwidth allocation for all users.

It can be realized that the dynamic programming algorithm has a worst-case complexity of $\mathcal{O}(\bar{P}^{max}|R||M|\bar{C}^{max})$ or $\mathcal{O}(\lfloor QP^{max}\rfloor|R||M|\bar{C}^{max})$. Therefore, given a power limit, fixed number of quantization steps and the number of levels in the code tree, the algorithm is guaranteed to run in polynomial time in the problem size - a very attractive property for algorithms for real-time applications. Moreover, the complexity formula means that we can trade the computational complexity against solution quality by varying the number of quantization steps.

8.5 Numerical Results

The WCDMA network used in our numerical experiments consists of 21 cells and 1050 active users. We have 30 DSCH users in the serving cell to be scheduled. The amount of power available for the DSCH is $P^{max} = 14$ W of maximal Node B power of 20 W. The remaining power in the serving cell is used for dedicated channels and common control channels. Since we only consider one cell at a time for scheduling, the other cells in the network are treated as interferers that transmit with 20 W. Due to multipath propagation the users in a cell will experience intra-cell interference. The SIR threshold γ_i is set to 3 dB and the thermal noise is 10^{-13} W. The chip rate is set to 3.84 Mcps and the spreading factors available are in the range from 4 to 256. The parameter $C^{max} = 3$ which corresponds to the maximum number of code trees with top node SF=4 that the DSCH can handle. The bit rates on the physical channel are determined by the SF and for the SF {4 8 16 32 64 128 256} we have the following bit rates {960 480 240 120 60 30 15} kbps. Since the DSCH can be scheduled every 10 ms (100 times per second) we assign a number of bits which is 1 % of the physical channel bit rates. The utility function in the calculations represents the number of bits transmitted in the TTI, $u_{ij} = R_j, \forall j \in R, \forall i \in M$. We assume that the Node B data buffers are full and the data buffers in the mobile terminal are unlimited in size, which means that the users can take full benefit of the highest bit rate in the set R.

We used an optimization package [46, 71] (AMPL/CPLEX) to find the optimum of our bandwidth allocation problem with its original (fractional) coefficients. We then applied the dynamic programming algorithm with the quantization step $Q = 100, 200, \ldots$, until the algorithm finds the same optimal solution. The results are illustrated in Figure 8.1. The figure shows, on a cell-by-cell basis, the optimal utility in transmitted bits and the number of quantization steps. The maximum utility is 14550 transmitted bits for cell 13 and the least is 10050 bits for cell 11 during a TTI of 10 ms. In addition, except for cell 9,15,18 and 21, the optimal utility is found with less than 1000 quantization steps. The cell that needed most quantization steps is cell 21 with 1700 steps to reach the optimal value.

The computational complexity can be traded against solution optimality in the dynamic programming algorithm. We show this trade-off for cell 18 in Figure 8.2, where the cell utility for various values of Q is plotted. The maximal possible utility (optimum to the problem with the original coefficients) is shown using a dotted line. We observe that most of the maximum possible utility is obtained for several hundreds quantization steps. Even for a low granularity with as few as 100 steps, the solution is only a few percent from the optimum.

Because the total utility equals the sum of the transmitted bits on the physical channel assigned to the users in the TTI, it can be expected that, at optimum,

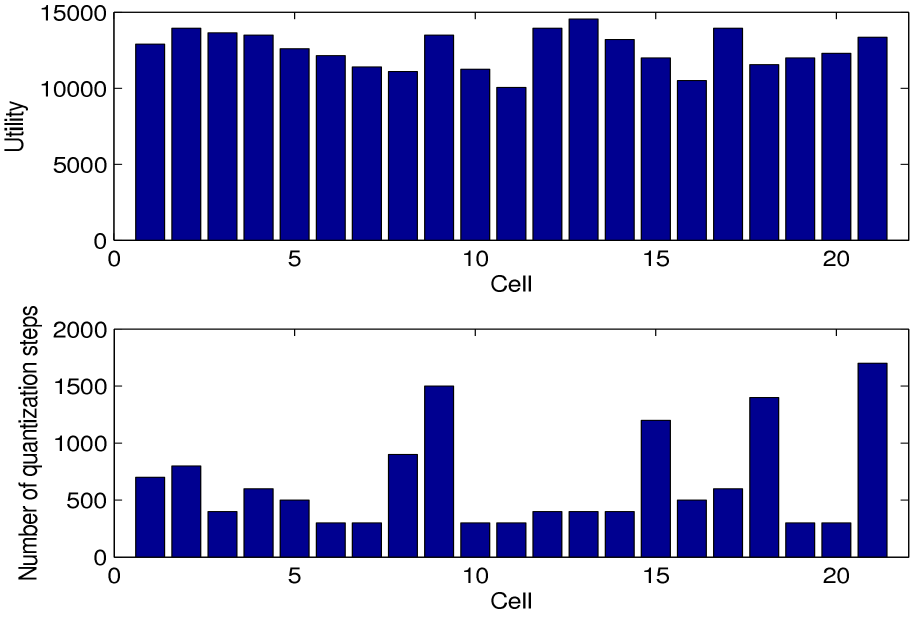

Figure 8.1: Optimal cell utility with associated number of quantization steps.

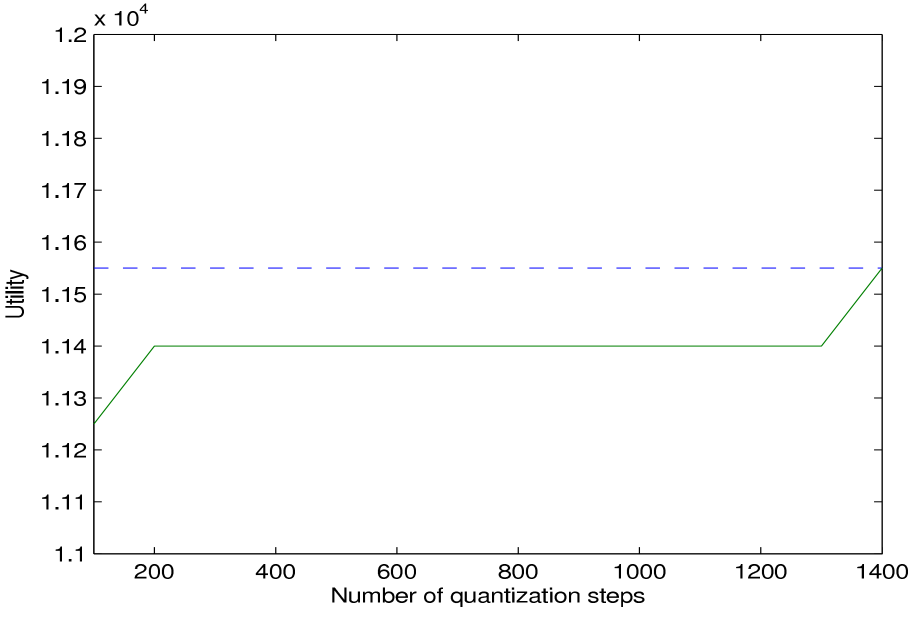

Figure 8.2: Cell utility and the number of quantization steps.

the system allocates power to a few users (those with a good channel condition), while denies bandwidth to the others. This is examined in Figure 8.3, which shows the transmission rate of the users in cell 18. Among 30 users in the cell, only 4 are allocated a positive rate. Thus users experiencing high path loss and interference are penalized in bandwidth allocation, if the utility function is solely based on the efficiency in resource utilization (throughput). If, however, the utility function also includes terms that reflect user fairness, we may expect a more even bandwidth allocation among the users.

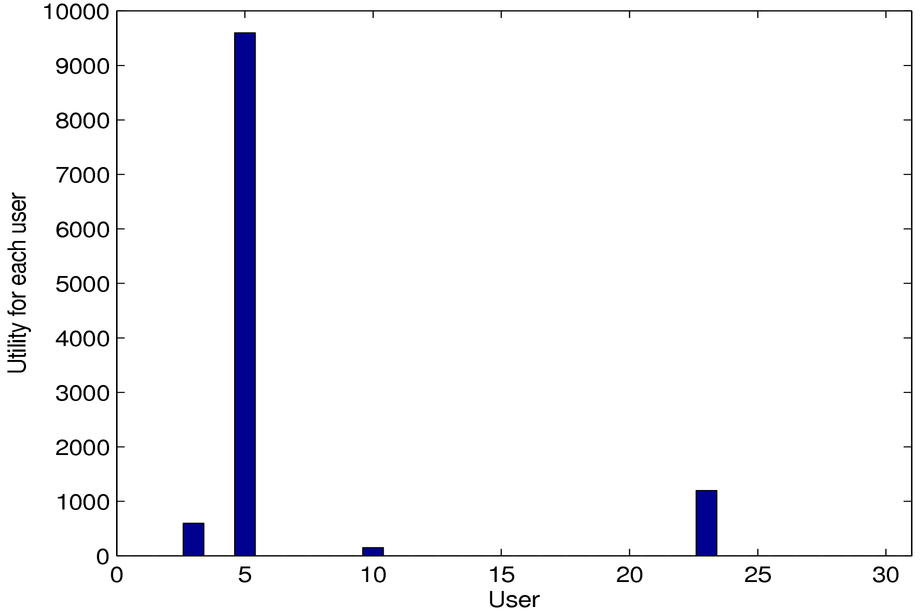

Figure 8.3: Cell utility for the users in cell 18.

8.6 Conclusions

In this chapter an optimization model and a dynamic programming solution technique for bandwidth allocation in the DSCH for a WCDMA network are presented. The bandwidth allocation problem is formulated using one multiple-choice knapsack model per cell. The objective is to maximize a utility function which corresponds to the bit rate. The model takes intra- as well as inter-cell interference into account. The solution method for the bandwidth allocation problem is dynamic programming. The dynamic programming algorithm has a pseudo-polynomial complexity. The predictability in performance makes the algorithm suitable for being used in an on-line (or closely on-line) environment. The model has two contraints that must be considered in the dynamic programming solution approach, power limitation constraint and code limitation constraint. The power constraint has fractional coefficients and a continuous right-hand side, therefore quantization is needed. The number of quantization steps needed for solving the optimization model to optimality with dynamic programming is for the most cells less than 1000. The cell which required most quantization steps to reach an optimal solution needed 1700 steps. Since the utility function used considers only the bit rates assigned, resources were assigned to users with the best radio environment.

Chapter 9

Scheduling the High Speed Downlink Packet Access Channel

This chapter presents an optimization model for the HSDPA channel. The problem considered is to assign channelization codes and bit rates to users in TTIs. The model also considers the effect of code multiplexing that means that more than one user is assigned resources in a TTI. Two simple greedy algorithms are used for solving the optimization problem.

9.1 Overview of the HSDPA Concept

The HSDPA concept has been designed to improve the downlink data throughput. The idea behind the HSDPA is to extend the WCDMA system with a channel that is flexible and can be scheduled fast. The HSDPA concept does not use OVSF and power control, instead adaptive modulation and coding and hybrid automatic repeat request procedures are introduced. A description of HSDPA can be found in [5, 6, 67, 112].

The HSDPA concept requires three new channels in the physical layer structure. The new channels are High Speed Downlink Shared Channel (HS-DSCH), High Speed Shared Control Channel (HS-SCCH) and in the uplink the High Speed Dedicated Physical Control Channel (HS-DPCCH).

The HS-DSCH has a TTI of 2 ms that provides a short communication time between the Node B and mobile terminal during scheduling. The SF is fixed for HSDPA and is set to 16 for the HS-DSCH. The maximum number of codes that can be allocated to a user is 15, but depending on the mobile terminal capability, individual terminals can receive 5, 10 or 15 codes. The total number of channelization codes with SF 16 is 16 but one branch is reserved for HS-SCCH and Release '99 common and dedicated channels, see Figure 9.1. The multicode transmission increases the bit rate by splitting the bits of a transmission block to the channelization codes in a TTI. A data stream with a high bit rate is divided into a number of lower data rate streams. All these streams are transmitted in parallel and synchronized to avoid time delay between each other. In HS-DSCH the OVSF is replaced by Adaptive Modulation and Coding (AMC). The HS-DSCH can change modulation and coding as fast as every 2 ms. Except for the original Quadrature Phase Shift Keying (QPSK) modulation scheme, the HS-DSCH can also handle 16 Quadrature Amplitude Modulation (QAM). In AMC coding and modulation are selected from a defined set. The channel condition is measured at the receiver and is signaled back to Node B. In bad channel conditions, a lower modulation order and/or a lower coding rate is selected. The HSDPA concept also introduces Hybrid Automatic Repeat Request (HARQ) procedure that allows the mobile terminal to request retransmission of erroneously received transport blocks. In HSDPA, retransmissions can be controlled in the Node B instead of the RNC, leading to shorter delay with packet data operation.

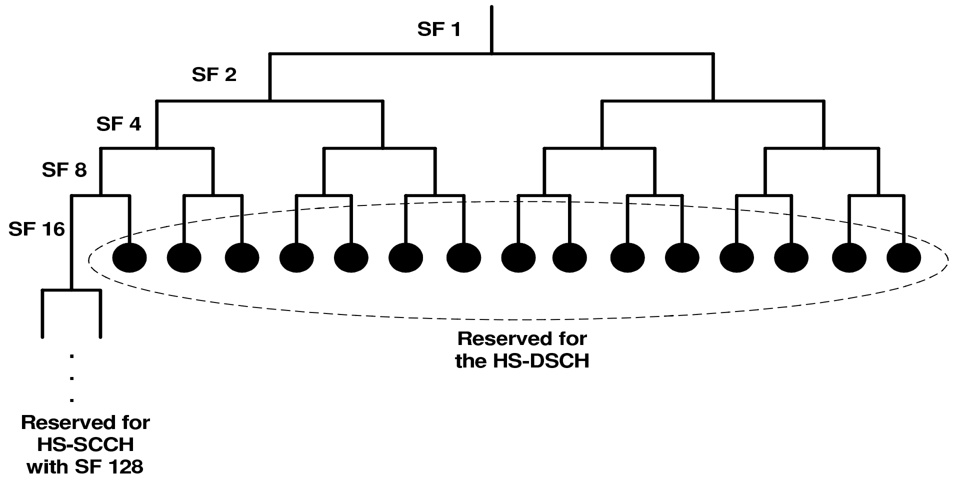

Figure 9.1: The code tree for the HS-SCCH and HS-SCCH.

The HS-SCCH carries information necessary for HS-DSCH modulation. The bits transmitted on the control channel are information about which codes to despread, mobile terminal category, and also validity information for the HARQ process. The uplink HS-DPCCH carries the acknowledge information for the physical layer retransmission and also the quality feedback information to be used in the Node B scheduler to determine to which mobile terminal to transmit and at which data rate.

Compared to a mobile terminal that only handles the 3GPP Release '99 packet data channels, an HS-DSCH mobile terminal must contain processing for the HARQ, multicode processing and also ability to handle the signaling channels HS-SCCH and HS-DPCCH. The mobile terminals that support HSDPA are divided into 12 different categories with maximum data rates from 0.9 to 14.4 Mbps, see [4]. The idea behind the categories is to allow different levels of terminal complexity. 10 of the mobile terminal categories support both 16 QAM and QPSK and 2 categories supports only QPSK. There is also a difference in HARQ buffer sizes between some categories. Another difference is the minimum inter-TTI, which determines how often a mobile terminal can be scheduled. The different minimum inter-TTIs are 1,2 and 3, e.g. an inter-TTI of 3 means that the mobile terminal can be scheduled at most every third TTI. The HSDPA concept does allow several users to share the code resource during a TTI, i.e. code multiplexing. If several users are sharing the codes in a TTI, one HS-SCCH per user must be established. For users with HSDPA mobile terminals that can handle only 5 codes, the code resource would not be used efficiently if only one user is scheduled during a TTI. In these cases it can be fruitful to multiplex several users in a TTI.

9.2 Previous Research

The performance of the HSDPA channel has been studied in several papers. In [98] simulations of the HSDPA concept are done by using maximum CQI and Round Robin schedulers. In [48, 114] HSDPA performance for TCP traffic scenarios are treated. Investigations of different scheduling strategies from simulators are presented in [14, 15, 80, 87, 107]. In [99] streaming applications over the HSDPA is treated. Performance of the HSDPA and DCH are investigated in [34, 115]. In [146] a workflow is presented for optimizing the balance between Release '99 services and HSDPA. In [18] an analytical model is proposed for evaluating HSDPA capacity. An approach to minimize the power consumption instead of maximizing the throughput is proposed in [111]. Code multiplexing is investigated in [85, 91]. In [85] the users are divided into 3 groups which share 5 codes each, i.e. up to 3 users can be multiplexed in a TTI. In [91] two different simulation cases are considered for code multiplexing. The first simulation case is a semi-static scenario where the channelization codes are divided into groups with an equal number of codes. These groups are assigned to the users that will be multiplexed. The second simulation case is dynamic code sharing, where the best user takes the number of codes the channel condition allows and the remaining users share rest of the codes.

9.3 Optimization Model

The optimization model to be presented schedules the users for one of the cells (serving cell) while treating the others as interferers. We use $M = \{1, \ldots, |M|\}$ to denote the set of users connected to the cell. The transmission rates for a single code are in the set $R = \{1, \ldots, |R|\}$, and R_j is the rate of $j, j \in R$. The number of channelization codes for the HS-DSCH is denoted by a set $K = \{1, \ldots, |K|\}$. The total bit rates when using k codes are $kR_j, j \in R, k \in K$. The task is to assign channelization codes and bit rates to the users in M while taking into account the code and power limitations. The scheduling model for this task is a linear integer optimization model.

The model uses the following variables

$$x_{ijk} = \begin{cases} 1 & \text{if user } i \text{ is assigned } k \text{ codes with bit rate } j \\ 0 & \text{otherwise.} \end{cases}$$

The parameter used in the model is

$p_{ijk} =$ the power for user i when using k codes with bit rate j.

For the serving cell the scheduling model can be stated as the following linear integer optimization model, HSDPA Bandwidth Allocation Problem (HSBAP),

$$\begin{array}{llll}
[HSBAP] & & & \\
\max & \sum_{i \in M} \sum_{j \in R} \sum_{k \in K} u_{ijk} x_{ijk} & & (9.1) \\
\text{s.t.} & \sum_{i \in M} \sum_{j \in R} \sum_{k \in K} (p_{ijk} + P^{hsscch}) x_{ijk} \leq P^{hsdpa} & & (9.2) \\
& \sum_{j \in R} \sum_{k \in K} x_{ijk} \leq 1, & i \in M & (9.3) \\
& \sum_{i \in M} \sum_{j \in R} \sum_{k \in K} k x_{ijk} \leq C^{max} & & (9.4) \\
& \sum_{j \in R} \sum_{k \in K} k x_{ijk} \leq C_i^{limit}, & i \in M & (9.5) \\
& x_{ijk} = 0/1, & \forall i, j, k & (9.6)
\end{array}$$

In the optimization model, constraint (9.2) states that the power consumed for the HSDPA users can not exceed the constant value P^{hsdpa}. The parameter P^{hsscch} is the power assigned to the signaling channel, HS-SCCH. The constraint (9.3) assigns at most one code/bit rate combination to a user i. The number of channelization codes that can be used in a cell is limited to C^{max}, which is stated by constraint (9.4). Constraint (9.5) assures that some mobile terminal categories can only handle C_i^{limit} number of channelization codes.

When scheduling the HSDPA users in a cell, the utility function plays a very important role. One type of scheduler is the Round Robin (RR), which assigns the users to a TTI in a cycle. As an example assume that we want to schedule 10 users to TTIs, the RR scheduler assigns user number 1 to TTI number 1 and user number 2 to TTI number 2 and so on. When all 10 users are assigned a TTI, the scheduler cycles and assigns TTI number 11 to user 1. All users get the same number of TTIs assigned to download data. The RR scheduler assigns the maximum kR_j combination to a user given the power limitation. Another scheduler that uses the bit rate as utility function coefficient is the Maximum Channel Quality Index (max CQI). The utility function coefficient for max CQI scheduling is

$$u_{ijk} = kR_j, i \in M. \tag{9.7}$$

The max CQI scheduler prioritizes the user with the best own cell interference to other cell interference ratio in a TTI. This scheduling scheme will never assign resources to users that suffer from a bad radio environment.

Another commonly used scheduling scheme is Proportional Fair (PF). The PF scheduling scheme uses a scheduling metric which is a ratio between the bit rate and the average smoothed throughput for a user [66]. The average smoothed throughput in TTI t is calculated by using a smoothed low pass filter [13, 107]. Initially the scheduler assigns resources to the users experiencing the best radio environment. When a user gets resources assigned from the network the throughput will increase which lead to a lower probability to be assigned resources in the next TTI. This forces the scheduler to assign resources to users that suffer from a bad radio environment. The utility function coefficient for PF is a logarithmic function of the average smoothed throughput. In Section 9.6 the PF scheduling metric is used in the optimization model HSBAP as a utility function coefficient. The utility function coefficient is

$$u_{ijk} = \frac{kR_j}{R_i^{thr}}, \tag{9.8}$$

R_i^{thr} is the throughput for user i, which is the accumulated assigned bit rates kR_j in all the previous TTIs during the simulation.

9.4 AMC Data

The input data for AMC is based on I_{or}/I_{oc} ratio for a user. The factor I_{or} is the own cell power received by the mobile terminal and I_{oc} is the total received interference from other cells plus noise at the mobile terminal. The fading trace gain is included in the I_{or} parameter. An example of I_{or}/I_{oc} ratios for the 10 HSDPA users in cell 1 are shown in Figure 9.2. The calculated I_{or}/I_{oc} ratios for the users in the serving cell are compared to eight concave function curves taken from [67]. The function curves contain the average user throughput per code as a function of power allocated per code and I_{or}/I_{oc} for a pedestrian-A 3 km/h model, see Figure 9.3. The values of I_{or}/I_{oc} for the function curves are {-2 0 2 4 6 8 10 15} dB. The bottom curve corresponds to a I_{or}/I_{oc} of - 2 dB and the curve at the top corresponds to 15 dB. The maximum power per channelization code is set to 10 W. The I_{or}/I_{oc} ratio is calculated for a user and is rounded to the nearest valid I_{or}/I_{oc} function curve value. The average throughput values are scaled down to suit inter-TTI of one, i.e. the average throughput values per code are divided by 500 (scheduling every 2 ms gives 500 TTIs per second). The set R consists of the scaled average throughput values. Applying k codes to a user gives that the bit rates in R and power per code increases with a factor k since all channelization codes for a user have the same modulation and coding in a TTI.

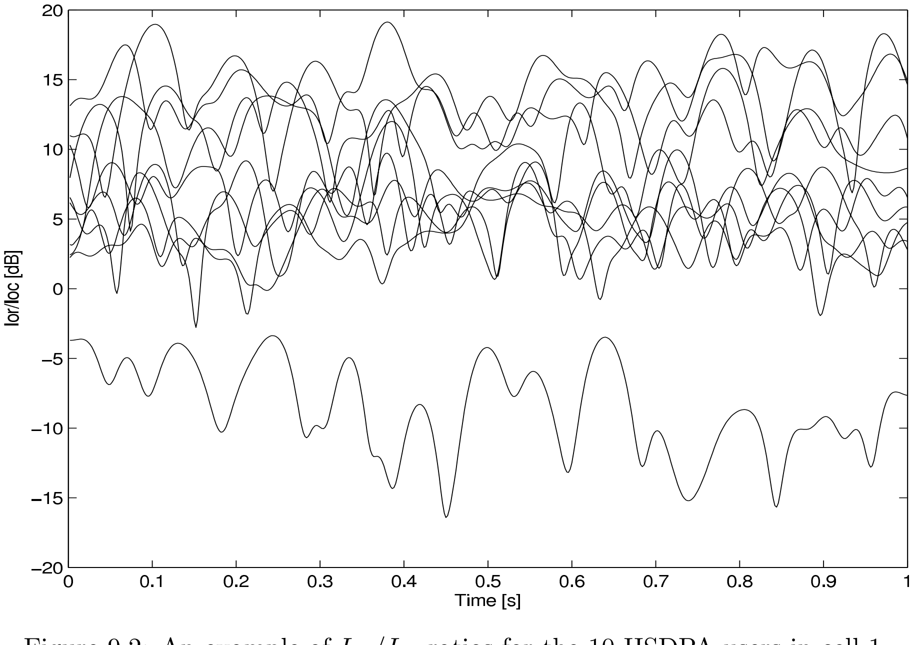

Figure 9.2: An example of I_{or}/I_{oc} ratios for the 10 HSDPA users in cell 1.

9.5 Two Greedy Algorithms

The scheduling model for the HSDPA is as mentioned in Section 9.3 a linear integer optimization model. One way to find a feasible solution to the scheduling

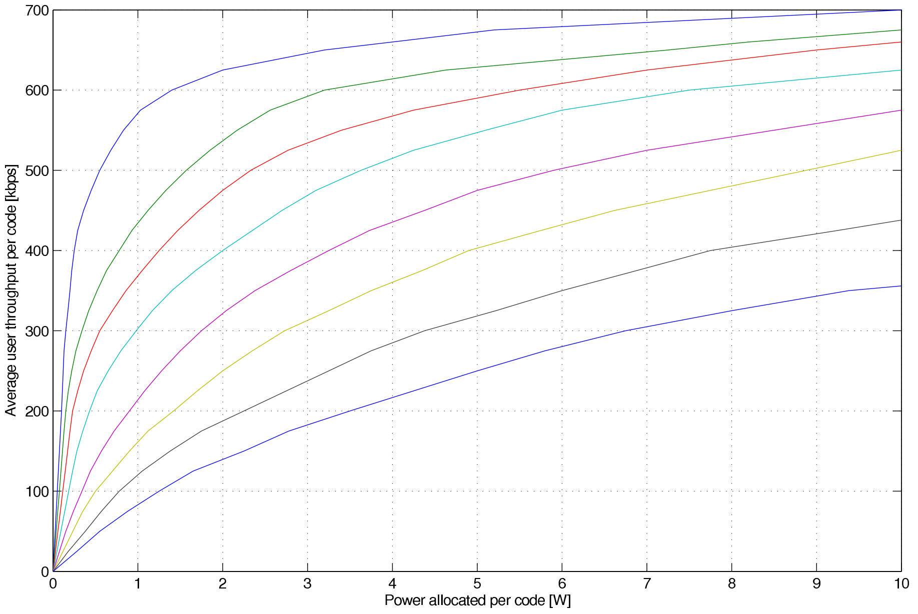

Figure 9.3: Input data for the HSDPA scheduling model. The functions curves are taken from [67].

model in TTI t is to apply a greedy algorithm. This Section contains two simple greedy algorithms, GREEDY 1 and GREEDY 2.

The GREEDY 1 algorithm chooses the largest feasible utility function coefficient u_{ijk} for a user and assigns the corresponding bit rate/channelization code combination kR_j to user i. When the first coefficient is determined, the algorithm continues to choose the maximum utility coefficient given the remaining resources of channelization codes and power. The pseudo code for the GREEDY 1 algorithm applied to scheduling model HSBAP is summarized in the following steps:

Algorithm: GREEDY 1

1. Calculate the utility function coefficients in TTI t.
2. Set $P^{remain} = P^{hsdpa}$ and $C^{remain} = C^{max}$.
3. While $P^{remain} > P^{hsscch}$ and $C^{remain} > 0$
   (a) Determine the maximum utility function coefficient u_{ijk} for $i \in M, j \in R$ and $1 \leq k \leq C_i^{limit}$.
   (b) If $p_{ijk} + P^{hsscch} \leq P^{remain}$ and $k \leq C^{remain}$.

i. Set $C^{remain} = C^{remain} - k$ and $P^{remain} = P^{remain} - P^{hsscch} - p_{ijk}$

ii. Assign k channelization codes and bit rate R_j to user i.

iii. Set all the utility function coefficients for user i to 0, to prevent user i to be assigned several bit rates.

(c) Else set $u_{ijk} = 0$ preventing that utility function coefficient to be chosen during the next loop.

4. Increment the time to $t = t + 1$. Go to step 1.

The AMC data to the HSDPA problem (see section 9.4) are concave curves. Assume that we are using the scheduling metric from equation (9.8) as utility function coefficient. The AMC input data function curves are scaled with the throughput for the users in TTI t, i.e. for each user we have a utility per code curve. To evaluate the behavior of the GREEDY 1 algorithm we assume the number of rates in the set R can be chosen to be very large and equally distributed. Assume that we choose the maximum utility function coefficient u_{ijk} and still have power and channelization codes left to assign another user. Since the number of rates in the set R is very large it should be possible to increase R_j to consume the maximum available power. This leads to a higher utility function value, which means that the original utility function coefficient cannot be maximum. Another way to look at the problem is the following. The initial assumption is that the maximum utility function coefficient u_{ijk} chosen does not consume all power or channelization codes. Since function curves are concave doubling the power/code gives less than the double utility. For the amount of power consumed by u_{ijk} it is possible to achieve a higher utility function value by assigning all C_i^{limit} channelization codes with less power per code. Since we have a very large number of equally distributed rates in the set R it should be possible to increase the rate until all power is consumed. This leads to a higher utility function value and no resources are left. The set R used in the numerical experiments contains 30 different bit rates leading to multiplexing in the 5 code simulation scenario, see Section 9.6.3.

The reason to develop the GREEDY 2 algorithm is to have an algorithm that multiplex users in the 15 codes case for the input data given. The GREEDY 2 algorithm is based on the GREEDY 1 algorithm with the extension to consider all combinations of two users. The pseudo code for the GREEDY 2 algorithm is

Algorithm: GREEDY 2

1. Calculate the utility function coefficients in TTI t. Store the utility function coefficients in a list Q.

2. Determine the maximum utility function coefficient u_{ijk} in Q.

3. If $p_{ijk} + P^{hsscch} \leq P^{max}$ and $k \leq C^{max}$ choose $u_{tot}(t) = u_{ijk}$ as the maximum utility and assign k channelization codes and bit rate R_j to user i. Else exclude u_{ijk} from the list Q and go to step 2.

4. For all $i, l \in M$, $j, m \in R$ and for all $1 \leq k \leq C_i^{limit}, 1 \leq n \leq C_l^{limit}$

   (a) Calculate $u^{two} = u_{ijk} + u_{lmn}$

   (b) If $p_{ijk} + P^{hsscch} + p_{lmn} + P^{hsscch} \leq P^{max}$, $k + n \leq C^{max}$ and $u^{two} > u_{tot}(t)$

      i. Assign bit rate of zero to all users.

      ii. Assign k channelization codes and rate R_j to user i and n channelization codes and bit rate R_m to user l.

      iii. Set $u_{tot}(t) = u^{two}$.

5. Increment the time to $t = t + 1$. Go to step 1.

9.6 Numerical Results

In this section we present the computational results when scheduling the HSDPA users. The maximum number of codes that can be assigned to all users in a TTI is 15 codes. Since not all categories of mobile terminals can handle the maximum number of channelization codes, computations have also been done when the number of codes is restricted to 5.

9.6.1 Assumptions

The network is the same as used in Chapter 8 and consists of 21 cells and has a path gain matrix for 1050 users. During the scheduling 10 users with lowest path loss are assigned to the serving cell. The other cells in the network are treated as interferers and transmit with a power of 20 W. The own cell interference is based of a Node B power of 20 W. The signaling channel HS-SCCH is transmitting with a constant value of $P^{hsscch} = 1.5$ W. In the case of code multiplexing when several HS-SCCH are used, all transmit with 1.5 W each. The power used for HSDPA is $P^{hsdpa} = 14$ W. Other types of traffic such as speech and common control channels use the remaining Node B power. The effect of thermal noise is set to a constant value of 10^{-13} W. To take into account Rayleigh fading a pedestrian-A 3 km/h fast fading trace affects the HSDPA users that will be scheduled. The fast fading trace contains data for 30 seconds and is split between the 10 HSDPA users, each user gets a 3 second part of the fading trace. The average throughput values (kbps) presented in Section 9.6.2 and 9.6.3 are the average number of assigned bits/second for a 3 seconds simulation. The mobile terminals considered in the simulations have a minimum TTI-interval of 1 which means that scheduling can be done every 2 ms. The data buffers in the mobile terminals are assumed to be unlimited in size. We assume that all users remain in the cell during the simulation. The users are assumed to download best-effort data and the Node Bs have full data buffers in each TTI, i.e. a user can take benefit of the highest bit rates. The AMC data used is calculated using HARQ with soft combining [67]. When the scheduler uses the history of the assigned TTIs all previous TTIs are weighted equally during the simulation, i.e. no low pass filtering is applied because of the short simulation time. The set R contains 30 scaled bit rates taken from Figure 9.3.

9.6.2 Simulating the 15 Channelization Codes Case

In this scenario we assume that all HSDPA terminals to be scheduled can handle 15 channelization codes. Comparing three different scheduling strategies, max-CQI, RR and the utility function coefficient given by equation (9.8), gives that the first has the highest average throughput. The reason for this is that max CQI is prioritizing the users with the best I_{or}/I_{oc} ratio, leading to only a few users are assigned resources from the network. A comparison between the three different scheduling methods is shown in Figure 9.4. In Figure 9.4 the max CQI assigns resources only to 4 users (users number 2, 3, 5 and 7) and has an average cell throughput of 7815 kbps, RR has an average cell throughput of 4576 kbps. When using the utility function coefficient from equation (9.8) the average throughput is 5690 kbps.

Numerical results for all the 21 cells in the network have been calculated and are shown in Table 9.1 there each row in the table corresponds to a cell. The utility function coefficient used is the metric ratio from equation (9.8) that is also used in Section 9.6.3. The second column in Table 9.1 contains the average throughput when the scheduling model is solved by the commercial optimization tool CPLEX [46, 71]. The third and fourth columns are the average throughput for GREEDY 2 and GREEDY 1 respectively. The fifth column is the percentage gain in average throughput between the two greedy algorithms. The last column in Table 9.1 is the percentage of TTIs that are multiplexed by several users when the HSBAP model is solved by the GREEDY 2 algorithm. The number of TTIs that is multiplexed in the CPLEX solution case is very similar compared to the GREEDY 2 solutions. The difference is at most 0.5 %.

The commercial optimization tool CPLEX solves the scheduling model to optimality in each TTI. The solutions show that up to two users are multiplexed in several of the TTIs. Comparing the CPLEX solutions with the GREEDY 2 solutions shows that the difference in average throughput is small. The computational times for CPLEX are many times longer compared to the GREEDY 2 algorithm. In several cells the GREEDY 2 algorithm gives exact the same solution as the CPLEX solution approach, i.e. both algorithms choose the same

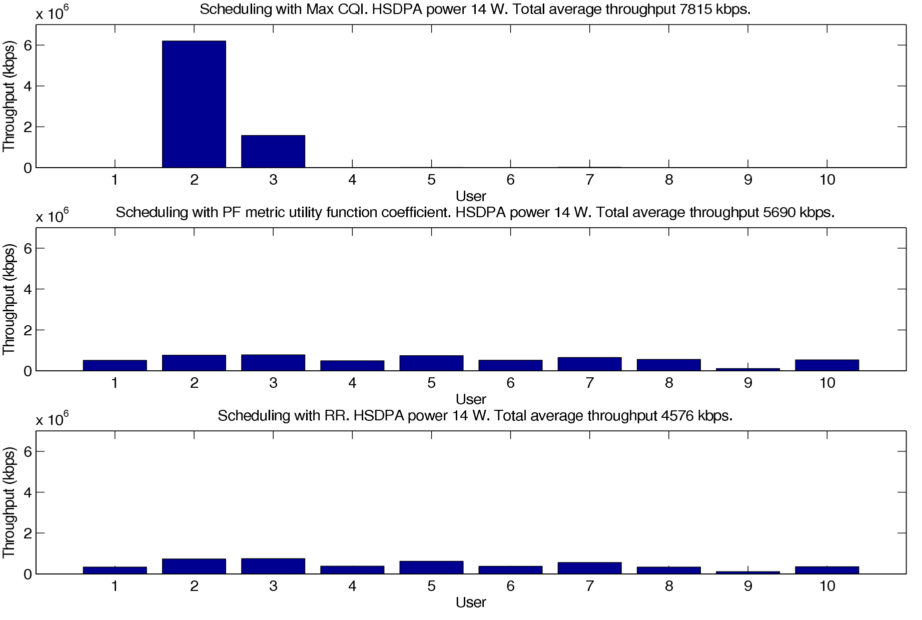

Figure 9.4: The average throughput when using different scheduling strategies with 10 users and 14 W of HSDPA power. Every mobile terminal can handle 15 codes in a TTI. The solution algorithm used is GREEDY 1.

users to schedule in each TTI. In some of the cells a small variation in average throughput between CPLEX and GREEDY 2 occurs. In the beginning of the simulation there exists multiple optima, and it is possible that GREEDY 2 and CPLEX chooses different user/users to schedule. Since the utility function is the ratio between the bit rate and throughput the optimal choice in the next TTI for GREEDY 2 and CPLEX can differ. It could be possible that the total average throughput for the GREEDY 2 algorithm exceeds the average throughput for the CPLEX solution.

In the numerical results for the GREEDY 1 algorithm only one user was assigned resources in each TTI. In the AMC data for the HSBAP model the maximum of kR_j will occur when the number of codes k is 14 or 15 depending on the I_{or}/I_{oc} ratio. In the case when only 14 codes are consumed there is not enough power left to set up a second HS-SCCH and multiplex another user in the TTI. A comparison between the GREEDY 1 and GREEDY 2 algorithms give that the latter has 0.1 to 9.3 % higher average throughput, column five in Table 9.1. However the GREEDY 2 algorithm is not guaranteed to give a higher average throughput compared to the GREEDY 1 algorithm. The average throughput for the whole network is 4713 kbps for GREEDY 1, 4861 kbps for GREEDY 2 and 4864 kbps for CPLEX.

Cell	Average throughput (kbps)			Gain (%)	TTI (%)
	CPLEX	GREEDY 2	GREEDY 1		
1	5787	5785	5690	1.7	46.4
2	5783	5783	5581	3.6	75.8
3	5331	5331	5171	3.1	54.7
4	5553	5553	5529	0.4	32.3
5	4904	4904	4717	3.9	45.1
6	5140	5138	4865	5.6	78.3
7	4827	4827	4443	8.6	82.5
8	4151	4149	3796	9.3	49.7
9	5006	5006	4885	2.5	61.5
10	4753	4753	4642	2.4	44.2
11	3752	3749	3726	0.6	13.1
12	4715	4715	4646	1.5	29.3
13	4638	4638	4616	0.5	17.9
14	2132	2126	2124	0.1	1.3
15	5397	5397	5158	4.6	77.5
16	3101	3101	2938	5.5	22.2
17	5269	5269	5226	0.8	32.7
18	5199	5199	5049	3.0	45.7
19	5210	5209	4976	4.7	69.3
20	5337	5333	5081	5.0	74.6
21	6114	6107	6065	0.7	22.4

Table 9.1: Numerical results when the mobile terminals can handle 15 codes. The power allocated for the HSDPA is 14 W.

9.6.3 Simulating the 5 Channelization Codes Case

Since not all categories of mobile terminals can handle 15 codes per TTI we investigate a scenario when the users have mobile terminals which can only handle 5 codes per TTI. The restriction of using only 5 codes/user is done by setting the parameter $C_i^{limit} = 5, i \in M$ in constraint (9.5) in HSBAP. Numerical results for all the 21 cells in the network have been calculated. The second column in the Table 9.2 contains the average throughput when the scheduling model is solved by the commercial optimization tool CPLEX. The third and fourth columns are the average throughput for GREEDY 2 and GREEDY 1 respectively. The last column is the percentage gain in average throughput between the two greedy algorithms. The solutions when solving the model with CPLEX show that all of the TTIs are multiplexed. The computational results show that 2 to 3 users are multiplexed and 10 to 15 channelization codes are assigned.

The GREEDY 2 algorithm can only multiplex two users in each TTI and consequently assign at most 10 channelization codes instead of 15 which results in a significant lower average throughput. The GREEDY 2 algorithm multiplexes users in almost all TTIs for the cells. Due to the input data the GREEDY 1 algorithm multiplexes users in several TTIs. The number of TTIs with multiplexed users in a cell varies between 11.7 % to 39.1 %. The average throughput for the whole network is 4370 kbps for CPLEX, 3507 kbps for GREEDY 2 and finally 2473 kbps for GREEDY 1.

9.7 Conclusions

The scheduling model assigns channelization codes and bit rates to users to maximize a utility function. The scheduling model is capable to handle code multiplexing, which permits several users to be assigned resources in one and the same TTI. The HSDPA uses AMC which is modeled by simulation data taken from [67]. The AMC input data consists of eight function curves which are the average bit rate per code as a function of power per code and I_{or}/I_{oc}. The scheduling model is solved with greedy algorithms and by a commercial optimization software in every TTI of 2ms.

The average throughput values presented in Section 9.6 are optimistic compared to [68], due to that we choose to assign the users with the lowest path loss to the serving cell and also the assumptions that the buffers of the Node Bs are full and the buffer in the mobile terminals are unlimited in size. Another aspect that affects the average throughput values in an optimistic way is that if a user has an I_{or}/I_{oc} ratio below -2 dB it will be rounded up to the -2 dB function curve in the AMC model.

Cell	Average throughput (kbps)			Gain (%)
	CPLEX	GREEDY 2	GREEDY 1	
1	5301	4231	2913	45.2
2	5213	4103	2967	38.3
3	4859	3793	2728	39.0
4	4910	3958	2883	37.3
5	4364	3513	2401	64.8
6	4753	3663	2635	39.0
7	4462	3367	2358	42.8
8	3722	2967	2009	47.7
9	4723	3719	2670	39.3
10	4141	3430	2332	47.1
11	3306	2812	2018	39.3
12	4240	3477	2367	46.9
13	4016	3397	2373	43.1
14	1683	1596	1274	25.3
15	4850	3818	2769	37.9
16	2750	2386	1591	49.9
17	4600	3798	2588	46.8
18	4714	3708	2604	42.4
19	4736	3739	2600	43.8
20	4844	3802	2686	41.5
21	5591	4417	3172	39.2

Table 9.2: Numerical results when the mobile terminals can handle 5 codes. The power allocated for the HSDPA is 14 W.

The numerical results show that up to two users are multiplexed in several TTIs when the mobile terminals can handle all the 15 codes. This results in that the solutions from the commercial optimization tool CPLEX and the GREEDY 2 algorithm are close. In the numerical results for the GREEDY 1 algorithm only one user was assigned bit rate and channelization codes in each TTI. Comparing the results from the GREEDY 1 and GREEDY 2 algorithms the latter only give a small increase in average throughput. If we consider mobile terminals that are restricted to handle maximum 5 channelization codes, it is preferable to multiplex up to three users in some TTIs. The GREEDY 2 algorithm has only a capability to multiplex two users and it can be clearly seen that allowing three multiplexed users gives a significant gain in average throughput. The GREEDY 1 algorithm performs less well due to the small number of assigned channelization codes. Extending the GREEDY 2 algorithm to handle 3 multiplexed users could easily be done but instead it would be more interesting to take advantage of the optimization method in Section 8.4, i.e. dynamic programming.

Chapter 10

Future Research

The contents of Part I, Part II and Chapter 8 in Part III of the thesis are based on published articles. Since the publication of these articles, other researchers have made some interesting work. The goal of this chapter is not to cover all the contributions that have been done. The focus is instead on some interesting articles that are connected to the work presented in this thesis and also highlight some interesting research fields for future work.

10.1 Scheduling in STDMA Ad Hoc Networks

An extension of the node and link assignment formulations presented in Part I is done by the authors of [59]. Instead of minimizing the STDMA frame the authors focus on traffic sensitivity and maximize the network throughput. The problem is solved with a column generation approach similar to the one used in Part I of the thesis. In [60] traffic sensitive scheduling algorithms and strategies are presented.

In the field of ad-hoc networks there have been intense research. One area that is of interest is joint optimization. In [75] a dual decomposition method is presented for the scheduling and routing problem. Joint optimization algorithms can also include the effect of power control. Power control is important to reduce the power consumption and interference. In [90, 137] mixed linear integer programming models for the joint problem of scheduling and power control are presented. In [74, 76] a nonlinear column generation technique is used to solve the joint scheduling, routing and power control problem. In [97] the authors present a centralized joint scheduling, routing and power control algorithm. Also a discussion of a distributed implementation of the algorithm is included.

Distributed algorithms are important and interesting for future investigations.

10.2 Frequency Assignment in Frequency Hopping GSM Networks

The work presented in Part II of this thesis has been extended by the authors of [105]. In the article methods to pregenerate sets of frequency lists of fixed length with desirable properties before the algorithm runs are presented. Two simulated annealing algorithms are presented for assigning the pregenerated frequency lists.

GSM is extended by EDGE to handle packet data with higher data rate. The authors of [78] investigate the impact of FH in EDGE. One of the conclusions drawn is that speech behaves differently than packet data traffic in interference limited networks. There are improvements in performance using FH for packet data despite a lower interference diversity gain.

The algorithms used in Part II of the thesis, solve the assignment problem heuristically. Future work is to derive effective bounding procedures for FAPH. Such methods are very useful for assessing the performance of the heuristics presented and are therefore interesting topics for further studies. It is also interesting to investigate assignment procedures with a mixed speech and packet data scenario.

10.3 Scheduling Packet Data in WCDMA Networks

The research directed towards the Release '99 channels include the assignment of OVSF codes, e.g. [33, 126, 133, 144]. The majority of the articles about scheduling in WCDMA are treating the high speed data packet channels. For the HSDPA scheduling model presented in Chapter 9 the next logical step is to apply optimization methods such as dynamic programming.

Given the improvements in efficiency and services provided by HSDPA, the next step would be to apply similar concepts to the uplink. High Speed Uplink Packet Access (HSUPA), also known as Enhanced Uplink, represents the latest 3GPP Release 6 technology and aims to provide faster packet data in the uplink. Multimedia services are popular and likely to grow as high-quality cameras become standard in most mobile terminals. Many of the features and enhancements of HSUPA can be traced directly back to HSDPA. The principles and basic operation of HSUPA is explained in [68, 101, 113]. Evaluations of the system performance of HSUPA are presented in [39, 125, 148, 153]. Scheduling strategies are important for the HSUPA and should be a target for future research.

Bibliography

[1] GSM World, **http:\\www.gsmworld.com**.

[2] 3rd Generation Partnership Project, *Physical Channels and Mapping of Transport Channels onto Physical Channels, version 6.7.0*, **3GPP Technical Report 25.211, 2005**.

[3] 3rd Generation Partnership Project, *Spreading and Modulation (FDD), version 6.4.0*, **3GPP Technical Report 25.213, 2005**.

[4] 3rd Generation Partnership Project, *UE Radio Access Capabilities, version 6.7.0*, **3GPP Technical Report 25.306, 2005**.

[5] 3rd Generation Partnership Project, *Physical Layer Aspects of UTRA High Speed Downlink Packet Access, version 4.0.0*, **3GPP Technical Report 25.848, 2001**.

[6] 3rd Generation Partnership Project, *High Speed Downlink Packet Access : Physical Layer Aspects, version 5.0.0*, **3GPP Technical Report 25.858, 2002**.

[7] K.I. Aardal, A. Hipolito, C.P.M. van Hoesel and B. Jansen, *A Branch-and-Cut Algorithm for the Frequency Assignment Problem*, **Research Memorandum 96/011, Maastricht University, Netherlands, 1996.**

[8] K.I. Aardal, C. Hurkens, J.K. Lenstra and S. Tiourine, *Algorithms for the Radio Link Frequency Assignment Problem*, **Technical Report, Utrech University/Technische Universiteit Eindhoven, Netherlands, 1999**.

[9] E. Aarts and J.K. Lenstra (Editors), *Local Search in Combinatorial Optimization*, **Wiley & Sons, West Sussex, United Kingdom, 1997**.

[10] M.A. Abu El-Ata, *Evolution of Mobile Cellular Communication Systems - The Journey to UMTS*, **Proceedings of the 17th National Radio Science Conference, 2000**.

[11] K.M. Ahn and S. Kim, *Optimal Bandwidth Allocation for Bandwidth Adaptation in Wireless Multimedia Networks*, **Computers & Operations Research, Vol. 30, pp. 1917-1929, 2003**.

[12] L. Ahlin and J. Zander, *Principles of Wireless Communications, Second edition*, **Studentlitteratur, Lund, Sweden, 1998**.

[13] B. Al-Manthari, N. Nasser and H. Hassanein, *Fair and Efficient Channel Dependent Scheduling Algorithm for HSDPA System*, **Proceedings of the 2nd International Symposium on Wireless Communication Systems, pp. 238-241, 2005**.

[14] P. Ameigeiras, *Packet Scheduling And Quality of Service in HSDPA*, **PhD Thesis, Department of Communication Technology, Aalborg University, Denmark, 2003**.

[15] G. Aniba and S. Aissa, *Resource Allocation in HSDPA Using Best-Users Selection Under Code Constraints*, **Proceedings of the 61st IEEE Vehicular Technology Conference, pp. 319-323, 2005**.

[16] S.M. Allen, D.H. Smith and S. Hurley, *Using Lower Bounds for Minimum Span Frequency Assignment*, **S. Voss, S. Martello, I.H. Osman (Editors), Meta Heuristics: Advances and Trends in Local Search Paradigms for Optimisation, Kluwer Academic Publishers, pp. 191-204, 1999**.

[17] S.M. Allen, D.H. Smith and S. Hurley, *Lower Bounding Techniques for Frequency Assignment*, **Discrete Mathematics 197/198, pp. 41-52, 1999**.

[18] M. Assaad and D. Zeghlache, *On the Capacity of HSDPA*, **Proceedings of IEEE Global Telecommunications Conference, pp. 60-64, 2003**.

[19] P. Björklund, *Two Resource Optimization Applications in Wireless Networks*, **Linköping Studies in Science and Technology, Thesis No. 943, Linköping University, Sweden, 2002**.

[20] P. Björklund, P. Värbrand and D. Yuan, *Resource Optimization of Spatial TDMA in Ad Hoc Radio Networks : A Column Generation Approach*, **Proceedings of IEEE INFOCOM '03, pp. 818-824, 2003**.

[21] P. Björklund, P. Värbrand and D. Yuan, *A Dynamic Programming Technique for Downlink Bandwidth Allocation in WCDMA Networks*, **Proceedings of the 59th IEEE Vehicular Technology Conference, pp. 2007-2011, 2004**.

[22] P. Björklund, P. Värbrand and D. Yuan, *A Column Generation Method for Spatial TDMA Scheduling in Ad Hoc Networks*, **Ad Hoc Networks, Vol. 2, No. 4, pp. 405-418, 2004**.

[23] P. Björklund, P. Värbrand and D. Yuan, *Optimal Frequency Planning in Mobile Networks with Frequency Hopping*, **Computers and Operations Research, Vol. 32, pp. 169-186, 2005**.

[24] J.A. Bondy and U.S.R. Murty, *Graph Theory with Applications*, **The MacMillan Press, New York, USA, 1978**.

[25] R. Borndörfer, A. Eisenblätter, M. Grötschel and A. Martin, *Frequency Assignment in Cellular Phone Networks*, **Annals of Operations Research, Vol. 76, pp. 73-93, 1998**.

[26] R. Borndörfer, A. Eisenblätter, M. Grötschel and A. Martin, *The Orientation Model for Frequency Assignment Problems*, **Technical Report, Konrad-Zuse-Zentrum, Germany, 1998**.

[27] F. Box, *A Heuristic Technique for Assigning Frequencies to Mobile Radio Nets*, **IEEE Transactions on Vehicular Technology, Vol. 27, pp. 57-74, 1978**.

[28] C. Brechtmann, K. Eckert and G. Plöger, *Planning of GSM and PCN Mobile Radio Networks*, **Proceedings of IEEE International Conference on Information Engineering, pp. 231-235, 1993**.

[29] R.L. Brooks, *On Colouring the Nodes of a Network*, **Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 37, pp. 194-197, 1941**.

[30] A. Capone and M. Trubian, *Channel Assignment Problem in Cellular Systems: A New Model and a Tabu Search Algorithm*, **IEEE Transactions on Vehicular Technology, Vol. 48, pp. 1252-1260, 1999**.

[31] C. Carneheim, S.O. Jonsson, M. Ljungberg, M. Madfors and J. Näslund, *FH-GSM Frequency Hopping GSM*, **Proceedings of the 44th IEEE Vehicular Tecnology Conference, pp. 1155-1159, 1994**.

[32] D.J. Castelino, S. Hurley and N.M. Stephens, *A Tabu Search Algorithm for Frequency Assignment*, **Annals of Operations Research, Vol. 63, pp. 301-319, 1996**.

[33] C.M. Chao and Y.C. Tseng, *Dynamic Bandwidth Allocation for Multimedia Traffic with Rate Gurantee and Fair Access in WCDMA Systems*, **IEEE Transactions on Mobile Computing, Vol. 4, No. 5, pp. 420-429, 2005**.

[34] C.S. Chiu and C.C. Lin, *Comparative Downlink Shared Channel Performance Evaluation of WCDMA Release 99 and HSDPA*, **Proceedings of IEEE International Conference on Networking, Sensing and Control, pp. 1165-1170, 2004**.

[35] I. Chlamtac and A. Lerner, *A Link Allocation Protocol for Mobile Multi-hop Radio Networks*, **Proceedings of IEEE Global Telecommunications Conference, pp. 238-242, 1985**.

[36] A.E. Chou and V. Li, *Slot Allocation Strategies for TDMA Protocols in Multihop Packet Radio Networks*, **Proceedings of IEEE INFOCOM '92, pp. 710-716, 1992**.

[37] I. Cidon and M. Sidi, *Distributed Assignment Algorithms for Multi-hop Packet Radio Networks*, **IEEE Transactions on Computers, Vol. 38, No. 10, pp. 1353-1361, 1989**.

[38] G. Drozdy, M. Salmi, J. Niemelä, J. Rosenberg, H. Kokkoniemi and V. Hara, *Simulation of Frequency Hopping in a GSM Network*, **Proceedings of IEEE Personal Wireless Communications Conference, pp. 257-261, 1996**.

[39] E. Englund, Y.-P. Eric Wang, C. Edholm, J.F. Cheng, C. Cozzo, M. Edvardsson and K. Wang Helmersson, *High-Data-Rate Availability in WCDMA Enhanced Uplink Systems*, **Proceedings of the 62nd IEEE Vehicular Technology Conference, pp. 2477-2481, 2005**.

[40] S. Engström, T. Johansson and F. Kronestedt, *Multiple Reuse Patterns for Frequency Planning in GSM Networks*, **Proceedings of the 48th IEEE Vehicular Technology Conference, pp. 2004-2008, 1998**.

[41] A. Ephremides and T. Truong, *Scheduling Broadcasts in Multihop Radio Networks*, **IEEE Transactions on Communication, Vol. 38, pp. 456-460, 1990**.

[42] *GSM System Introduction*, **Student Text EN/LZT 123 3641 R2B Ericsson Radio Systems AB, Sweden, 1998**.

[43] Ericsson Telecom, Telia, *Att förstå telekommunikation 2*, **Studentlitterartur, Lund, Sweden, 1998 (in swedish)**.

[44] I. Forkel, T. Kriengshaiyapruk, B. Wagmann, and E. Schulz, *Dynamic Channel Allocation in UMTS Terrestrial Radio Access TDD Systems*, **Proceedings of the 53rd IEEE Vehicular Technology Conference, pp. 1032-1036, 2001**.

[45] C.E. Fossa and N.J. Davis, *A Dynamic Code Assignment Algorithm for Quality of Service in 3G Wireless Networks*, **Proceedings of IEEE Wireless Communications and Networking Conference, pp. 1-6, 2002**.

[46] R. Fourer, D.M. Gay and W. Kernighan, *AMPL - A Modeling Language for Mathematical Programming*, **Boyd and Fraser, Danvers, USA, 1993**.

[47] M. Frullone, G. Riva, P. Grazioso and G. Falciasecca, *Advanced Planning Criteria for Cellular Systems*, **IEEE Personal Communications, Vol. 3, No. 6, pp. 10-15, 1996**.

[48] A. Furuskär, S. Parkvall, M. Persson and M. Samuelsson, *Performance of WCDMA High Speed Packet Data*, **Proceedings of the 55th IEEE Vehicular Technology Conference, pp. 1116-1120, 2002**.

[49] N. Funabiki and Y. Takefuji, *A Parallel Algorithm for Broadcast Scheduling Problems in Packet Radio Networks*, **IEEE Transactions on Communications, Vol. 41, pp. 828-831, 1993**.

[50] A. Gamst, *Some Lower Bounds for a Class of Frequency Assignment Problems*, **IEEE Transactions on Vehicular Technology, Vol. 35, pp. 8-14, 1986**.

[51] A. Gamst, *Application of Graph Theoretical Methods to GSM Radio Network Planning*, **Proceedings of IEEE International Sympoisum on Circuits and Systems, pp. 942-945, 1991**.

[52] V.K. Garg and J.E. Wilkes, *Principles and Applications of GSM*, **Prentice Hall, New Jersey, USA, 1999**.

[53] P.C. Gilmore and R.E. Gomory, *A Linear Programming Approach to the Cutting Stock Problem*, **Operations Research, Vol. 9, pp. 849-859, 1961**.

[54] P.C. Gilmore and R.E. Gomory, *A Linear Programming Approach to the Cutting Stock Problem - part II*, **Operations Research, Vol. 11, pp. 863-888, 1963**.

[55] F. Glover, *Future Paths for Integer Programming and Links to Artificial Intelligence*, **Computers and Operations Research, Vol. 5, pp. 533-549, 1986**.

[56] J. Grönkvist, *Traffic Controlled Spatial Reuse TDMA in Multi-hop Radio Networks*, **Proceedings of the 9th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1203-1207, 1998**.

[57] J. Grönkvist, A. Hansson and J. Nilsson, *A Comparision of Access Methods for Multi-hop Ad Hoc Radio Networks*, **Proceedings of the 51st IEEE Vehicular Technology Conference, pp. 1435-1439, 2000**.

[58] J. Grönkvist, *Assignment Methods for Spatial Reuse TDMA*, **First Annual Workshop on Mobile and Ad Hoc Networking and Computing, pp. 119-124, 2000**.

[59] J. Grönkvist, J. Nilsson and Di Yuan, *Throughput of Optimal Spatial Reuse TDMA for Wireless Ad-Hoc Networks*, **Proceedings of the 59th IEEE Vehicular Technology Conference, pp. 2156-2160, 2004**.

[60] J. Grönkvist, *Interference-Based Scheduling in Spatial Reuse TDMA*, **Doctoral Thesis, Royal Institute of Technology, Sweden, 2005**.

[61] Ö. Gürbüz and H. Owen, *Dynamic Resource Scheduling for Variable QoS Traffic in W-CDMA*, **Proceedings of IEEE Global Telecommunications Conference, pp. 703-707, 1999**.

[62] Ö. Gürbüz and H. Owen, *Dynamic Resource Scheduling Schemes for W-CDMA Systems*, **IEEE Communications Magazine, Vol. 38, No. 10, pp. 80-84, 2000**.

[63] M.K. Gurcan, O.O. Oyefuga, S. Kahn and B.O. Lijofi, *A Dynamic Programming Algorithm for Downlink Throughput Optimisation in Wideband CDMA Systems*, **Proceedings of the 3rd International Conference on 3G Mobile Communication Technologies, pp. 484-488, 2002**.

[64] B. Hajek and G. Sasaki, *Link Scheduling in Polynomial Time*, **IEEE Transactions on Information Theory, Vol. 34, Issue 5, pp. 910-917, 1988**.

[65] G. Heine, *GSM Networks: Protocols, Terminology, and Implementation*, **Artech House Publishers INC., Norwood, USA, 1999**.

[66] P.A. Hosein, *QoS Control for WCDMA High Speed Packet Data*, **Proceedings of the 4th International Workshop on Mobile and Wireless Communications Network, pp. 169-173, 2002**.

[67] H. Holma and A. Toskala (Editors), *WCDMA for UMTS: Radio Acess For Third Generation Mobile Communications*, **John Wiley & Sons, West Sussex, United Kingdom, 2004**.

[68] H. Holma and A. Toskala (Editors), *HSDPA/HSUPA for UMTS: High Speed Radio Access for Mobile Communications*, **John Wiley & Sons, West Sussex, United Kingdom, 2006**.

[69] C.A.J. Hurkens and S.R. Tiourine, *Upper and Lower Bounding Techniques for Frequency Assignment Problems*, **Technical Report COSOR 95-34, Eindhoven University, Netherlands, 1995**.

[70] S. Hurley, D.H. Smith and S.U. Thiel, *FASoft: A System for Discrete Channel Frequency Assignment*, **Radio Science, Vol. 32, pp. 1921-1939, 1997**.

[71] ILOG, *Ilog cplex, user's manual, 2000.*

[72] A. Jain, *Cell Planning in Mobile Communication*, **Proceedings of IEEE Conference on Personal Wireless Communications, pp. 190-194, 1996**.

[73] B. Jaumard, O. Marcotte and C. Meyer, *Mathematical Models and Exact Methods for Channel Assignment in Cellular Networks*, in **Telecommunications Network Planning, B. Sansò and P. Soriano (Editors), Kluwer Academic Publishers, Boston, USA, 1999**.

[74] M. Johansson and L. Xiao, *Scheduling, Routing and Power Allocation for Fairness in Wireless Networks*, **Proceedings of the 59th IEEE Vehicular Technology Conference, pp. 1355-1360, 2004**.

[75] M. Johansson, L. Xiao and S. Boyd, *Simultaneous Routing and Resource Allocation via Dual Decomposition*, **IEEE Transactions on Communications, Vol. 52, No. 7, pp. 1136-1144, 2004**.

[76] M. Johansson and L. Xiao, *Cross-Layer Optimization of Wireless Networks using Nonlinear Column Generation*, **IEEE Transactions on Wireless Communications, Vol. 5, No. 2, pp. 435-445, 2006**.

[77] R.M. Joyce, L.J. Ibbetson and L.B Lopes, *GSM Data Service Performance Improvement Through the use of Slow Frequency Hopping*, **Proceedings of the 47th IEEE Vehicular Technology Conference, pp. 1872-1876, 1997**.

[78] T. Jönsson, P. de Bruin and S. Craig, *High Capacity Strategies for GSM/EDGE - Impacts on Data Traffic Performance*, **Proceedings of the 60th IEEE Vehicular Technology Conference, pp. 3429-3433, 2004**.

[79] H. Kaaranen, A. Ahtiainen, L. Laitinen, S. Naghian and V. Niemi, *UMTS Networks*, **John Wiley & Sons, West Sussex, United Kingdom, 2001**.

[80] M. Kazami and N. Wiberg, *Scheduling Algorithms for HS-DSCH in a WCDMA Mixed Traffic Scenario*, **Proceedings of the 14th IEEE Symposium on Personal, Indoor and Mobile Radio Communication, pp. 1485-1489, 2003**.

[81] M. Keller, *A Comprehensive Simulation Approach to study FH in the GSM System regarding the Physical Layer Level and the System Level*, **Proceedings of the 47th IEEE Vehicular Technology Conference, pp. 1852-1856, 1997**.

[82] S. Kirkpatrick, C.D. Gellat and M.P. Vecchi, *Optimization by Simulated Annealing*, **Science, Vol. 220, pp. 671-680, 1983**.

[83] D. Kim, *Rate-Regulated Power Control for Supporting Flexible Transmission in Future CDMA Mobile Networks*, **IEEE Journal on Selected Areas in Communications, Vol. 17, pp. 968-977, 1999**.

[84] D.I. Kim, E. Hossain and V.K. Bhargava, *Dynamic Rate and Power Adaption for Provisioning Class-Based QoS in Cellular Multirate WCDMA Systems*, **IEEE Transactions on Wireless Communications, Vol. 3, No. 5, 2004**.

[85] T.E. Kolding, *Link and System Performance Aspects of Proportional Fair Scheduling in WCDMA/HSDPA*, **Proceedings of the 58th IEEE Vehicular Technology Confernece, pp. 1717-1722, 2003**.

[86] A.M.C.A. Koster, *Frequency Assignment: Models and Algorithms*, **PhD Thesis, Proefschrift, Universiteit Maastricht, Netherlands, 1999**.

[87] T.E. Kolding, F. Frederiksen and P.E. Mogensen, *Performance Aspects of WCDMA Systems with High Speed Downlink Packet Access (HSDPA)*, **Proceedings of the 56th IEEE Vehicular Technology Conference, pp. 477-481, 2002**.

[88] T. Kriengshaiyapruk and I. Forkel, *Dynamic Allocation of Capacity in UTRA TDD Systems*, **Proceedings of the 2nd IEEE International Conference on 3G Mobile Communication Technologies, pp. 262-266, 2001**.

[89] F. Kronestedt and M. Frodigh, *Frequency Planning Strategies for Frequency Hopping GSM*, **Proceedings of the 47th IEEE Vehicular Technology Conference, pp. 1862-1866, 1997**.

[90] G. Kulkarni, V. Raghunathan and M. Srivastava, *Joint End-to-End Scheduling, Power Control and Rate Control in Multi-hop Wireless Networks*, **Proceedings of IEEE Global Telecommunications Conference, pp. 3357-3362, 2004**.

[91] R. Kwan, P. Chong, E. Poutiainen and M. Rinne, *The Effect of Code-Multiplexing on the High Speed Downlink Packet Access (HSDPA) in a WCDMA Network*, **Proceedings of IEEE Wireless Communication and Networking Conference, pp. 1728-1732, 2003**.

[92] W.K. Lai and G.G. Coghill, *Channel Assignment through Evolutionary Optimization*, **IEEE Transactions on Vehicular Technology, Vol. 45, pp. 91-95, 1996**.

[93] J.W. Lee, R.R. Mazumdar, and N.B. Shroff, *Downlink Power Allocation for Multi-Class CDMA Wireless Networks*, **Proceedings of IEEE INFOCOM '02, pp. 1480-1489, 2002**.

[94] C. Lee, J. Lehoczky, R. Rajkumar, and D. Siewiorek, *On Quality of Service Optimization with Discrete QoS Options*, **Proceedings of the 5th IEEE Real-Time Technology and Applications Symposium, pp. 276-286, 1999**.

[95] H. Lee, J. Yeo, S. Kim and S. Lee, *Time Slot Assignment to Minimize Delay in Ad-hoc Networks*, **Technical Report, Radio and Broadcasting Technology Lab, ETRI, Department of Industrial Engineering, Korea Advanced Institute of Science and Technology, South Korea, 2001**.

[96] T. Lee and S. Park, *An Integer Programming Approach to the Time Slot Assignment Problem in SS/TDMA Systems with Intersatellite Links*, **European Journal of Operational Research, Vol. 135, pp. 57-66, 2001**.

[97] Y. Li and A. Ephremides, *Joint Scheduling, Power Control, and Routing Algorithm for Ad-Hoc Wireless Networks*, **Proceedings of the 38th Hawaii International Conference on System Sciences, pp. 1-10, 2005**.

[98] R. Love, A. Ghosh, R. Nikides and L. Jalloul, *High Speed Downlink Packet Access Performance*, **Proceedings of the 53rd IEEE Vehicular Technology Conference, pp. 2234-2238, 2001**.

[99] M. Lundevall, B. Olin, J. Olsson, N. Wiberg, S. Wänstedt, J. Eriksson and F. Eng, *Streaming Applications over HSDPA in Mixed Service Scenarios*, **Proceedings of the 60th IEEE Vehicular Technology Conference, pp. 841-845, 2004**.

[100] M. Madfors, K. Wallstedt, S. Magnusson, H. Olofsson, P.O. Backman and S. Engström, *High Capacity with Limited Spectrum in Cellular Systems*, **IEEE Communications Magazine, Vol. 35, pp. 38-45, 1997**.

[101] D. Malladi, J. Damnjanovic, X. Zhang and S. Willenegger, *Introduction to WCDMA Enhanced Uplink*, **Proceedings of the 5th IEEE International Conference on 3G Mobile Communication Technologies, pp. 240-244, 2004**.

[102] S. Martello and P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*, **John Wiley & Sons, West Sussex, United Kingdom, 1990**.

[103] A. Mehrotra and M.A. Trick, *A Column Generation Approach for Graph Coloring*, **Informs Journal on Computing, Vol. 8, pp. 344-354, 1996**.

[104] H.N.B Meriem and S. Tabbane, *Simulations of Frequency Hopping Schemes with Power Control in a GSM Network*, **Proceedings of the 51st IEEE Vehicular Technology Conference, pp. 2380-2383, 2000**.

[105] J.N. Moon, L.A. Hughes and D.H. Smith, *Assignment of Frequency Lists in Frequency Hopping Networks*, **IEEE Transactions on Vehicular Technology, Vol 54, No. 3, pp. 1147-1159, 2005**.

[106] M. Moustafa, I. Habib, M. Naghshineh, and M. Guizani, *QoS-Enabled Broadband Mobile Access to Wireline Networks*, **IEEE Communications Magazine, pp. 50-56, 2002**.

[107] Y. Ofuji, A. Morimoto, S. Abeta and M. Sawahashi, *Comparison of Packet Scheduling Algorithms focusing on User Throughput in High Speed Downlink Packet Access*, **Proceedings of the 13th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1462-1463, 2002**.

[108] Oh S.-J. and K. M. Wasserman, *Adaptive Resource Allocation in Power Constrained CDMA Mobile Networks*, **Proceedings of IEEE Wireless Communications and Networking Conference, pp. 510-514, 1999**.

[109] S.-J Oh, T. L. Olsen, and K. M. Wasserman, *Distributed Power Control and Spreading Gain Allocation in CDMA Data Networks*, **Proceedings of IEEE INFOCOM '00, pp. 379-385, 2000**.

[110] H. Olofsson, J. Näslund and J. Sköld, *Interference Diversity Gain in Frequency Hopping GSM*, **Proceedings of the 45th IEEE Vehicular Technology Conference, pp. 102-106, 1995**.

[111] J. Olsson and N. Wiberg, *Power Efficient Scheduling for Cellular Packet Data*, **Proceedings of RadioVetenskap och Kommunikation, pp. 421-426, 2005**.

[112] S. Parkvall, E. Englund, P. Malm, T. Hedberg, M. Persson and J. Peisa, *WCDMA Evolved - High-Speed Packet-data Services*, **Ericsson Review, No. 2, pp. 56-65, 2003**.

[113] S. Parkvall, J. Peisa, J. Torsner, M. Sågfors and P. Malm, *WCDMA Enhanced Uplink - Principles and Basic Operation*, **Proceedings of the 61st IEEE Vehicular Technology Conference, pp. 1411-1415, 2005**.

[114] J. Peisa and E. Englund, *TCP Performance over HS-DSCH*, **Proceedings of the 55th IEEE Vehicular technology Conference, pp. 987-991, 2002**.

[115] K.I. Pedersen, T.F. Lootsma, M. Stöttup, F. Fredriksen, T.E. Kolding and P.E. Mogensen, *Network Performance of Mixed Traffic on High Speed Downlink Packet Access and Dedicated Channels in WCDMA*, **Proceedings of the 60th IEEE Vehicular Technology Conference, pp. 4496-4500, 2004**.

[116] D. Pisinger, *Algorithms for Knapsack Problems*, **PhD Thesis, Department of Computer Science, University of Copenhagen, Denmark, 1995**.

[117] J. Plehn, *Applied Frequency Assignment*, **Proceedings of the 44th IEEE Vehicular Technology Conference, pp. 853-857, 1994**.

[118] L. Pond and V. Li, *A Distributed Time-slot Assignment Protocol for Multihop Broadcast Packet Radio Networks*, **Proceedings of IEEE Military Communications Conference, pp. 70-74, 1989**.

[119] C. Prohazka, *Decoupling Link Scheduling Constraints in Multihop Packet Radio Networks*, **IEEE Transactions on Computers, Vol. 38, pp. 455-458, 1989**.

[120] C.R. Reeves (Editor), *Modern Heuristic Techniques for Combinatorial Problems*, **Blackwell Scientific Publications, Oxford, United Kingdom, 1993**.

[121] U. Rehfuess and K. Ivanov, *Comparing Frequency Planning Against 1x3 and 1x1 Re-Use in Real Frequency Hopping Networks*, **Proceedings of the 49th IEEE Vehicular Technology Conference, pp. 1845-1849, 1999**.

[122] P.S. Rha, *Frequency Reuse Scheme with Reduced co-channel Interference for Fixed Cellular Systems*, **IEEE Electronics Letters, Vol. 34, pp. 237-238, 1998**.

[123] T. Rintakoski, M. Kuulusa and J. Nurmi, *Hardware Unit for OVSF/Walsh/Hadamard Code Generation*, **Proceedings of IEEE International Symposium on System-on-Chip, pp. 143-145, 2004**.

[124] M.C. Ronchini and E. Gaiani, *Improvement of GSM System Performance due to Frequency Hopping and/or Discontinous Transmission*, **Proceedings of IEEE International Conference on Communications, pp. 1596-1600, 1998**.

[125] C. Rosa, J. Outes, K. Dimou, T.B. Sörensen, J. Wigard, F. Fredriksen and P.E. Mogensen, *Performance of Fast Node B Scheduling and L1 HARQ Schemes in WCDMA Uplink Packet Access*, **Proceedings of the 59th IEEE Vehicular Technology Conference, pp. 1635-1639, 2004**.

[126] A.N. Rouskas and I.I. Athanasakis, *Assignemet of Multiple OVSF Codes at the Downlink in WCDMA*, **Proceedings of the 15th IEEE International Symposium on Personal, Indoor and Mobile Communications, pp. 454-458, 2004**.

[127] J. Samuelsson, *Private Communications, 2000.*

[128] E. Sarkola, *Channel Assignment Methods in Frequency Hopping Cellular Radio Networks*, **Proceedings of the 8th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 771-775, 1997**.

[129] A. Sathyendran and P.J. Smith, *A Theoretical Study of the Performance Improvement in GSM Networks due to Slow Frequency Hopping*, **Proceedings of the 48th IEEE Vehicular Technology Conference, pp. 2207-2211, 1998**.

[130] J. Schiller, *Mobile Communications, Second edition*, **Addison Wesley, London, United Kingdom, 2003**.

[131] J.G. Sempere, *An Overview of the GSM system*, **Department of Electronic & Electrical Enginneering, University of Strathclyde, Glasgow, United Kingdom, 2000**.

[132] J. Shor and T.G. Robertazzi, *Traffic Sensitive Algorithms and Performance Measures for the Generation of Self-Organizing Radio Network Schedules*, **IEEE Transactions on Communications, Vol. 41, pp. 16-21, 1993**.

[133] D.N. Skoutas and A.N. Rouskas, *Efficient Selection of OVSF Codes for the Downlink DCH in WCDMA*, **Proceedings of the 59th IEEE Vehicular Technology Conference, pp. 1954-1958, 2004**.

[134] D.H. Smith, S. Hurley and S.U. Thiel, *Improving Heuristics for the Frequency Assignment Problem*, **European Journal of Operational Research, Vol. 107, pp. 76-86, 1997**.

[135] D.H. Smith, S. Hurley and S.M. Allen, *A New Lower Bound for the Channel Assignment Problem*, **IEEE Transactions on Vehicular Technology, Vol. 49, pp. 1265-1272**.

[136] D.S. Stevens and M.H. Ammar, *Evaluation of Slot Allocation Strategies for TDMA Protocols in Packet Radio Networks*, **Proceedings of IEEE Military Communications Conference, pp. 835-839, 1990**.

[137] J. Tang, G. Xue, C. Chandler and W. Zhang, *Link Scheduling with Power Control for Throughput Enhancement in Multihop Wireless Networks*, **IEEE Transactions on Vehicular Technology, Vol. 55, No. 3, pp. 733-742, 2006**.

[138] S. Tiourine, C. Hurkens and J.K. Lenstra, *An Overview of Algorithmic Approaches to Frequency Assignment Problems*, **Technical Report, Eindhoven University of Technology, Netherlands, 1995**.

[139] T. Toftegård Nielsen, J. Wigard and P. Mogensen, *On the Capacity of a GSM Frequency Hopping Network with Intelligent Underlay-Overlay*, **Proceedings of the 47th IEEE Vehicular Technology Conference, pp. 1867-1871, 1997**.

[140] T. Toftegård Nielsen, J. Wigard, P.H. Michaelsen and P. Mogensen, *Slow Frequency Hopping Solutions for GSM Networks of Small Bandwidth*, **Proceedings of the 48th IEEE Vehicular Technology Conference, pp. 1321-1325, 1998**.

[141] T. Toftegård Nielsen, J. Wigard, P.H. Michaelsen and P. Mogensen, *Resource Allocation in Frequency Hopping PCS1900/GSM/DCS1800 Type of Network*, **Proceedings of the 49th IEEE Vehicular Technology Conference, pp. 209-214, 1999**.

[142] T. Toftegård Nielsen, J. Wigard, S. Skjaerris, C.O. Jensen and J.Elling, *Enhancing Network Quality using Baseband Frequency Hopping, Downlink Power Control and DTX in a Live GSM Network*, **Proceedings of the 9th IEEE International Symosium on Personal, Indoor and Mobile Radio Communications, pp. 905-909, 1998**.

[143] T. Toftegård Nielsen and J. Wigard, *Performance Enhancements in a Frequency Hopping GSM Network*, **Kluwer Academic Publishers, Boston, USA, 2000**.

[144] Y.R. Tsai and L.C. Lin, *Quality Based OVSF Code Assignment and Reassignment Strategies for WCDMA Systems*, **Proceedings of IEEE International Conference on Wireless Networks, Communications and Mobile Computing, pp. 238-243, 2005**.

[145] G.W. Tunnicliffe, A. Sathyendran and A.R. Murch, *Performance Improvement in GSM Networks due to Slow Frequency Hopping*, **Proceedings of the 48th IEEE Vehicular Technology Conference, pp. 1857-1861, 1997**.

[146] J. Voigt, J. Dessiner, J. Hubner, D. Hunold and S. Möbius, *Optimizing HSDPA Performance in the UMTS Network Planning Process*, **Proceedings of the 61st IEEE Vehicular Technology Conference, pp. 2384-2388, 2005**.

[147] L.C. Wang, G.L Stuber and C.T Lea, *Architecture Design, Frequency Planing, and Performance Analysis for a Microcell/Macrocell Overlaying System*, **IEEE Transactions on Vehicular Technology, Vol. 46, pp. 836-848, 1997**.

[148] K. Wang-Helmersson, E. Englund, M. Edvardsson, C. Edholm, S. Parkvall, M. Samuelsson, Y.P. Wang and J.F. Cheng, *System Performance of WCDMA Enhanced Uplink*, **Proceedings of the 61st IEEE Vehicular Technology Conference, pp. 1427-1431, 2005**.

[149] J. Wigard, P. Mogensen, J. Johansen and B. Vejlgaard, *Capacity of a GSM Network with Fractional Loading and Random Frequency Hopping*, **Proceedings of the 7th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 723-727, 1996**.

[150] J. Wigard, P. Mogensen, P.H. Michaelsen, J. Melero and T. Halonen, *Comparison of Networks with Different Frequency Reuses, and Power Control using the Effective Frequency Load Approach*, **Proceedings of the 51st IEEE Vehicular Technology Conference, pp. 835-839, 2000**.

[151] R. Wyrwas and J.C. Campbell, *Radio Topology Design with Slow Frequency Hopping for Interference Limited Cellular Systems*, **Proceedings of the 41st IEEE Vehicular Technology Conference, pp. 88-93, 1991**.

[152] L. Xu, X. Shen, and J.W. Mark, *Dynamic Bandwidth Allocation with Fair Scheduling for WCDMA Systems*, **IEEE Wireless Communications, Vol. 9, No. 2, pp. 26-32, 2002**.

[153] M. Zhang, T. Chen and J. Hu, *Throughput-based and Power-based Load Control in HSUPA*, **Proceedings of IEEE International Conference on Wireless Communications, Networking and Mobile Computing, pp. 969-973, 2005**.

Appendix A

Calculations of Interference Parameters

A.1 Co-channel Interference

In this section of the appendix the co-channel interference parameter a_{ij}^{co} for the FAPH is calculated. Consider two cells A and B, where A is the interferer and cell B is the serving cell. Cell B has a risk of dropping calls due to the interference from cell A. Consider only the interference from the TCHs. The TCH-STRX in the interfering cell A is denoted STRX I and for the serving cell B, STRX S, see Figure A.1.

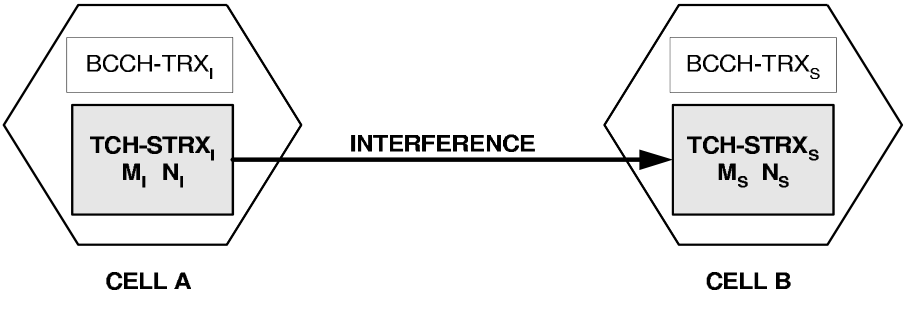

Figure A.1: Interference between the TCH-STRXs in two cells.

The following notation will be used:

M_I	$=$	MAL used for STRX I
M_S	$=$	MAL used for STRX S
N_I	$=$	Number of TRXs used in STRX I
N_S	$=$	Number of TRXs used in STRX S

The interference value C_{IS}^{co} is equal to the amount of traffic (in Erlang) that STRX S fails to serve if all TRXs in STRX S are suffering from co-channel interference from cell A all the time.

Assume that for a single TRX q in STRX S, the corresponding amount of traffic lost is C_{IS}^{co}/N_S. STRX S uses $|M_S|$ frequencies and as the hopping is synchronized, a certain TRX $q \in$ STRX S will spend its time evenly on all frequencies in M_S. The portion of time TRX q suffers from co-channel interference from the interfering cell A is $\frac{|M_I \cap M_S|}{|M_S|}$, which is equal to the probability that it uses any of the frequencies allocated in the serving cell B. Assume that TRX q in the serving cell is using a frequency which is assigned to both M_I and M_S. TRX q will only be interfered if one of the TRXs in STRX I is using the same frequency at the same time. Since any of the TRXs in STRX I uses M_I frequencies, the probability is $\frac{1}{|M_I|}$ for every TRX in STRX I. If all N_I TRXs in the interfering cell are taken into account, the probability is $\frac{N_I}{|M_I|}$. For a TRX $q \in$ STRX S the probability of being interfered from cell A is $\frac{N_I}{|M_I|}\frac{|M_I \cap M_S|}{|M_S|}$. The expected amount of lost traffic for TRX q is C_{IS}^{co}/N_S, taking the sum over all TRXs in STRX S gives

$$C_{IS}^{co}\frac{N_I}{|M_I|}\frac{|M_I \cap M_S|}{|M_S|}. \tag{A.1}$$

If we take the FH gain factor into consideration, the expected co-channel interference is

$$C_{IS}^{co}\frac{N_I}{|M_I|}\frac{|M_I \cap M_S|}{|M_S|}g(|M_I|,|M_S|). \tag{A.2}$$

Definition: *Consider a pair of STRXs p and q with corresponding MALs M_i and M_j, the scaling parameter for co-channel interference is*

$a_{ij}^{co} = \frac{|M_i \cap M_j|}{|M_i||M_j|}g(|M_i|,|M_j|)$.

The parameter a_{ij}^{co} takes values in the interval [0,1], where zero indicates no co-channel interference att all. A value of unity is obtained when both $|M_i|$ and $|M_j|$ have length one and contain the same frequency. It can be realized that a_{ij}^{co} is a natural extension of the corresponding parameter in the classical frequency assignment formulation. The only difference is that it is defined on MAL pairs instead of frequency pairs.

A.2 Adjacent-channel Interference

The same type of arguments apply to adjacent-channel interference. Using the same notation, cell A is the interfering cell with MAL M_I and N_I number of TRXs, and cell B is the serving cell with MAL M_S and N_S TRXs.

The interference value C_{IS}^{adj} is equal to the amount of traffic (in Erlang) that STRX S fails to serve if all TRXs in STRX S are suffering from adjacent-channel interference from cell A all the time.

As in the co-channel interference case, it is reasonable to assume that one TRX q in the serving cell B will lose C_{IS}^{adj}/N_S traffic if it suffers from adjacent-channel interference all the time from cell A. Consider now that TRX q will spend its time evenly on the $|M_S|$ frequencies available in the MAL for STRX S. The fraction of time TRX q suffers from interference by an adjacent frequency in cell A is

$$\frac{|(f : f \in M_S, f+1 \in M_I \text{ or } f-1 \in M_I)|}{|M_S|}. \tag{A.3}$$

Frequencies adjacent to f are denoted $f-1$ and $f+1$, higher adjacent frequencies are not taken into consideration. The notation $|(f : f \in M_s, f+1 \in M_I \text{ or } f-1 \in M_I)|$ is the number of frequencies for which TRX q is suffering from adjacent-channel interference. The probability of using one of the frequencies in STRX S is $\frac{1}{|M_S|}$.

Assume that $f \in M_S$ and $f+1 \in M_I$, then it is not allowed that $f-1 \in M_I$ since the internal separation for a MAL requires at least two frequencies in between, i.e. a separation of 3. For a frequency $f \in M_S$ there can be at most one adjacent frequency in M_I or vice versa. Using the internal separation requirement of 3 in both M_s and M_I, the following expressions are equivalent.

$$|(f : f \in M_S, f+1 \in M_I \text{ or } f-1 \in M_I)| \quad = \tag{A.4}$$
$$|(f : f \in M_S, f+1 \in M_I)| + |(f : f \in M_S, f-1 \in M_I)| \tag{A.5}$$

and

$$|(f : f \in M_I, f+1 \in M_S \text{ or } f-1 \in M_S)| \quad = \tag{A.6}$$

$$|(f : f \in M_I, f+1 \in M_S)| + |(f : f \in M_I, f-1 \in M_S)|. \tag{A.7}$$

At a certain time slot, there will be at most one TRX in STRX I that can cause adjacent interference. Since any of the TRXs in STRX I uses $|M_I|$ frequencies, the probability is $\frac{1}{|M_I|}$ for every TRX in STRX I. The expression now yields

$$\frac{|(f : f \in M_S, f+1 \in M_I \text{ or } f-1 \in M_I)|}{|M_I||M_S|} = \tag{A.8}$$

$$\frac{|(f : f \in M_I, f+1 \in M_S)| + |(f : f \in M_I, f-1 \in M_S)|}{|M_I||M_S|}. \tag{A.9}$$

Expression (A.9) is the probability that TRX q in the serving cell receives adjacent interference from one TRX in STRX I. Taking the sum over all TRXs in the interfering cell A results in the probability that TRX q receives adjacent interference from STRX I

$$N_I \cdot \frac{|(f : f \in M_I, f+1 \in M_S)| + |(f : f \in M_I, f-1 \in M_S)|}{|M_S||M_I|}. \tag{A.10}$$

The expected adjacent interference for TRX q is C_{IS}^{adj}/N_S, taking the sum over all TRXs in STRX S gives

$$C_{IS}^{adj} N_I \frac{|(f : f \in M_I, f+1 \in M_S)| + |(f : f \in M_I, f-1 \in M_S)|}{|M_S||M_I|}. \tag{A.11}$$

Scaling the expression with the FH gain factor $g(|M_I|, |M_S|)$ results in the total adjacent interference between the interfering cell A and the serving cell B

$$C_{IS}^{adj} N_I \frac{|(f : f \in M_I, f+1 \in M_S)| + |(f : f \in M_I, f-1 \in M_S)|}{|M_S||M_I|} g(|M_I|, |M_S|). \tag{A.12}$$

Definition: Consider a pair of STRXs p and q with corresponding MALs M_i and M_j, a scaling parameter for adjacent-channel interference is $a_{ij}^{adj} = \frac{|(f : f \in M_i, f+1 \in M_j)| + |(f : f \in M_i, f-1 \in M_j)|}{|M_i||M_j|} g(|M_i|, |M_j|)$.

Appendix B

Estimating the Parameter M_{ij}

If a link (i, j) is not active the SIR requirement must be redundant. This is done by the parameter M_{ij} in equations (3.5) and (3.12). The probability is low that all potential interferers are transmitting simultaneously. The reason for this is that the interferers to node j must avoid the type 1 and type 2 collisions and fulfill the SIR requirement in equation (2.2) as well. Type 1 collision occurs when a node transmits simultaneously to another node which is also transmitting in the same time slot. Type 2 collision occurs when a node is receiving packets from two other nodes in the same time slot. As an example, assume that a communication link (i, j) is established between the nodes i and j. The nodes k, m and n are interferers to node j. Assume that links (k, m) and (n, m) exist. If node k is transmitting to node m, nodes m and n can not transmit during that time slot. In Figure B.1, the link (i, j) is active and only one of the nodes k, m and n is allowed to transmit in the same time slot.

The reason for reducing the value of the M_{ij} parameter is that we will obtain a stronger formulation - from the point of view of the LP relaxation. To calculate a lower M_{ij} value for a link (i, j), an optimization model can be used. The objective is to maximize the possible interference under conditions that avoid collisions of type 1 and type 2, and SIR is above the threshold γ_1. For communication to be established between nodes k and m, the SIR constraint must be fulfilled. We introduce the following variables

$$w_k = \begin{cases} 1 & \text{if node } k \text{ is transmitting} \\ 0 & \text{otherwise.} \end{cases}$$

The interfering node k can not be the same node as in the original communication

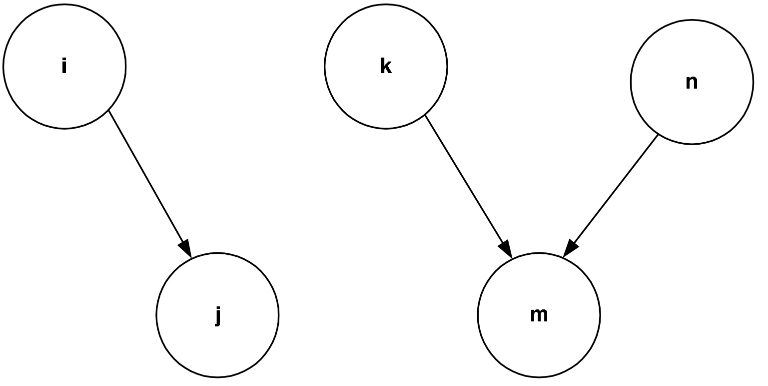

Figure B.1: Node j can only be interfered by one of the nodes k, m and n.

link (i, j), and we therefore define the set K as

$$K = \{k \in N : k \neq i, k \neq j\} \tag{B.1}$$

To calculate an upper bound for M_{ij} the following model can be used

$$\begin{aligned}
\max \quad & \sum_{k \in K} \frac{P_k/N_r}{L_b(k,j)} w_k && \text{(B.2)}\\
\text{s.t.} \quad & w_k + \sum_{n \in K:(n,k) \in A} w_n \leq 1, k \in K && \text{(B.3)}\\
& \frac{P_k/N_r}{L_b(k,m)} w_k + \gamma_1(1 + M_{km})(1 - w_k) \geq \gamma_1(1+ \\
& \sum_{n \in K:(n,m) \in A, n \neq k} \frac{P_n/N_r}{L_b(n,m)} w_n), \forall (k,m) \in A, k, m \in K && \text{(B.4)}\\
& w_k = 0/1, k \in K && \text{(B.5)}
\end{aligned}$$

The factor M_{km} can be set to the worst case scenario

$$M_{km} = \sum_{n \in K:(n,m) \in A, n \neq k} \frac{P_n/N_r}{L_b(n,m)}. \tag{B.6}$$