Title Practical Design of Optimal Wireless Metropolitan Area Networks: Model and Algorithms for OFDMA Networks

Name Gordejuela Sánchez, Fernando

PRACTICAL DESIGN OF OPTIMAL WIRELESS METROPOLITAN AREA NETWORKS: MODELS AND ALGORITHMS FOR OFDMA NETWORKS

Fernando Gordejuela Sánchez

Ph.D.

2009

UNIVERSITY OF BEDFORDSHIRE

UNIVERSITY OF BEDFORDSHIRE

PRACTICAL DESIGN OF OPTIMAL WIRELESS METROPOLITAN AREA NETWORKS: MODELS AND ALGORITHMS FOR OFDMA NETWORKS

by

Fernando Gordejuela Sánchez

A thesis submitted to the University of Bedfordshire, in partial fulfilment of the requirements for the degree of Ph.D.

October 2009

Practical Design of Optimal Wireless Metropolitan Area Networks: Models and Algorithms for OFDMA Networks

F. GORDEJUELA SÁNCHEZ

ABSTRACT

This thesis contributes to the study of the planning and optimisation of wireless metropolitan area networks, in particular to the access network design of OFDMA-based systems, where different parameters like base station position, antenna tilt and azimuth need to be configured during the early stages of the network life. A practical view for the solution of this problem is presented by means of the development of a novel design framework and the use of multicriteria optimisation. A further consideration of relaying and cooperative communications in the context of the design of this kind of networks is done, an area little researched.

With the emergence of new technologies and services, it is very important to accurately identify the factors that affect the design of the wireless access network and define how to take them into account to achieve optimally performing and cost-efficient networks. The new features and flexibility of OFDMA networks seem particularly suited to the provision of different broadband services to metropolitan areas. However, until now, most existing efforts have been focused on the basic access capability networks. This thesis presents a way to deal with the trade-offs generated during the OFDMA access network design, and presents a service-oriented optimization framework that offers a new perspective for this process with consideration of the technical and economic factors.

The introduction of relay stations in wireless metropolitan area networks will bring numerous advantages such as coverage extension and capacity enhancement due to the deployment of new cells and the reduction of distance between transmitter and receiver. However, the network designers will also face new challenges with the use of relay stations, since they involve a new source of interference and a complicated air interface; and this need to be carefully evaluated during the network design process.

Contrary to the well known procedure of cellular network design over regular or hexagonal scenarios, the wireless network planning and optimization process aims to deal with the non-uniform characteristics of realistic scenarios, where the existence of hotspots, different channel characteristics for the users, or different service requirements will determine the final design of the wireless network. This thesis is structured in three main blocks covering important gaps in the existing literature in planning (efficient simulation) and optimisation. The formulation and ideas proposed in the former case can still be evaluated over regular scenarios, for the sake of simplicity, while the study of latter case needs to be done over specific scenarios that will be described when appropriate. Nevertheless, comments and conclusions are extrapolated to more general cases throughout this work.

After an introduction and a description of the related work, this thesis first focuses on the study of models and algorithms for classical point-to-multipoint networks on Chapter 3, where the optimisation framework is proposed. Based on the framework, this work:

- Identifies the technology-specific physical factors that affect most importantly the network system level simulation, planning and optimization process.
- It demonstrates how to simplify the problem and translate it into a formal optimization routine with consideration of economic factors.
- It provides the network provider, a detailed and clear description of different scenarios during the design process so that the most suitable solution can be found. Existing works on this area do not provide such a comprehensive framework.

In Chapter 4:

- The impact of the relay configuration on the network planning process is analysed.
- A new simple and flexible scheme to integrate multihop communications in the Mobile WiMAX frame structure is proposed and evaluated.
- Efficient capacity calculations that allow intensive system level simulations in a multihop environment are introduced.

In Chapter 5:

- An analysis of the optimisation procedure with the addition of relay stations and the derived higher complexity of the process is done.
- A frequency plan procedure not found in the existing literature is proposed, which combines it with the use of the necessary frame fragmentation of in-band relay communications and cooperative procedures.
- A novel joint two-step process for network planning and optimisation is proposed.

Finally, conclusions and open issues are exposed.

DEDICATION

I dedicate this thesis to my family, especially to my parents, Fernando Gordejuela and Liboria Sánchez, my sister, Verónica Gordejuela-Sánchez, and my girlfriend Susana Gómez Escalante.

ACKNOWLEDGEMENTS

I would like to thank my Director of Studies, Professor Jie Zhang, and my supervisory team, Dr Alpar Juttner and Professor Gordon Clapworthy, for their guidance and support in my research work.

The encouragement, support and friendship of students and staff both in the Centre for Wireless Network Design (CWiND), the Institute for Research in Applicable Computing and elsewhere in the University of Bedfordshire, as well as the support of my girlfriend and family have helped me through the different stages in the pursuit of this Ph.D. degree.

The excellent research environment, high-motivating spirit and hard-working culture found in CWiND made me a big impact from the very beginning. This is one of the key aspects that have helped CWiND to become one of leading research groups in wireless communications in the world, and with any doubt, it will continue like this in the future.

I would like to thank the reviewers, Dr Enjie Liu and Dr Jianhua He, for their valuable comments.

This research is funded by the EU FP6 Marie Curie "RANPLAN-HEC" project on Automatic 3G/4G Radio Access Network Planning and Optimisation - A High End Computing Approach. Grant number MEST-CT-2005-020958.

DECLARATION

I declare that this thesis is my own unaided work. It is being submitted for the degree of Ph.D. at the University of Bedfordshire.

It has not been submitted before for any degree or examination in any other University.

Name of candidate: Fernando Gordejuela Sánchez

Signature:

Date:

24th October 2009

CONTENTS

LIST OF ACRONYMS

2G	2nd Generation of mobile networks
3GPP	Third Generation Partnership Project
AMC	Adapting Modulation and Coding
API	Application Programming Interface
AWGN	Additive White Gaussian Noise
BB	Branch and Bound
BE	Best Effort
BLER	Block Error Rate
BS	Candidate Sector/Base Sector
CapEx	Capital Expenditure
CDMA	Code Division Multiple Access
CINR	Carrier to Interference Plus Noise Ratio
CQI	Channel Quality Indicator
CS	Candidate Site
DFP	Dynamic Frequency Planning
DL	Downlink
EDGE	Enhanced Data rates for GSM Evolution
ErtPS	Extended real-time Polling Service
ESA	Evolutionary Simulated Annealing
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FFT	Fast Fourier Transform
FUSC	Full Usage of SubCarriers
FSA	Fixed Subchannel Allocation
GA	Genetic Algorithm
GPRS	General Packet Radio Service
GS	Greedy Search
GSM	Global System for Mobile communications
HSDPA	High Speed Downlink Packet Access
HSPA	High Speed Packet Access
HSUPA	High Speed Uplink Packet Access
ILP	Integer Linear Programming
KPI	Key Performance Indicators
LOS	Line of sight
LTE	Long Term Evolution
LUT	Look-Up Tables
MCS	Modulation and Coding Scheme
MS	Mobile Station
MIC	Mean Instantaneous Capacity
MILP	Mixed Integer Linear Programming
NLOS	Non line of sight
nrtPS	non-real-time Polling Service

OFDMA	Orthogonal Frequency-Division Multiple Access
OFUSC	Optional Full Usage of SubCarriers
OpEx	Operational expenditure
PMP	Point to Multipoint
PUSC	Partial Usage of SubCarriers
QoS	Quality of Service
R-D	Relay-Destination
RB	Resource Blocks
RFP	Radio Frequency Planning
RRM	Radio Resource Management
RS	Relay Sectors
rtPS	real-time Polling Service
S-D	Source-Destination
SA	Simulated Annealing
SB	Scheduling Block
SC-FDMA	Single Carrier Frequency Division Multiple Access
SD	Standard Deviation
SLS	System Level Simulation
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TS	Tabu Search
UGS	Unsolicited Grant Service
UL	Uplink
VBS	Visual Basic Script
VNS	Variable Neighbourhood Search
WCDMA	Wideband Code Division Multiple Access
WiMAX	Wireless Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Networks

LIST OF SYMBOLS

λ	Weighting value in cost function
η_{sub}	Efficiency factor (subcarriers/subchannel)
η_{sub}	Subcarriers per subchannel
η_{Be}	Bearer efficiency
ε	Losses due to the combination of signals
η_s	Number of symbols in the access zone
η_{Cij}	Portion of zone that is being interfered by the neighbour sector
An	Set of years of a return period in a business plan
a	A particular year in a business plan
A	Maximum number of years in a business plan
b	Region
B	Maximum number of regions
Bc	Value of the infrastructure economic parameter related to a site
Be	Bearer
Bs	Value of the infrastructure economic parameter related to a sector
Cs	Set of candidate sites
$C_{m,j,r}$	Received power from sector j on resource r at user m
$C_{i,pilot}$	Received pilot power from sector i
CostAdj	Adjustment factor
Cu	Subset of criteria related to the user CINR levels
$Cu_{h'}$	Criterion related to the user CINR levels
CQI	Channel Quality indicator
Cr	Subset of criteria related to the RS CINR levels
$Cr_{h'''}$	Criterion related to the RS CINR levels
$d_{m,h}$	If user m is evaluated by the cost function related to criterion h
D_i	Required subchannels per sector
F_{pen}	Penalty function in cost function
F_{costs}	Aggregate of the installation and annual maintenance fee for the sites and sectors
f_{costs}	Installation and annual maintenance fee for a site or a sector
f_{pen}	Penalty value of a single user
$f^{(max)}$	Maximum value of the penalty function
$F_{sampling}$	Sampling frequency of the system
f	Cost function
g	Interfering user
G	Set of design criteria
G_h	A design criterion

Ga	Antenna gains
H	Maximun number of design criteria
H′	Maximum number of criteria related to the user CINR levels
H″	Maximum number of criteria related to the user throughput
H‴	Maximum number of criteria related to the RS CINR levels
i	Serving sector
I_m	Interference on a certain user m
$Ie_{i,j,m}$	Interference event
j	Interfering sector
k	Subchannel
K	Total number of subchannels
L_{jm}	Path loss between user m and station j
L_j , L_m	Feeding losses
Ms_i	Set of users attached to a sector i
m	User
M_i	Total number of users attached to a sector i
N_z	Total number of sectors
n	Network solution
N_{pilot}	Noise power from pilot signal
n_b	Maximum number of CSs installed in a region b in the final solution
N	Set of economic parameters related to a user
N_v	Economic parameter related to a user
N′	Subset of economic parameters related to a user
N_r	Noise power of resource r
n_0	Thermal noise
Nf_{eq},	Equipment noise figure
N_{usedSC}	Number of considered sub-carriers
$N_{totalSC}$	Number of total sub-carriers
$N(Sf_{u,n})$	Neighborhood of search front at iteration u with n solutions
n_{Ki}	Maximum number of subchannels allowed in the UL upper zones (in transparent relay mode)
Of	Optimal front
Profit	Calculated profit
ProfitEst	Estimated profit
Pc	Set of economical parameters related to a site
Pp	Economical parameter related to a site
P	Maximum number of economical parameters related to a site
Ps	Set of economic parameters related to a sector
Pq	Economic parameter related to a sector
$P_{j,r}$	Power applied by j in r in dBm
Q	Maximum number of economical parameters related to a sector

r	resource
R	Total number of resources
Ra_n	Rank of a solution n
Rb	Set of regions
Rl_m	Requested slots by user m
Rc_m	Requested capacity by user m
Rs_i	Set of RSs attached to a BS i
Sc	Set of subcarriers
Sl	Set of WiMAX slots or LTE scheduling blocks
Sh	Set of subchannels
Sr	Set of resources
S_T	Sensibility of the terminal
So_n	Number of solutions that dominate n
S_z	Set of sectors
s_i	Usage of a sector i
s_z	Usage of a site z
$Sf_{u,n}$	Search front at iteration u with n solutions
Se_m	Slot efficiency of user m
$Se_{m'}$	Slot efficiency of the first hop of user m
s	Solution for subchannel allocation
s0	Random solution
sb	Best solution
sn	Neighbour solution
T_m	Throughput of user m
$t_{m,h}$	Value of criterion h in user m
$T^{(min)}$	Threshold in which the penalty function starts decreasing
$T^{(max)}$	Threshold in which the penalty function starts increasing
Tg	Subset of criteria related to the user throughput
$Tg_{h''}$	Criterion related to the user throughput
Te	Temperature
$u_{j,r}$	Usage of resource r by station j
u	Optimisation algorithm iteration
$v_{m,r}$	Usage of resource r by user m
V	Maximum number of economic parameters related to a user
w	Relay sector (RS)
W_i	Maximum number of RSs attached to a BS i
W	Interference restriction matrix
$w_{i,j}$	Element restriction matrix
x	subcarrier
X	Total number of subcarriers
$x_{i,j}$	Indicates whether sector i uses subchannel k or not.
y	WiMAX slot or LTE scheduling block
Y	Total number of WiMAX slots or LTE scheduling blocks
$y_{i,j,k}$	Indicates if both sectors i and j use subchannel k
z	Candidate site
Z	Total number of candidates sites

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction to Wireless Network Design

There are several factors that influence the design of optimal wireless metropolitan area networks (WMAN). The network can be considerably improved if the network solution found is designed to achieve optimal performance, and also reliable and cost-effective Mobile WMAN networks. This chapter provides a description of the most important design parameters in this process, and an introduction to the general aspects of radio network planning and optimisation. This is followed by a description of the main technology specific issues affecting the wireless network design process and the focus of this thesis.

1.1 The Wireless Network Design process

The number of combinations of network elements and parameters that can be configured during the design process (e.g. antenna tilt, azimuth, station location, power, radio resource management (RRM) parameters) constitutes the solution space. The size of this space determines the degree of complexity of finding appropriate solutions. In WMAN scenarios the number of options is high, so it is very unlikely that the optimal network configuration can be found using a manual method.

1.1.1 Design Parameters

Each configurable parameter in the network can be optimised. However, the network design process needs some inputs:

- Related to the particular scenario (from a map)

- User density
- Traffic demands

After the desired or expected inputs are evaluated, the network designer may need to configure a number of parameters. The main parameters that need to be configured during diverse stages of the WMAN network life are listed below:

- Site location and number

The network provider needs a solution for the number and location of sites from a given set of available candidate sites (CSs). In the example of Figure 1.1, 20 CSs are spread over the example scenario, where each candidate site has different associated costs. The solution selected will determine the infrastructure costs and also the performance of the network.

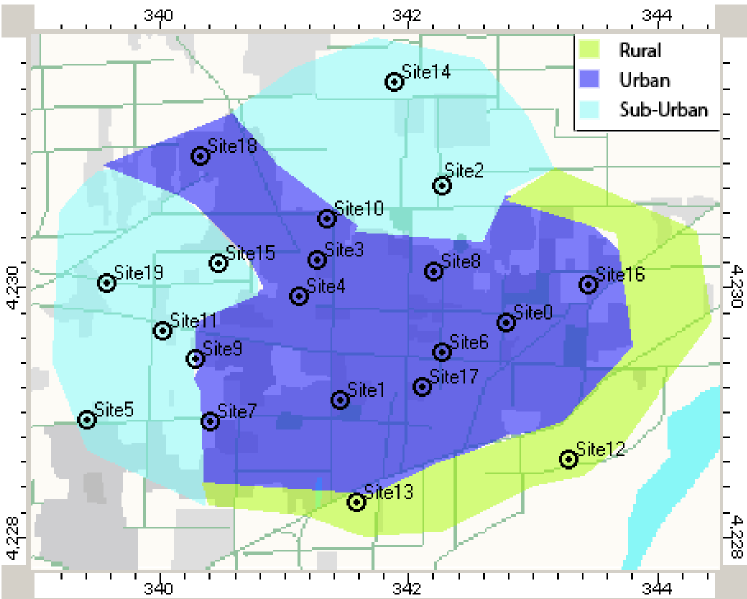

Figure 1.1. Example of a map with three different user densities and 20 candidate sites.

Figure 1.1 also shows an example of traffic maps with three different user densities. In this case, a rural user density corresponds to 10 users/km^2, urban to 40 users/km^2 and suburban to 20 users/km^2 [1].

- Antenna based parameters:
  - Antenna type and its radiation characteristics. The network designer may have many options available in the market. Table 1.1 shows information about beamwidth and gain of some options often used in WMAN network design. Many other configurations are possible.

Table 1.1. Example of antenna configuration.

Beamwidth	Gain
33deg	21dBi
65deg	18dBi
70deg	17dBi
90deg	14.5dBi
90deg	17dBi
120deg	14.5dBi
Omnidirectional	11dBi

  - Antenna tilt and azimuth. The main lobe of the vertical or horizontal plane radiation pattern of an antenna can be aimed to different directions.

    In the case of the vertical plane movement, the simplest way is mechanical beam tilt, where the antenna is physically mounted in such a manner as to lower the angle of the signal on one side. More common is electrical beam tilt, where the phasing between antenna elements is tweaked to make the signal go down (usually) in all directions [2].

    The optimal design of these parameters aims to reduce intercell interference and improve the station and mobile station (MS) power utilisation, while ensuring certain coverage constraints [3].

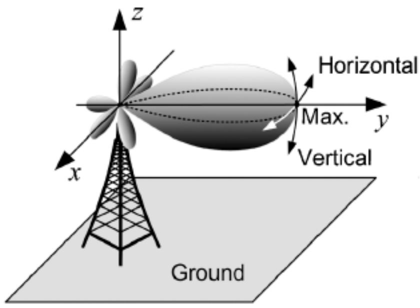

Figure 1.2. Antenna tilt and azimuth [3].

- Antenna height. This parameter depends on the area characteristics and it is related to the site location.

- Pilot power and other powers

Pilot signal is used by the system to relate the most appropriate serving cells to certain MSs. Pilot pollution generally is the result of putting too many cell sites too close together at too high transmit power. The optimal design of this parameter and other powers aims to reduce intercell interference and improve the base station and MS power utilisation, which may result in better quality of service (QoS) and larger battery life of the MSs.

- RRM parameters, system scheduling and frequency planning

Radio resources are allocated by the stations to the MSs in order to satisfy their QoS requirements, while taking the traffic and channel conditions into account.

Furthermore, the system can schedule users according to different strategies, for example, if users with better channel conditions are given priority, the system throughput will normally be higher. Other strategies consider different degrees of fairness.

The bandwidth resources available are limited regardless of the WMAN technology used, and therefore the fine tuning of RRM parameters and the selection of a suitable scheduling strategy determine the performance of the whole system. In addition to this, the set of available resources in WMANs can be reused, and frequency planning strategies may help improve the performance of multicarrier systems. The frequency planning procedure aims to reuse the available bandwidth by using different portions of it within some distance in order to avoid intracell interference, more information is provided in Chapter 5.

- Handover parameters and neighbour cell relations

The nature of this kind of parameters strongly depends on the particular WMAN technology used to provide service, and the way the handover is performed. In any case, an optimal neighbour list and handover parameters avoid unnecessary computation in the network, and therefore extra costs and bad MS performance.

The needed change and configuration of the above parameters during the network design have different costs, for example the cost of sending a technician to a certain station location is higher than the simple change by software of the handover parameters. This fact should be taken into account in this process. Basically, there are two methods of parameter classification related to wireless access network design [4]:

- Parameters which can be remotely controlled, and parameters which require action of human personnel at site location.
- Parameters which changes have only local influences, and parameters which influences other cells.

1.1.2 Introduction to Wireless Network Design

The wireless network design process is normally performed in several stages, listed in Figure 1.3, that consider the specific nature of the configurable network parameters.

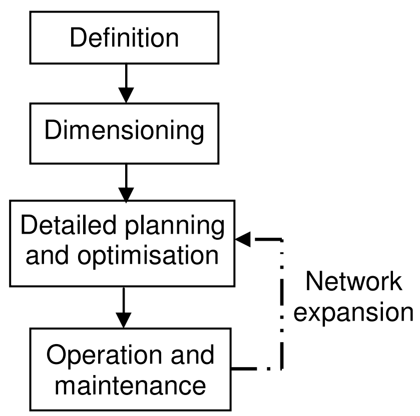

Figure 1.3. Network life phases.

The first stage of the network deployment is the actual definition of the network, also known as preplanning phase. In this phase different criteria is defined, including description of the expected traffic, services, network topology, and deployment scenarios, as well as coverage, throughput and system capacity requirements, relay type, and usage model.

The preplanning phase is followed by a complete network planning and optimization process which consists of other three main phases: dimensioning, detailed network planning and optimization, and operation and maintenance. Each of the three network planning and optimization phases needs to perform a process shown in Figure 1.4. The concept of planning involves finding the estimated quality of the network performance for a given configuration, whereas optimization refers to finding the configuration for which the network quality is optimal. There are several methods and algorithms in the existing literature to perform the latter task that will be described in Chapter 2.

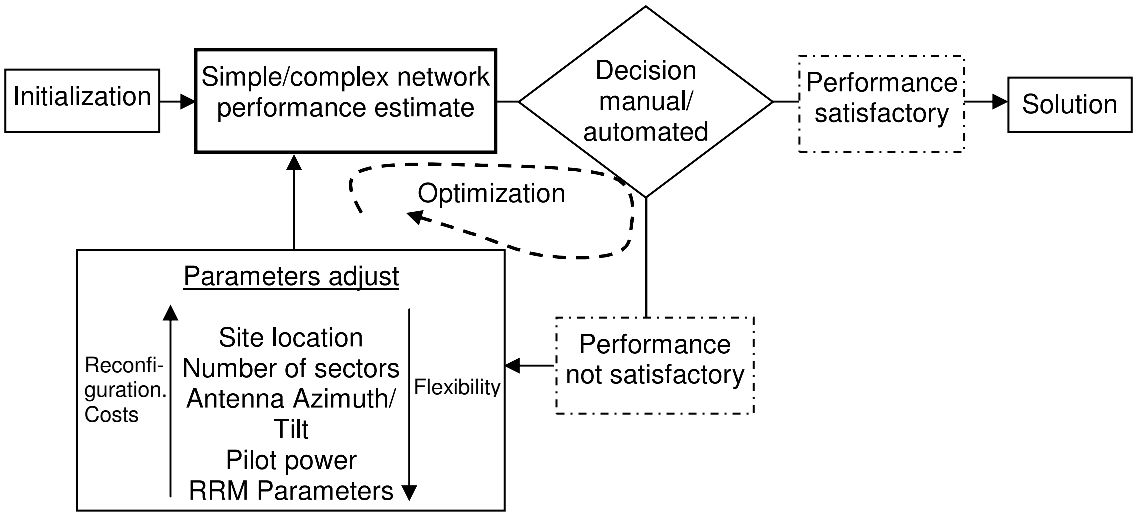

Figure 1.4. Network planning and optimisation process.

In the network dimensioning phase, a first manual network adjustment based on data collection from a digital map and propagation model tuning can be done. Pilot power, coverage range and an estimate of the number of stations are calculated based on carrier to interference plus noise ratio (CINR) predictions.

A following detailed network planning and optimization phase normally makes use of a tool for capacity calculation to predict the network performance of each configuration tested for different parameters with high reconfiguration costs e.g. site location-number, tilt or azimuth. The tool should support traffic, propagation, Radio Resource Management models, and other parameters. The objective of this tool is to calculate Key Performance Indicators (KPI) that can represent the quality of a certain network configuration.

The optimization algorithm test different configurations for different parameters - e.g. station location from a set of candidate sites, antenna tilt or azimuth - that need to be discrete, and the resolution chosen will determine the size of the solution space. The algorithm stops when the cost function based on a series of penalty functions that evaluate different KPIs and network costs, meet the needs of the service provider. Different criteria can be considered in the cost function.

$$\min\{\lambda_1 \cdot F_{pen_1} + \lambda_2 \cdot F_{pen_2} + \lambda_3 \cdot F_{pen_3} + \lambda_4 \cdot F_{pen4} + \ldots\} \tag{1.1}$$

Different strategies can be set by the network designer by tuning the weighting values (λ_1, λ_2, λ_3, λ_4) in order to give more or less preference to different objectives in the cost function [5].

The processes described in the detailed planning and optimization phase are considered off-line since the optimization loop does not include the real network. As mentioned, a tool for network simulation is used instead as it would be infeasible to iteratively reconfigure some parameters that need physical changes. The last operation and maintenance phase can make use of on-line optimization by using real measurement data, and optimize parameters that can be reconfigured by software, such as RRM, powers or handover parameters. The difference is that the KPIs to be penalized in such functions would correspond to real network data instead of the simulation procedure.

The objective of the operation and maintenance phase is to respond to changes in the traffic, architecture or network demands, and further self-healing and self-optimization procedures can be applied.

There are several methods to obtain real network data. Drive test measurements can be performed to check the signal quality over the planned area. Also, network statistics -

such as handovers or dropped calls - and uplink (UL) measurement reports from the MSs can be employed.

Note that a further detailed network planning and optimization can be performed again in the event of a network expansion (Figure 1.3). For example, when more stations may be needed, or if a malfunctioning of the network appear due to a bad network definition and dimensioning.

1.2 Challenges in Wireless Metropolitan Area Networks

The use of optimization techniques for wireless network design is not new, but as new technologies emerge it is very important to accurately identify which factors affect the planning and optimization process, define how to take them into account, and when to penalize them or not under different circumstances.

1.2.1 GSM, GPRS and EDGE

GSM (Global System for Mobile communications) was the first wireless network that was deployed worldwide and made a big impact on the provision of mobile telephony. It is considered the second generation (2G) of mobile networks after some initial studies on this kind of networks. The initial market target was to serve voice users, although some data transfer was also possible.

As it can be seen in Figure 1.5, the spectrum is divided in several channels, and each of these can handle many users. This is done using frequency division multiple access (FDMA) and time division multiple access (TDMA).

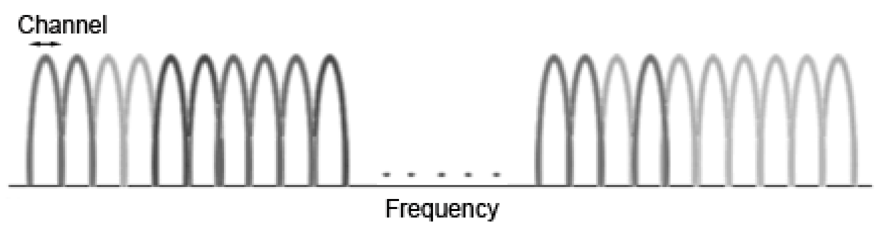

Figure 1.5. GSM Spectrum.

During the initial stages of the implementation of this kind of technology, different concepts such as cell sectorization or frequency reuse were introduced.

In GPRS and EDGE (2.5G), voice circuits are substituted by packet circuits, are different enhancements in this kind of networks increase the achievable throughput.

Network operators faced some radio network design issues for the first time, taking into account elements such as site location and number, antenna configuration, coverage or frequency planning. The moderate bit rates requested by users made the radio network planning and optimisation problem a young but growing topic.

1.2.2 WCDMA and HSPA

UMTS (Universal Mobile Telecommunication System) (3G) was the next disruptive technology. It was created to handle the increased traffic requirements (DL = 350 kbps). It is based on WCDMA (Wideband Code Division Multiple Access), where a series of spreading, scrambling codes and further filtering help the system accommodate all users into the whole available spectrum (see Figure 1.6). The concept of cell breathing, which relates number of users over the same bandwidth and coverage, need to be taken into account [6].

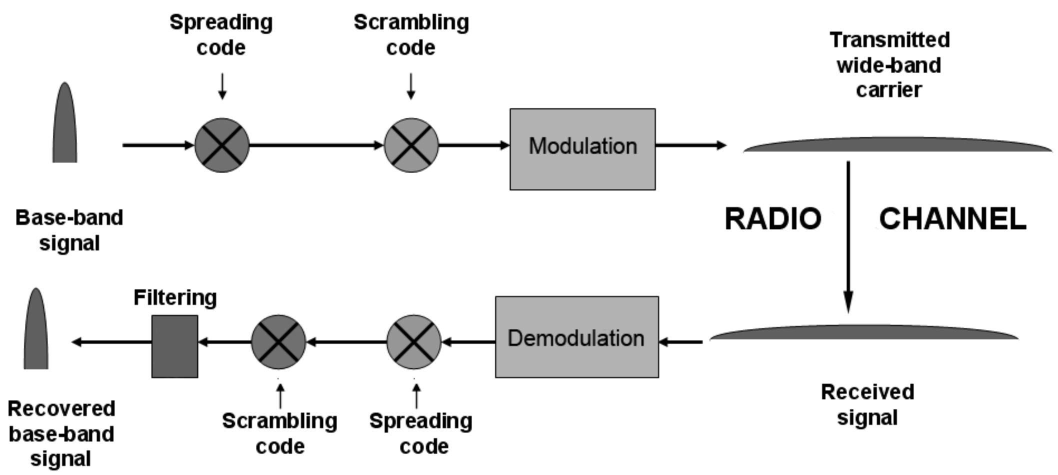

Figure 1.6. UMTS signal spread [6].

Some new and existing network elements may be considered, such as the use smart antennas or repeaters. And concepts such as power control or scheduling have more importance, making the network more complex and difficult to configure. As a result, network planning and optimisation becomes the main challenge of this kind of networks.

This is more critical in HSPA (High Speed Packet Access) networks (3.5G) [7], where the main target market is the data user. New network features are introduced to deal with the increased traffic requirements in both DL and UL (HSDPA - High Speed Downlink Packet Access and HSUPA - High Speed Uplink Packet Access). The main ones are hybrid automatic repeat-request (HARQ), adaptive modulation and coding (AMC), fast packet scheduling, multi-code operation, fast power control and soft handover. These features may be considered in different parts of the network planning and optimisation process.

1.2.3 WiMAX and LTE

Technologies based on OFDMA (Orthogonal Frequency-Division Multiple Access) such as WiMAX (Wireless Interoperability for Microwave Access, [8]) or LTE (Long Term Evolution, [9]) are considered pre-4G or 3.9G technologies.

Two versions of the WiMAX are the most extended within this standard, the fixed and mobile deployment [10]. The first one (IEEE802.16d) is based on a fixed FFT OFDM (Orthogonal Frequency Division Multiplexing) physical layer, while the mobile version of WiMAX (IEEE802.16e) is based on a variable fast Fourier transform FFT OFDMA physical layer. The standard is quite open, for example mechanisms such as QoS and bandwidth allocation are defined in the standard, but others such as scheduling or resource allocation are left to free implementation by the vendors.

The MAC layer of Mobile WiMAX is able to handle different applications with different QoS. It is divided in two sub-layers that classify the incoming packets to a proper connection with the specified QoS, and perform functionalities such as packing, scheduling and resource allocation. Five different types and priorities of service have been defined by the standard: UGS (Unsolicited Grant Service), rtPS (real-time Polling Service), ErtPS (Extended real-time Polling Service), nrtPS (non-real-time Polling Service) and BE (Best Effort).

The option chosen by most of the vendors is time division duplex (TDD) in the PHY layer of Mobile WiMAX, due to its scalable bandwidth [11]. Figure 1.7 shows the TDD frame structure of Mobile WiMAX frame, which is based on OFDMA. Frequency division duplex (FDD) is also supported.

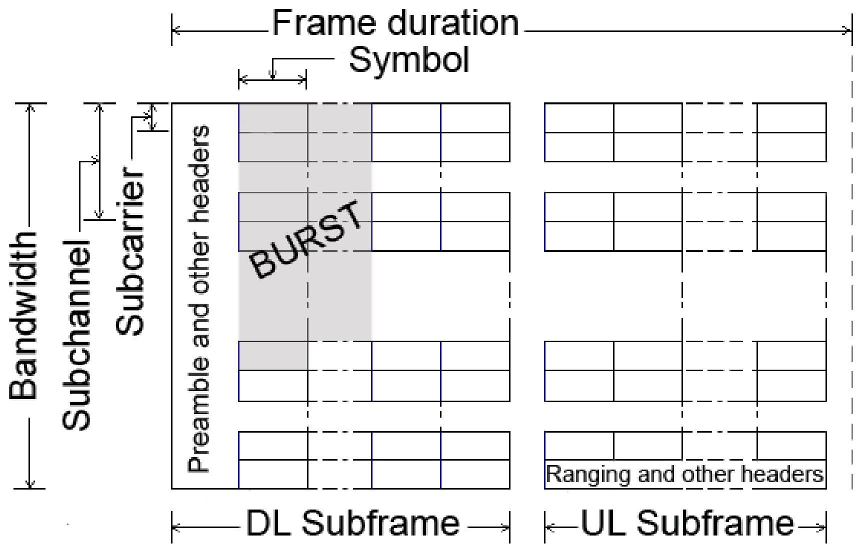

Figure 1.7. Mobile WiMAX TDD OFDMA frame and an example of a user burst.

The available bandwidth is formed by orthogonal subcarriers, which are divided into several groups called subchannels. In the time domain, the frame is divided in several symbols. The slot is the minimum frequency-time resource that can be allocated by the base station. A single user is allocated a contiguous set of frequency-time resources with the same modulation and coding scheme (MCS), known as burst.

The number and exact distribution of the subcarriers that constitute a logical subchannel in Mobile WiMAX depend on the subcarrier permutation mode. For example, the subcarriers that form a subchannel can be adjacent to each other, which is more desirable for beamforming. This allows the system to exploit multiuser diversity by allocating the most appropriate resources from the frame to users according to their particular radio propagation conditions. The subcarriers constituting a subchannel can also be distributed throughout the frequency band, which provides better frequency diversity [11]. This latter case is used in the DL/UL PUSC (Downlink/Uplink Partial Usage of SubCarriers) mode, in which two symbols are used per slot. This mode is mandatory in Mobile WiMAX. The frame can be divided in several zones for each mode, and situate the burst of the user to a particular zone according to channel estimation or other criteria. The optional modes are FUSC (Full Usage of SubCarriers), OFUSC (Optional Full Usage of SubCarriers) and Band AMC (Adapting Modulation and Coding). The TDD frame can be fragmented to accommodate different zones (see Figure 1.8).

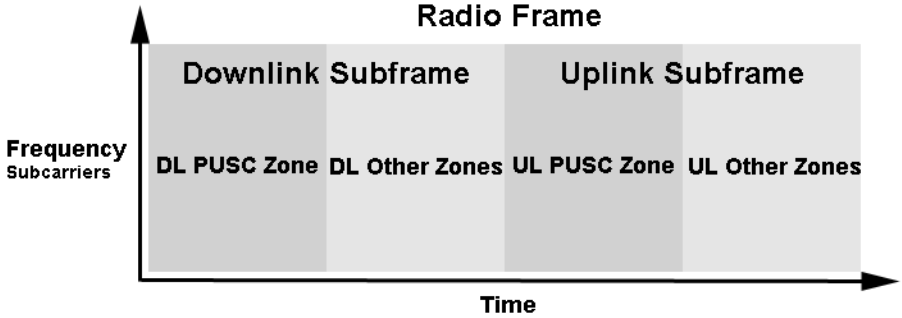

Figure 1.8. TDD frame fragmentation.

LTE describes the standardization work by the Third Generation Partnership Project (3GPP) to define a new high-speed radio access method for mobile communication systems. LTE networks have high complexity, and for the radio access parts it use multiple access schemes on the air interface: OFDMA in DL and SC-FDMA (Single Carrier Frequency Division Multiple Access) in UL, briefly described below.

SC-FDMA is a single carrier multiple access technique which has similar structure and performance to OFDMA. It uses single carrier modulation and orthogonal frequency multiplexing using DFT-spreading in the transmitter and frequency domain equalization in the receiver.

In both cases, subcarriers are grouped into resource blocks (RBs) of 12 adjacent subcarriers with an intersubcarrier spacing of 15 kHz [9]. Each RB has a time slot duration of 0.5 milliseconds, which corresponds to 6 or 7 symbols. Figure 1.9 shows the resource grid. The smallest resource unit that a scheduler can assign to a user is a scheduling block (SB), which consists of two consecutive RBs, spanning a subframe time duration of 1 millisecond 8 [12]. One LTE radio frame consist of 10 ms. All SBs belonging to a single user can be assigned to only one MCS in each scheduling period.

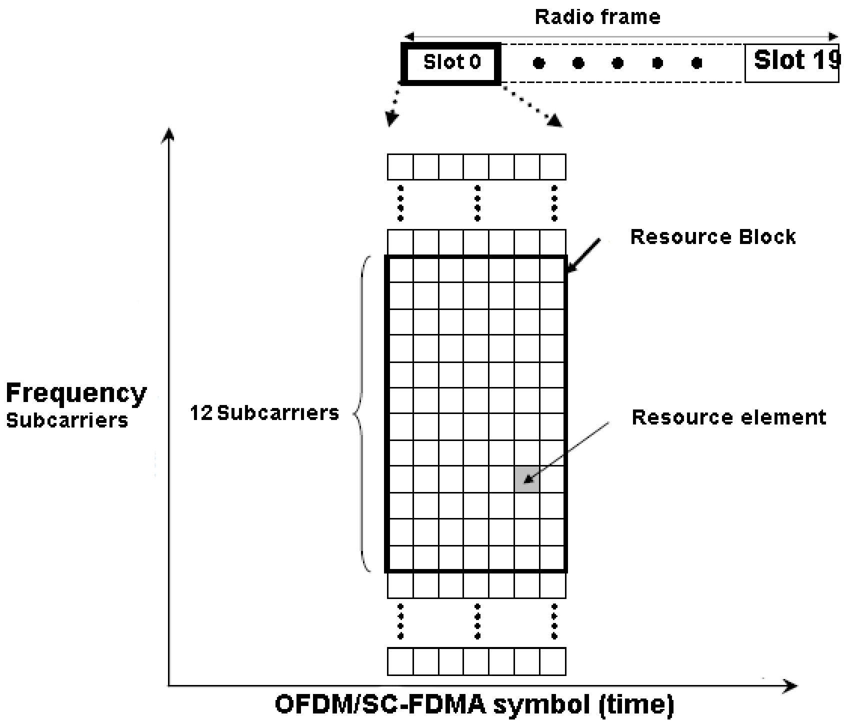

Figure 1.9. LTE resource grid.

In short, OFDMA makes an efficient use of the system capacity due to the exploitation of frequency or multiuser diversity. The impact of the two dimensional allocation described at a system level depends on the bandwidth demand, QoS requirements and channels conditions of a certain user.

Establishing a reliable connection for the users in OFDMA-based networks is not as easy as just assigning enough transmission power for each resource. The reason is that there will probably be a certain level of interference between users depending on the position and the number of stations and the resources available for each of them. Within the same frequency band, some users interfere each other and some others do not.

Each CS (Cs = {1, ...,z...,Z}) in the system may have several sectors (known as base sectors - BS), with some associated costs. BSs belonging to the same CS, will share the costs related to the same location z (BSs: S_z = {1, ...,i,j...,N_z}). In a system with a set of users associated to the active ones: Ms_i = {1, ...,m,g...,M_i}, let us introduce the following notations:

- Set of subcarriers: Sc = {1, ..., x...,X}.

- Set of WiMAX slots or LTE SB: Sl = {1, ..., y..., Y}.

The number of subchannels can be calculated from the efficiency factor η_{sub}, which is the subchannel efficiency expressed in (subcarriers/subchannel). As a result we can define the following:

- Set of subchannels: Sh = {1, ..., k..., K}

  where: $K = \lceil X/\eta_{sub} \rceil$

- Set of resources: Sr = {1, ..., r..., R}

  where: $R = K \cdot Y$

The DL interference on a certain user (I_m) attached to a BS (i) is the sum of the power received from other interfering BSs (j) that are using the same resources ($C_{m,j,r}$), as (1.2) indicates. Depending on the channel conditions and the service requirements, the users will require different number of resources and will be interfered differently, and thus diverse situations may be created.

$$I_m = \sum_{r \in Sr} \left(\sum_{z \in Cs} \sum_{j \in S_z, j \neq i} C_{m,j,r} \cdot u_{j,r} \right) \cdot v_{m,r} \tag{1.2}$$

where:

$$v_{m,r} = \begin{cases} 1 & \text{If resource r is used by user m} \\ 0 & \text{Otherwise} \end{cases} \tag{1.3}$$

$$u_{j,r} = \begin{cases} 1 & \text{If resource r is used in station j} \\ 0 & \text{Otherwise} \end{cases} \tag{1.4}$$

1.3 Focus of this Thesis

The work of this thesis is focused on a detailed planning and optimization phase, where a Greenfield (no previous infrastructure) or expansion scenario is assumed, as described in Section 1.1.2. Off-line optimisation with network simulation are needed in this procedure for parameters that need physical changes to be configured, as it would be infeasible to iteratively reconfigure some parameters in a real network that have high configuration costs, such as BS location. Also for this reason, existing self-optimisation procedures that involve a non-centralized configuration of the network will not be within the scope of this work.

Note that, the detailed planning and optimization phase determines the initial configuration of the radio access physical infrastructure (or the expansion of an existing one). It would involve further investment from the network operator if more changes had to me made on certain parameters. Therefore, this process makes a big influence on the network design from the economical point of view, which will be considered in this work.

Regarding the technology to be studied, this work focuses on OFDMA-based networks, especially on Mobile WiMAX and also with some examples on LTE, where the particular matrix-like interference model should be taken into account when planning and optimizing the WMAN network. This kind of systems are capacity-limited, and make a dynamic use of bandwidth according to the throughput requirements of each service, and several time-frequency slots are allocated to fulfil their needs as described. This fact leads to a quite different simulation compared to 2G (voice-oriented system) and 3G networks (interference-limited system), where the nature of interference is totally different and also the way the situation changes for the CINR and throughput values in diverse network configurations. This affects the network planning and optimization differently.

As new technologies emerge, it is important to identify the factors that affect the design of the wireless access network. However, until now, most existing efforts have been focused on the basic access capability networks. This work presents a way to deal with the trade-offs generated during the OFDMA access network design, and presents a service-oriented optimization framework.

With the introduction of relay stations criteria like coverage extension and capacity enhancement are improved due to the deployment of new cells and the reduction of distance between transmitter and receiver. However, the network designers will also face new challenges, since they involve a new source of interference and a complicated air interface.

Chapter 2 provides a description of the related work and a more detailed motivation of this thesis. Then the study of models and algorithms for classical point-to-multipoint networks and the proposed optimisation framework is presented in Chapter 3. In Chapter 4, this work analyses the impact of the relay configuration on the network planning process, and the evaluation of a new frame structure for this network topology. The analysis of the optimisation procedure with the addition of relay stations and the derived higher complexity of the process is done in Chapter 5. Finally, Chapter 6 presents the conclusions and open issues.

Chapter 2: Wireless Network Planning and Optimisation Methods

In order to successfully compete with other existing and future wireless, cellular and wire-line services, the wireless network designers need to fully consider the specific technical constraints that influence the whole design process. This chapter provides an introduction to the existing wireless network planning and optimisation methods by describing different existing approaches and highlighting gaps in the literature.

2.1 Network Performance Estimation

This section focus on the planning concept introduced in Section 1.1.2 and Figure 1.4. This is an essential stage as it is necessary to predict the network performance of each tested configuration by the optimisation stage.

2.1.1 Measurement Based Performance Estimation

Real network measurements can be used to estimate the network performance during the network optimisation process. It can be used in off-line or on-line optimisation during network operation and maintenance, but normally not with parameters with high reconfiguration costs. An example can be found in [13], where handover parameters are optimised.

It can also be used in conjunction with network simulation in order speed up calculations as described in [14].

In general, the main advantages of measurement based optimisation are that the calculations are based on real network data, not on a defined model, and it is quite accurate. The main disadvantages are the limited spectrum of data available, the fact that data is normally vendor-dependent, and that it is difficult to implement.

2.1.2 Simulation Based Performance Estimation

The use of efficient models for simulation is critical for network design tools. The algorithms used in the optimisation stage normally require thousands of SLSs (System Level Simulations) to find an optimal network solution [15].

- Analytical Modelling

The network performance can be analytically analyzed by using mathematical formulas, and assigning constant values to some expected actions. It is normally used for comparison of computer models or to be used in conjunction with them, which drastically reduces the computing time. An example of the latter case can be found in [16], where the authors validate a simulation model for WiMAX Mesh networks.

- Link and System Level Simulators

In link level simulation, the behaviour of the radio link is studied in detail, considering a complete simulation of aspects such as modulation, coding or fast fading. System level simulation takes into account the behaviour of the whole network, considering factors such as mobility, interference or propagation loss in a larger temporal scale (attenuation and shadowing) [11]. Link and system level simulations may be executed independently, and the interaction between them is done by using interfaces called Look-Up Tables (LUTs). The LUTs are a set of tables that represent the results of the link level simulation, relating to the quality experienced in the radio link connection, such as CINR, with a determined parameter of quality, like the Block Error Rate (BLER). This set of tables serves as input for the system level simulation by extracting from them the radio link information for each user in the system. This procedure represents a simple and efficient way of including the effect of the transmission channel into system level simulations, thus reducing the computational load.

- Static and Dynamic System Level Simulators

Static simulators estimate the network performance of a snapshot (a single time moment). On the other hand, dynamic simulators allow the network to run as a function of time or events, accurately modelling mobility of users, full RRM procedures and handover parameters. This option is accurate but not computationally efficient.

The tools used for planning and optimisation should be able to support intensive SLS. Specialized simulators for this purpose that balance the tradeoff accuracy-computational load are needed in this process.

However, mobile WiMAX network planning is often done in previous works with tools similar to those used in 2nd generation (2G) cellular networks with a few adaptations [17]. All detailed and specific link-level performance factors affecting the SLS, such as the impact of the parallel transmission in OFDMA, adaptive modulation and coding or WiMAX services should be considered in a computational efficient SLS.

The WiMAX SLS Forsk Atoll [1] is a non-open source commercial software. It provides an extensive model of a Mobile WiMAX network, but important weaknesses can be found e.g. the interference prediction. Other commercial planning tool for Mobile WiMAX SLS presented in [18] and [19] are ATDI and EDX Signal Pro respectively. This software contains a powerful graphical environment for predictions, but the computational load of its predictions is too high and does not support traffic modelling for actual network simulation. A Mobile WiMAX SLS is proposed in [17] where the authors develop a PHY abstraction methodology. This solution assumes a dynamic simulator only prepared for performance evaluation. There is also available different research software for network simulation such as NS-2 [20] or OPNET [21]. Nevertheless, due to its complexity they are more suitable for link level performance [2], [1].

2.2 Optimisation Techniques

The radio network optimization problem is NP-hard; hence it is not possible to find a theoretical optimum in polynomial time [22], [23]. Different methods can be applied to solve the optimization problem presented in form of Mixed Integer Program [5] that minimizes the cost function (1.1).

2.2.1 Metaheuristics

Metaheuristics are algorithms based on a search within the solution space. They seem a common approach in related works to solve the network planning and optimization problem as they can provide close to optimal solutions in reasonable time [5][24]. They do not guarantee the optimality of the solution, but they often perform very well in practice when the number of variables becomes large.

Base station location, antenna azimuth and tilt optimization have been widely investigated with the analysis of these methods, as main part of advanced network planning and optimization tools, especially in the case of UMTS, CDMA based networks [5][23][25][26][27].

Few works specialized on mobile WiMAX [28][29] and LTE [30][31] have been found. They focus on special cases or self healing/self optimization capabilities of an already deployed network, and do not consider a precise formulation and computationally efficient methodology on multiobjective problems.

In the following we describe some metaheuristics widely used in the related literature.

- GA (Genetic Algorithm)

GA facilitates searching across a problem space via provided recombination and selection operators in which the problem states are altered [32] [33] . Mainly, crossover and mutation operators are used to recombine child solutions, and selection and replacement rules are utilized to evolve a population of solutions.

- SA (Simulated Annealing)

SA explores across the whole search space of the problem under taken throughout a simulated cooling process. Every state promoted for the next step is a consequence of a probabilistic rule, which provides with a better state of solution or creates a perturbation to prevent the possible local minima. An example can be seen in [34].

- VNS (Variable Neighbourhood Search)

In the recent VNS [15] solutions are manipulated through two nested loops in which the core one alters and explores via two main functions so called shake and local search. The outer loop works as a refresher reiterating the inner loop.

- TS (Tabu Search)

In TS (Tabu Search), the main idea is to arrange moves to the best of the neighbourhood regardless of the solution quality of the state undertaken [35]. A Tabu list that stores the sequence of moves done is used to avoid loops during the search. The algorithm does the non-Tabu moves within the neighbourhood structure by selecting other configurations.

2.2.2 Greedy Algorithms

A variety of less sophisticated Greedy search algorithms are used obtain good approximate solutions within a reduced amount of computing time. Information about insertion and removal based algorithms can be found in [36]. These algorithms run incomparably faster than metaheuristics, which is a desirable feature for on-line solution, but may have poor performance in complex problems with big solution spaces.

2.2.3 Integer Linear Programming

Integer Linear Programming (ILP) is used to determine the best solution of optimization problems, given a list of requirements represented as linear equations. These methods guarantee the optimality of the solution, or the proximity to the optimal solution. However, they could be time consuming and computationally expensive.

An example of ILP is Branch and Bound (BB) [37]. Discrete programming models for cellular network optimization are proposed. The problem is to select an optimal set of BS locations from a given pool of configurations. However, this solution assumes quite a general situation for several technologies and has not considered some economical (BS costs) and technical constraints.

2.2.4 Multiobjective Optimization

The weighting of the different objectives in the cost function (1.1) may result in a difficult task during the network planning and optimization process, especially when the complexity of the network is high and more decisions need to be made, with objectives that incur in different trade-offs. For example, a final design with low infrastructure costs may have worse performance than another with more active stations.

The optimization problem can be solved by using the solutions of the Pareto front [38], where the network provider must select a posteriori the most appropriate ones according to some policies. The Pareto optimal solutions are called non-dominated. Each of the obtained solutions will represent a certain optimal trade-off between the different factors in the cost function as it is not possible to improve one of them without worsening the others.

Figure 2.1 illustrates the solutions in an example of a two-objective minimization problem, where the stars represent the set of non dominated Pareto optimal solutions of rank 1. The rank of a solution n is defined by $Ra_n = 1 + So_n$, where So_n is the number of solutions by which n is dominated in the set of feasible solutions. To obtain the solutions of Pareto rank Ra_n, the solutions of rank $Ra_n - 1$ have to be removed. The solutions are iteratively calculated by an optimization algorithm in three main stages:

- Search front expansion. The algorithm searches for neighbour solutions in the current search front. For example, a neighbour solution is obtained by removing/changing the position of one BS, or by changing the antenna azimuth/tilt.
- Update of the optimal Pareto front. The non-dominated solutions of the neighbourhood are selected.
- Selection of the new search front. This is done according to the optimization algorithm methodology. Many of the algorithms described in section can be adapted to multicriteria optimisation.

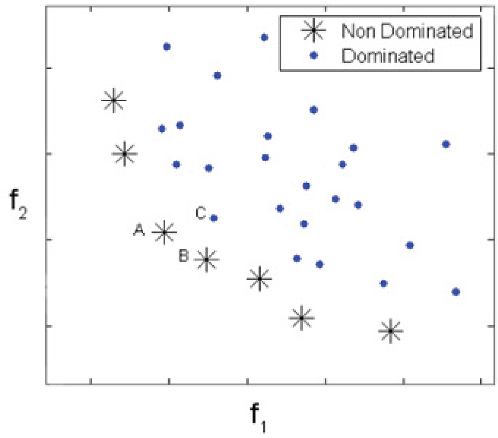

Figure 2.1. Example of a 2 function minimization problem with dominated and non-dominated solutions. Solution C has become dominated with the addition of solution A and B

In [38], the authors first present a practical case of multiobjective optimization framework on wireless networks, as described in [39]. This is a study on WLAN technology (802.11g), where the authors do not take into account network simulation and other important issues that influence the design of WMAN networks. In this kind of networks the complexity is higher, and the objectives of the cost function should not be fully correlated in order to offer valuable information about the different solutions found. In addition, the Pareto front may contain a high number of non-dominated solutions.

2.3 Wireless Access Network topology

The network topology has a direct impact on the complexity of finding an appropriate solution. Multihop relaying networks involve a more complicated air interface than other kind of cellular networks, and need to be rigorously analyzed during the process of network design to achieve an outstanding performance.

2.3.1 Point to Multipoint

Most of the works in the existing literature focus on the traditional PMP topology for cellular networks, where a BS provides direct service to MS in line of sight (LOS) or non line of sight (NLOS) conditions. The main contribution on the network planning and optimisation field was done by [5], where the authors study the optimal base station location problem considering technology critical and specific issues of UMTS such as power control. The degree of complexity of the problem is evaluated and several options to solve the problem are described in the form of mixed integer linear programming problem and the use of heuristics. Some important further contributions to this work where done by the mentioned works [23][37], that give a more mathematical and solution-oriented approach. Some other examples are the works mentioned in Section 2.2.1. Also, in [15][40][41] some mathematical models for UMTS radio network planning are presented, and different single-objective optimization strategies are analysed in a comparative investigation. In those works, the BS location problem is modelled as a simplified problem, and parameter tuning for different metaheuristics is presented.

With regard to OFDMA networks, a deep economic study of the problem is made in [28], but only a few physical constraints are taken into account, such as the size of the cells and the number of subchannels used. Others like the network performance or interference prediction are not considered. In [29], four different algorithms including Tabu search are compared for automatic cell planning in Mobile WiMAX networks. However, the authors mainly focus on the economic problem without describing how to calibrate the different factors in the cost function. This model does not consider most of the effects of the particular features of WiMAX on the planning process.

2.3.2 Multihop Communications

The situation is more delicate for the multihop option, where the network complexity is higher. Multihop communications at WMAN level were introduced in the 802.16a version of the WiMAX standard as mesh mode, and it will be integrated in the mobile version (802.16e) in the upcoming release 802.16j as relay topology [42]. Figure 2.2 shows the difference between the PMP, mesh and relay topologies.

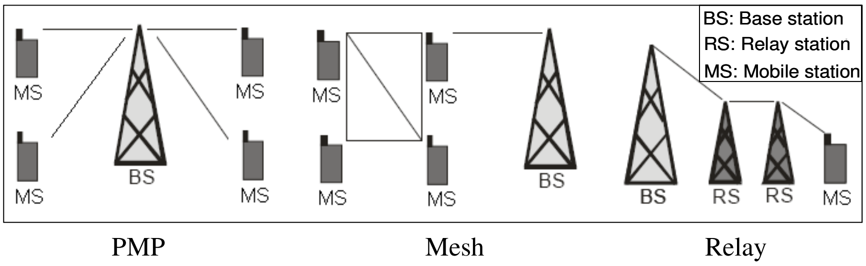

Figure 2.2. Different topologies in WiMAX.

- Mesh

At the moment of writing this work the mesh mode has become obsolete within the WiMAX standard. The main problem in the mesh approach is that the standard PMP MAC (Medium Access Control) frame structure needs to be replaced in order to allow direct communication between MSs. The Relay mode in this standard (802.16j) will overcome this problem by allowing just tree-based multihop communication without making any change in the PMP frame or any modification in the MSs, and thus achieving full compatibility for the MAC procedures.

In [28], a deep economic study of the problem and comparison between PMP and Mesh is done. The authors in [43] propose a special case of 802.16a mesh topology design of 3G backhaul network, in which many factors of the generalised version of the problem are not considered, e.g. interference or site costs. Some other works like [44] study how throughput per cell decreases faster than linearly with the increasing number of cells served by a single wired BS, or how the planning a dense mesh network is coupled with routing the traffic towards the BS in [45].

- Relay

The multihop relaying configuration aims to deal with challenging radio propagation characteristics and low CINR at the cell edge with the introduction of relay sectors (RSs). RSs forward the information from some MSs to the legacy BSs that are connected to the core network. The BSs will provide resources to the RSs and other MS to which are directly connected.

Figure 2.3 shows different usage scenarios for RSs. The RS is able to work in LOS and NLOS propagation conditions, and supports MS handover. The use of RSs can help overcome the dependency on wired backbones, and therefore can reduce the infrastructure costs. It can also extend the coverage beyond the BS range, increase the quality for indoor coverage, deal with shadowing in urban scenarios, balance the network load, help the terminals to save battery and it may increase the overall system capacity, which can be particularly useful in hotspots. It can also provide broadband access in emergency situations or to locations in rural and developing areas where broadband is currently unavailable.

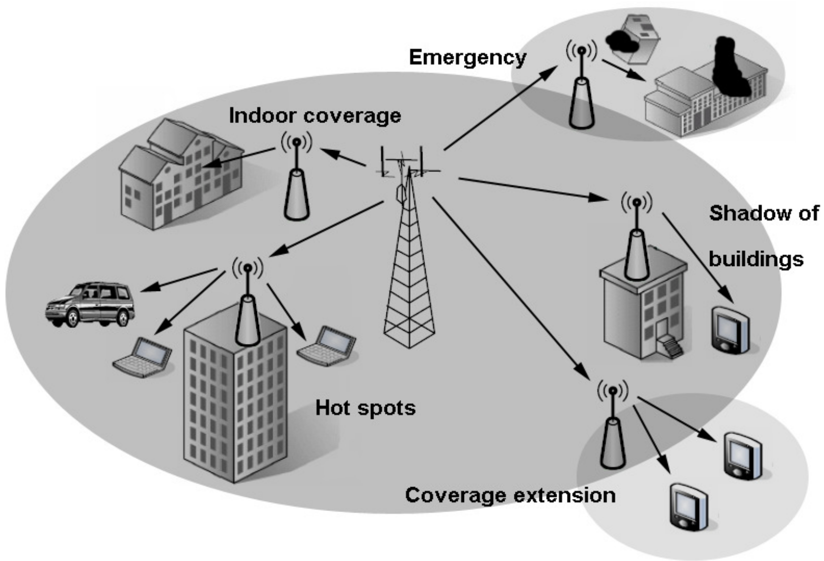

Figure 2.3. Example of usage scenarios for BS and RS.

The concept of cooperative relaying is based on one signal that can be received and forwarded by multiple stations to the same MS, which combines the signals received from the RSs, and possibly from the BS, in order to take advantage of channel diversity gain.

In this kind of topology, the number of parameters to be optimised is bigger, and also the number of constrains. As a result, in a RS enhanced network the solution space is much larger since decisions about the number, position and configuration of the RS have to be made in addition to the PMP parameters. Also the estimation of the network performance is more complex, and if inaccurate, it can lead to non-optimal solutions. New constrains have to be considered, such as the number of hops in the communication or how the link BS-RS is performed.

Some methods and concepts to integrate multihop communication into the WMAN level are currently being proposed and studied. These methods involve some modifications in RSs, and are similar to concepts previously proposed for OFDM based technologies such as HiperLan/2 [46] or the Fixed version of WiMAX [47][48], where the authors describe the "subframe concept" that implies no special standard options. Some methods has also been sketched out for WiMAX relay in [49], where a scheme to improve the standard efficiency is proposed, by reducing the heading payload. These options are also being considered by the IEEE 802.16j task group [42].

The literature on OFDMA of relay networks in the field of planning and optimisation is not extensive, nor very complete. A first approach was done in [50], where the authors solve the WiMAX optimization problem as a MILP problem with some assumptions to simplify the process for a tree topology. Some works like [51], [52] study the particular case of a small system (with one RS in the former case), and analytically evaluated. In [53], the authors focus on the analysis of optimisation algorithms in a generalised view of the problem. Others like, [54], [55] are geometrical approaches for regular scenarios for one cell or a small hexagonal network.

2.3.3 Picocells and Femtocells

As mentioned, the use of RSs can help increase indoor coverage. Another possibility to achieve this is to increase the power of the outdoor base stations or to add more cells. However, this is expensive, and sometimes unfeasible due to high interference or health regulation issues. In addition, there are other specialised options, which consist of adding antennas directly inside the buildings:

- Distributed antenna systems
- Picocells (Indoor base stations - installed by the operator)
- Femtocell (Home base station - installed by the user)

The installation of these systems is normally done for specific purposes in an already deployed network, and therefore their configuration is normally done manually or with self-configuration procedures.

Some recent studies point that the use of femtocells can be substantial in the future [31]. On the other hand, the configurable parameters of these devices do not normally need physical reconfiguration or high costs. The reason is that they are installed by the user with omnidirectional pattern. Some self-optimisation methods for RRM parameters can be applied, but these will take place in the network operation and maintenance stage.

2.4 Frequency Planning

There are some further important challenges in multihop networks. A major concern for network designers is the management of interference, which can degrade the performance of the system considerably. In multihop networks this is a critical issue, since the addition of RSs constitutes a new source of interference, and therefore its power, frequency and antenna configuration must be carefully selected to avoid interference.

In relation to this topic, some approaches on multicarrier technologies such as GSM suggested the use of radio frequency planning (RFP) in order to mitigate interference and enhance the system performance [56] [25]. RTP has been traditionally performed in an operation and maintenance stage after an initial detailed network planning and optimisation. It aims to reduce system interference level by assigning channels to different cells, and reused within some distance.

Similarly, OFDMA supports reconfiguration of the subchannel usage with different frequency reuse schemes in order to mitigate intercell interference. Some of these strategies are proposed in the existing literature also on OFDMA multihop communications. However, they only consider fixed reuse patterns or static assignment of subchannels in regular or hexagonal scenarios [57][58][59], and do not describe a joint formulation with other parameters such as antenna tilt/azimuth. Another basic approach is shown in [60] for SDR networks, with many assumptions related to the particular nature of the technology studied.

A traditional frequency planning procedure performed as subchannel planning has also been proposed to OFDMA PMP networks [36]. In such case, the use to a certain subset of subchannels is assigned to different sectors. However, with the addition of RSs, the corresponding BS frames need to be fragmented in many cases in order to provide resources to both serving and relayed stations. RS and BS may also interfere each other, resulting in a complex situation not yet fully formulated and evaluated in changing environments.

2.5 Motivation of this Thesis

Traditionally, the wireless network planning and optimization process has been focused on the basic access capability of networks. This thesis offers a new perspective for this process and presents optimization framework that describes a way to deal with the trade-offs generated during the WMAN access network design, with consideration of technical and economic factors. SLS and radio access network optimisation are integrated in an efficient technology-specific and a service-oriented process.

An access-oriented design (used in current important works, e.g. [61]), may not be a good solution because of the following reasons. Some of these designs may not be efficient for areas with high number of users and with different service demands, which often result in a bottleneck for the rest of the network. In addition, for many existing designs that are developed to gain basic access, the customer may not be able to obtain the desired quality-of-service (QoS). Also, the access-oriented design may not be fair to the provider who develops the infrastructure, because the service provider only earns the access fee, which is usually paid monthly and is relatively low compared to the deployment cost.

Multihop relaying networks involve a more complicated air interface than other kind of cellular networks. This work will also explain the impact of the new elements of the WMAN access network architecture, the relay stations, in the network design process. With the use of examples for capacity and coverage planning the thesis will highlight the complexity of the design process in this kind of networks. A frequency planning procedure for this kind of networks will be presented.

With the increasing need of supporting new services and applications for different scenarios or multiple objectives, the problem of wireless network architecture design becomes too large in scope to be handled efficiently with a single technique. This work presents a multiobjective optimization framework that provides a clear and comprehensive description of different options and solutions to achieve an optimal network configuration.

Chapter 3: Multiobjective Optimisation Framework for WMAN Design

As described in Chapter 2, few works specialized on mobile WiMAX [28][29] and LTE [30][31] have been found. They focus on special cases or self healing/self optimization capabilities of an already deployed network, and do not consider a precise formulation and computationally efficient methodology on multiobjective problems. In [38], the authors first present a practical case of multiobjective optimization framework on wireless networks. This is a study on WLAN technology, where the authors do not take into account network simulation and other important issues that influence the design of WMAN networks.

This chapter introduces a multiobjective optimisation framework that represents a financial model reflecting the cost and revenue of OFDMA-based networks over suitable return periods in order to deal with different tradeoffs. Adding or removing a BS incurs in costs that cover the installation (CapEx – Capital Expenditure) and annual management (OpEx – Operational expenditure). This aims to offer the network provider with a practical and flexible way of optimizing the network in the network design stage that is focus of this thesis, where economical factors have a big impact on the parameters to be configured.

3.1 Radio Access Network Planning and Optimization: A Service-Oriented and Technology-Specific View

In the wireless network design process it is very important to accurately identify which factors affect the planning and optimization process, define how to take them into account, and when to penalize them.

3.1.1 Objectives of Network Planning and Optimisation

The service-oriented and technology-specific objectives need to be reflected in the cost function (1.1). Each technology may have specific features, but this cost function should have factors that reflect the quality of a certain network configuration. These factors are known as KPIs. The following KPIs affect the design process in different ways when optimizing an OFDMA-based network.

- Coverage

An MS is covered if the received pilot signal can be demodulated by the terminal. Therefore, as indicated in (3.1), the received pilot power ($C_{i,pilot}$) must be higher than the sensibility of the terminal (S_T) plus the noise power (N_{pilot}). The MS checks the signal from the best server ($i \in S_z$), from the subset of installed CSs ($z \in Cs', Cs' \subset Cs$).

$$C_{m,i,pilot} \geq S_T + N_{pilot} \qquad \textbf{(3.1)}$$

A cellular network optimization methodology that is only based on coverage and network costs is described in [37] as a general solution for wireless networks optimization. The AMC feature, which is present for example in Mobile WiMAX, is not considered in such methodology. However, it plays an important role in the coverage procedure of this kind of networks. The inverse relationship between data rate (throughput) and BS coverage range, together with limitations on range imposed by transmit power, may result in deployments with large numbers of BSs to cover a given area. This results in a trade-off between throughput, coverage, network cost and interference which has to be considered and balanced when optimizing an OFDMA-based network.

- Radio access infrastructure costs

The aggregate of the installation and annual maintenance fee for the sites and sectors (F_{costs}) are represented in the cost function (1.1). The integer variables s_i and s_z in (3.2) indicate if site i and sector z are selected in the solution.

$$F_{\cos ts} = \sum_{z \in Cs} \left(f_{\cos ts}(z) \cdot s_z + \sum_{i \in S_z} f_{\cos ts}(i) \cdot s_i \right) \tag{3.2}$$

where:

$$s_i = \begin{cases} 1 & \text{If sector i is used in the solution} \\ 0 & \text{Otherwise} \end{cases} \tag{3.3}$$

$$s_z = \begin{cases} 1 & \text{If site z is used in the solution} \\ 0 & \text{Otherwise} \end{cases} \tag{3.4}$$

- Interference

As mentioned in section 1.2.3, the level of interference between users depends on the position and the number of BSs and the resources available for each of them. Within the same frequency band, some users interfere each other and some others do not. The particular interference model of OFDMA should be taken into account when planning and optimizing a network. Full network simulation is needed, instead of estimation of adjacent cell overlapping estimation as described in [62], [56].

There are many options for interference avoidance in OFDMA-based networks. For example in a Mobile WiMAX network with a 1x3x1 frequency scheme [11], i.e. frequency re-use factor 1 and three sectors per site, the total number of subchannels and symbols in the frame are available to all sectors, and the interference is separately calculated in all of them. The orthogonal nature of OFDMA can deal with interference. Another common configuration is 1x3x3, where the OFDMA frame is fragmented in three parts and allocated to the three sectors of each site. This reduces the amount of resources available at each sector, but it aims to reduce interference and therefore increase CINR levels, and therefore, possibly higher user throughputs. In addition, as mentioned in Section 2.4, a frequency planning procedure in the form of subchannel planning has been proposed to OFDMA networks, where the use to a certain subset of subchannels is restricted to different sectors, also with the aim of reducing interference.

The way resources are reused in the technology studied play an important role in the network design process. For example, it can happen that an increase in the number of

BSs reduces the general interference level of the system, while it is possible that increases in other kind of networks in the same situation. This should be accurately modelled in the SLS and calibrated in the cost function.

In relation to this, some assumptions can be taken in order to simplify the network planning and optimisation procedure. In the example described in Figure 1.1, the 20 candidate sites of the scenario can be grouped into four different areas where services need to be provided (Figure 3.1). Only one or none candidate site is selected from a region to install a BS. With this simplified network scenario, the optimization procedure of the algorithm turns to be more affordable. Moreover, significant solutions are not expected to be dismissed when using this method, since it sets an upper bound in the number of BS installed and thus, helps avoid testing solutions with high interference levels or high costs. The process of selecting the number and the appropriate areas corresponds to the first manual adjustment of the network dimensioning stage described in Section 1.1.2.

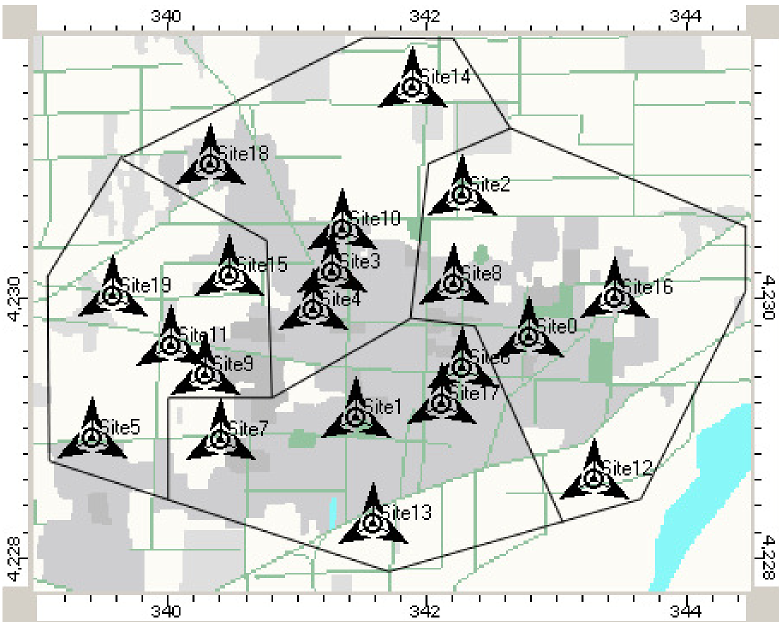

Figure 3.1. Example of candidate sites grouped in four planning regions.

Therefore, in the network dimensioning, the whole area under consideration can be divided into several regions (Rb = {1,...,b...,B}), and each region allows one CS to be installed in the final solution, such that:

$$\sum_{z \in Cs} s_z \leq B \quad \textbf{(3.5)}$$

In order to specify the maximum number of CSs to be installed in a region in the final solution, let us define n_b as the maximum number of CSs that can be installed in a region b in the final solution. In the general case it can be assumed that:

$$\sum_{z \in Cs} s_z \leq \sum_{b \in Rb} n_b \quad \textbf{(3.6)}$$

- Quality of Service

The effective throughput per user can be the quality performance indicator considered in the network design process, as it is directly related to the user experience. Other interesting performance indicators for the QoS are packet loss and delay. However, these factors also depend on the quality of the channel and there is no need to evaluate them separately to have a description of the user QoS. Most importantly, the simulation considered will be static as it is described in next section, and no real time indicators are possible. The effective throughput of the user has to meet a target throughput defined by a certain service, as shown in (3.7).

$$T_m \geq T_m^{(\min)} \quad \textbf{(3.7)}$$

where a certain user m should have at least a throughput $T_m^{(\min)}$ to be satisfied, which is determined by the service class used [11].

The choice of the factors and criteria to be used in the cost function is critical in multiobjective wireless network optimization, and the calibration must be specific to the technology used. The reason is that different technologies can show different performance under the same situations when changing the configuration or the position and number of BSs, and this has to be evaluated in the optimization framework.

3.1.2 Formulation of the Problem

The algorithm of the optimization stage (see Figure 1.4) test different configurations for different parameters, and stops when the cost function based on some costs and a series of penalty functions based on different KPIs ($G = \{G_1, G_h ..., G_H\}$) meet the needs of the service provider. Different criteria h can be considered in the cost function as KPI (obtained in the planning stage).

The cost function (3.8) summarizes the objectives described in the previous section in a three-objective minimization problem: infrastructure costs, user CINR (represented as Cu), throughput (represented as Tg), and therefore G = {Cu, Tg}

$$\min\{\lambda \cdot F_{\cos ts} + F_{pen(Cu)} + F_{pen(Tg)}\} \quad \textbf{(3.8)}$$

The first objective, F_{costs}, represents the aggregate of the installation and annual maintenance fee for the stations as described by (3.2) and no special penalty function is applied.

$F_{pen(Cu)}$ represents the wireless connection by penalizing the effective CINR perceived over all resources uses by each MS, and it includes the coverage and interference avoidance objectives described in previous section. Note that, different MSs can have the same ”connectivity”, by having similar average CINR over their set of resources, but they may require different number of resources (slots on Mobile WiMAX or SBs on LTE) and thus, different final throughput. Therefore, a separate indicator – $F_{pen(Tg)}$ – is needed in the cost function, which ensures a certain throughput, and thus a quality of service (QoS) to the MSs with different services.

The different criteria values (Cu, Tg) are evaluated over all the users, with the use of penalty functions calibrated by the network designer. A binary variable $d_{m,h}$ indicates if user m is evaluated by the cost function related to criterion h that has a value of $t_{m,h}$ (see (3.9)). The setting up of this variable for the penalization of different users is done in the preplanning phase according to the network designer estimate. The selection is then done by software in the framework used.

$$F_{pen(G_h)} = \sum_{z \in Cs} \left(\sum_{i \in S_z} \sum_{m \in Ms_i} f_{pen(G_h)}(t_{m,h}) \cdot d_{m,h} \right) \quad \textbf{(3.9)}$$

where,

$$d_{m,h} = \begin{cases} 1 & \text{If user m is evaluated by the cost function related to criterion h} \\ 0 & \text{Otherwise} \end{cases} \quad \textbf{(3.10)}$$

A threshold-based function avoids the excessive influence of very good or very bad users. Maximal penalty is applied when the value of the criterion $t_{m,h}$ from user m is smaller than a lower bound T_{min} and no penalty exists when it is higher than a threshold T_{max}. In between these two values the function is linearly decreasing as (3.11) and Figure 3.2 indicates.

$$f_{pen(G_h)}(t_{m,h}) = \begin{cases} 0 & if\,(T_{m,h}^{(\max)} < t_{m,h}); \\ f_{m,h}^{(\max)} \cdot \left(\dfrac{t_{m,h} - T_{m,h}^{(\max)}}{T_{m,h}^{(\min)} - T_{m,h}^{(\max)}} \right) & if\,(T_{m,h}^{(\min)} < t_{m,h} \le T_{m,h}^{(\max)}); \\ f_{m,h}^{(\max)} & if\,(t_{m,h} \le T_{m,h}^{(\min)}); \end{cases} \quad \textbf{(3.11)}$$

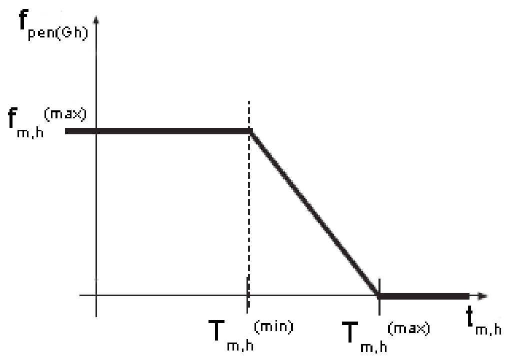

Figure 3.2. Threshold based penalty function for minimization.

Different strategies can be set by the network designer by tuning the values (λ, $f^{(max)}$, $T^{(min)}$, $T^{(max)}$) in order to give more or less preference to different objectives in the cost function (3.8).

By minimizing the described cost function, high costs are penalized as well as a bad performance. By considering the three objectives (Costs, CINR and Throughput), the system:

- Minimizes the number of BSs;

- Ensures coverage for users;
- Maintains separation between BSs;
- Ensures certain QoS.

The combination of both analytical analysis and static simulation can be considered to calculate the values of the cost function and thus, improve the speed of the network optimization process.

For the system-level simulation part of this study – which is iteratively called by the optimization algorithm – a software called Forsk's Atoll [1] SLS is used. Using an application programming interface (API) based on Visual Basic Script (VBS) and C++, this Mobile WiMAX and LTE simulation tool is integrated into the optimization framework according to the scheme shown in Figure 3.3. The optimization algorithm evaluates a cost function that contains KPIs representing the quality of the tested network at each iteration, and updates the optimal network configuration.

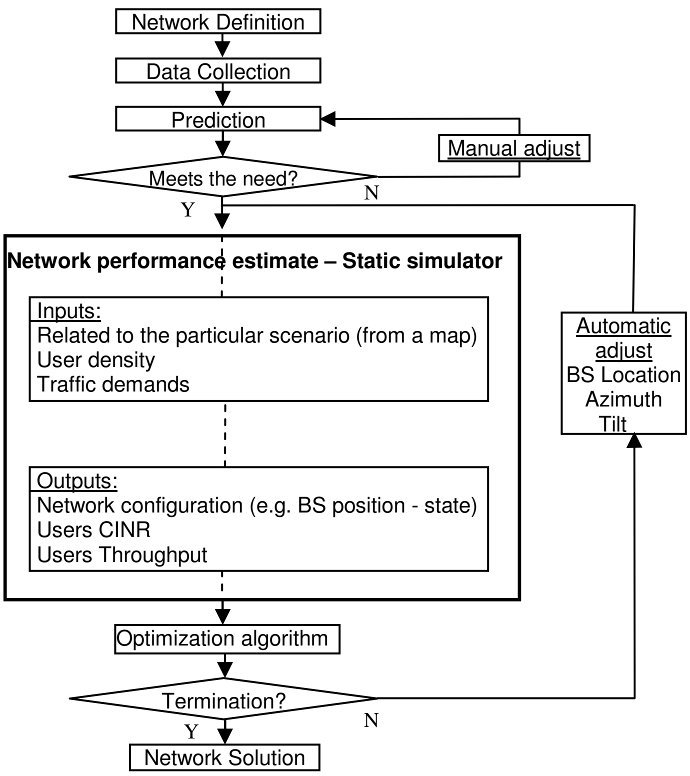

Figure 3.3. Optimisation framework.

3.1.3 Planning Criteria for OFDMA Networks - An Economic Perspective

As mentioned, the network deployment stage focussed in this work is known as detailed network planning and optimisation, and a Greenfield scenario (no previous infrastructure) or a network expansion is assumed. The influence of economic costs on this stage is high, since the parameters to be configured have high reconfiguration costs (Figure 3.3). An economic perspective can be given to the design activity as follows:

$$Profit = ProfitEst - CostAdj \tag{3.12}$$

In this model, some expected profits can be estimated during the network definition phase, and represented in ProfitEst in (3.12). This factor includes ideal estimations for the profit in a business case and it considers license acquisition costs, and installation and maintenance costs of the core network. The optimization process tries to minimize a cost adjustment factor, CostAdj, which includes access network costs (F_{costs}) and some penalty costs for the performance of a particular network configuration ($F_{pen(Gh)}$ see (1.24)). These penalties balance the ideal estimations done in ProfitEst.

$$\min\{CostAdj\} = \min\left\{ F_{\cos ts} + \sum_{h \in G} F_{pen(G_h)} \right\} \tag{3.13}$$

The problem now is to select candidate sites from each region to install BSs with a certain antenna configuration such that the traffic capacity and the number of covered MSs are maximized with the lowest installation cost and a resilient network.

Infrastructure costs, CINR (Cu) and throughput (Tg) are considered in (3.14)) in a three objective minimization problem. The different criteria in the cost function can be grouped according to the main KPIs. Different subsets of criteria related to the particular KPI nature will determine different constraints to the problem, e.g. different throughput in services or different channel profiles, which are shown in (3.15), (3.17), (3.18) (3.19), and described below. Another example is when, the MSs will be either in transmission (UL) or reception (DL) state and the designer may want to penalize differently these two states.

$$\min\left\{\sum_{z\in Cs}\left(f_{\cos ts}(z)\cdot s_z+\sum_{i\in S_z}f_{\cos ts}(i)\cdot s_i\right)+\ldots\right.$$

$$+\sum_{z\in Cs}\left(\sum_{i\in S_z}\sum_{m\in Ms_i}\sum_{h'\in Cu}f_{pen(Cu_{h'})}(CINR(dB)_{m,h'})\cdot d_{m,h'}\right)+\ldots \quad \textbf{(3.14)}$$

$$\left.+\sum_{z\in Cs}\left(\sum_{i\in S_z}\sum_{m\in Ms_i}\sum_{h''\in Tg}f_{pen(Tg_{h''})}(T(kbps)_{m,h''})\cdot d_{m,h''}\right)\right\}$$

subject to:

$$\sum_{z\in Cs}s_z\le\sum_{b\in Rb}n_b \qquad n_b=1 \quad \textbf{(3.15)}$$

$$N_z\cdot s_z-\sum_{i\in S_z}s_i\ge 0 \qquad \forall z \quad \textbf{(3.16)}$$

$$1\le\sum_{h'\in Cu}d_{m,h'}\le 2 \qquad \forall m \quad \textbf{(3.17)}$$

$$1\le\sum_{h''\in Tg}d_{m,h''}\le 2\cdot Ser_{\max} \qquad \forall m, Ser_{\max} \quad \textbf{(3.18)}$$

$$s_z, s_i, d_{m,h'}, d_{m,h''}\in\{0,1\} \qquad \forall m,z,i,h',h'' \quad \textbf{(3.19)}$$

- Radio access infrastructure costs criterion

f_{costs} represents the aggregate of CapEx and OpEx in a radio access network. This method represents a financial model over suitable return periods. The calculation of the infrastructure parameters of sites and sectors is made from a set of economical parameters related to a site Pc = {P_1, ..., P_p..., P_P} and related to a sector Ps = {P_1, ..., P_q..., P_Q} over a A return period from a business plan (A $\in$ An, An = {1, ..., a...,A}). The radio access infrastructure costs of a certain location with at least one sector active are shown in (3.20).

$$f_{\cos ts}(z)=\left(\sum_{p\in Pc}\sum_{a\in An}Bc_{a,p,z}\right)\cdot s_z+\sum_{i\in S_z}\left(\left(\sum_{q\in Ps}\sum_{a\in An}Bs_{a,q,i}\right)\cdot s_i\right) \quad \textbf{(3.20)}$$

where Bc and Bs indicate the value of the infrastructure economic parameter Pp and Pq at the year a. Note that different sectors may use the same site. Constrains (3.15) indicate that only one CS can be installed per region, and (3.16) forces each CS to have at least one BS.

The way the costs are defined in (3.20) differs from (3.14) as the former represents the general case, where neighbour CSs may have different number of BSs installed.

Table 3.1 shows one example for 5 years showing an example of Bc and Bs parameters. This economic study has been done for a typical Mobile WiMAX CS/BS costs found in [29] and [47]. Other examples can be found in [63] and [64]. The cost associated to BSs located in dense populated areas, where more traffic is expected, may be higher due to a more expensive renting fee or higher capacity equipment.

Table 3.1. Example of Economic Parameters For Radio Infrastructure Over 5 Year Return Period [29], [47].

Year		**1**	**2**	**3**	**4**	**5**
Base site CapEx (k€)	Installation	30	26.6	23.6	21.2	19
	Establishment	0.75	0.75	0.75	0.75	0.75
Base site OpEx (k€)	Maintenance	1.5	1.33	1.18	1.06	0.95
Sector CapEx (k€)	Installation	5	4.35	3.78	3.29	2.87
	Establishment	0.07	0.07	0.07	0.07	0.07
Sector OpEx (k€)	Maintenance	0.25	0.22	0.19	0.17	0.14
	Rental	0.48	0.48	0.48	0.48	0.48

Each parameter varies over the return period (five years here), and both equipment pricing and installations (CapEx) and management costs (OpEx) decreases with time.

- CINR criterion (Cu)

$f_{pen(Cuh')}$ represents the wireless connection by penalizing the effective CINR perceived by the MSs over all resources used. The thresholds $T^{(min)}$ and $T^{(max)}$ can be set to the CINR thresholds for the minimum and maximum MCS in the system since below the minimum bearer it is not possible any communication, and the maximum MCS is considered the best case in terms of the system resource utilization efficiency.

The assumptions about the expected users made by the network provider in the network definition, such as the user profile and speed, influence the final network design. Different MSs profiles may require different costs for the system due to the different channel conditions. Therefore, the network designer may penalize each of them differently (note that with different channel conditions apply different CINR thresholds to get a certain MCS). As a result, a subset of criteria related to the user CINR levels may apply for different situations ($Cu = \{Cu_1, Cu_{h'} ..., Cu_{H'}\}$, $Cu \subset G$). Table 1.5 shows the $T^{(min)}$ and $T^{(max)}$ thresholds for three different user profiles in a Mobile WiMAX system [65], [11].

Table 3.2. Mobile WiMAX CINR Thresholds.

Modulation	**CINR (dB) Fixed user** (Cu_1)	**CINR (dB) Pedestrian user** (Cu_2)	**CINR (dB) 50K/h user** (Cu_3)
QPSK1/12 ($T^{(min)}_{m,h'}$)	2.88	4.89	6.37
64QAM3/4 ($T^{(max)}_{m,h'}$)	17.50	21.60	22.40

Since each user is in UL or DL (or both), and in certain channel conditions at each snapshot in a static simulation, constraint (3.17) indicates that a maximum of two penalty functions are applied per user.

The maximum penalty $f^{(max)}$ can be set in economic terms related to users. For example, in a user with low CINR, a penalty related to the user connection tariff and annual subscription may be applied.

Some of the user specific service and economic parameters are shown in Table 3.3. Each parameter varies over the return period.

Table 3.3. Example of Economic Parameters for the Users Over a 5 year Return Period [29], [47].

Year	**1**	**2**	**3**	**4**	**5**
Annual subscription (€)	500	440	386	337	293
Connection Tariff (€)	50	45	41	37	33
Mb allowance (€)	500	450	410	380	360
Other charges (€)	62	56	50	45	41

In general, the value for the maximum penalty can be calculated as the aggregate of a subset of economic parameters for one user related to a certain performance indicator G_h ($N' \in N$, $N = \{N_1,...,N_v...,N_V\}$) over a A return period (see (1.31)). With the appropriate selection of the elements of subset N′, the network provider will be able to give more importance to certain performance indicators.

$$f_{m,h}^{(\max)} = \sum_{v \in N'} \sum_{a \in An} Mc_{a,v,m} \quad \textbf{(3.21)}$$

where Mc indicates the value of the user economic parameter N_u at the year a.

- Throughput criterion (Tg)

$f_{pen(Tg_{h''})}$ is the penalty function for the MSs throughput. $T^{(min)}$ and $T^{(max)}$ can be set to the maximum and minimum throughput request for each service . The use of this factor in the cost function provides a flexible and service-oriented network design, in which different business plans, based on certain services for some areas, can be applied.

A subset of criteria related to the user throughput levels may apply for different services and situations (DL or UL) ($Tg = \{Tg_1, Tg_{h''} ..., Tg_{H''}\}$, $Tg \subset G$). Table 3.4 shows the $T^{(min)}$ and $T^{(max)}$ thresholds for four different services. Constraint (3.18) indicates that the MS will use at least one service DL or UL, and a maximum of services, Ser_{max} (4 in this case) in both UL and DL which can be penalized in different ways.

Table 3.4. Example of Service Requirements Thresholds [1].

Name	Maximum Throughput (DL) (kbps) $T_{m,h''}^{(max)}$	Minimum Throughput (DL) (kbps) $T_{m,h''}^{(min)}$	Maximum Throughput (UL) (kbps) $T_{m,h''}^{(max)}$	Minimum Throughput (UL) (kbps) $T_{m,h''}^{(min)}$
Web Browsing (Tg_1)	128	64	64	32
FTP Download (Tg_2)	1,000	0	100	0
Video Conference (Tg_3)	64	64	64	64
VoIP (Tg_4)	12.2	12.2	12.2	12.2

The maximum penalty value can be configured with user economic parameters, e.g. a bad QoS will be reflected in losses in the Mb allowance and other charges related to services (See Table 3.3). Therefore, $f^{(max)}$ is also calculated as in (3.21), and represents an estimation of the losses that the lack of certain QoS in a user mean for the total financial model. Besides the values presented in Table 3.3, the network designer may also penalize each service differently.

3.2 Single Cost Function Optimisation

The multiobjective problem can be solved as a single-objective optimisation problem by combining all aspects of the cost function.

3.2.1 Simulation Configuration on Heuristic Optimization

In this case all factors of the cost function need to be accurately weighted, i.e. to choose λ, $f^{(max)}$ values in the network costs factor and each performance penalty function, to give more or less importance compared to the rest of indicators. The problem is that there will probably exist some trade-offs between the different performance factors and costs. As a result, finding the right weight coefficients for a cost function is not simple and can be a problem in big instances, where several objectives are optimized. The network provider must have a clear idea about the exact preferences, and the penalization values can be set iteratively in a trial and error method until a reasonably good solution is reached.

In the case of having one of the factors in the cost function (3.14) with much higher values than the rest, the final result will be determined by this objective. Compared to the usual values found in the penalty values of the KPIs, the network costs factor F_{costs}, normally has more weight, and therefore should be normally decreased by a factor $\lambda<1$. After an appropriate network solution is found, the value can be returned to $\lambda<1$ in order to obtain the estimation described by (3.12).

The study heuristic algorithms is a topic well documented throughout many years, not only in the existing literature related to this work, but also in other disciplines such as operations research. For this reason, the improvement of these algorithms or proposition of new ones is limited nowadays. However, the choice and calibration of the optimization method to be used in an optimization framework have an important impact on the final performance. In the following, a statistical comparative investigation not yet found in the existing literature in this kind of networks is presented.

Based on the optimization platform described above, this work evaluates the scenario shown in Figure 1.1. All 3-sector antennas are installed with the azimuth being 0 degrees offset from the North and 0 degrees down tilt. The path loss model used is COST 231 Hata with a correction factor for frequency at 3,500 MHz [66]. A rectangular area of 5km × 4km is considered to contain 20 candidate base stations with 3-sector antennas. The number of MS is around 350 for each snapshot. The rest of parameters are shown in Table 3.5. The channel bandwidth used is 10MHz, and the services are shown in Table 3.4.

The simulator takes multiple snapshots in a Monte Carlo routine [22] to statistically observe the network behaviour. Each snapshot around 320 MSs need to transmit or receive at the same time with different service requirements and they are randomly distributed in this area. All MSs have a traffic activity factor of 1.0. The number of snapshots taken every iteration of the optimization routine is 100, which showed to be enough for similar problems [41].

Table 3.5. Simulation Parameters.

Parameter	Value	Parameter	Value
Candidates sites	20	CPE Antenna Gain	0dBi
Carrier Frequency	3.5GHz	CPE Antenna Pattern	Omni
Channel Bandwidth	10MHz	CPE Antenna Height	1.5m
BS TX Power	43dBm	CPE Noise Figure	5dB
BS Antenna Gain	18dBi	CPE Cable Loss	0dB
BS Antenna Height	30m	σ(Shadow Fading)	8dB
BS Noise Figure	4dB	Intra BS correlation	0.7
BS Cable Loss	3dBm	Inter BS correlation	0.5
CPE Tx Power	23dBm	Path Loss Model	COST231-Hata

There are several factors that influence the comparative investigation of several heuristic algorithms. The process of algorithm parameter tuning is customized for different problems and is normally done empirically. The time that the algorithm is running together with the algorithm configuration will determine the performance of the whole optimisation framework.

Four different heuristic algorithms have been set up with a parameter configuration that shows reasonable performance for different scenarios in mobile WiMAX. A comparative investigation is done against a less sophisticated Greedy Search (GS) algorithm. The network performance is optimised with respect to the BS location by using and evaluating four heuristic algorithms GA, ESA, VNS and TS in a single-optimization configuration. ESA corresponds to a SA enhanced with evolutionary operators [67].

A set of B = 4 regions of interest have been defined and each one has five possible candidate sites. This grouping helps to set the neighborhood of each solution in the algorithm. Candidate BS locations and the planning regions are represented in Figure 3.1. We have implemented the economic study presented in Section 3.1.3. The threshold values $T^{(min)}$ and $T^{(max)}$ used for the Cu criteria are similar to the minimum and maximum values shown in Table 3.2, and extracted from the software used [1]. The threshold values for the Tg criteria are set according to Table 3.4, depending on the user state, i.e. if it is transmitting (UL) or receiving (DL) and service required. The values used for F_{costs} and penalty $f^{(max)}$ are taken from the ones presented in Table 3.1 and Table 3.4. There are variations for the infrastructure costs depending on the location of the BS. For the numerical results we have assumed three different kind of candidate sites according to the environment. The clutter information from the map provides information about the area that the BS is installed, and a random generation of economical values in different ranges is made from the example values $Bc_{a,p,z}$. The ranges are $[0.5 \cdot Bc_{a,p,z}, 1.5 \cdot Bc_{a,p,z}]$, $[1 \cdot Bc_{a,p,z}, 2 \cdot Bc_{a,p,z}]$ and $[1.5 \cdot Bc_{a,p,z}, 2.5 \cdot Bc_{a,p,z}]$ for the most expensive sites.

3.2.2 Numerical Results in a Comparative Investigation on Heuristic Optimization

Figure 3.4 and Figure 3.5 shows the initial non-optimized solution, and the best solution found after 50 runs in the five evaluated algorithms (which accounts for 250 optimisations). The best signal level, and user I + N (Interference + Noise) values are displayed for one snap-shot of the simulation routine. Figure 3.5 shows how the general interference levels are improved in the best solution, especially in the central area, while still providing good signal level. This results in higher CINR values.

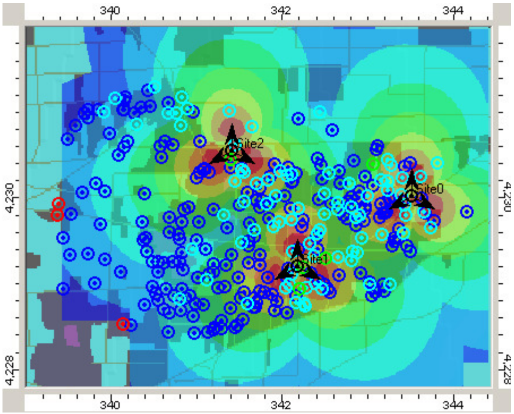

Figure 3.4. Initial non-optimized solution.

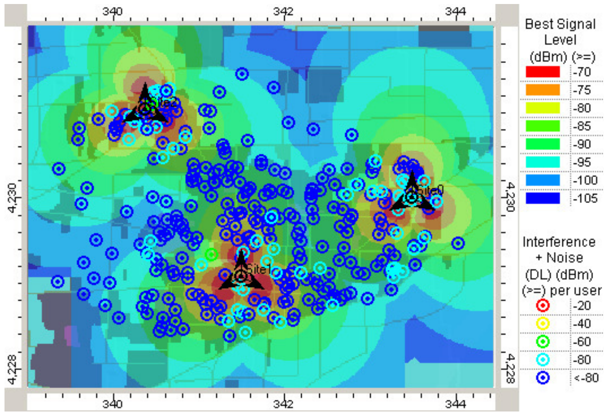

Figure 3.5. Best solution found.

Since the performance of each algorithm depends on the tuning, it has been calculated the cumulative probabilities of the results of 50 runs in order to statistically analyse the distribution of the best cost values, as shown in Figure 3.6 and Figure 3.7.

As described in Section 2.2.1, each algorithm has a different internal structure with different loops and counters. Therefore, they are compared in terms of running times over the same computer. This time does not include simulation time, which depends on the size of the network (number of BSs, users), and for this reason it is not considered.

One iteration of TS includes one full simulation of the network (100 snapshots). For the scenario shown in the non-optimized solution, and a 2GHz Intel Core 2 Duo – 2GB RAM computer, it takes 62s. As a result, with a 1000 iteration optimisation in TS (assuming the same simulation time in all solutions), the process would take 17.22 hours. This time may increase exponentially in a larger network with a larger solution space (that also may need more iterations in TS), which justifies the use of metaheuristics against a brutal search approach.

The resulting cost shown is relative to the worst solution found. Hence, a relative cost of 0.67 (67%) means a 33% improvement with respect to the initial non-optimised solution shown in Figure 3.4.

TS has the highest probability to get the best solution, and the performance of VNS is slightly worse in the case of 150s running time (around 1000 iterations in TS). In this case, the algorithms can get a result less than 0.68 with a probability of 0.44. GA and ESA can get a result which is less than 0.68 with a probability of 0.28 and 0.3, which proves a significant difference with TS and VNS. On the other hand, GS provides the worst result as expected.

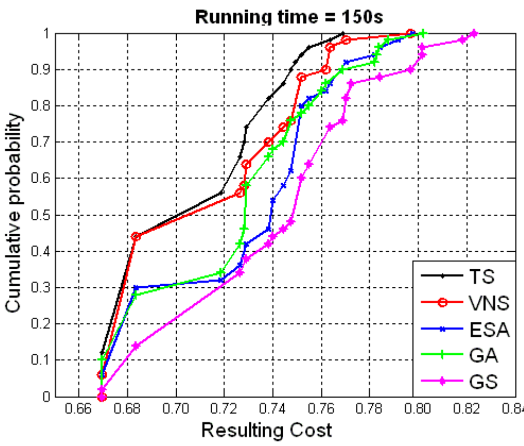

Figure 3.6. Cumulative probability of the results (t=150s).

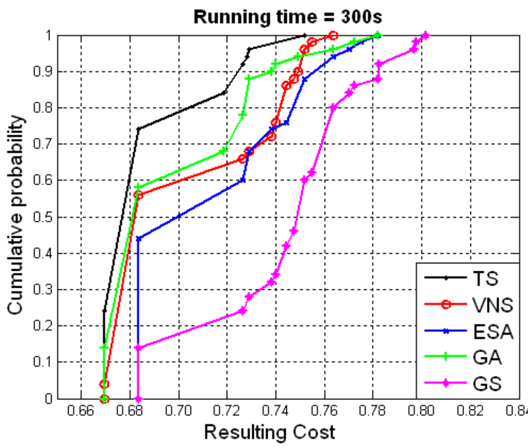

Figure 3.7. Cumulative probability of the results (t=300s).

As 150s seems a short time to get reasonable solutions out of search algorithms, another series of experimentation for longer time (300s) is needed (Figure 3.7). All the algorithms but the GS, have quite improved their performance with the addition of running time. TS is still the best algorithm as it has increased to 0.74 the probability of getting a result less than 0.68. GA raises this probability to 0.59, similarly to VNS with 0.57, although the improvement of the latter is not as good as the experimented by GA. On the other hand, ESA have been outperformed by TS, VNS and GA, but it is still comparable with them. In addition, note that the results produced by TS are no bigger than 0.75 at all.

Figure 3.8 indicates how increasing even the running time (450s) improves the performance of the algorithms but no significantly, except in the case of ESA that presents a much better performance. This concludes that both TS, GA and especially ESA greatly improve with more iterations, and the performance significantly remains in favour of TS over the rest of heuristics. It can be observed that the convergence speed of GA and ESA is too slow to reach the near optimum solution with a short time. GA and

ESA has more parameters to be tuned comparing to TS and VNS, which makes difficult to keep good performance for different network instances.

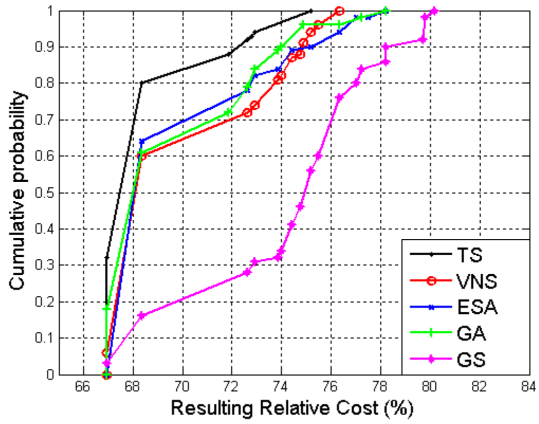

Figure 3.8. Cumulative probability of the results (t=450s).

Figure 3.9 indicates the means and the standard deviations (SD) of the results. It shows that TS achieves the lowest means and SD values. This means that, TS is the most robust algorithm for this kind of optimization problem, while VNS, ESA and GA do not provide so robust performance, but better than GS, which is expectedly the worst one in both measures as it is an algorithm with a high probability to get stuck in the local optima [29]. Another conclusion can be drawn is the contribution of longer running time to the improvement of the performance. All but GS show improvement, and GA and ESA show a great change and seem to perform much better with a high number of iterations. GS cannot get much benefit in robustness for more searches.

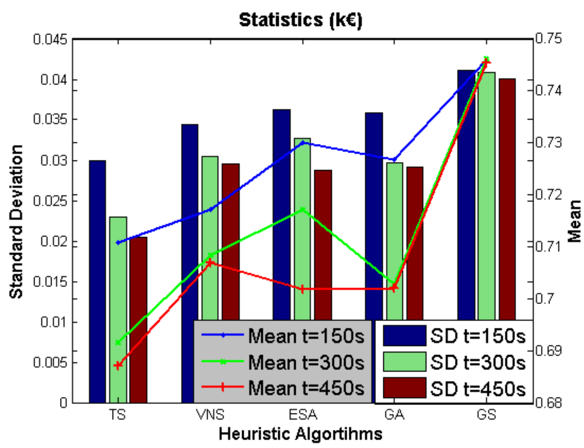

Figure 3.9. Standard deviations and means of the results.

Obviously, if the optimisation time is increased, results will tend to group together. In the small instance shown, around 2000s in optimisation and several weeks of simulation time are needed to achieve similar results with the different algorithms. Also, due to the small solution space considered (chosen to introduce the model proposed), the differences in the final results are not drastic. Again, these times would increase exponentially in a larger network with a bigger solution space, which justify the use of metaheuristics against a brutal approach.

Note that it would be difficult to consider the variation of all of the algorithm parameters that can be tuned in a comparative study. Therefore, the aim of this investigation on Heuristic optimisation has been to show the wide range of solutions that may be obtained, and to highlight the complexity of the design process as the operator needs to estimate diverse parameters. The following sections will evaluate the model proposed from the new multiobjective perspective.

3.3 Multiobjective Optimisation

The multiobjective problem can also be solved by using the solutions of the optimal Pareto front and the user must select aposteriori the most appropriate ones according to some policies.

Many algorithms can be adapted to multicriteria optimization. In the following, a detailed description of the procedure for multiobjective TS is done. The algorithm minimize the previously defined criteria radio access infrastructure costs, CINR, and throughput, and try to improve the current search front $Sf_{u,n}$ composed of n solutions at every iteration u by calculating the neighbouring solutions. A neighbouring solution is a non-Tabu change of the state of the network by adding, removing or moving a CS/BS or changing any other parameter of the network configuration.

The total neighbouring set is the aggregate of the neighbourhoods of the n solutions of the search front and it is stored in the short term memory created by the algorithm. Similarly to the standard algorithm, the Tabu list is maintained in the long term memory.

Figure 1.14 sketches out this method. The optimal front Of is updated at the end of each iteration u by adding all the non dominated solutions of the neighbourhood of the search front. The previous solutions of the Of that have become dominated with the new ones (rank $Ra_n > 1$) have to be removed. The new search front $Sf_{u+1,n}$ is created randomly from previous solutions, and thus introducing diversity in the search process. In addition, the new solutions n + 1 created from solution n are stored in the Tabu list, which is also updated at the end of each iteration u. The algorithm stops after a certain number of iterations defined by the user u_{max}, and the most representative solutions from the Pareto front are selected.

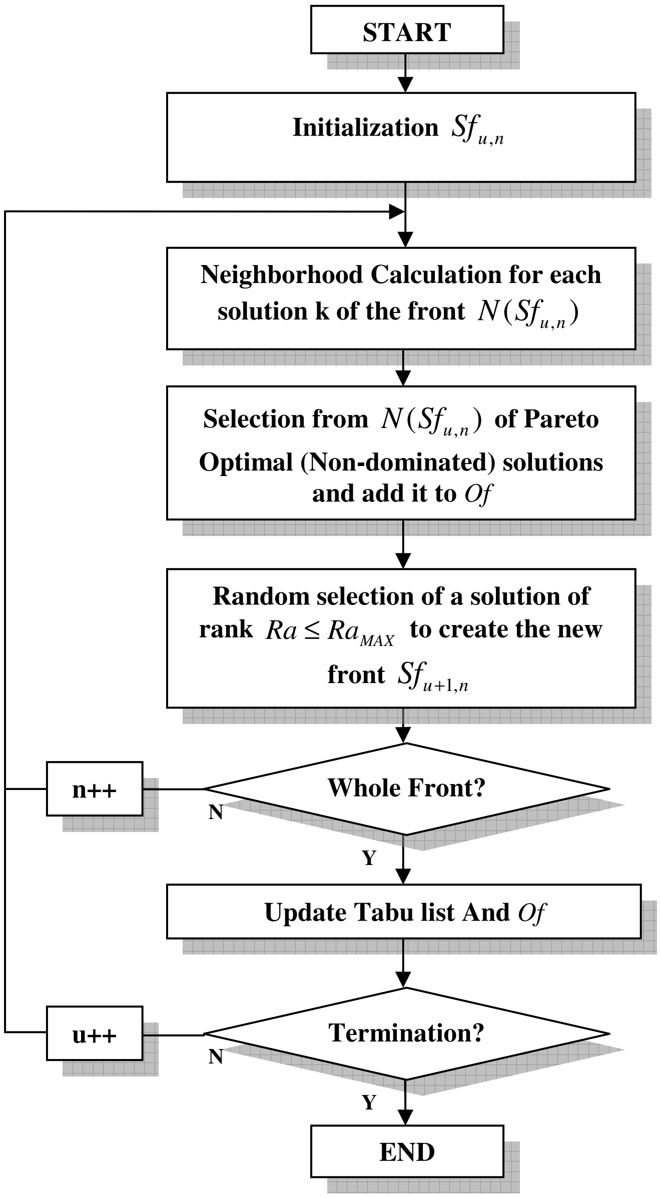

Figure 3.10. Multiobjective Tabu Search algorithm.

The network provider has to calibrate the cost function also in this method to solve the multiobjective optimization problem. The main advantage is that the network provider has more flexibility in the Pareto front method since the decisions have to be made a-posteriori by choosing one solution from the final set offered, and also the information for every factor of the cost function is provided. Each factor in the cost function may have any value, which will be represented in the optimal front.

- Mobile WiMAX Network Design

In the following, an example of multicriteria network design process has been performed for a Mobile WiMAX network. The simulation parameters are set according to Table 3.5 and Section 3.2. The threshold values $T^{(min)}$ and $T^{(max)}$ used for the Cu criteria are similar to the minimum and maximum values shown in Table 3.2. The exact values are extracted from the software used [1], and shown in Table 3.6. It presents the Mobile WiMAX bearer to access different MCS, and a shadow fading margin is considered. In order to consider the effect of different channels, a margin of 3 dB is applied to each bearer in this software.

Table 3.6. Mobile WiMAX bearer thresholds [1].

Best Bearer	CINR
1	0 = $T^{(min)}$
2	2.5
3	6.5
4	9
5	13.25
6	17.4
7	17.8
8	19.8
9	23.3 = $T^{(max)}$

Every time the algorithm tests a set of candidate sites to install BSs, the SLS is called to calculate the cost function values. A threshold is set, u_{max}=1600, for the maximum number of SLSs iteratively called. When the threshold is met, the algorithm stops immediately. The Tabu list size and Ra_{max} are chosen empirically during the initial algorithm tuning. The Tabu list size is typically close to the size of the neighbourhood. A final set of Of = 10 solutions is selected.

The planning algorithm looks for both the number of BSs and their locations, obtaining a 3-D representation of different states of the network. Every time the algorithm tests a set of candidate sites to install BSs, the SLS is called to calculate the cost function values. The optimal front is represented in Figure 3.11.

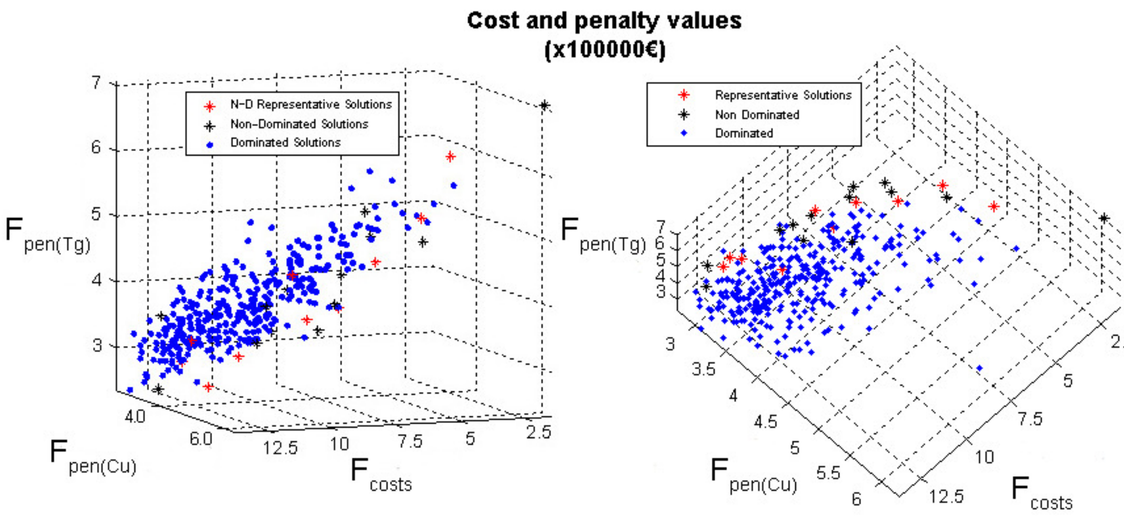

Figure 3.11. Non dominated and dominated solutions of the Pareto front in Multiobjective Tabu search. Some representative solutions are selected from the non dominated set.

The non dominated solutions from the Pareto front represent optimal tradeoffs in the network, and some representative solutions must be selected in order to clearly visualize the main choices for the network provider. A final set of Of = 10 solutions is selected. Table 3.7 shows the number of CSs used in each solution together with the costs of the network considering (F_{costs}), penalty costs related to the network performance ($F_{pen(Cu)}$ and $F_{pen(Tg)}$), and the adjustment costs (CostAdj).

Table 3.7. Cost and Penalty Values of the 10 Most representative solutions of the Pareto Front (×100000€).

Solution	*Number of installed CSs*	F_{costs}	$F_{pen(Cu)}$	$F_{pen(Tg)}$	*CostAdj*
1	4	9.722	3.271	2.884	15.877
2	4	11.163	2.73	2.733	16.626
3	2	5.855	3.332	3.552	12.739
4	1	3.578	4.892	6.111	14.581
5	2	3.707	4.114	5.031	12.852
6	3	5.059	3.781	4.319	13.159
7	4	10.843	2.752	3.038	16.633
8	3	7.226	3.376	3.414	14.016
9	3	10.188	2.755	2.336	15.279
10	3	7.465	3.157	4.058	14.680

The results obtained show always positive values for $F_{pen(Cu)}$ and $F_{pen(Tg)}$. The reason is that they model a real system and balance the ideal assumptions made in ProfitEst for the performance of the users.

In general, the results show that the more CSs are placed, the lower the throughput and CINR penalty values, but the more expensive the network infrastructure costs. The options with more CSs would be suitable to a network provider whose policy is to be fair with all the users targeted and try to provide good service to most of them. The drawback is that a higher initial investment should be done.

The options with fewer CSs used are normally cheaper in the infrastructure costs and therefore, in the initial investment. However, the user performance is worse and this may be reflected in the total costs. For example, solution 4 with 1 CS, has higher total costs than solution 8 with 3 CSs due to poor performance. The use of fewer CSs would be a good option for network providers that prefer a slow deployment of the network, and plan a future expansion of it according to real user demands.

In order to analyze the different results, Table 3.9, Figure 3.12, Figure 3.13, Figure 3.14 and Figure 3.15 show the effective CINR levels of one snapshot for a non-optimized solution, and solutions 3, 9 and 10. Table 3.8 shows the cost function for a random dominated solution.

Table 3.8. Cost and Penalty Values of the of a Dominated Solution (×100000€).

Solution	*Number of installed CSs*	F_{costs}	$F_{pen(Cu)}$	$F_{pen(Tg)}$	*CostAdj*
Dominated	3	9.109	9.230	3.809	22.149

Table 3.9. Number of users at different effective CINR levels.

Solution	CINR < 0 dB	CINR >= 0 dB	CINR >= 6.5 dB	CINR >= 17.4 dB	CINR >= 23.3 dB
Dominated	94	60	95	37	34
3	80	38	97	51	54
9	53	50	107	44	66
10	64	48	99	45	64

During each snapshot the simulator randomly distributes the MSs, and CINR and throughput values are calculated. The assumptions made by the network provider in the initial planning phase about the users, such as the user profile and speed, influence this calculation since the system may determine different permutation zones in the RRM process for users with different channel frequency responses. For example, in the Band AMC permutation mode, where subcarriers constituting a subchannel are adjacent to each other, the system can get multiuser diversity gain by allocating subchannels to users with the highest CINR in that subchannel. However, this is more suited for fixed and low-mobility users with a large coherence bandwidth of the channel. On the other hand, users with higher mobility can take advantage from frequency diversity that other permutation schemes benefit, e.g. PUSC. As a result, some decisions about the expected scenario can influence the system setting up and implementation, specially affecting some areas of the Mobile WiMAX standard that are left to providers to define, like the setting up of different permutation zones in RRM.

Band AMC and PUSC zones are used in our Monte Carlo simulator [1], and the profile of the users is randomly generated as stationary, pedestrian or 50 km/h users. The users are allocated to each zone depending on a quality threshold, and the maximum speed of the user. No segmentation is applied for PUSC. In the system-level simulator used, the effect of a particular permutation zone in the RRM is modelled with different gains. In addition, the user need a certain number of slots depending on the zone allocated, and this will lead to a certain throughput value. This process is iterative in order to calculate the general interference level, and so it is the estimation of the size of each frame zone to suit the traffic demands.

In Figure 3.12, it can be observed that the general CINR levels of the non-optimized solution are low. As per the location of the CSs, most of the users receive strong signal but also have high interference levels. As Table 3.9 indicates, a high number of users are penalized as they are not able to connect to the network or just get the minimum bearer for communication. The reason is that the successful users need more resources to meet the QoS requirements, and therefore the OFDMA frame is filled by fewer users compared to other optimized solutions. As it can be seen in Table 3.8, the total costs are much worse than in any other solution in Table 3.7. This confirms that a manual network configuration based on the prediction of the signal level is not enough, and network simulation is necessary to predict a real picture of the situation. The information offered by the cost function highlights the importance of an optimization procedure, as it can provide valuable information about several objectives even in small-medium sized scenarios.

Table 3.7 shows that in solution 3 (Figure 3.13), a lower number of serving CSs results in lower infrastructure costs and the network still provides good connectivity and QoS. On the other hand, there are a number of users with low CINR levels concentrated at the edge of the coverage area, where they may have a poor channel and also receive similar signal levels of two BSs for the same resource of the OFDMA frame. As already mentioned, this can be a suitable option for network providers that plan a future expansion of the network according to the real user demands. The addition of one more CS with the particular position of the BSs in solution 9 (Figure 3.14), gives more flexibility to the RRM strategies of the system, increasing the number of successful users (Table 3.9). This is one of the solutions least penalized on performance Table 3.7 but with higher initial investment needed on the infrastructure costs.

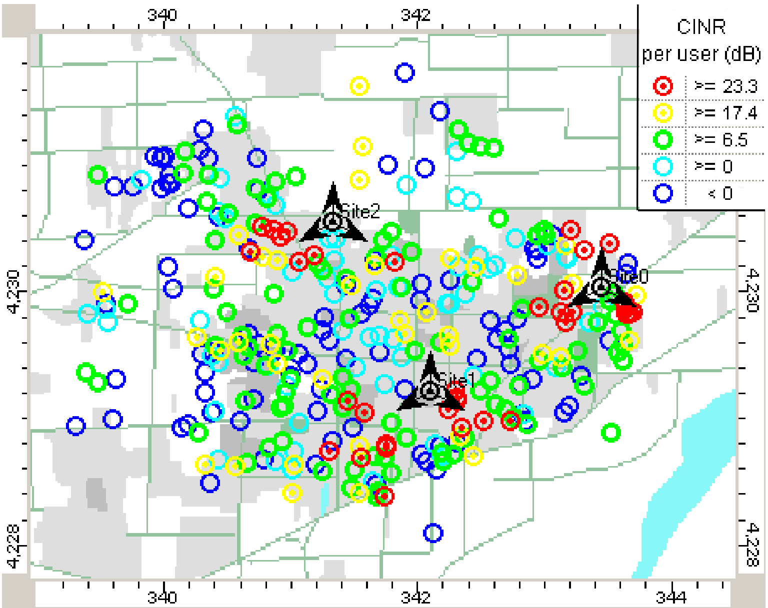

Figure 3.12. CINR levels for each user in one snapshot of a non-optimized solution.

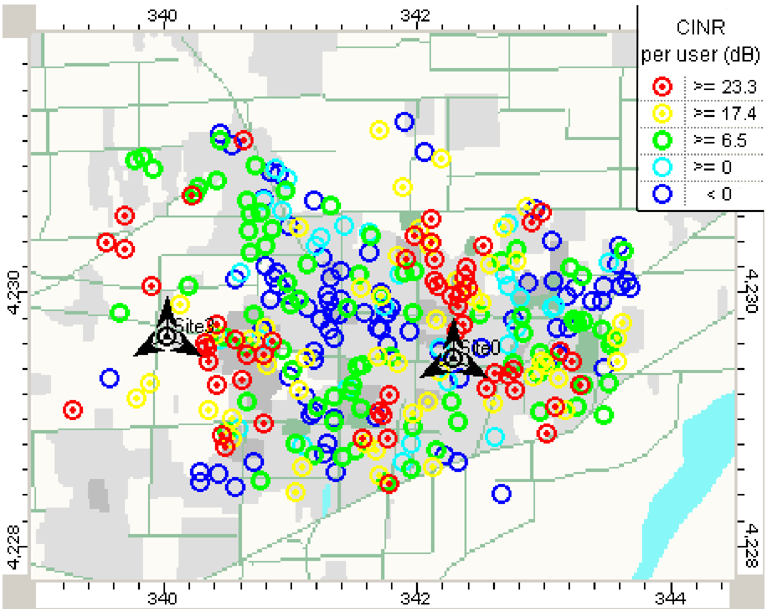

Figure 3.13. CINR levels for each user in one snapshot of solution 3.

The limits shown in Table 3.9 represent some of the limits taken from Table 3.6 for static users. This represents s a worse case for users with higher mobility. For example a 50km/h user may not get the minimum bearer for communication for 0 dB <CINR<6.5 dB.

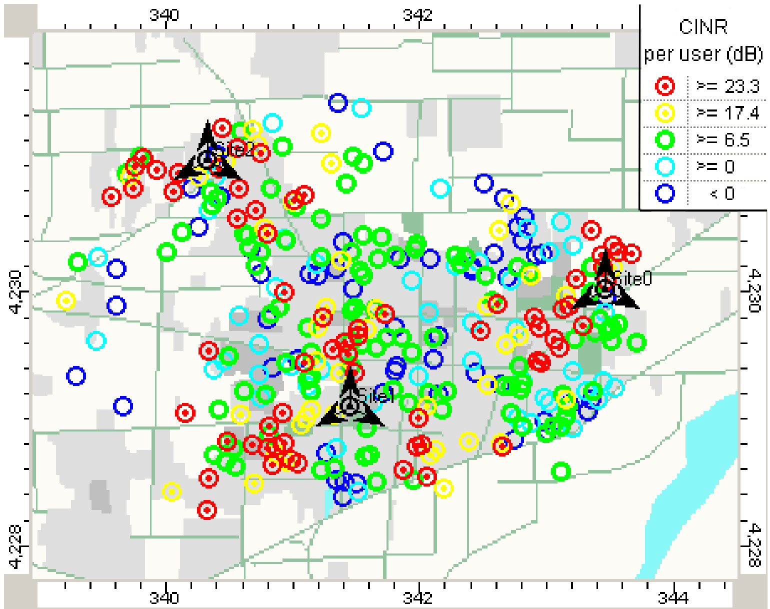

Figure 3.14. CINR levels for each user in one snapshot of solution 9.

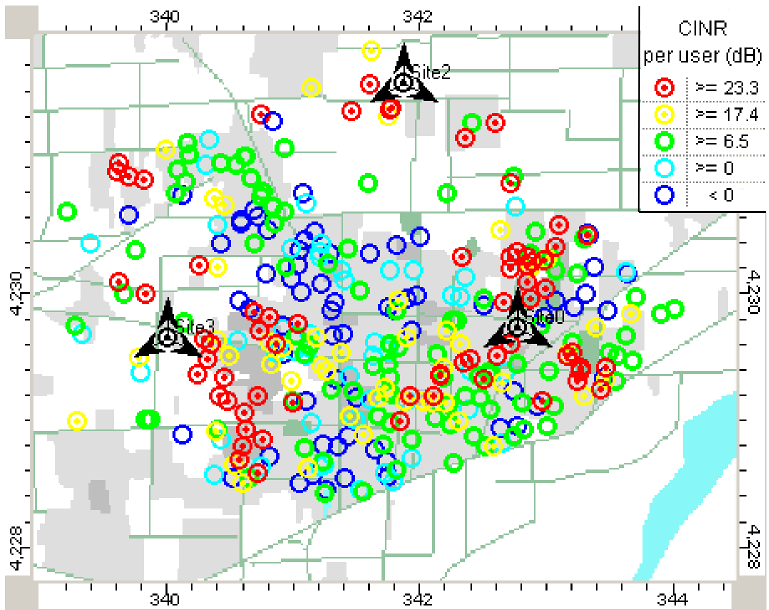

Figure 3.15. CINR levels for each user in one snapshot of solution 10.

Investigating the results between solution 3 and 9 further, we can observe in Table 3.7 that the improvement of the throughput criterion (Tg) is bigger than the improvement in the CINR criterion. As described in Section 1.2.3, there are several factors involved in the allocation of network resources that lead to different bit rates for the users. The final value for the user throughput depends on the CINR level received, the burst size and also on the number and position of BSs, since this will influence on how the resources are reused. In the case of solution 9, with the location of three CSs in that particular position, not only the interference levels are improved, but also the signal is stronger for more users, resulting in a considerably higher throughput levels. The price of the third CS brings the total cost to a higher value compared to solution 3. Penalization is performed in different ranges and values, and depend on the network provider criteria as some objectives can be given more or less preference. Due to this fact diverse situations can be created. We can observe a different effect in solution 10 (Figure 3.15), in which the addition of a third CS slightly improves the CINR value but the throughput criterion is even worsened compared to solution 3. Both solutions have similar CINR levels but, as Table 3.9 indicates, some new users can gain connection to the network because of the existence of one more CS, and therefore some penalty values related to connection tariff or annual subscription are reduced. On the other hand, the throughput per user decreases specially in the MSs that are far from the CS and get low CINR values in those areas. Users that require services that are bandwidth consuming are more penalized and thus, this option would be suitable for network providers that prefer to ensure connectivity to as many users as possible in a wide area than providing high bit rates, while maintaining low infrastructure costs. In fact, we can see that solution 10 is part of the Pareto front due to the low infrastructure costs (F_{costs}) compared to other solutions with better performance.

These results show that this kind of optimization problem is indeed a multiobjective problem, since a lot factors have to be taken into account and they generate tradeoffs with different effects under different configurations. The advantage of this multicriteria view is that the network provider does not have to analyze every solution in depth; the data shown in Table 3.7 offers enough information to choose a suitable solution for the network provider strategy.

- Large Solution Spaces – An LTE Radio Access network design Case

The process of selecting the most representative solutions from the Pareto front can be difficult to perform manually by using graphical representation (Figure 3.11), especially when the solution space is large or there are more than 3 objectives. Instead, some algorithms like the Sharing function presented in [39] can be used. Basically, the aim is to pick at least as many solutions as tradeoffs represented in the Pareto front. The function assigns to each solution a weight proportional to the density of solutions in its neighbourhood. This sharing function is computed knowing the distance between solutions in the evaluation function space. Each sharing set represents a family of solutions expressing the same order of tradeoffs between the criteria. The maximum cost for each function should be taken into account by dismissing the solutions not fulfilling this requirement.

In the following, an LTE radio access network design case is presented. The carrier frequency is 2 GHz and the channel bandwidth 10 MHz. In order to present a larger solution space, antenna tilt and azimuth are configured, as well as BS position. In addition, CSs may be installed with some BSs inactive, which do not compute costs. The number of CSs that may be placed in a region is n_b=2. CSs are the same that in Figure 3.1.

Traffic density is increased by including dense urban traffic maps. Three different types of traffic pattern have been used to simulate the uneven distribution of the users across the scenario in a dense urban, urban and suburban traffic: with a density of 80 user/km^2, 40 user/km^2 and 10 user/km^2, respectively that account for approximately 700 users per snapshot. The MSs have different service requirements (Web browsing, FTP download, VoIP and H.263 video) and their profile is generated as stationary, pedestrian or 50 km/h users.

The threshold values $T^{(min)}$ and $T^{(max)}$ used are extracted from the software used [1], and shown in Table 3.10. A shadow fading margin is considered. A margin of 2.5 dB is applied to each bearer in this software in order to consider the effect of different channels. As a result, channels may present different values, and the results shown in Figure 3.16 and Figure 3.17 indicate whether all users within a snapshot are successful or not according to the particular service requirements.

Table 3.10. LTE bearer thresholds for static users [1].

Best Bearer	effective CINR
1	-6.5
2	-4
3	-2.6
4	-1
5	1
6	3
7	6.6
8	10
9	11.4
10	11.8
11	13
12	13.8
13	15.6
14	16.8
15	17.6

Different configurations are searched by a Tabu Search, which gets optimal values for Sector position and number, antenna tilt and azimuth - in intervals of 10^0 -. The economic parameters for sites/sectors CapEx and OpEx (F_{costs}), have been decreased with λ=0.7 in order to observe the incidence on the final results.

An automated selection of representative solution allows more accuracy to find different tradeoffs. Therefore, a final set of five representative solutions may be selected for the three-objective scenario, and shown in Table 3.11.

Table 3.11. The 5 Most Representative Solutions of the Pareto Front (×100000€).

Solution	*Number of installed BSs*	F_{costs}	$F_{pen(Cu)}$	$F_{pen(Tg)}$	*CostAdj*
1	10	3.56	2.41	1.29	7.26
2	12	4.60	1.15	2.67	8.42
3	9	3.02	2.27	1.96	7.25
4	8	2.24	3.59	3.86	9.69
5	12	5.40	1.87	1.95	9.22

Figure 3.16 shows which users fulfil the service requirements in a single snapshot in solution 1. In this solution, a low number of sectors results in medium infrastructure

costs and low throughput penalties. Dense urban areas and indoor coverage benefit from this configuration. On the other hand, there are a number of MSs with low CINR levels concentrated at the edge of the coverage areas, where they may have a poor channel and also high interference. Again, this solution can be a suitable option for service providers that plan a future expansion of the network according to the real user demands.

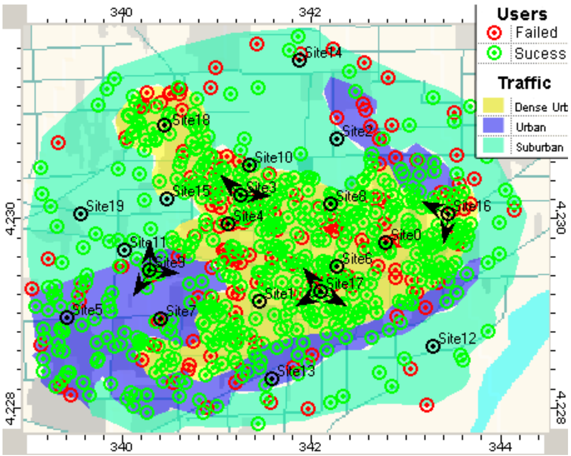

Figure 3.16. Snapshot in solution 1.

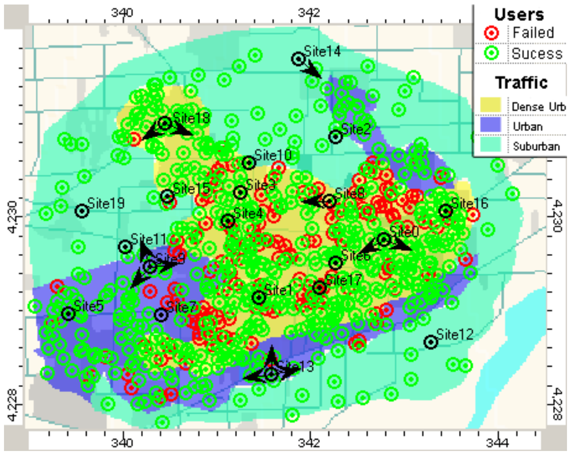

Figure 3.17. Snapshot in solution 2.

In solution 2 (Figure 3.17) some new users can gain connection to the network because of the addition of more sectors, and therefore some penalty values related to connection are reduced. Note also that more sites are used. On the other hand, the throughput per user decreases specially in the MSs that are in dense urban area. Users that require services that are bandwidth consuming are more penalized and thus, this option would be suitable for network providers that prefer to ensure connectivity to as many users as possible than providing services with high bit rates.

Solution 1 represents a throughput enhancement strategy, while solution 2 characterizes a coverage extension strategy. Other solutions in Table 3.11 represent other tradeoffs. Solution 3 and 4 can reduce infrastructure costs at the expense more penalties in “connection” values, or even more degraded performance in the latter case. Solution 5 shows the best user performance, but with much higher infrastructure costs.

Since each factor is processed separately, the main effect of having reduced weight (λ) in the F_{costs} factor is the influence in the selection of the final processing of CostAdj and ProfitEst that the operator may need.

In order to fully understand the differences between solution 1 and 2, Table 3.12 shows the distribution of users in CINR levels. The figures are similar. We can observe that the number of users that cannot access MCS for communication (CINR < -6.5dB) is slightly lower in solution 2. However, the main difference with solution 1 is the high number of users that have low CINR levels that may only allow low MCSs. These users need more SBs to fulfil their service requirements and the frame is filled faster, thus, leaving other users with the same situation with lower throughput levels (higher penalties in throughput).

Table 3.12. Number of users at different effective CINR levels.

Solution	CINR < - 6.5 dB	CINR >= - 6.5 dB	CINR >= 6.6 dB	CINR >= 11.4 dB	CINR >= 17.5 dB
1	130	157	149	138	113
2	122	179	146	132	108

In this case, on a 2GHz Intel Core 2 Duo – 2GB RAM, the simulation time of one iteration of multiobjective TS is around 100s (as it depends on the particular solution tested by the algorithm). As a result, with a 1600 iteration optimisation in multiobjective TS (assuming the same simulation time in all solutions), the process takes 44.44 hours.

3.4 Conclusions

To achieve optimally performing and cost-efficient networks, the design the problem should be simplified and translated it into an optimization routine with consideration of technology specific and economic factors. This chapter presents a service-oriented optimization framework that offers the service provider a comprehensive, detailed and clear description of different scenarios during the network architecture design process so that the most suitable solution can be found.

The selection of the heuristic algorithm for the optimisation stage has an important impact on the final performance of the framework. The comparison of different methods is difficult, since they all have different parameters to configure and the final performance depends on the tuning done. Based on a standard calibration, which empirically performs reasonably well in different scenarios, a statistical comparison has been presented. The single-function optimisation experiments show that the analysed heuristics produce different performance depending on the running time. TS achieved the best performance and outperformed the other heuristics, while the worst performance appeared with the GS algorithm, which was expected. It has been found that GA and ESA have more parameters to be tuned comparing to TS and VNS, which makes difficult to keep good performance for diverse instances and running times.

A multiobjective problem can be solved as a single-objective optimisation process by combining all aspects of the cost function. It can also be solved by using the solutions of the optimal Pareto front and the user must select a-posteriori the most appropriate ones according to some policies. Although calibration has to be done in both methods, the main advantage in the latter method is that the network provider has more flexibility, since the decisions is done by choosing one solution from the final set offered. Also the information for every factor of the cost function is provided.

A Multiobjective Tabu Search procedure and the use of the Sharing function are proposed for the framework. Scenarios on Mobile WiMAX and LTE with different degree of complexity have been presented.

The results of the scenarios analyzed in this chapter highlight the complexity of the design process, as the operator needs to estimate a number of parameters, but also shows the benefits of the described optimization framework, which can provide valuable help to the network designer even in small-medium sized scenarios. The final figures rely on a network simulation tool and a previous economic study of several factors. Nevertheless, a multiobjective perspective can provide a good assistance when choosing a solution in any case by showing the trends of different configurations.

Different strategies can be adopted according to the penalty function calibration: throughput enhancement, coverage extension or a conservative strategy in terms of initial investment on initial costs.

Contrary to a traditional access-oriented design, the service-oriented perspective presented is more efficient dealing with areas with different number of users or with irregular service demands, which may result in a bottleneck for the rest of the network.

Some more performances analysis can be done in this framework, for example different services may be used in different areas and penalised differently.

In addition, other network life phases may be considered and integrated in the framework as a future work, in which network measurements and other parameters besides BS location or antenna tilt/azimuth may be configured.

Chapter 4: Efficient OFDMA Capacity Estimations in a Multihop Relaying Environment

The new relay-based architecture that is being developed for Mobile WiMAX and LTE [42], [9], imposes a demanding performance requirement on relay stations. The frame structure used for transmission plays an important role in the design of multihop networks as it manages the channel access and therefore determines the resulting network capacity. Some methods and concepts to integrate multihop communication into the WMAN level are currently being proposed and studied. These methods involve some modifications in RSs, and are similar to concepts previously proposed for other technologies [46] [47] and [48]. On the other hand, some commercial planning tools provide an extensive model of a Mobile WiMAX network, but important weaknesses can be found e.g. the interference prediction. These software tools may also contain a powerful graphical environment for predictions, but the computational load of its predictions is too high and does not support traffic modelling for actual network simulation.

This chapter firstly proposes a new simple and flexible concept to integrate multihop communications into the mobile WiMAX frame structure and compare this method with other basic concepts already proposed for the upcoming IEEE 802.16j release of the WiMAX standard. Secondly, it introduces efficient capacity calculations that allow intensive system level simulations in a multihop environment for the purpose of network design. Finally, network capacity is evaluated within a new simulation model for various configurations and services allowing interesting observations about OFDMA radio efficiency. Numerical results and discussion highlight the advantages and drawbacks of the selected strategies and provide a deeper understanding on the range extension aspect of a relay network.

4.1 Properties of relay based networks

4.1.1 Introduction

The multihop relaying configuration aims to deal with challenging radio propagation characteristics and low CINR at the cell edge with the introduction of RSs. RSs forward the information from some MSs to the legacy BSs that are connected to the core network. The BSs will provide resources to the RSs and other MS to which are directly connected.

A major concern for network designers is the management of interference [42], which can degrade the performance of the system considerably. In multihop networks this is a critical issue, since the addition of RSs constitutes a new source of interference, and therefore its power, frequency and antenna configuration must be carefully selected to avoid interference.

There are many transmit power configuration for RSs. Some studies evaluate the economics of having many RSs inside a BS at a low power, against the possibility of installing one or two RSs per BS at similar power, obtaining diverse results [68], [69]. The main problem in the former case is the difficulty of finding appropriate site locations for a high number of RSs, which incur in more costs related to rental and maintenance. This thesis will analyse the latter case since its focus is on the detailed planning and optimisation stage, where parameters with high costs get an initial configuration. An example is when a network with only BSs needs to be expanded (see Figure 1.3) due to an increase of capacity demand or a coverage extension requirement.

The RSs normally work in half-duplex mode, which means the station does not receive and transmit using the same channel simultaneously. In addition, the RS can operate in two possible schemes, depending on how it processes the received signal: Amplifying and Forwarding (AF), where the RS just amplifies and retransmits symbols, and Decoding and Forwarding (DF), where the RS demodulates and decodes the received signal before retransmission.

The concept of cooperative relaying is based on one signal that can be received and forwarded by multiple stations to the same MS, which combines the signals received from the RSs, and possibly from the BS, in order to take advantage of channel diversity gain (see Figure 4.1). The cooperation is considered as source diversity if identical signals are transmitted simultaneously in time and frequency by both BS and RSs, or transmit diversity if space-time codes are used [69].

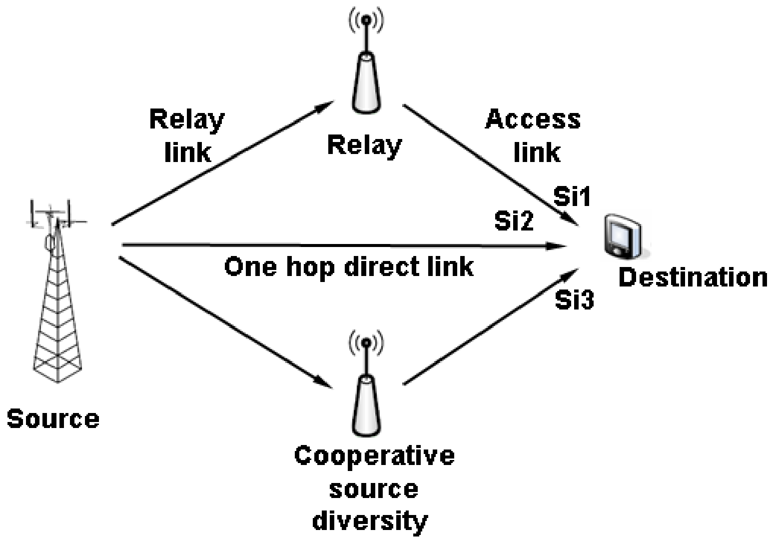

Figure 4.1. Cooperative relaying.

Another way of cooperation is soft handover, which is performed in UMTS [6] by means of macro diversity between different BSs. However, it is agreed that inter-BS soft handover will not be included in 3GPP LTE [9]. Although it is mentioned within the IEEE 802.16e-2005 standard, it is left as a future optional mode left for the vendors' implementation.

4.1.2 Technology-specific features affecting the network planning and optimization process

In relay topology whenever the BS transmits, it is doing so in the downlink (DL). Whenever the MS transmits, it is doing so in the uplink (UL). The RS, however, must transmit and receive in both DL and UL. In order to accommodate these communications the radio frame has to be split in several zones.

With the frame split, fewer slots will be available for the users of each zone and therefore the efficiency decreases. However, in average, the users will receive a stronger signal, and therefore higher CINR levels. As a result, higher level MCSs may be used and thus fewer slots would be needed to achieve a certain throughput. In addition, the interference may decrease depending on how the separation of resources is done in the frame division.

When cooperative transmission is considered, the transmissions instances are repeated in each cooperative path or frame zone (in both BS and RS), and therefore the pool of available resources is reduced due to the redundancy. The gains that the cooperative transmission may provide will depend on the particular scenario, and it should overcome this drawback if is taken into account.

More than three hops in the communication (two relay stations) is not practical due to the low radio resource efficiency and possible impairments in the communication such as delay [69], [70]. Figure 4.2 shows the frame structure for two hop communications and time domain forwarding. For more hops, the BS should reserve one zone to accommodate each relay link. The length of each zone of the frame will change to allow different MCSs depending on the link quality and channel conditions.

Frequency domain forwarding and other duplex configurations are also possible [68], but in this chapter we will focus on structures derived from the described scenario in Figure 4.2 [69]. In general, some of the frame pieces are allocated to the access links and some others to the relay ones in order to avoid interference from the BS-MSs and RS-MSs communication (access links) to BS-RS (relay link). However, the designer still needs to deal with other interferences, and different frame structures described in next section have been proposed in the existing literature. Note that in case that more than one two-hop RS is attached to a BS, the RS access and relay links may be placed consecutively in RS frame in any case.

In practice, the relays often operate in decode-and-forward mode because the MAC layer requires data contained in the header and subheaders to operate, and the PHY layer is generally unaware of the difference between the headers and the payload data.

Therefore, the data may need several frames to be relayed by the RS. However, an amplify-and-forward operation is also allowed, where the RS demodulates and deinterleaves its signal, and then immediately interleaves and modulates it for transmission without decoding. In this case, the order of the RS frame zones showed in Figure 4.2 need to change, and an example from the IEEE 802.16j is shown in next section.

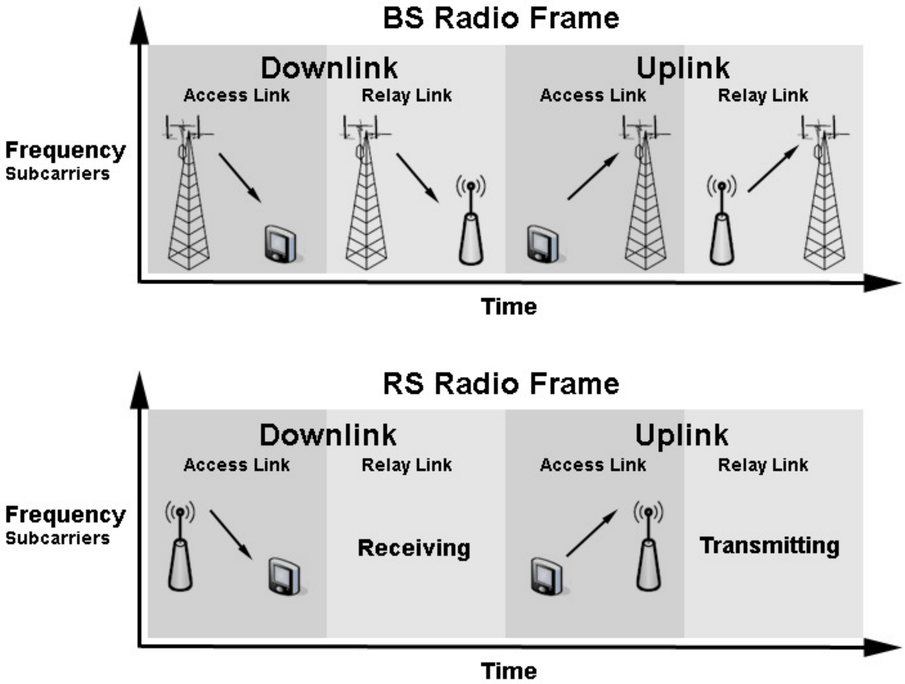

Figure 4.2. Time forwarding.

4.1.3 Frame Structure Design for IEEE 802.16j

It is requirement of IEEE 802.16j a backward compatibility with MSs in IEEE 802.16e to ensure the implementation and transparency of the system. In this section a new simple and flexible concept to integrate multihop communications in the mobile WiMAX frame structure is proposed and compared to other basic concepts already proposed for the upcoming IEEE 802.16j release.

- Non-Transparent Relay

In this mode, the frame is divided into several pieces with similar structure than in IEEE 802.16e in order to keep the system as general as possible. Some of these pieces are allocated to the access links and some others to the relay links to avoid interference from the BS-MSs and RS-MSs communications to BS-RS. However, the designer still needs to deal with the interference between BS-MSs and RS-MSs. Figure 4.3 shows the frame structure proposed in the IEEE 802.16 draft [42] for two hop communications. For more hops, the BS should reserve one zone to accommodate each relay link. The length of each zone of the frame will change to allow different MCSs depending on the link quality and channel conditions.

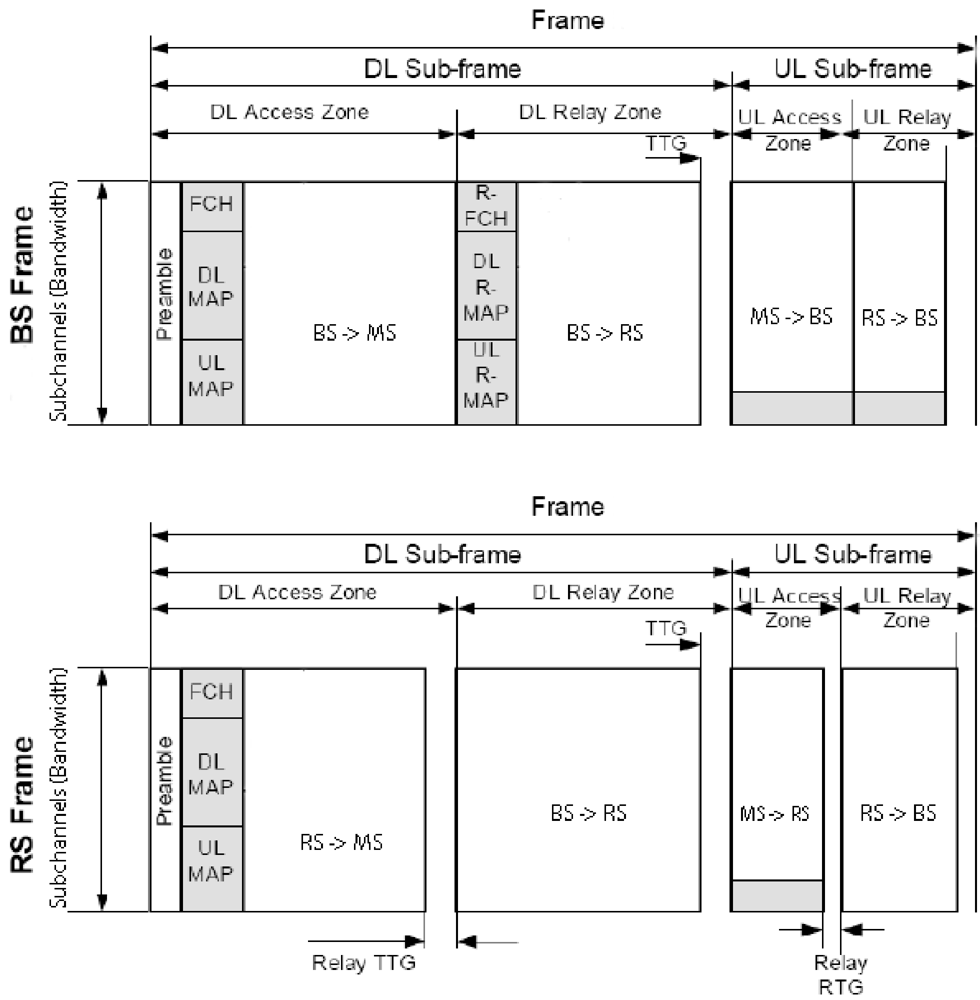

Figure 4.3. Frame structure for non-transparent relay stations [42].

The main feature of this mode is that BSs and RSs transmit the same kind of preamble as that defined in the IEEE 802.16e, which enables different users to conduct network entry and synchronization over the BS or the RS. The FCH, DL-MAP and UL-MAP [11] in the relay frame structure have the new access information belonging to the relay part in the same frame. With this structure, the RS provides coverage extension for the BS since MSs recognize the non-transparent RS as a regular BS. Depending on which station creates MAPs (BS or RS itself), the scheduling and resource allocation type will be centralized or distributed. If cooperative transmission is considered, BSs may also transmit to the same MSs in its access zone.

The relays need operate in decode-and-forward mode because the MAC layer requires data contained in the header and subheaders to operate, and the PHY layer is generally unaware of the difference between the headers and the payload data. Therefore, the data may need several frames to be relayed by the RS. An exception is described in the next scheme.

- Transparent Relay

The main difference with the previous mode is that the transparent RS does not transmit preamble, FCH and DL-/ULMAP to the MSs. The MS never recognizes the transparent RS, which has to synchronize with the BS to allow communication. As a result no coverage extension is possible. This scheme is also proposed in the IEEE 802.16 draft [42].

In this case, the RS acts as a MS for the BS, in which the DL subframe is split into two zones. Figure 4.4 shows the transparent relay frame structure. In the DL access zone the BS sends data to the MSs and the subordinate RSs. However, The RS transmits to MSs in the following transparent zone - instead of the access zone - in order to avoid interference between users of both sectors. The BS may transmit cooperatively to the same MSs in this frame zone.

In this mode, the interference from BS-MSs and BS-RS to RS-MSs is avoided at the expense of less bandwidth usage than the previous scheme. However, this scheme is designed for throughput enhancement and less number of users at the cell edge compared to the non-transparent mode. Note that inter-BS and inter-RS cell interference still has to be managed. This mode allows less control signal latency and lower overhead than the non-transparent mode. As a result, transparent relays are less sophisticated and therefore less expensive.

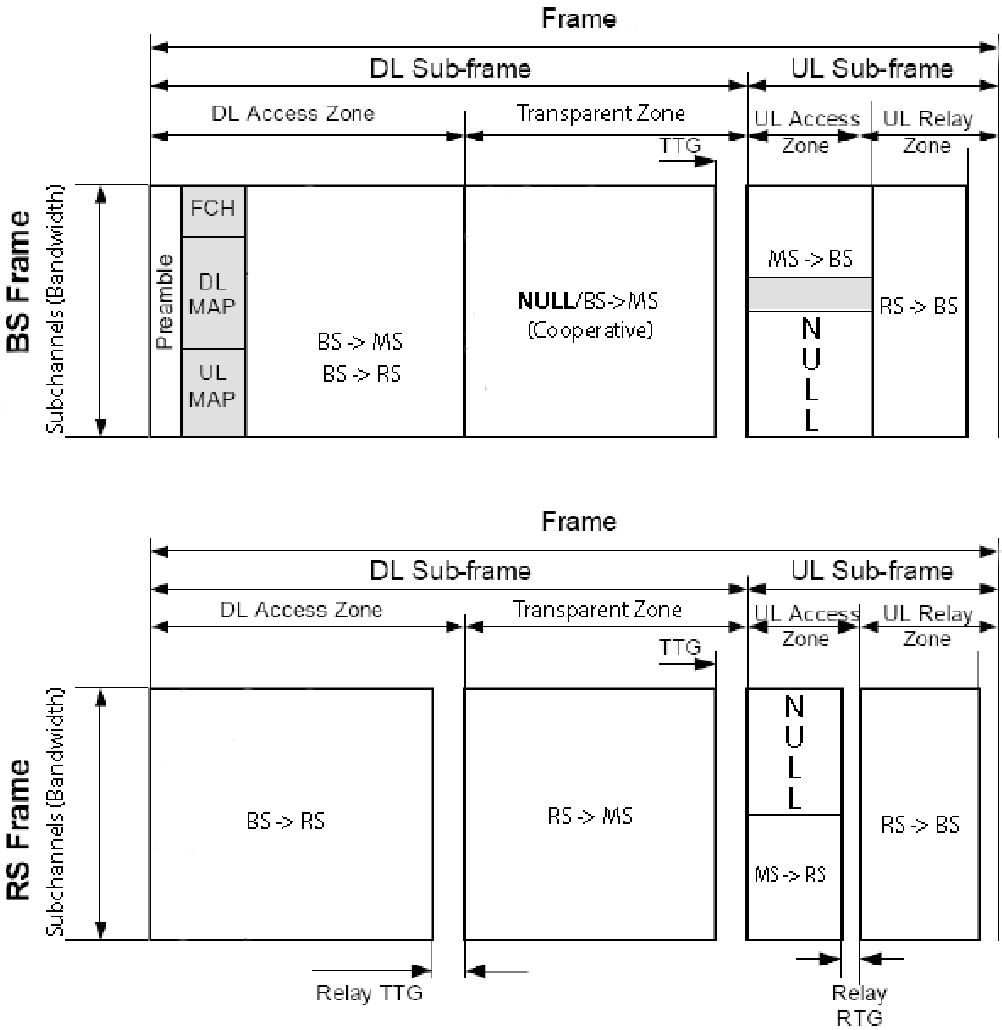

Figure 4.4. Frame structure for transparent relay stations [42].

In a transparent two-hop mode, an amplify-and-forward operation would be allowed in a direct relay mode, where the RS demodulates and deinterleaves its signal, and then immediately interleaves and modulates it for transmission without decoding. In this case, the RS may be able to receive and transmit the same data in the same frame. A decode-and-forward mode is however still preferred in the IEEE 802.16j draft to fully exploit the mobile WiMAX MAC capabilities.

- Dynamic fragmentation

In this subsection, a new simple and flexible frame structure design for Mobile WiMAX relay networks is proposed (Figure 4.5). A similar frame structure to the concept proposed in this section has been previously used for OFDM (Orthogonal Frequency-Division Multiplexing) based technologies such as HiperLan/2 [46] or the Fixed version of WiMAX [47], [48], where the authors describe a subframe concept that implies no special standard options. Another related concept is the one known as the beacon concept [71] also applied in HiperLan/2.

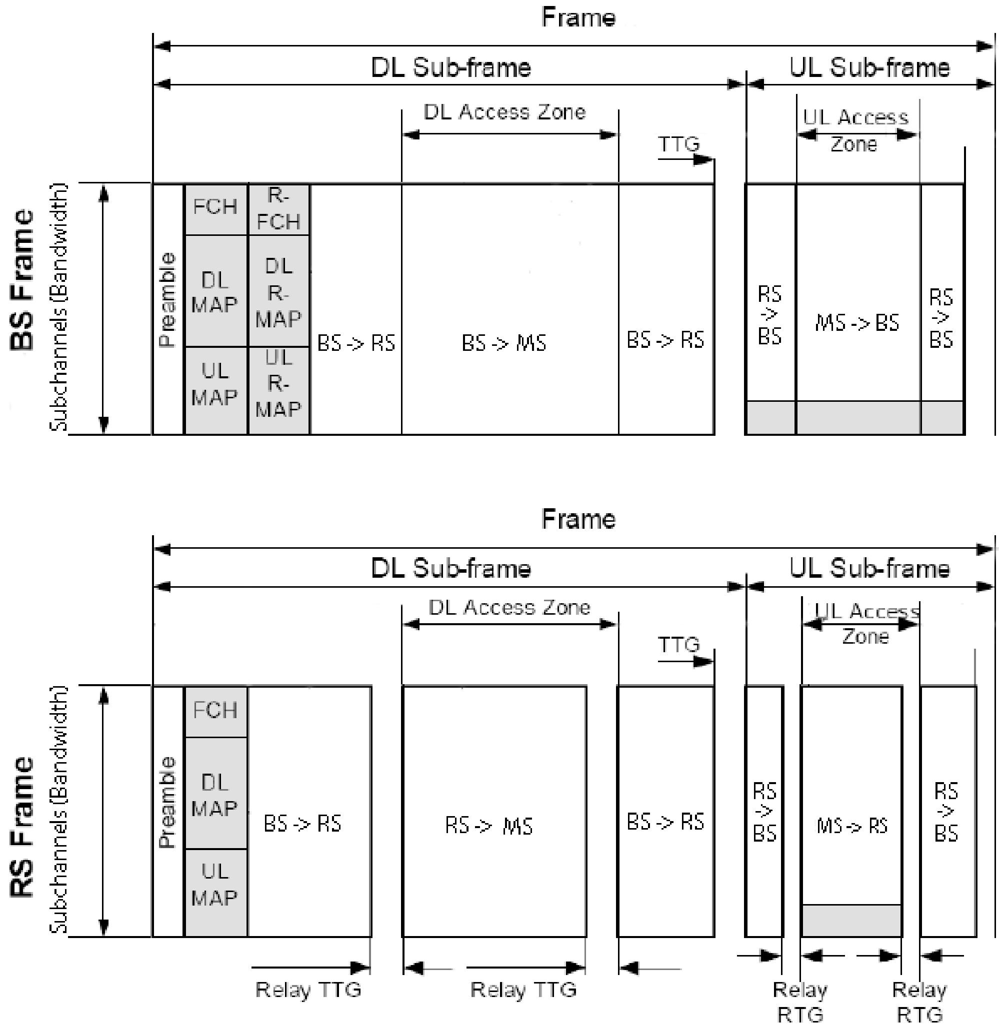

Figure 4.5. Frame structure for the dynamic relay mode.

In this case, subchannel segmentation is also used to avoid interference between the links BS-MSs and BS-RS (Access links) to RS-MSs (Relay Link). The DL and UL access zone will be dynamically located within the BS DL and UL subframe according to the RRM strategies to allow efficient scheduling. The Relay zones will fill the frame area not used.

The RSs and MSs are dynamically controlled by the BS and the RS still maintains the backward compatibility. The RS behaves according to the BS schedule and it receives, and transmits during bursts scheduled for the corresponding hop. DL-MAP and UL-MAP and other management messages will indicate the access and synchronization for the relay users as Figure 4.5 indicates for two hops.

In order to calculate an appropriate schedule, the BS needs to know which connections of the entire relay enhanced cell are active, which QoS requirements they have, and which connections request bandwidth. Having received the broadcast information, the RS filters it and forwards only the relevant subset of information to its subcell [47]. Also in this case it is assumed that the BS and RS are synchronised.

The difference with the previous methods is that the RRM methods will have more flexibility and some improvement in the performance can also be expected for some situations due to multiuser diversity [11] or the system scheduling. On the other hand, the interference avoidance with other BSs and RSs will be performed differently as the location of the relay zones will be performed randomly. The overall system performance will vary depending on the total available resources, number of users and service requirements.

- Out of band relay

This method is a basic configuration that does not apply frame fragmentation, and the RS uses the same band as the BS. Figure 4.6 shows the structure. The backhaul communications can be done through other channels, which increase the costs in this scheme, or using the same channel and directive antennas, also known as not controlled relay. This configuration is simpler than the previous ones.

This configuration may be appropriate for large bandwidths or when the use of subchannel segmentation is not suitable.

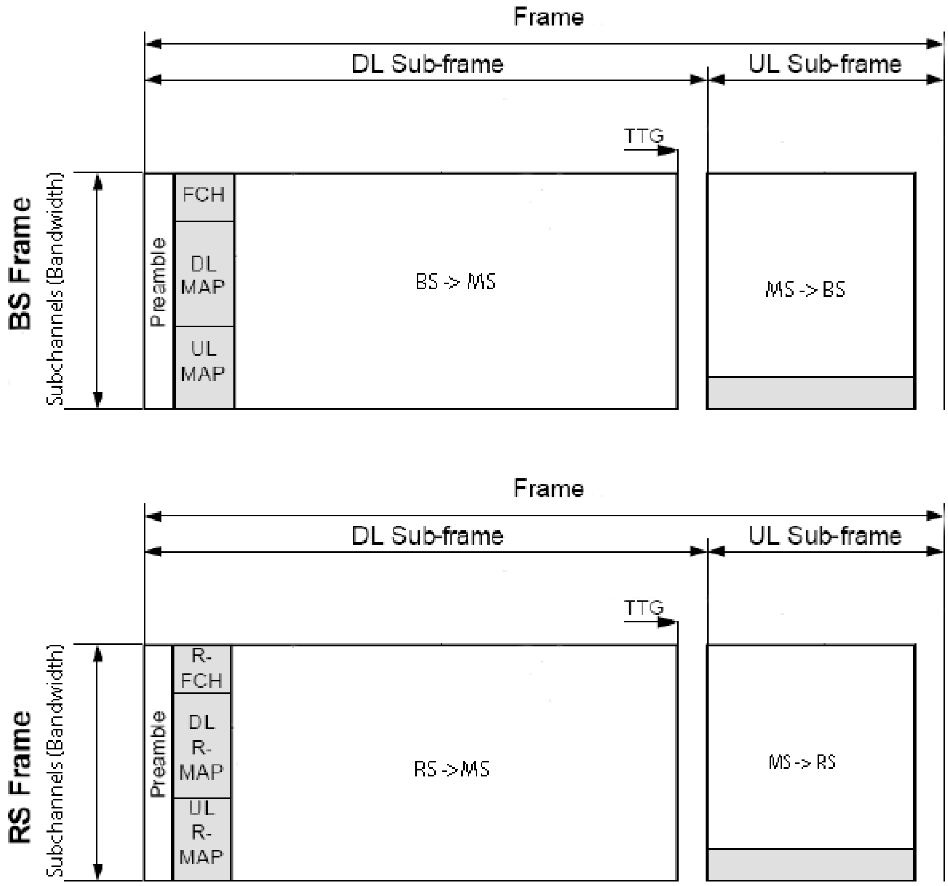

Figure 4.6. Out of band relay mode.

4.2 Efficient System level simulation methods

As mentioned in Section 1.1.2, the use of network simulation is needed in the detailed network planning and optimization phase. In addition, efficient calculations are needed in order to allow the intensive SLS required by the optimisation algorithm. This section introduces a simulation framework and efficient capacity calculations that allow intensive system level simulations in a multihop environment.

4.2.1 A System Level Simulation Framework

Static Monte Carlo simulation is often used in cellular network planning and optimization due to the high computational load that dynamic simulation would require in an iterative process that may need thousands of simulations. This kind of simulation is snapshot-based and represents an instant of the network performance with fixed position of MSs. The simulator takes multiple Monte Carlo snapshots to statistically observe the network behaviour.

The simulation procedure is shown in Figure 4.7. Within in each Monte Carlo snapshot, the throughput is calculated iteratively until the performance of the system converges.

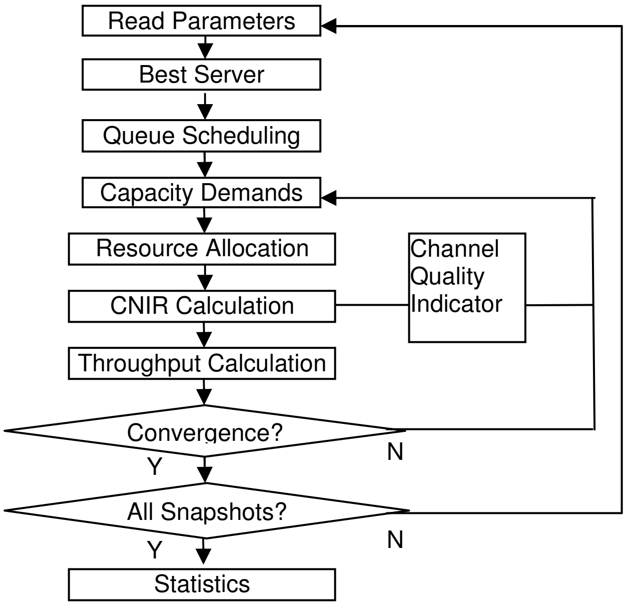

Figure 4.7. SLS procedure.

Since the resources used by one user in one sector may interfere with other users, the system converges when it is stable in the selection of MCS for communication and assignment of resources. In order to fulfil the service requirements of the user, the channel quality indicator (CQI) helps to set the adequate MCS for communication according to the performance obtained in the previous iteration. Link level simulation results are imported in the form of LUT for throughput calculation.

Network time synchronization is necessary in order to minimize multiaccess interference, as well as for the proper performance of handovers. Without proper timing, transmission instants would vary between different Sectors. This could lead to the UL period of some Sectors overlapping with the downlink period of some others, increasing the level of intercell interference. On the other hand, sectors may have different access zone lengths according the traffic requirements, and this may affect the interference in several ways [72].

The model has been built on a C++ framework and Matlab graphical interface. This platform can import and manage many data structures and modules. The propagation model selection is critical in the efficiency of the final simulation process. Empirical, semi-empirical, Ray tracing like or FDTD (Finite Difference Time Domain) like models may be chosen. There is generally a tradeoff accuracy-complexity, and the network designer may choose one model according to its needs. More information can be found at [73].

4.2.2 Efficient capacity calculations

The number of required resources is approximated in the capacity demand stage after the MS queue scheduling. The number of requested slots Rl_m for each user m of sector i is calculated, as shown in (4.1), from the requested capacity from the service Rc_m (kbps) and slot efficiency Se_m (kbps).

$$Rl_m = \frac{Rc_m}{Se_m} \tag{4.1}$$

As (4.2) indicates, Se_m depends on the number of subcarriers per subchannel η_{sub}, and the bearer efficiency $\eta_{B(CQIm)}$ in terms of bits/symbol. The bearer efficiency depends on the MCS assigned to the user, and it is selected according to the CQI feedback from previous iterations. T_{frame} is the frame duration.

$$Se_m = \frac{\eta_{sub} \cdot \eta_{Be(CQI_m)}}{T_{frame}} \tag{4.2}$$

In order to consider the relay link information, a set of active RSs attached to a sector i can to be defined as a subset of S_z ($Rs_i = \{1, ...,w...,W_i\}/ Rs_i \subseteq S_z$). Therefore, a sector w can be located in CS z' ($w \in S_{z'}$), and be a relay of a sector i ($w \in Rs_i$), situated in a different CS location z ($i \in S_z$).

Each RS can act as a user, and the RRM should identify it in order to allocate the appropriate resources in the OFDMA frame according to the relay mode selected. As a result, the RS Se can be defined as:

$$Se_w = \frac{\eta_{sub} \cdot \eta_{Be(CQI_w)}}{T_{frame}} \tag{4.3}$$

After the resource allocation is performed, CINR and throughput are calculated for each resource, r. In order to obtain these values, interference calculations need to be performed.

A certain user m, whose best server is i, is interfered in DL by other stations (j), only if, i and j are using the same resource for transmission within a distance smaller than the system re-use distance. The final interference suffered by the user m will be the sum of all the interference rays coming from neighbouring base stations (j).

$$I_{m,r}^{(DL)} = \sum_{z \in Cs} \left(\sum_{j \in S_z, j \neq i} \left(P_{j,r} \cdot G_j \cdot L_j \cdot L_{j,m} \cdot G_m \cdot L_m \right) \cdot u_{j,r} \right) \tag{4.4}$$

where, ‘m’ indicates the interfered user; ‘i’ is the server base station; ‘j’ are interfering stations; ‘r’ is the studied resource; $P_{j,r}$ is the power applied by j in r in dBm; and L_{jm} is the path loss between user m and station j; Ga stands for antenna gains and L_j and L_m stand for feeding losses. The integer variable $u_{j,r}$ indicates if cell j is using slot r, and it takes into account which zones in the OFDMA frame can be used by each station (access zone and relay zones). Note that the power of all resources within a certain subchannel will be the same. Note that linear units must be used, and that the term losses refer to gain values smaller than 0.

Since Rs_i is a subset of S_z, all stations (BSs and RSs) are taken into account in the DL interference calculation in (4.4).

In the UL methodology, the final interference suffered by the user m is the sum of all the interference rays coming from users in neighbouring base stations (j) plus RSs using the same resources.

$$I_{m,r}^{(UL)} = \sum_{z\in Cs}\left(\sum_{j\in S_z, j\neq i}\sum_{g\in Ms_j}\left(P_{g,r}\cdot Ga_g\cdot L_g\cdot L_{g,i}\cdot Ga_i\cdot L_i\right)\cdot v_{g,r} + \ldots \right.$$
$$\left. + \sum_{j\in S_z}\sum_{w\in Rs_j}\left(P_{w,r}\cdot Ga_w\cdot L_w\cdot L_{w,i}\cdot Ga_i\cdot L_i\right)\cdot v_{w,r}\right) \quad \textbf{(4.5)}$$

where, 'm' and 'g' indicate the interfered and interfering user; 'i' is the server base station; 'j' are interfering stations; 'w' is a RS; 'r' is the studied resource; $P_{g,r}$, $P_{w,r}$ is the power applied by g or w in r in dBm; and L_{gi} , L_{wi} is the path loss between user g or RS w and station i; Ga stands for antenna gains and L_w, L_i and L_g stand for feeding losses.

Note that all BSs/RSs and MSs are considered in the interference calculation. The reason behind this is that the access and relay zones may not have the same size in inter-BS/RS interference modelling. The integer variables $v_{g,r}$, $v_{w,r}$ indicate if user g, or RS w are using slot r in UL.

By using the interference model, the CINR of user m, either in DL or UL, can be calculated as follows:

$$CINR_{m,r} = \frac{C_{m,r}}{I_{m,r} + N_r} \quad \textbf{(4.6)}$$

A RS w, can be considered as a user to a BS.

$$CINR_{w,r} = \frac{C_{w,r}}{I_{w,r} + N_r} \quad \textbf{(4.7)}$$

where, C and I are the received power in dBm of the carrier and interference signals, respectively, and N is the noise power in dBm of the resource r studied. The carrier signal power can be estimated by using:

$$C_{m,r}^{(DL)} = P_{i,r}\cdot Ga_i\cdot L_i\cdot L_{i,m}\cdot Ga_m\cdot L_m \quad \textbf{(4.8)}$$

$$C_{m,r}^{(UL)} = P_{m,r}\cdot Ga_m\cdot L_m\cdot L_{m,i}\cdot Ga_i\cdot L_i \quad \textbf{(4.9)}$$

The noise, which is composed of thermal noise n_0 and the equipment noise figure, Nf_{eq}, can be approximated by (4.10), and (4.11).

$$N_r = n_0 + Nf_{eq} \tag{4.10}$$

$$n_0 = -174\frac{dBm}{Hz} \cdot 10\log(F_{sampling} \cdot \frac{N_{usedSC}}{N_{total SC}}) \tag{4.11}$$

where, $F_{sampling}$ represents the sampling frequency of the system, while N_{usedSC} and $N_{totalSC}$ represent the number of considered and total sub-carriers, respectively. Different metrics can be used to calculate the effective CINR of one subchannel from its subcarriers. In this case, MIC (Mean Instantaneous Capacity) is used [11].

In order to take into account the effects of the radio channel at small temporal scales, LUTs from link-level simulations are used. The use of LUT is critical in network design in order to obtain an efficient network planning and optimization process. Figure 4.8 shows an example of LUT, where an AWGN (Additive White Gaussian Noise) channel is used.

The DL/UL throughput of a MS m in the last hop is calculated, as shown in (4.12), from the BLER value which is a function of the bearer Be used and the CINR level of the studied resource r.

$$T_m = \sum_{r \in Sr} Se_m \cdot \left(1 - BLER_{(Be(CQI_m),CNIR_{m,r})}\right) \cdot v_{m,r} \tag{4.12}$$

Each bearer corresponds to a certain modulation and coding scheme (the values for Mobile WiMAX are shown in Table 4.1). The bearer used for transmission is determined according to the CQI of the MS in a previous iteration, and it is represented by CINR measurements over the whole burst [11].

In the previous hop the integer variable $v_{m', r}$ indicate which slots from the RS relay link ($CINR_{w,r}$) belong to user m (and represented as m')

$$T_{m'} = \sum_{r \in Sr} Se_w \cdot \left(1 - BLER_{(Be(CQI_w),CNIR_{w,r})}\right) \cdot v_{m',r} \quad \textbf{(4.13)}$$

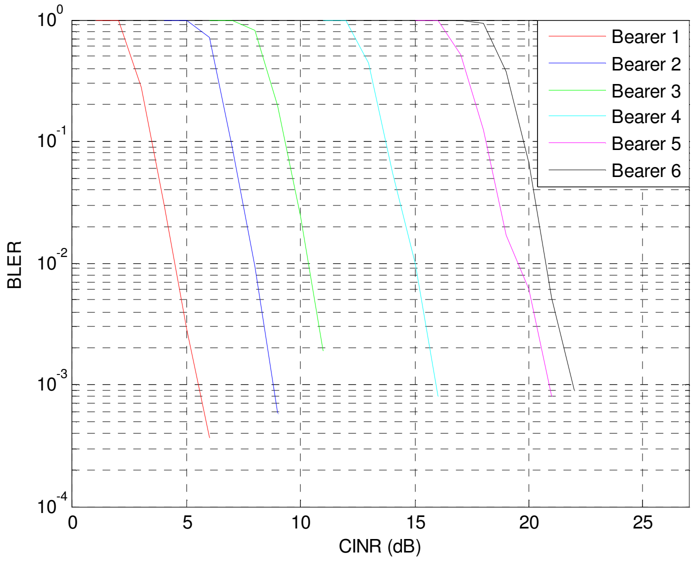

Figure 4.8. BLER values for througput calculation [65].

Table 4.1. Mobile WiMAX Burst Profiles Used.

Radio Bearer index	Modulation	Channel Coding Rate	Bearer Efficiency, η_{Be} (bits/subcarrier)	CINR Threshold (dB)
1	QPSK	0.5	1	2.88
2	QPSK	0.75	1.5	5.74
3	16QAM	0.5	2	8.79
4	16QAM	0.75	3	12.22
5	64QAM	0.67	4	15.88
6	64QAM	0.75	4.5	17.50

In decode-and-forward multihop communications, the throughput is calculated separately in each link since the signal is decoded and encoded again. The final value perceived by the user is determined by the weakest link. Therefore in a two-hop communication the final throughput is:

$$T_m = \min\left(T_{m'}, T_m\right) \quad \textbf{(4.14)}$$

When cooperative communications are considered, the DL signal from other paths may contribute with some gain, which can be modelled according to the CINR values of the cooperative paths in a LUT. A simple but common case is when a MS receives signal from both RS and BS. The LUT relates the CINR of the BS-MS ($CINR_{m,r}^{(S-D)}$) link with a certain gain on the $CINR_{m,r}^{(R-D)}$ value. The number of transmission resources is assumed to be the same in both paths.

The gains may be estimated in previous link-level simulations or a method representing the gain of traditional combining techniques [68], such as the one shown in (4.15):

$$CINR_{m,r} = \varepsilon^2 \cdot (CINR_{m,r}^{(R-D)} + CINR_{m,r}^{(S-D)}) \quad \textbf{(4.15)}$$

where ε represents the losses due to the combination of signals (ε = [0.2 – 0.3] dB).

4.3 Network capacity evaluation

This chapter focus on the network planning part of the process described in Section 1.1.2. The objective is to obtain an efficient SLS with accurate results. A following optimisation stage will make use of the SLS results in order to obtain an optimal configuration. Next chapter will focus on that stage that aims to deal with the non-uniform characteristics of realistic scenarios. In this section some of the frame structures described in section 4.1.3 are evaluated over a regular scenario.

The SLS methodology presented for two-hop Mobile WiMAX networks makes use of Monte Carlo simulations. As a result, some dynamic processes like handovers are not accurately simulated. Static MSs are created with a fixed position each iteration of the process, and the results of multiple iterations with different positions is analysed statistically.

Based on the model described above, this section evaluates a uniform Mobile WiMAX network scenario formed by seven 3-sector sites and four relay sectors, having a total of 25 sectors. The scenario used is shown in Figure 4.9. The RS configuration shown has been selected as it has been a common configuration of RSs and repeaters in the evaluation of hexagonal and regular scenarios in UMTS [6]. The selection of one RS attached to one BS at high transmit power corresponds to the focus on the detailed planning and optimisation stage, where parameters with high costs get an initial configuration. An example is when a network with only BSs needs to be expanded due to an increase of capacity demand or a coverage extension requirement.

A rectangular area of 4 km × 4 km is considered, where 250 MSs are randomly distributed in this area for each snapshot performed in the Monte Carlo routine.

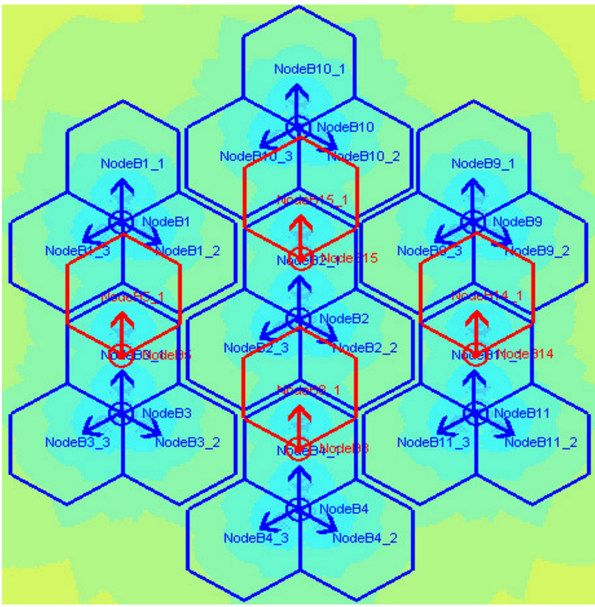

Figure 4.9. Simulation scenario.

The validation of this model presents some difficulties since at the moment of producing this manuscript there is no other IEEE 802.16j SLS specialised on network planning and optimisation. The PMP version of the methodology presented was validated in [73], and the processing times were similar than a network planning and optimisation state-of-the-art PMP tool [1]. The results for in-band non-transparent relay mode produce similar results than other works such [68], [48].

Results for both the number of users and throughput are discussed. Non-transparent mode, dynamic fragmentation mode and only PMP configuration (without RS consideration) are evaluated. Since transparent RSs has an specific purpose, which only provides throughput enhancement to users also covered by the BS, this mode will be compared to the other modes within non-uniform and specific scenarios in Chapter 5.

The parameters used for the simulation are shown in Table 4.2. The non controlled relay performs the backhaul communications over the same channel.

Table 4.2. Simulation Parameters.

Parameter	Value	Parameter	Value
PMP Sectors	21	BS/RS Antenna Height	40/30m
Relay Sectors	4	CPE Antenna Gain	0dBi
Carrier Frequency	3.5GHz	CPE Antenna Pattern	Omni
Channel Bandwidth	10MHz	CPE Antenna Height	1.5m
DL:UL Ratio	1:1	CPE Noise Figure	5dB
Permutation Scheme	AMC	CPE Cable Loss	0dB
BS/RS TX Power	43/30dBm	Min Service TP	64Kb/s
BS/RS Antenna Gain	18dBi	Max Service TP	92Kb/s
BS/RS Antenna Beam	90º	σ(Shadow Fading)	8dB
BS/RS Antenna Tilt	3º	Path Loss Model	COST231- Hata
BS/RS Noise Figure	4dB	Traffic Map Density	30-90u/Km^2
BS/RS Cable Loss	3dB	CPE Tx Power	23dBm

Resources in subchannels are allocated in a maximum throughput strategy described in [73]. In practice, this method implies a non-fair use of resources as it is achieved by serving first users with high CINR values, and therefore users requiring fewer resources. However, this does not affect the comparison done in the following as they are equally applied to all schemes. As it will be indicated in the future work, a joint study of RRM and frame fragmentation may be interesting.

As described, in Figure 4.3, Figure 4.5 and Figure 4.6 overhead is considered within the different kind of fragmentations. They normally require only 1-2 symbols in the domain time, which is considered in every scheme.

Four different simulations have been performed to check how the capacity of the network changes when the channel bandwidth varies. All the parameters have been kept constant for these simulations (with a type of service for all the users of 64 kbps - web browsing), except the channel bandwidth, which has been modified using the following values: 1.25MHz, 5MHz, 10MHz and 20MHz. The results of the simulations are shown in Figure 4.10 and Figure 4.11.

In a system with a low bandwidth (1.25MHz), the number of successful users is low due to the small number of available resources. The improvement in the CINR values provided by the relay configuration is slightly bigger in the dynamic fragmentation mode compared to the non-transparent mode. In a bandwidth of 5 MHz the difference is significantly more favourable in the new proposed mode. The reason is that in this scenario, where the objective is to cover users with low CINR, and therefore users that need a lot of slots to fulfil their requirements, interference is efficiently managed locating the access zones dynamically, as resources are scarce. The non-controlled relay offers bad performance due to high interference levels between access zones and relay links in the serving BSs.

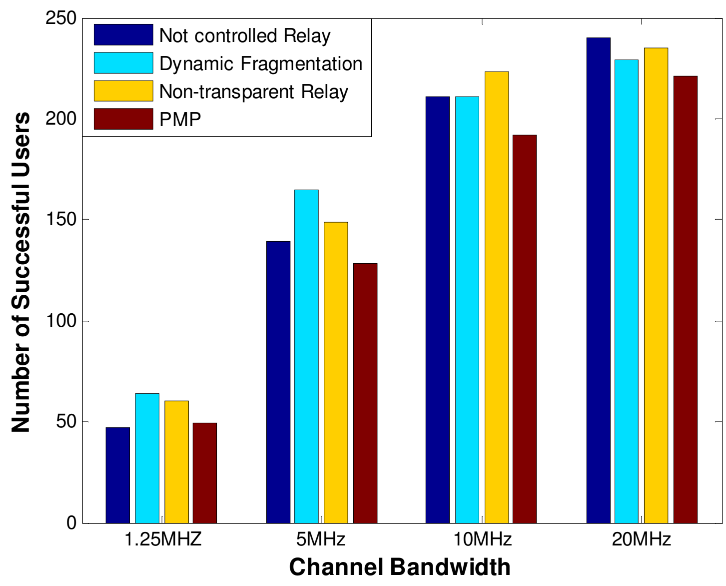

Figure 4.10. Capacity Vs Channel Bandwidth.

As the bandwidth increases to 10MHz more resources can be allocated and the total number of successful MSs increases. The best relay strategy is the non-transparent relay since inter-BS interference is more efficiently avoided in relay links with the use of more resources and bigger frame zones.

In a system with a large bandwidth (20MHz), the number of successful users is large due to the large number of available resources. The difference with respect to a 10MHz bandwidth is not important because almost the total number of users was already successful with this kind of scenario. Regarding the relaying methods, the number of blocked users is already low in the PMP case, and the slight improvement of relay configuration in dynamic fragmentation and non-transparent mode is due to the users who previously were beyond coverage. Note that the non controlled relay configuration provides good performance in this configuration due to the big number of resources available.

In Figure 4.11, it can be observed how the throughput of the network increases, when the bandwidth increases. The proposed scheme for relay configuration provides the best performance for 5 MHz bandwidth and it is outperformed by the rest of relay configurations as the bandwidth is increased.

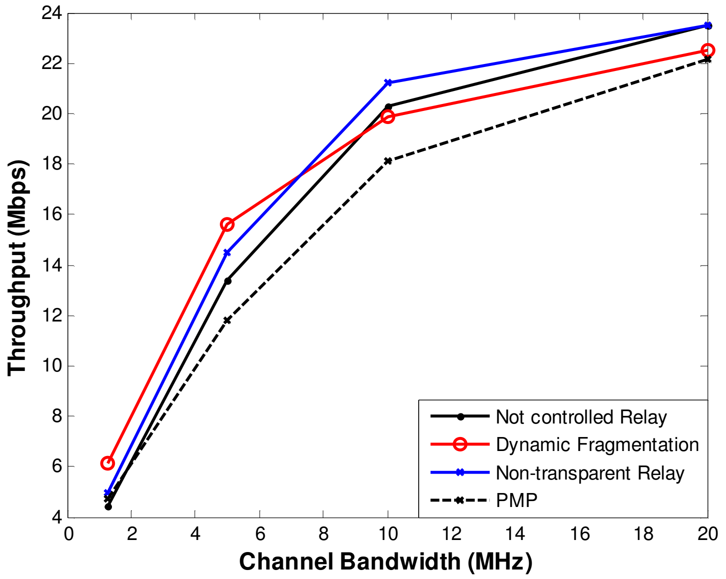

Figure 4.11. Throughput Vs Channel Bandwidth.

In general, when the bandwidth increases the interference decreases since there is less probability that the resources are reused in several cells at the same time. As a result, the capacity of the system is increased. The addition of RSs provides higher power to the users and also helps to decrease the interference levels with the frame fragmentation allowing higher CINR values. Higher CINR in the different bursts of a user means fewer slots to maintain a certain throughput in this system and thus, more users that can be served. The final QoS perceived by the MSs will depend on two links, BS-RS and RS-MSs. However, note that the former one normally presents very good performance due to the location chosen for the RS in this scenario.

Three different simulations have been done to check how the capacity of the network changes when the service bit rate requirement varies. All the parameters have been kept constant for these three simulations (with a bandwidth of 5MHz), except the service throughput requirement for all the users, which has been modified using the following values: 12.2Kbps (VoIP), 64Kbps (Video-Conference) and 128Kbps (Web-Browsing). The results are shown in Figure 4.12 and Figure 4.13.

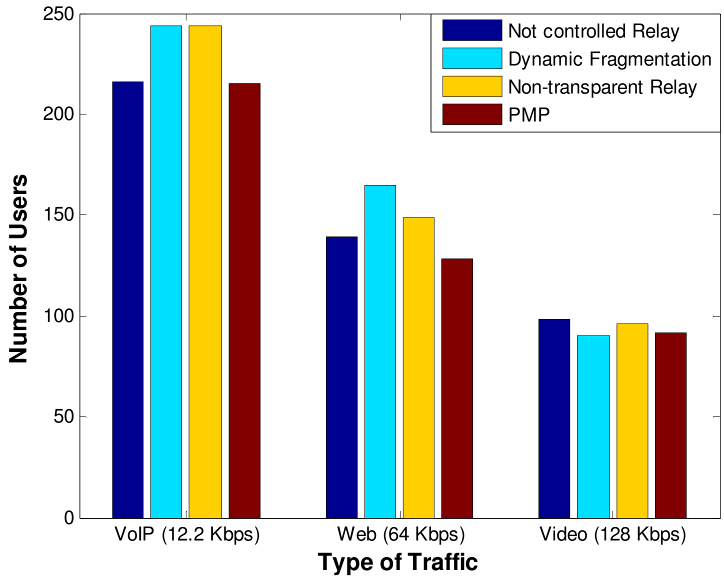

Figure 4.12. Capacity Vs Service Requirement.

In a service with a low throughput requirement, like in the case of VoIP, the number of successful users is large due to the existence of more resources available than resources requested in the system. In this case, the improvement due to the addition of RSs is bigger due to the greater flexibility for the RRM procedures, which are able to satisfy more users, and both dynamic fragmentation and non-transparent mode have similar performance.

With higher throughput requirement (web browsing), where the number of slots needed by each user is bigger, the resources will be more efficiently managed in a dynamic way as the frame may be filled faster.

In a service with a high throughput requirement (such as the video-conference service) the number of successful users in the system is low due to the high number of resources needed by the MSs. The number of users fitting within the frame is smaller and the probability of interference is larger. Also, for this reason, the improvement of the relay configuration is not important.

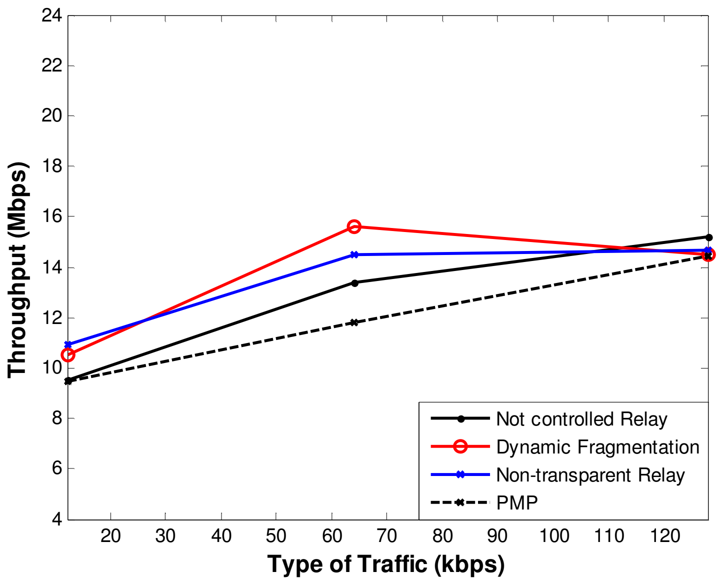

Figure 4.13. Throughput Vs Service Requirement.

In Figure 4.13, it can be observed how in some cases, the throughput of the network increases when the service bit rate requirement increases. The scheme for dynamically controlled relay configuration provides the best performance when the service requirement is web browsing (64 kbps).

Further comparing Figure 4.12 and Figure 4.13, it can be observed how the capacity in terms of users decreases and the system throughput is increased (or maintained) when increasing the service requirement. The reason is that from a system throughput perspective it is more favourable to have few users supporting a high bit rates than a lot of users with a low bit rates, which just satisfy their requirements.

These results highlights the complexity of OFDMA relay-based network design process, where the operator needs to estimate a number of parameters such as the number of users to serve, the services that will be used or the number, position and configuration of BSs and RSs.

4.4 Conclusions

This chapter proposes and evaluates a new simple and flexible scheme to integrate multihop communications in the mobile WiMAX frame structure in a comparison with another basic concept already proposed for the upcoming IEEE 802.16j release of the standard. In addition, it introduces efficient capacity calculations that allow intensive system level simulations in a multihop environment for the purpose of network design. Network capacity is evaluated for various configurations and services allowing interesting observations about OFDMA radio efficiency.

The addition of RS will normally improve the performance of the system, but this improvement should be evaluated for every situation, since new RSs may incur in some tradeoffs or more costs. The proposed scheme for relay configuration is designed to avoid interference in the time domain by taking advantage of the necessary frame fragmentation to integrate in-band relay links. The dynamic allocation of access zones is more beneficial in networks with low-medium bandwidth, where interference avoidance is more critical. In the scenario evaluated, the relay links are set to offer optimum performance by locating the RSs in the direction of BSs maximum gain.

The results obtained show small gains related to the regular/hexagonal nature of the scenario, chosen to evaluate the efficient simulation model presented. However, final results depend on the particular configuration on the network, which shows the complexity of relay-based network design process, where the operator needs to estimate a number of parameters to get a satisfactory network performance. A following optimization process is considered in next chapter, and it aims to deal with the non-uniform characteristics of realistic scenarios, where the existence of hotspots, different channel characteristics for the users, or different service requirements will determine the final design of the wireless network.

The selection of one RS per BS corresponds to the focus on the detailed planning and optimisation stage, where parameters with high costs get an initial configuration, like the case of a network expansion need due to an increase of capacity demand or a coverage extension requirement. However, as a future work, more RSs may be considered per BS and also more hops in the communication. New methods and evaluation under different scenarios and network configurations can be done. The effect of different scheduling algorithms and resource allocation methods can be studied on the described platform for the purpose of WMAN network design.

Chapter 5: Planning and Optimization of OFDMA Multihop Relaying Networks

There are several aspects that influence the design of wireless networks. As new technologies and services emerge, it is very important to accurately identify the factors that affect this process and define how to take them into account to achieve optimally performing and cost-efficient networks. This chapter explains the impact of relays on the network architecture design and presents examples of deployment cost analysis for different services and propagation conditions in the context of Mobile WiMAX networks. Information about the quality and reliability of different strategies is provided and the model introduced in Chapter 1 is considered.

Some approaches on multicarrier technologies such as GSM suggested the use of radio frequency planning in order to mitigate interference and enhance the system performance [56] [25]. RTP has been traditionally performed in an operation and maintenance stage after an initial detailed network planning and optimisation. Similarly, OFDMA supports reconfiguration of the subchannel usage with different frequency reuse schemes in order to mitigate intercell interference. However, traditionally they only consider fixed reuse patterns or static assignment of subchannels in regular or hexagonal scenarios [57][58][59], and do not describe a joint formulation with other parameters such as antenna tilt/azimuth.

This chapter establishes the fundamentals of a new approach that integrates the configuration of BS/RS location, tilt, azimuth with an optimal frequency assignment in OFDMA networks - in the form of subchannel planning in frequency domain - according to the traffic and the radio channel conditions in a novel two-step process for network design.

5.1 Multihop Relaying Networks Design

5.1.1 Deployment of Relay-Based Networks

In general, the deployment of a relay-based network depends on the particular scenario. However, there are different strategies that can be adopted according to the service provider needs, illustrated in Figure 5.1. RSs can help to extend the BS coverage by placing them close to the edge of the coverage area.

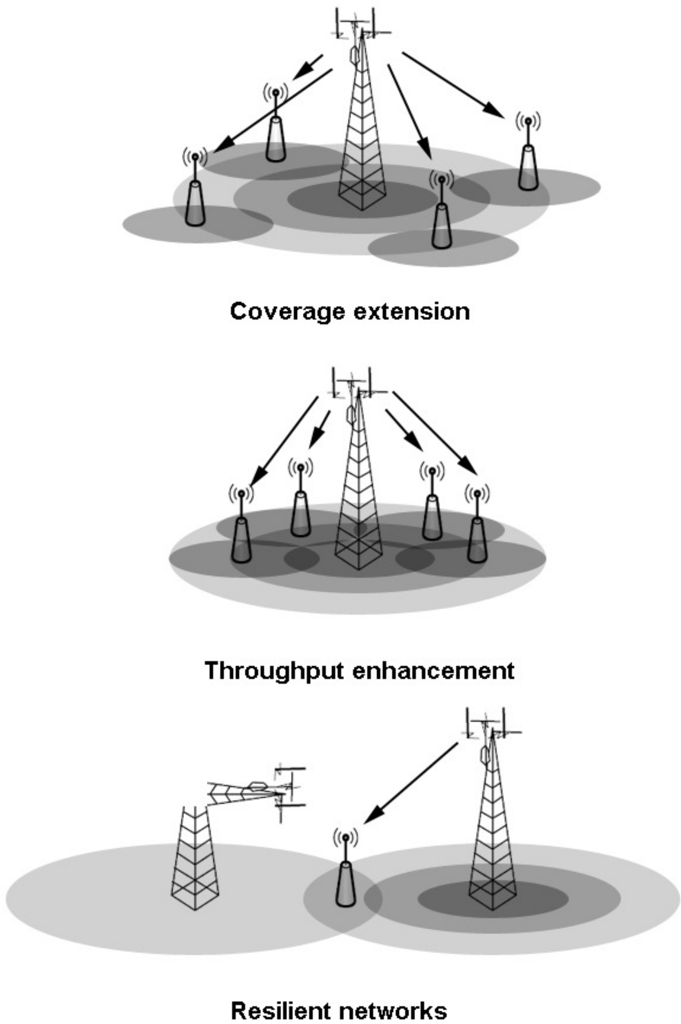

Figure 5.1. Scenarios for different strategies in BS and RS position.

On the other hand, today's OFDMA networks supports a variety of MCS [11], [9], and the system can select the most robust schemes for poor radio propagation characteristics (i.e. low CINR), normally in regions close to the cell edge. The most robust MCSs generate low throughput levels, and therefore the use of additional RSs in these areas can help to increase the CINR, and also the throughput.

Another aspect to take into account in a relay-based wireless network is the network resilience, due to a higher probability of having a failure in a multihop communication. Different strategies can reflect diverse abilities of the network to recover from an event of having a problem in a station or a link failure.

5.1.2 OFDMA Frequency Planning

There are several ways of performing a frequency planning in OFDMA networks. In the fixed subchannel allocation (FSA), the whole channel is divided in several segments, and each one is assigned to a different sector. This scheme simplifies the radio frequency planning design, since the operator only needs to assign segments to sectors. In the case of three sectors per site, this procedure mitigates intercell interference by reducing the probability of slots collision by a factor of 3. However, the sector capacity is also reduced by a factor of 3. This method is suitable for regular/hexagonal scenarios [74].

Dynamic frequency planning (DFP) in the form of subchannel planning has been recently been proposed for OFDMA in the existing literature (e.g. [36], [74]), where the use to a certain subset of subchannels are assigned to different sectors and can be adapted to the environment conditions. In other words, DFP takes advantage of multiuser channel and traffic diversity to adjust the subchannel allocation. Basically, DFP can be divided into two categories: centralized DFP, where the subchannel allocation decision is made by a central controller; and distributed DFP, where the subchannel allocation decision is made by MSs or BSs independently.

The aim of frequency planning is to restrict the use of the subcarriers to a certain subset of subchannels in each sector of the network. This subchannel restriction may involve a loss of cell resource utilization efficiency. However, the objective of this procedure is to reduce the interference and collisions, thus having higher CINR levels that should enhance the system performance.

As described in Chapter 2, the OFDMA frame need to be fragmented in order to accommodate in-band relay links. This chapter proposes the joint use of time domain forwarding and traditional frequency planning in the form of subchannel planning.

5.2 A Two-Step Method for the Optimization of OFDMA Multihop Networks

5.2.1 An Optimisation Framework for Joint Station Configuration and Frequency Planning

This section proposes a centralized DFP scheme procedure that can be performed in the detailed planning and optimization phase as an inner optimization loop, with an optimal assignment of subchannels for each configuration tested. The subchannel usage can be configured in a new network, or an expansion of an existing one with new BRs/RSs to react against the emergence of new hotspots or different traffic during seasons. A distributed DFP scheme would only be suitable in the operation and maintenance phase of an already deployed network, where self-configuration techniques may be more suitable [31].

The subchannel restriction of the frequency plan should be combined with the use of the necessary frame fragmentation of in-band relay communications and cooperative procedures. The different size of access zones need to be considered.

In centralized DFP, the system intercell interference can be characterized by a coupling matrix (see [36], [74]), without the need of full simulation of RRM procedures. As a result, the optimization becomes an efficient process, where an optimization algorithm iteratively goes through different possible solutions for subchannel assignments aiming to minimize the global interference represented in the mentioned matrix. The difference of this optimisation procedure with respect to optimisation of other kind of parameters is that the initial interference characterisation between any two sectors is not changing for every subchannel solution tested, and therefore, the same matrix can be used.

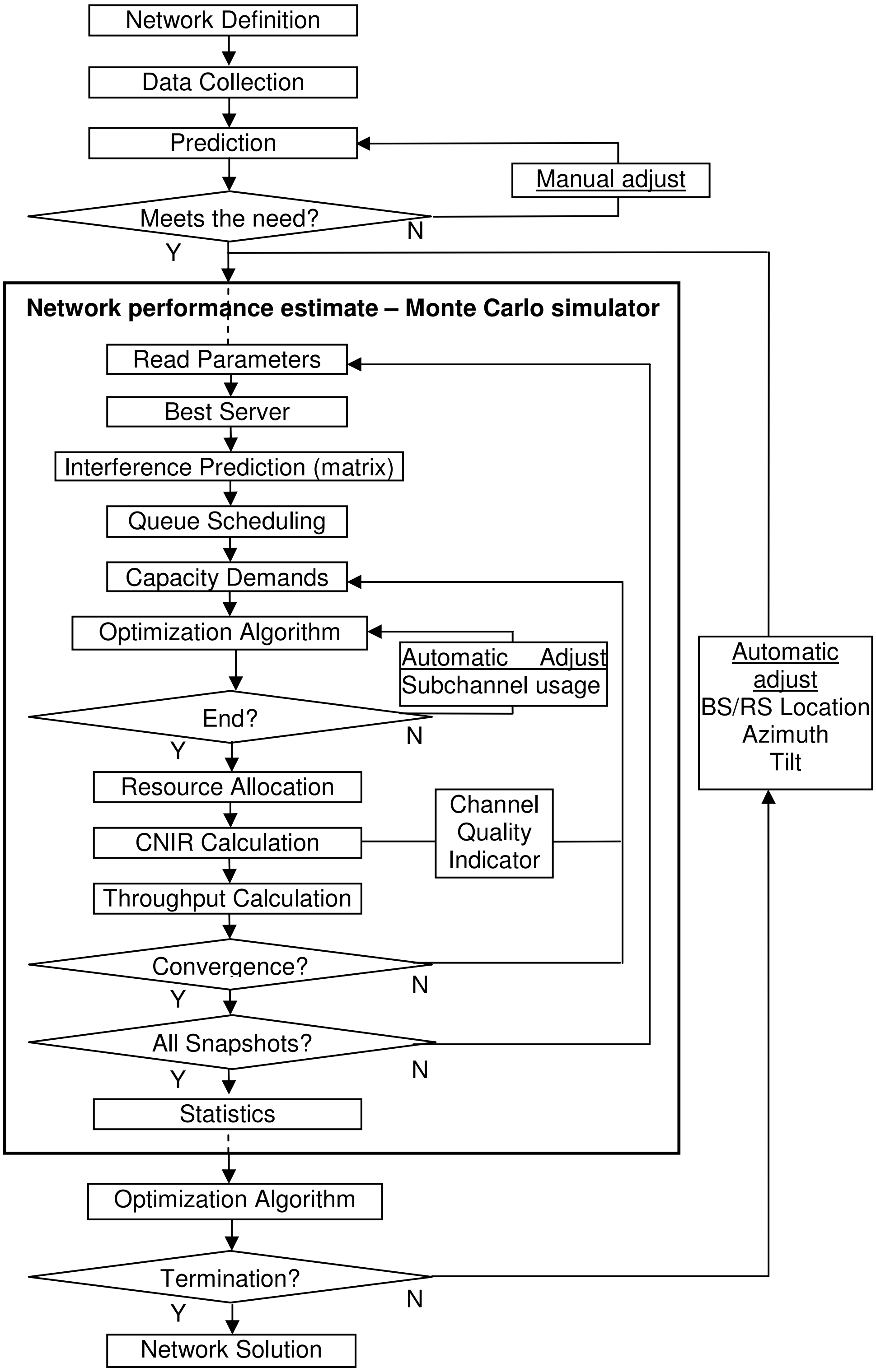

Figure 5.2. Optimisation framework.

The MS service requirements and the MCS estimated from the CQI are needed to compute the number of required subchannels per cell. The internal subchannel configuration is done in two stages, interference prediction and capacity demands integrated in the simulation model described in Chapter 4. Both inner and outer optimisation loop are described in the following section, and sketched in Figure 5.2.

5.2.2 Inner Optimization Loop: Frequency Planning

The objective of the frequency planning procedure described in this section is to offer an optimal assignment of subchannels for each configuration tested of other parameters (e.g. BS/RS position) in order to achieve a best network solution. The assignment of subchannels of the final solution found should be applied when the network is deployed.

At each iteration of the SLS process, the first automatic adjustment of the framework is done. After the system is initialized, the best server of each user is identified and a number of MSs are associated to a sector (M_i). This is done by checking the pilot power strength of all stations, including RSs.

- Interference Prediction

The interference prediction stage characterizes the intercell interference of the network, represented by the restriction matrix W of size $(Z \times N_z) \times (Z \times N_z)$, in which $w_{i,j}$ represents the intercell interference between sectors i, j.

Let us introduce the concept of Interference Event (Ie). There is an interference event between two sectors, i (server) and j (neighbour) over a user m ($Ie_{i,j,m} = 1$), if the power level of the carrier signal (from i to a served user m) is smaller than the power level of a neighbouring interfering signal (from j to the same user) plus a given threshold. Some cases do not account for interference. The first case is cooperative communications between RS and the serving BS, since the MS must be able to identify and combine the different signals to obtain diversity gain. The other is between RS and serving BS access zones in transparent relays (due to frame fragmentation).

The threshold T_{hres} is a protection margin against interference and it is set empirically by the operator according to its planning targets. A small value in this threshold involves a more conservative approach by assuming that neighbouring users are more likely to interfere. However, reasonable variations in this term does not significantly affect the optimal subchannel allocation, as it affects the whole system modelling. T_{hres} is set to 12 dB, which shows a good performance in a Mobile WiMAX system [36]. The following pseudocode shows the procedure for DL:

```
for(z = 1; z ≤ Z; z + +)                                   //All candidate sites
    for(i = 1; i ≤ N_z; i + +)                             //All sectors
        for(z = 1; z ≤ Z; z + +)                           //All candidate sites
            for(j = 1; j ≤ N_z; j + +)                     //All neighbour sectors
                for(m = 1; m ≤ Ms_i; m + +)                // All mobile stations
                    if( C^(DL)_m,i,k < I^(DL)_m,j,k + T_hres(12dB)) then
                    {
                        if(m ∈ Ms_i && m ∈ Ms_j) then      //Cooperative
                            continue                       //communications
                        If (transparent && ( i ∈ Rs_j || j ∈ Rs_i )
                            Continue                       //Frame Fragmentation
                        Ie_i,j,m = 1
                    }
```

The modelling considers that the same subchannel k is used. The following expressions show the calculations needed ((5.1), (5.2)).

$$I^{(DL)}_{m,j,k} = P_{j,k} \cdot Ga_j \cdot L_j \cdot L_{j,m} \cdot Ga_m \cdot L_m \quad \textbf{(5.1)}$$

$$C^{(DL)}_{m,i,k} = P_{i,k} \cdot Ga_i \cdot L_i \cdot L_{i,m} \cdot Ga_m \cdot L_m \quad \textbf{(5.2)}$$

where, ‘m’ indicate the interfered user; ‘i’ is the server base station; ‘j’ are interfering stations; P is the power applied in dBm; and L_{im} is the path loss, Ga stands for antenna gains and L for feeding losses. Note that linear units must be used, and that the term losses refer to gain values smaller than 0.

As shown below, the UL methodology is similar. The main difference is that the interference is calculated from interfering users and RS in neighbour cells, as shown in (5.3), (5.4). The main problem in this procedure is that the position of the real MS interfering in UL after RRM is not known. In UL, the worst case scenario is adopted, where a neighbour MS close to the network edge is characterized as the UL interferer. However, in reality, the position of the UL interferer may be at further distance due to the radio resource allocation of the neighbour cell. This fact can be mitigated by regulating the parameter Thres.

```
for(z = 1; z ≤ Z; z + +)                                    //All candidate sites
    for(i = 1; i ≤ N_z; i + +)                              //All sectors
        for(z = 1; z ≤ Z; z + +)                            //All candidate sites
            for(j = 1; j ≤ N_z; j + +)                      //All neighbour sectors
                for(m = 1; m ≤ Ms_i; m + +)                 //All mobile stations
                    for(g = 1; g ≤ Ms_j; g + +)             // All neighbour MSs
                    {
                        if( C^(UL)_{m,i,k} < I^(UL)_{m,j,k} + Thres(12dB)) then
                        {
                            if(m ∈ Ms_i && m ∈ Ms_j) then   //Cooperative
                                continue                    //communications
                            If (transparent && ( i ∈ Rs_j || j ∈ Rs_i )
                                continue                    //Frame Fragmentation
                            Ie_{i,j,m} = 1
                            break                           //Interferer found
                        }
                    }
```

The following expressions show the calculations needed:

$$I^{(UL)}_{m,j,k} = P_{g,k} \cdot Ga_g \cdot L_g \cdot L_{g,i} \cdot Ga_i \cdot L_i \quad \textbf{(5.3)}$$

$$C^{(UL)}_{m,i,k} = P_{m,k} \cdot Ga_m \cdot L_m \cdot L_{m,i} \cdot Ga_i \cdot L_i \quad \textbf{(5.4)}$$

where, ‘m’ and ‘g’ indicate the interfered and interfering user; ‘i’ is the server base station; ‘j’ are interfering stations; ‘w’ is a RS; P is the power applied in dBm; and L_{gi} is the path loss, Ga stands for antenna gains and L for feeding losses.

The system intercell interference is modelled by W, as (5.5) indicates, in terms of percentage of time that users of certain cells interfere each other considering that the same subchannel is used, and taking into account the requested user capacity by the MS Rc_m, e.g. 12.2 kbps for VoIP services or 64 kbps for video. This gives a service-oriented perspective to the problem since users with more throughput demands will generally need more resources, and therefore have more probability of being interfered. The fact that some users need more resources than other due to the channel conditions is still not considered in this estimation.

The power level of the server, neighbouring cells and other information is reported using measurement reports within the OFDMA radio frame.

$$w_{i,j} = \frac{\sum_{m \in Ms_i} Ie_{i,j,m} \cdot Rc_m}{\sum_{m \in Ms_i} Rc_m} \quad \textbf{(5.5)}$$

The matrix W models the interference statistically in a non-regular scenario. Interference between MSs in access zones is avoided by restricting the use of k subchannels to different stations. This subchannel allocation procedure only makes sense in colliding access zones, like in the non-transparent relay mode or inter-RS interference mitigation in the transparent zone as it is described in next section. The dynamic fragmentation mode proposed in Chapter 4 will not be considered in this procedure.

DL and UL procedures have to be performed separately, resulting in different matrixes W, and therefore different solutions for subchannel allocation in different DL and UL frame zones. Relay zones are fully utilised to increase the subchannel allocation degree of freedom in the access zones and adapt to the particular traffic loads of different cells. RRM and the selection of the appropriate relay zone size will manage the relay zones interference, which is less critical in this zone.

- Capacity Demand

After the queue scheduling, the optimal number of required subchannels per sector D_i is approximated in the capacity demand stage. As mentioned, this work proposes a subchannel allocation taking into account the traffic demands in the access zones in both BS and RS. The reason is that relay links are expected to be more reliable due to an optimal position of the RSs during the design process, and MSs communications are more likely to be affected by changes in the traffic density or channel conditions.

Network time synchronization is necessary in order to minimize multiaccess interference, as well as for the proper performance of handovers. Without proper timing, transmission instants would vary between different Sectors. This could lead to the UL period of some sectors overlapping with the downlink period of some others, increasing the level of intercell interference. On the other hand, sectors have different access zone lengths according to the traffic requirements. Also, some sectors may not have associated RSs and therefore MSs attached to this sector use the whole frame as an access zone and need less subchannels, D_i. This is represented in Figure 5.3 and should be taken into account in the formulation of the problem. The figure represents in grey colour the available resources in the OFDMA frame. The number of symbols in the access zone (η_{Si}) may be different in several stations.

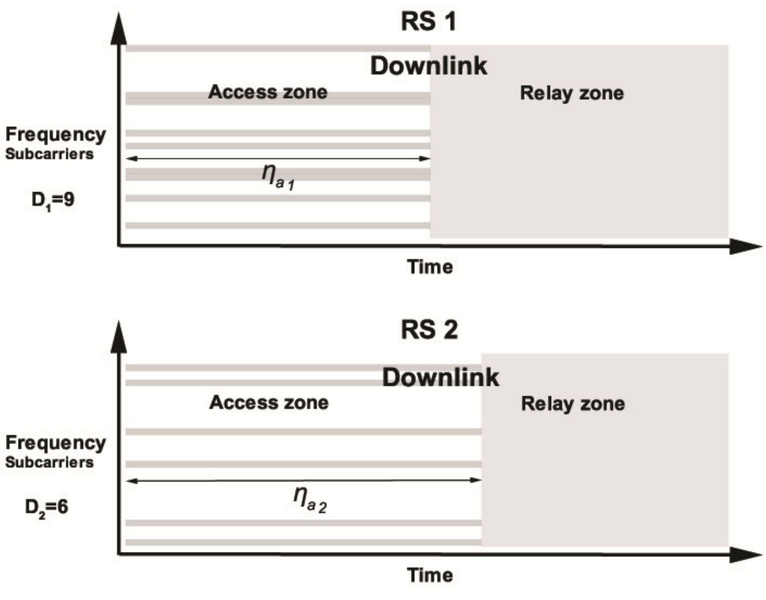

Figure 5.3. Example of DL frequency planning with access zone consideration.

As introduced in (4.1) and (4.2), the number of requested slots Rl_m for each user m of sector Si is calculated, from the requested capacity Rc_m (kbps) and slot efficiency

Se_m (kbps). The average symbol efficiency depends on the average MCS selected by the users within the sector. It can be obtained from previous frames.

The requested subchannels Di are calculated in two different ways depending on the relay mode selected:

- Non-transparent relay. D_i in DL/UL for both BS and RS is calculated by dividing sum of all Rl_m over al MSs within the sector i - which includes cooperative communications - by the number of symbols in the access zone per subchannel η_{Si}:

$$D_i = \frac{1}{\eta_{Si}} \cdot \sum_{m \in Ms_i} RL_m \qquad \textbf{(5.6)}$$

  A factor η_{Cij} represents the portion of zone that is being interfered by the neighbour sector (see (5.7)), and will be used to weight W.

$$\eta_{Ci,j} = \frac{\min(\eta_{Si}, \eta_{Sj})}{\eta_{Si}} \qquad \textbf{(5.7)}$$

- Transparent relay. When transparent relays are used, BSs use the access zone for MSs attached to it, and also for relay links. The transparent zone is used for cooperative communications. RSs use the transparent zone for the access of its own users Figure 4.4. The frequency planning procedure is therefore useful to avoid interference inter-RS, inter-BS and inter-RS-BS (for BSs with no RS attached using the whole frame).

  As a result, MS attached to the RSs should not be considered in the DL BS access zone. Instead, resources for the relay links have to be included for D_i calculation.

$$D_i = \begin{cases} \frac{1}{\eta_{Si}^{(DL)}} \cdot \left(\sum_{m \in Ms_i'} Rl_m + \sum_{w \in Rs_i} \frac{\sum_{g \in Ms_w} Rl_g \cdot Se_g}{Se_w} \right) & \text{if DL \&\& i=BS} \\ \frac{1}{\eta_{Ti}} \cdot \sum_{m \in Ms_i} RL_m & \text{if DL \&\& i=RS} \\ \frac{1}{\eta_{Si}^{(UL)}} \cdot \sum_{m \in Ms_i} RL_m & \text{if UL} \end{cases} \quad \textbf{(5.8)}$$

where:

$$Ms_i' = Ms_i - \left(Ms_i \cap \left(\bigcup_{w \in Rs_i} Ms_w \right) \right) \quad \textbf{(5.9)}$$

when DL and BS. Since the transparent zone in the DL BS frame is only used for cooperative communications no frequency plan is done for it. The DL RS transparent zone is calculated considering number of symbols in the transparent zone per subchannel η_{Ti}.

The factor η_{Cij} represents the portion of zone interfered in different options, shown in (5.10).

$$\eta_{Ci,j} = \begin{cases} \dfrac{\min\left(\eta_{Si}^{(DL)}, \eta_{Sj}^{(DL)}\right)}{\eta_{Si}^{(DL)}} & \text{if DL \&\& (i=BS, j=BS)} \\ \dfrac{\min\left(\eta_{Ti}, \eta_{Tj}\right)}{\eta_{Ti}} & \text{if DL \&\& (i=RS, j=RS)} \\ \dfrac{\max\left(\eta_{Si}^{(DL)} - \eta_{Sj}^{(DL)}, 0\right)}{\eta_{Si}^{(DL)}} & \text{if DL \&\& (i=BS, j=RS)} \\ \dfrac{\max\left(\eta_{Ti} - \eta_{Tj}, 0\right)}{\eta_{Ti}} & \text{if DL \&\& (i=RS, j=BS)} \\ \dfrac{\min\left(\eta_{Si}^{(UL)}, \eta_{Sj}^{(UL)}\right)}{\eta_{Si}^{(UL)}} \cdot \dfrac{\min\left(n_{Ki}, n_{Kj}\right)}{n_{Ki}} & \text{if UL \&\& (i=BS, j=BS)} \\ \dfrac{\min\left(\eta_{Si}^{(UL)}, \eta_{Sj}^{(UL)}\right)}{\eta_{Si}^{(UL)}} \cdot \dfrac{\max\left(n_{Ki} - n_{Kj}, 0\right)}{n_{Ki}} & \text{if UL \&\& (i=BS, j=RS) || (i=RS, j=BS)} \end{cases} \quad \textbf{(5.10)}$$

where n_{Ki} indicates the maximum number of subchannels allowed in the UL upper zones (vertical fragmentation in transparent UL relay mode).

Once that the inputs has been obtained, the optimization routine can be defined as a mixed integer programming problem, where the objective is to find the optimal solution that minimizes the given cost function representing the overall network interference:

$$\min \sum_{z \in Cs} \sum_{i \in S_z} \sum_{z' \in Cs} \sum_{j \in S_{z'}} \sum_{k \in Sh} \frac{w_{i,j} \cdot \eta_{Ci,j}}{D_i \cdot D_j} \cdot y_{i,j,k} \quad \textbf{(5.11)}$$

subject to:

$$\sum_{k \in Sh} x_{i,k} = D_i \qquad \forall i,k \quad \textbf{(5.12)}$$

$$x_{i,k} + x_{j,k} - 1 \leq y_{i,j,k} \qquad \forall i,j,k \quad \textbf{(5.13)}$$

$$y_{i,j,k} \geq 0 \qquad \forall i,j,k \quad \textbf{(5.14)}$$

$$x_{i,k} \in \{0,1\} \qquad \forall i,k \quad \textbf{(5.15)}$$

The constrains are set similarly to [72], where $x_{i,j}$ is a binary variable that indicates whether sector i uses subchannel k or not. Constraint (5.12) imposes that sector Si must use D_i subchannels. Inequality (5.13) and (5.14) force that in an optimal solution, $y_{i,j,k}$ = 1 if and only if both sectors i and j use subchannel k, and $y_{i,j,k}$ = 0 otherwise. Finally, the cost function is the sum of the interference between all pair of sectors i and j taking into account all the frequencies k.

The interference restrictions $w_{i,j}$ are divided in (5.11) by the number of used subchannels D_i and D_j in both interfered sectors i, j , and thus, the bigger the number of subchannels per sector, the smaller the chance of interference.

Sectors have different access zone lengths according to the traffic requirements. Also, some sectors may not have associated RSs and therefore MSs attached to this sector use the whole TDD frame and need less subcarriers, D_i. In general, the matrix w_{ij} represents interference events, but as shown in (5.11) it is weighted accordingly with D_i, D_j, which takes into account the frame zone length. The variable $y_{i,j,k}$ will indicate in which subchannels in both sectors interfere each other.

As mentioned, a special case is the transparent RS, which do not use relay zone in DL. In this case, frequency planning is mainly useful in managing intercell interference between RSs (which may be the most frequent situation with the addition of many of them), and between BSs, while the interference management between RS and other BSs with different access zone lengths - or no RSs attached - is less efficient. However, due to the fact that this mode does not provide coverage extension, this point is not as important as in the non-transparent mode. In addition, note that another RRM constraint in UL is the number of available subchannels to be allocated to BS and RS access zones.

5.2.3 Outer Optimization Loop: Station Configuration

In the outer loop, the optimization algorithm tests different configurations for different parameters that have higher reconfiguration costs in a real network evaluation – station location and number, tilt or azimuth - and that need a full simulation of RRM procedures. As previously mentioned, such parameters need to be discrete, and the resolution chosen will determine the size of the solution space. This optimization is derived from the process shown in Chapter 3 in (3.14)-(3.19).

A new factor may be considered in the multiobjective cost function. The RS reliability $F_{pen(Cr)}$ penalizes different degrees of network resiliency. A failure in the multihop communication is more likely to happen for weak relay links connections.

Thus, CINR levels in these links can provide information to the service provider about RS reliability. Note that the performance in this criterion is penalized over the BS-RS link, and therefore over the subset of RSs. A binary variable $d_{m,h'''}$ indicates if user m is evaluated by the cost function related to criterion h''', which may represent a subset of criteria related to RS resiliency ($Cr = \{Cr_1, Cr_{h'''} ..., Cr_{H'''}\}$, $Cr \subset G$), for example when the CINR values of the relay link are penalized differently in various channel conditions. With the addition of multihop communications, the problem becomes a four objective minimization process:

$$\begin{aligned}\min\Bigg\{ &\sum_{z\in Cs}\left(f_{\cos ts}(z)\cdot s_z + \sum_{i\in S_z} f_{\cos ts}(i)\cdot s_i \right) + \dots \\ &+ \sum_{z\in Cs}\left(\sum_{i\in S_z}\sum_{m\in Ms_i}\sum_{h'\in Cu} f_{pen(Cu_{h'})}(CINR(dB)_{m,h'})\cdot d_{m,h'} \right) + \dots \\ &+ \sum_{z\in Cs}\left(\sum_{i\in S_z}\sum_{m\in Ms_i}\sum_{h''\in Tg} f_{pen(Tg_{h''})}(T(kbps)_{m,h''})\cdot d_{m,h''} \right) + \dots \\ &+ \sum_{z\in Cs}\left(\sum_{i\in S_z}\sum_{w\in Rs_i}\sum_{h'''\in Cr} f_{pen(Cr_{h'''})}(CINR(dB)_{w,h'''})\cdot d_{w,h'''} \right)\Bigg\}\end{aligned} \tag{5.16}$$

subject to:

$$\sum_{z\in Cs} s_z \le \sum_{b\in Rb} n_b \qquad n_b = 1 \tag{5.17}$$

$$N_z\cdot s_z - \sum_{i\in S_z} s_i \ge 0 \qquad \forall z \tag{5.18}$$

$$1 \le \sum_{h'\in Cu} d_{m,h'} \le 2 \qquad \forall m \tag{5.19}$$

$$1 \le \sum_{h''\in Tg} d_{m,h''} \le 2\cdot Ser_{\max} \qquad \forall m, Ser_{\max} \tag{5.20}$$

$$\sum_{h'''\in Cr} d_{w,h'''} \le 2 \qquad \forall w \tag{5.21}$$

$$s_z, s_i, d_{m,h'}, d_{m,h''}, d_{w,h'''} \in \{0,1\} \qquad \forall m, w, z, i, h', h'', h''' \tag{5.22}$$

As described in Chapter 3, f_{costs} represents the aggregate of CapEx and OpEx in a radio access network. Constraint (5.17) indicates that only one CS can be installed per region, and (5.18) forces each CS to have at least one BS or RS. $f_{pen(Cuh')}$ represents the wireless connection by penalizing the effective CINR perceived by the MSs over all resources used. The effective CINR of the user is computed by using a MIC metric. Constraint (5.19) indicates that a maximum of two penalty functions are applied per user, since each user is in UL or DL (or both), and in certain channel conditions at each snapshot in a static simulation. $f_{pen(Tgh'')}$ is the penalty function for the MSs throughput, and constraint (5.20) indicates that the MS uses at least one service DL or UL, and a maximum of services, Ser_{max} in both UL and DL which can be penalized in different ways.

The new factor, $f_{pen(Crh''')}$, penalizes different degrees of network resiliency. In the event of having a failure in a RS, other RSs may cover the new empty area by extending the coverage of an alternative station. In addition, the RSs previously attached to the faulty station need to reconnect to a new one. These procedures incur in reconfiguration costs that are considered in this objective, and an example is shown in Table 5.1.

Table 5.1. Example of Economic Parameters for RS Reconfiguration.

Year		1	2	3	4	5
RS OpEx (k€)	Reconfiguration	0.25	0.22	0.19	0.17	0.15

During the network planning some failure events can be simulated for each configuration. A failure in the multihop communication is more likely to happen for weak relay links connections. Note in (5.16) that the performance in this criterion is penalized over the BS-RS link, and therefore over the subset of RSs. Similarly to the Cu criterion (in Table 3.2), fixed RS profiles can be penalized according to the LUTs provided, adjusting $T^{(min)}$ and $T^{(max)}$ to the CINR values to for the minimum and maximum MCS. Constraint (5.22) indicates that different penalties can be applied in both UL and DL.

5.3 Considerations on Planning and Optimisation of OFDMA Relay-based networks

The network designer may choose to perform a frequency plan only DL. The DL subchannel solution can be assumed to be also appropriate for UL. The reasons behind this are:

- As mentioned, the process has to be repeated twice, DL and UL, in order to provide subchannel plans in both cases. This may not be efficient for network designs with high computational load, which doubles the process time.
- In the UL Ie and W calculation, the worst case scenario is adopted, where a neighbouring MS close to the network edge is selected as the UL interferer. Some average position of users within the cell could be assumed, although this can lead to a less effective interference characterisation with respect to DL.
- The objective of the frequency planning procedure described in this section is to offer an optimal assignment of subchannels for each configuration tested of other parameters (e.g. BS/Rs position), in order to achieve the best network solution.
- The primary use of current OFDMA networks is data services, which are asymmetrical, i.e., more demands on the downlink than uplink.

5.3.1 Effects of frequency planning on high-cost parameter optimisation

In the following, an example of network expansion scenario has been performed over a rectangular area of the city of Munich (3.3km × 2.3km). The used scenario is a non-regular network with a channel bandwidth of 10 MHz, composed of 9 fixed and active BSs and 60 candidate RSs. In this area, an MS density of 50 MS/km^2 and two hot spots adding 20 MS/km^2, account for around 510 MSs per snapshot of the Monte Carlo routine. The MSs of the main traffic map use one of the services described in Table 3.4, while each MS added in the hotspot traffic maps use high demanding services, i.e.

video-conferencing and web browsing services. The rest of parameters are similar to standard simulation values found in [68][11], and shown in Table 5.2.

Table 5.2. Simulation Parameters.

Parameter	Value	Parameter	Value
Candidates sites	20	CPE Antenna Gain	0dBi
Carrier Frequency	3.5GHz	CPE Antenna Pattern	Omni
Channel Bandwidth	10MHz	CPE Antenna Height	1.5m
BS/RS TX Power	43/30dBm	CPE Noise Figure	5dB
BS/RS Antenna Gain	18dBi	CPE Cable Loss	0dB
BS/RS Antenna Height	40/30m	Snapshots	100
BS/RS Noise Figure	4dB	Density wide area MSs	50 user/km^2
BS/RS Cable Loss	3dBm	Density of hotspot MSs	70 user/km^2
CPE Tx Power	23dBm	Path Loss Model	Ray Launching

Different configurations are searched in both the inner and outer optimization loop. The outer loop gets optimal values for antenna tilt and azimuth in intervals of 10 degrees, and RS position from a given set of candidate sites, with the optimal assignment of subchannels calculated in the inner loop. Figure 5.4 shows the pilot power and user status of a single snapshot in the solution found.

Hotspots and indoor coverage in the central area benefit from this configuration. However, there are a number of MSs with low CINR levels concentrated at the edge of the coverage areas, where they may have a poor channel quality. This solution is an example of throughput enhancement and network reliability, since RSs are well covered. Other strategies with different degrees of fairness for users or risky configurations for the operators can be set by tuning the values ($f^{(max)}$, $T^{(min)}$, $T^{(max)}$) in the penalty function.

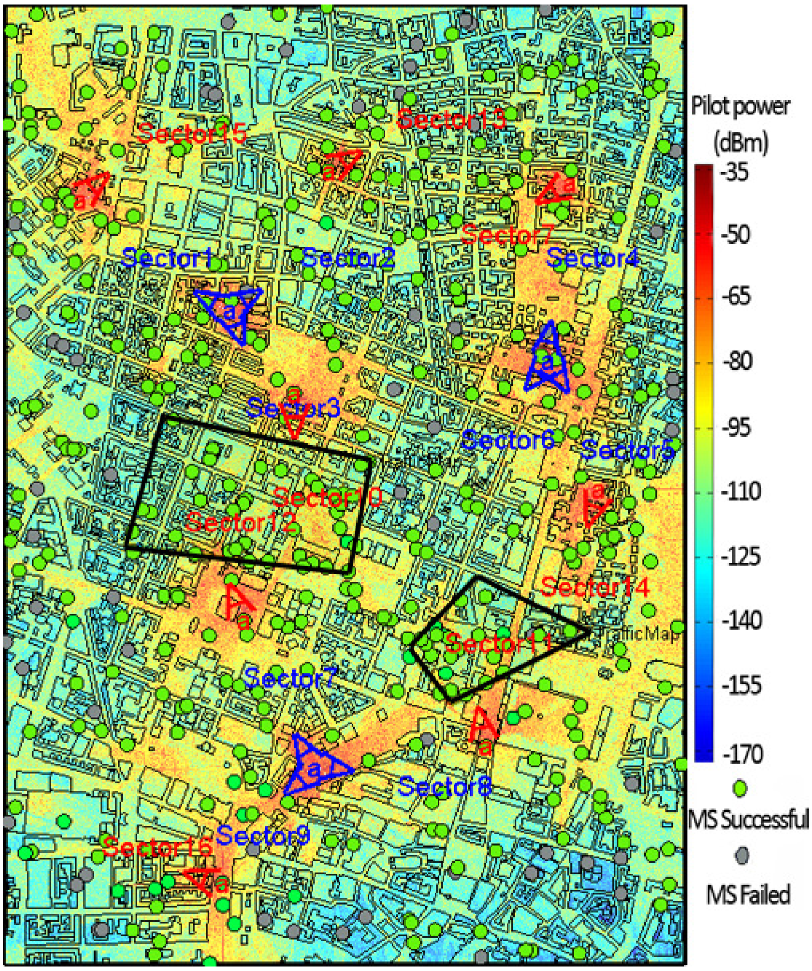

Figure 5.4. Optimal solution.

One snapshot of a solution obtained without the use of the inner optimization loop is shown in Figure 5.5. Instead, the frequency planning process has been performed after the optimization of tilt/azimuth and RS position. In this case, the throughput per user decreases specially in the MSs that are in the hotspots. The reason is that the tilt/azimuth and RS position set were calculated for a frequency reuse 1, and the subsequent automatic frequency planning has more difficulties to efficiently allocate more subchannels to sectors with large traffic demands due to the lower CINR levels with such antenna configuration.

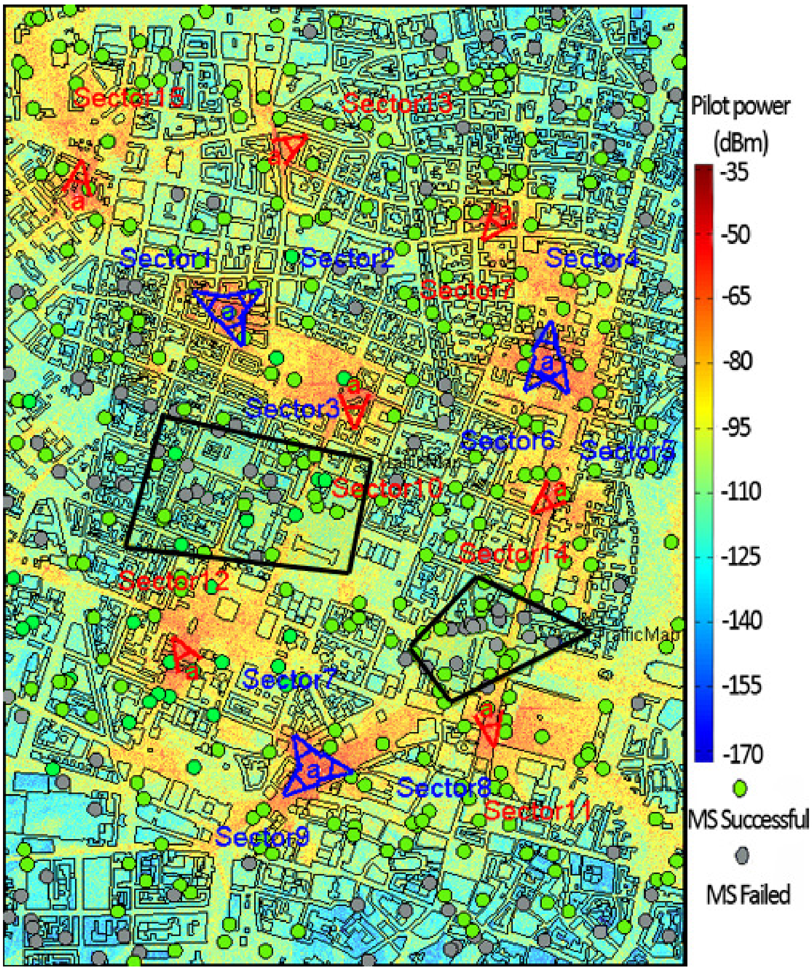

Figure 5.5. Solution without Inner optimization loop.

The tilt/azimuth and RS position values have been set to avoid interference between sectors as much as possible. For example, compared to the optimal solution in Figure 5.4, it can be observed that sectors avoid pointing at each other. Also, in some cases, the RS positions are not set optimally for the relay link. As a result, the network is less resilient. Since the final throughput perceived by many MSs is the result of the communication of two links, the total throughput in this system is lower. The system throughput in this case is 853 Mbps while in the optimal solution is 1125 Mbps.

Figure 5.6 shows the distribution of users at different CINR levels. We can observe that the number of users that cannot get the minimum MCS for communication (CINR < 0) is sensibly higher in the solution without the inner loop. However, the main problem in this solution is the high number of users that have low CINR levels that only allow low MCSs. This is due to the inefficient interference management. These users

(CINR>0) will need large bursts to fulfil their service requirements and the OFDMA frame will be filled faster, thus leaving other users with the same situation in an "unsuccessful" state. This effect is critical in the two hotspots were the traffic is higher.

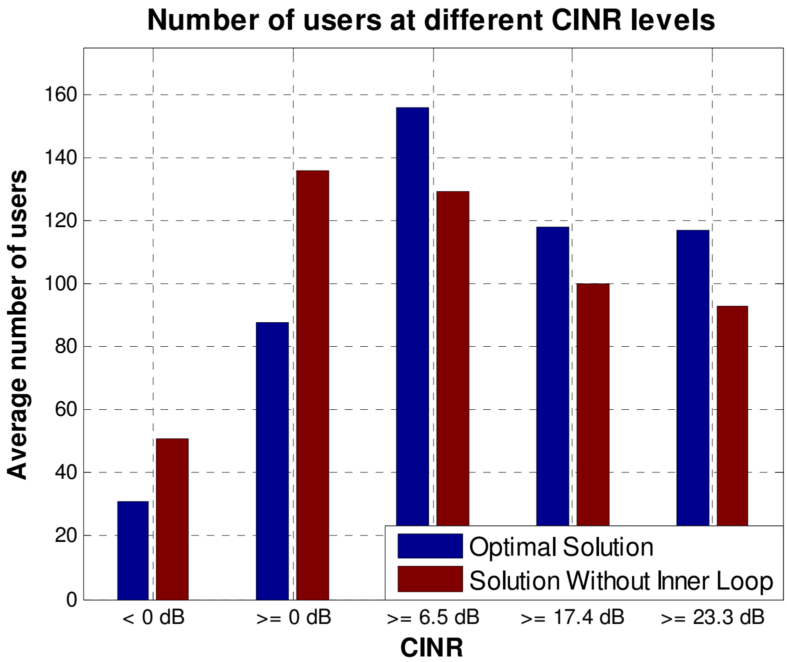

Figure 5.6. Number of users at different CINR levels.

5.3.2 Multicriteria Wireless Network Architecture Design in Multihop Environments

In the following, an example of multicriteria network design process has been performed over a rectangular area of the city of Munich in a relay network expansion scenario to contain 60 candidate RSs, and a previous BS infrastructure, which is kept fixed.

Different configurations are searched by a Simulated Annealing algorithm in the inner optimisation loop. As discussed in Chapter 3, the extended version of algorithm may need an accurate fine tuning to get good performance for different scenarios. In this case an standard implementation of SA is used due to the multiplie possibilities of customisation during the mentioned tuning process. In the outer loop, the described

Multiobjective Tabu Search procedure gets optimal values for RS position and number, antenna tilt and azimuth by using the method of Pareto front.

The rest of simulation parameters are set similarly to Table 5.2 and Chapter 3. Table 5.3 shows a final set of five representative solutions selected for the four-objective scenario.

Table 5.3. Cost and penalty values of 5 representative solutions of the Pareto Front (×100000€).

Solution	*Number of installed RSs*	F_{costs}	$F_{pen(Cu)}$	$F_{pen(Tg)}$	$F_{pen(Tg)}$	*CostAdj*
1	6	4.47	2.53	1.21	2.02	10.23
2	8	5.08	1.25	2.63	2.68	11.64
3	6	3.89	1.7	3.19	3.28	12.06
4	8	5.03	2.66	1.46	1.75	10.90
5	5	3.85	2.93	3.14	2.64	12.56

Figure 5.7 shows one snapshot of solution 1 and 2. Since different MSs may have different service requirements, the picture illustrates which MSs are successful.

In solution 1, the lower number of RSs results in low infrastructure costs and also provide low throughput penalties. Hotspots and indoor coverage in the central area benefit from this configuration. On the other hand, there are a number of MSs with low CINR levels concentrated at the edge of the coverage areas, where they may have a poor channel and also high interference. This solution can be a suitable option for service providers that plan a future expansion of the network according to the real user demands.

Note that the penalization is performed in different ranges and values, and depends on the network provider criteria as some objectives can be given more or less preference, and due to this fact diverse situations can be created.

In solution 2 some new users can gain connection to the network because of the addition of more RS, and therefore some penalty values related to connection tariff or annual subscription are reduced. On the other hand, the throughput per user decreases specially in the MSs that are in the two hotspots. Users that require services that are bandwidth consuming are more penalized and thus, this option would be suitable for network providers that prefer to ensure connectivity to as many users as possible in a wide area than providing services with high bit rates.

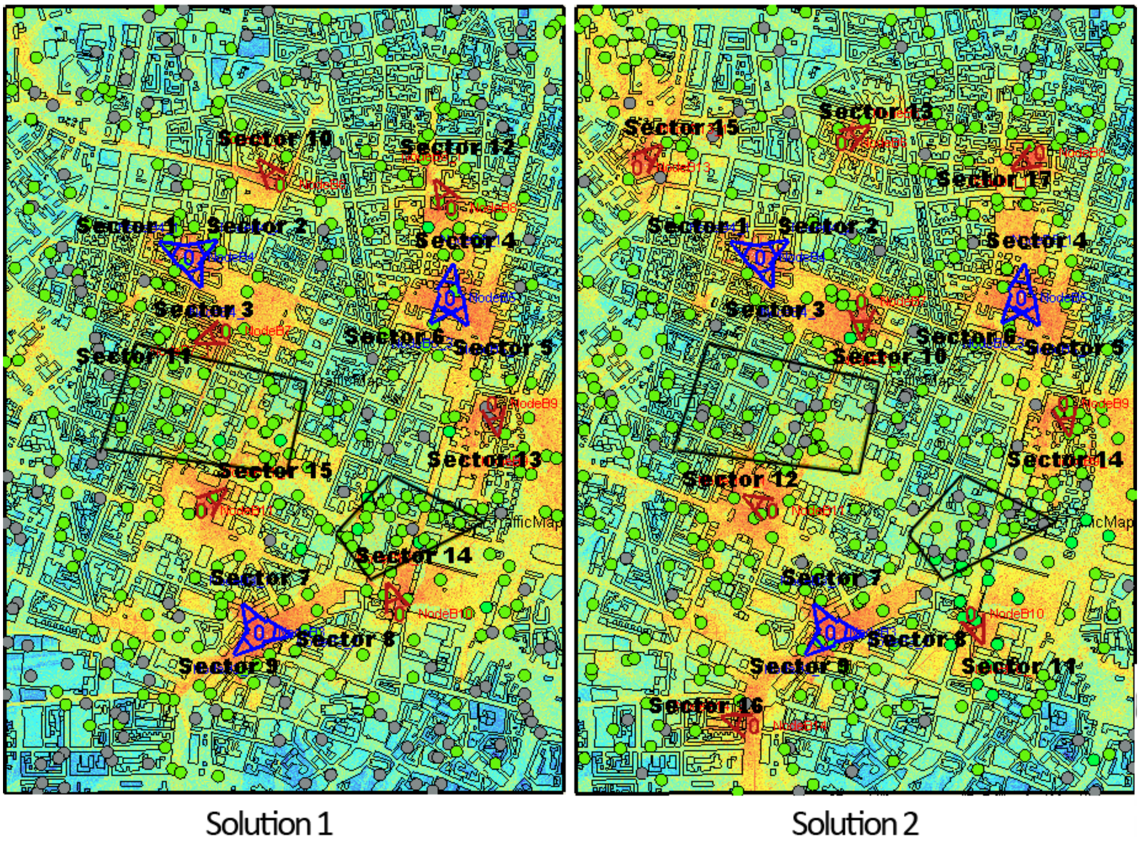

Figure 5.7. Two solutions from the Pareto front.

Solution 1 represents a throughput enhancement strategy, and also a reliable network, while solution 2 characterizes a coverage extension strategy. Other solutions in the Pareto front represent situations that incur in different costs, which should be evaluated by the network designer.

5.3.3 Frequency Planning in IEEE 802.16j Networks

The multiobjective optimisation example in Section 5.3.2 has been carried out using non-transparent relays and SA as a single-objective procedure for the internal optimisation loop. The selection of the relay mode and the tuning of the optimisation algorithm corresponded to a standard mode recommended in the standard [42], and an empirical calibration. However, as shown in Chapter 3 and Chapter 4, the appropriate selection of these parameters may impact the final network performance.

This section evaluates different options over the two solutions described in Section 5.3.2, and analyse their influence on the network design process and final performance.

- Introduction to SA for Optimal Subchannel Allocation

The optimization algorithm is a key process the evaluation of an optimization framework. As mentioned, the time that the algorithm is running and the algorithm configuration will determine the performance of the whole optimization framework.

In SA, every state promoted for the next step is a consequence of a probabilistic rule, which provides with a better state of solution or creates a perturbation to prevent the possible local minima [56]. The probabilistic process is defined by an acceptance probability function $P(f(s),f(sn),T_e)$. This function depends on the cost function of the current and the neighbour solutions f(s), f(sn) and on a factor called temperature T_e. The temperature is a global parameter that decreases with time in an exponential fashion dictated by the annealing factor. The initial temperature and the annealing factor are the parameters tuned by the operator. SA is detailed in the following pseudo-code:

```
begin
  s = s0; fs = f(s)             // Random Solution
  sb = s; fsb = fs              // Best Solution
  Te = 0; u = 0                 // Initialize Temperature and Iterations
  while u < umax do
    u = u + 1                   // Increase iterations
    sn = neighbour(s)           // Select a neighbour
    fn = f(sn)                  // Compute its cost function
    if fn< fsb then             // Is this a new best so far?
        sb = sn; fsb = fn       // Yes, save it
        continue                // Go for another iteration
    if P(fs, fn, Te) then       // Should we move to it?
        sb = sn; fsb = fn       // Yes, move to it
  end
  Te* = annealing_factor        // Update Temperature
end
```

- Coverage Extension Scenario

A coverage extension scenario is characterized by RSs placed close to the edge of the coverage area. The objective of this strategy is mainly that the new introduced stations provide service to previously uncovered users. The scenario corresponds to Solution 2 in Section 5.3.2.

The restriction matrix approximates the inter-cell interference between sectors of the network in terms of percentage of interfering users. Table 5.4 shows the demand vector and restriction matrix for one snapshot in the coverage extension scenario with a traffic distribution of 50 users/km^2 plus hotspots. This matrix has been obtained for non-transparent relay configuration, which is suited for this scenario.

Table 5.4. Demand vectors and restriction matrix for non-transparent relays in a coverage extension scenario.

i	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	D_i
1	0	0	37.5	0	0	0	6.2	0	0	6.2	0	25	0	0	25	0	0	**21**
2	40	0	40	10	0	10	0	0	0	0	0	0	40	0	30	0	60	**8**
3	58.3	0	0	0	0	0	0	0	0	60	0	38.6	0	0	0	0	0	**11**
4	0	26.6	6.6	0	6.6	20	0	0	0	0	0	0	40	0	0	0	60	**9**
5	0	0	0	20	0	80	0	0	0	20	0	0	0	53	0	0	0	**7**
6	0	0	0	15.3	69.2	0	7.6	7.6	0	64.5	0	0	0	28.7	0	0	23.0	**10**
7	0	0	17.6	0	5.8	17.6	0	41.1	35.2	29.4	0	41.1	0	5.8	0	23.5	0	**25**
8	0	0	14.2	0	28.5	28.5	57.1	0	23.8	28.5	38.0	0	0	38.8	0	0	0	**29**
9	0	0	0	0	0	0	14.2	57.1	0	0	21.4	0	0	0	0	21.4	0	**12**
10	0	0	81.2	0	10.5	46.1	42.3	10.5	0	0	0	47.2	0	10.5	0	0	0	**24**
11	0	0	0	0	14.8	0	0	74.0	11.1	0	0	0	0	23.9	0	0	0	**35**
12	13.9	0	49.1	0	0	0	55.8	0	25.5	38.8	0	0	0	0	0	30.2	0	**46**
13	11.1	66.6	0	38.8	0	0	0	0	0	0	0	0	0	0	33.3	0	77.7	**13**
14	0	0	17.6	0	68.4	57.7	29.4	42.1	0	32.3	12.7	0	0	0	0	0	5.8	**43**
15	77.4	25.8	0	0	0	0	0	0	0	0	0	0	16.1	0	0	0	9.6	**40**
16	0	0	0	0	0	0	40.9	0	86.3	0	0	27.2	0	0	0	0	0	**18**
17	0	33.3	33.3	83.3	0	33.3	0	0	0	0	0	0	66.6	0	0	0	0	**7**

The weighted restriction matrix ($w_{ij} \times \eta_{Cij}$) characterizes the interference relationship between sectors, for example, the restriction is null when sectors are far away from each other. On the other hand, the interference is large between adjacent RS-BS, such as RS 10 and BS 3 or RS 15 and BS 1, since they reuse resources in the non-transparent mode. Also, note that some relay sectors such as RS 15, RS 14 and RS 12, have the largest demand vector as they cover large areas or the hotspots.

We can also observe that the influence between two sectors is not reciprocal in most cases. The reason is that the number of MSs attached to several sectors is different. For example BS 2 is affected 40% by BS 1, but not the opposite, since BS 1 demands are higher (D_1 = 21) than those for S_2 (D_2 = 8).

The tuning of SA must be customized for different problems and it is normally done empirically. In SA, the parameter that has more influence in the final performance is the annealing factor [56]. In order to fine-tune such parameter, the variation of the cost function with the number of iterations has been analyzed. The results for 50 users/km^2 plus hotspots and non-transparent relay configuration can be seen in Figure 5.8.

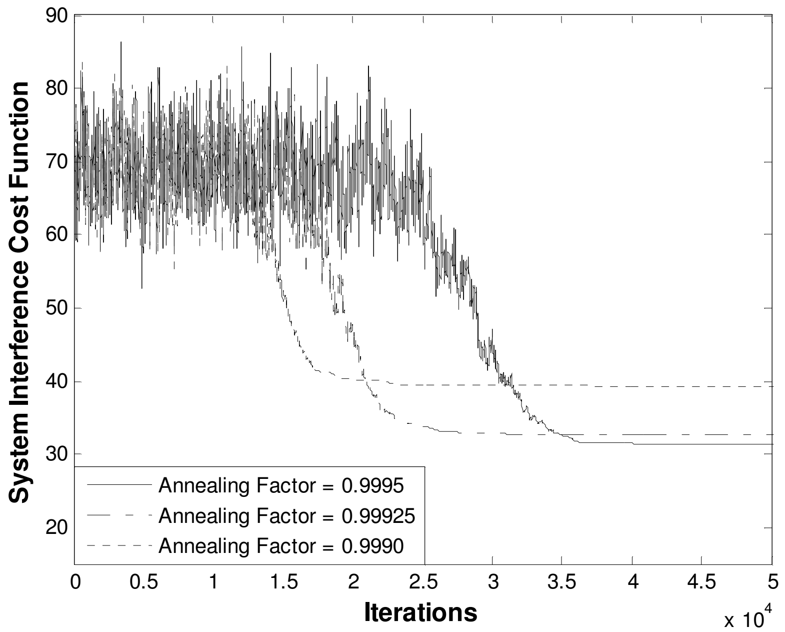

Figure 5.8. System interference cost function for non-transparent relays in a coverage extension scenario.

It can be observed from Figure 5.8 that when the annealing factor is smaller, the search obtains a worse optimization result but converges quicker, and therefore, there is a trade-off between accuracy and computing time. For this scenario Te=10^6 and the annealing factor is set to 0.99925, which shows good performance.

Figure 5.9 shows how the total throughput of the system changes when the user density varies (main traffic map). The two relay configurations and the use of frequency planning are studied over the coverage extension scenario.

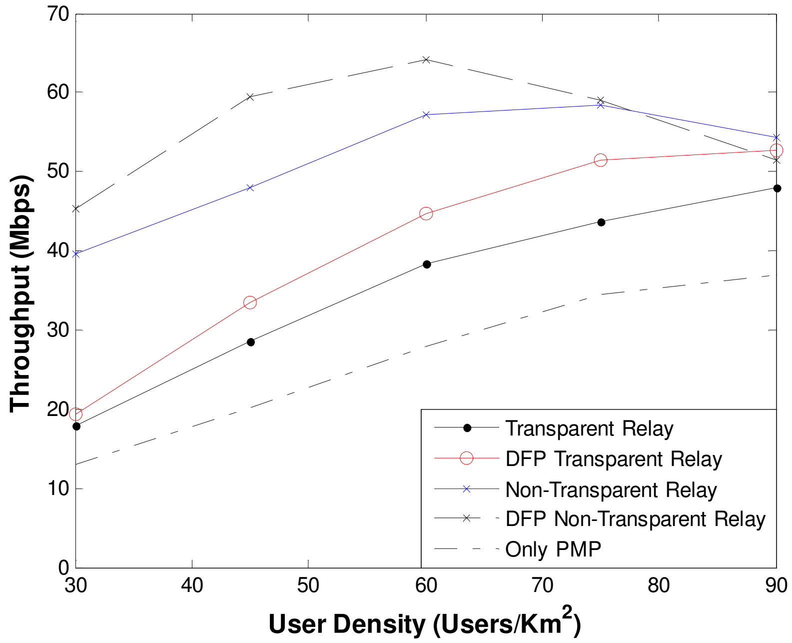

Figure 5.9. System throughput vs. user density in a coverage extension scenario.

In general, it can be observed that the non-transparent relay mode provides larger network throughput in this scenario than the transparent relay mode. The reason is that the former allows coverage extension, and some users that were out of the range of the BS can now gain access to the network, thus increasing the total throughput.

In a scenario with a low user density (30 users/km^2), the improvement in throughput provided by frequency planning is low, especially in the transparent mode, since there are lot of resources for few users. In this case, the interference levels were already low in the non-planned option, and the OFDMA features deal with interference.

As the user density increases to 45 and 60 users/km^2, the improvement in throughput provided by frequency planning is more significant. Some sectors with large traffic demands close to the hotspots are allocated more subchannels than neighbouring sectors with low traffic demands. It can be observed that this improvement is more important in the non-transparent relay as this mode reuses more traffic.

With a high user density (75-90 users/km^2), the number of subchannels demanded D_i is larger, and therefore, the frequency planning procedure has less possible combinations to offer interference mitigation. As a result, the probability of slot collision is larger, and the performance of the frequency planning is degraded. This is especially critical in the non-transparent mode, where the higher reuse of resources (including those for cooperative communications) results in high sector demands and intercell interference levels. In addition, note that the final throughput perceived by MSs attached to RSs depend on two links, BS-RS and RS-MSs, which contribute to the probability of interference.

As the user density increases in the traffic map, more MSs are attached to transparent RSs, but still less than considering non-transparent RSs that can extend coverage. All the MSs in transparent RSs can benefit in some degree by the cooperative communications, and the frequency planning procedure shows better performance than the non-transparent mode.

- Throughput Enhancement Scenario

A throughput enhancement scenario is characterized by RSs placed in key areas where hotspots and indoor coverage may benefit from the higher CINR values (Solution one in Section 5.3.2). The objective of this option is mainly to provide high bit rates rather than ensure connectivity in a wide area.

Table 5.5 shows the demand vector and weighted restriction matrix ($w_{ij} \times \eta_{Cij}$) for one snapshot in the throughput enhancement scenario for transparent relay configuration, which is suited for this scenario. The snapshot corresponds to a traffic distribution of 50 users/km^2 plus hotspots. It can be observed that the interference is higher in the inter-RS and inter-BS cases. Other cases are the interference between RS 15 and BSs 4, 5, or RS 10 and BS 1 due to the sectors' different access zone lengths related to the different traffic requirements; and also RS 15 and BS 6, which does not have associated RSs. The table shows that the number of subchannels required by each sector is high, since the OFDMA frame utilisation is lower.

Table 5.5. Demand vectors and restriction matrix for transparent relays in a throughput enhancement scenario.

i	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	D_i
1	0	32.7	24.5	0	0	0	3.2	0	0	8.8	0	0	0	0	0	**37**
2	80	0	5	20	0	0	0	0	0	0	0	0	0	0	0	**11**
3	62.9	7.4	0	0	0	0	11.1	0	0	0	0	0	0	2.6	1.3	**37**
4	0	5	0	0	22.5	12.5	0	0	0	0	0	0	0	0	0	**19**
5	0	0	0	20	0	75	5	35	0	0	0	0	0	2.1	2.4	**27**
6	0	20	47.5	15	77.5	0	32.5	30	0	0	0	0	0	28.4	29.1	**32**
7	0	0	19.2	0	3.8	23	0	46.1	19.2	0	0	0	0	0	0	**40**
8	0	0	17.3	0	23.9	28.2	54.3	0	17.3	0	0	0	0	0	0	**39**
9	0	0	0	0	0	0	40	32.7	0	0	0	0	0	0	0	**38**
10	12.5	0	0	0	0	0	0	0	0	0	0	54.5	0	0	0	**4**
11	0	0	0	0	0	0	0	0	0	15.3	0	0	0	0	50.5	**42**
12	0	0	0	0	0	0	0	0	0	25.9	0	0	0	0	0	**12**
13	0	0	0	0	0	0	0	0	0	0	5.5	0	0	11	0	**19**
14	0	0	4.3	0	5.1	28.1	0	0	0	0	0	0	12	0	0	**19**
15	0	0	3.2	0	6.4	35.3	0	0	0	0	41.5	0	0	0	0	**31**

The tuning annealing factor for 50 users/km^2 and hotspots, and transparent relay configuration can be seen in Figure 5.10.

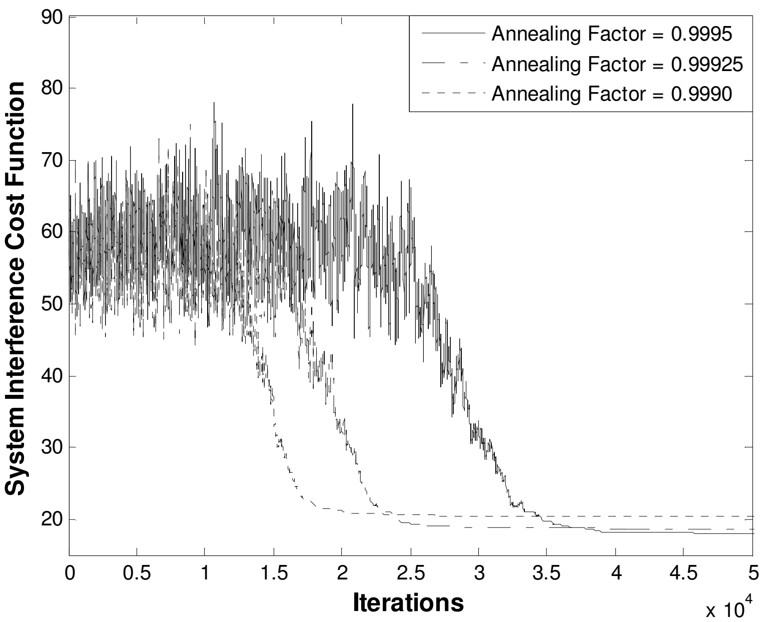

Figure 5.10. System interference cost function for transparent relays in a throughput enhancement scenario.

It can be seen that the obtained value of the cost function, which represents the system interference levels, is lower than in the non-transparent case. In addition, the evaluation of the previously selected values for the annealing factor show less impact on the final results than in the previous case.

Figure 5.11 shows how the total throughput of the system changes when the user density varies. The two relay configurations and the use of frequency planning are studied over the throughput enhancement scenario.

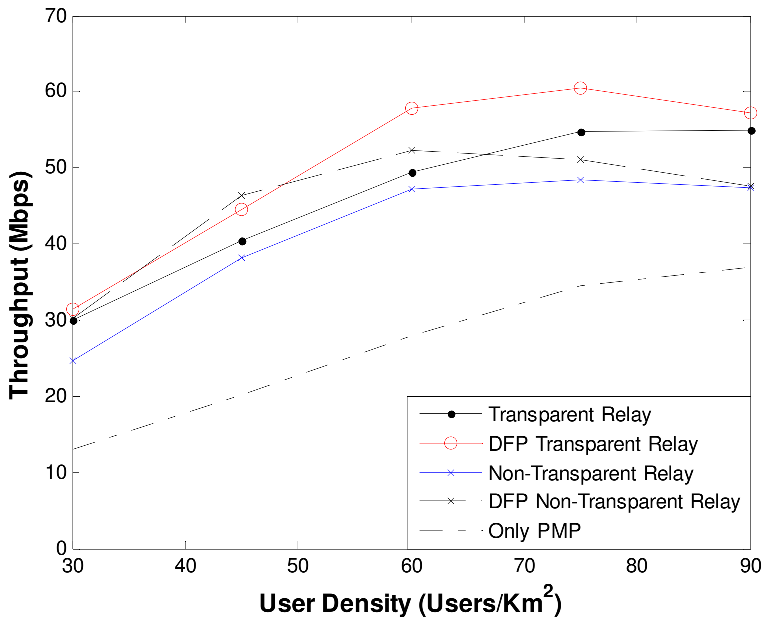

Figure 5.11. System throughput vs. user density in a throughput enhancement scenario.

In this case the transparent relay mode provides larger network than the non-transparent relay mode. The reason behind this is that RS are configured to point to the hotspots and the inner areas of the cells. Some sectors' main lobes are pointing to each other, and the separation of resources of the transparent mode allows a more efficient interference avoidance.

In a scenario with a low user density (30 users/km^2), the improvement in throughput provided by frequency planning is low, since there are lot of resources for few users. As the user density increases to 45 and 60 users/km^2, the improvement in throughput provided by frequency planning is more significant. Hotspot sectors with large traffic demands are allocated more subchannels than neighbouring sectors with low traffic demands. With a high user density (75-90 users/km^2), the number of subchannels demanded D_i is larger, and therefore, the frequency planning procedure has less possible combinations to offer interference mitigation. As a result, the probability of slot collision is larger, and the performance of the frequency planning is degraded. In this kind of scenario where RSs are closer to BSs, more MSs can benefit by cooperative communications, and the frequency planning procedure shows better performance than the transparent mode.

- Comparative Analysis of Transparent and Non-transparent Mode

Since both methods are designed for different purposes, a comparative analysis of their performance for the percentage of users that are able to read the pilot and preamble, is shown in Figure 5.12. Such number is higher in the non-transparent mode since it provides coverage extension.

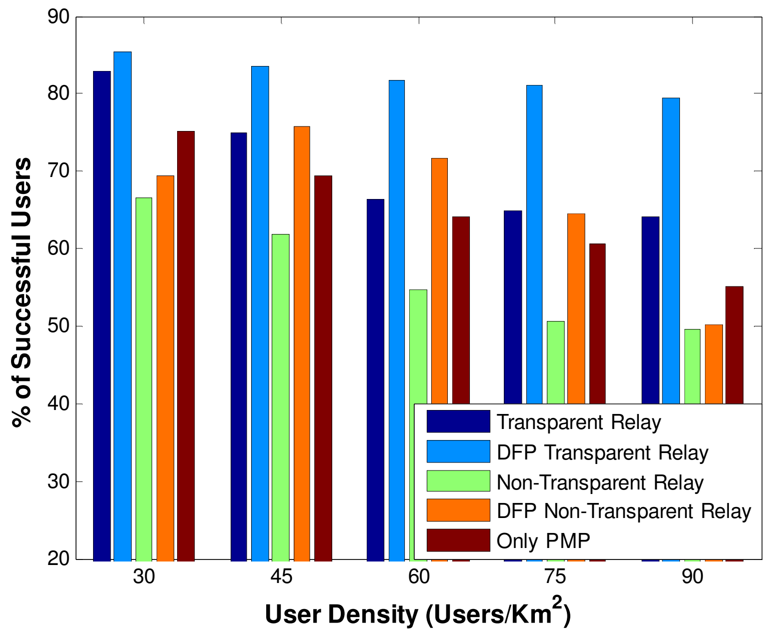

Figure 5.12. Percentage of successful users vs. user density.

It can be observed that the percentage of successful users from the initial target - from those users that are able to decode the preamble - is larger in the transparent relay. It is due to the fact that this mode has less traffic interfering within the same cell (between BS and RS). Also the dynamic allocation of resources is more flexible.

In Figure 5.12, it can also be observed how the improvement of frequency planning in transparent relay is quite stable with the traffic increase, while in the non-transparent mode, it is more variable because resources have to be reused more tightly to allow coverage extension.

5.4 Conclusions

This chapter explains the impact of relays on OFDMA network architecture and presents a service-oriented optimization framework for the design of this kind of networks. Relay networks involve a more complicated air interface than other kind of cellular networks, and need to be analyzed during the process of network design to achieve an optimal performance. Frequency planning and base station configuration are integrated in a framework that provides a wide perspective to the network designer.

This chapter proposes the joint use of time domain forwarding, cooperative communications and frequency planning with the optimisation of the configuration of parameters with high reconfigurations costs. The procedure offers an optimal assignment of subchannels for each configuration tested of other parameters (e.g. BS/RS position or antenna tilt/azimuth) in order to achieve a best network solution.

The results of the relay scenario analyzed in this article highlight the complexity of the design process, where the operator needs to estimate a number of parameters. The overall process completion, however, can be iterative. After a system analysis and calibration - as shown in Section 5.3.3 - a detailed network planning and optimisation procedure can be repeated with the tuned parameters until a satisfactory solution is found.

In general, the frequency planning strategy presented increases the throughput of the system. The final performance of both transparent and non-transparent modes may vary with different scenarios, since they are designed for different purposes. The improvement of frequency planning in the non-transparent mode is more variable with changes in the amount of traffic in the network, since resources have to be reused more tightly than in the transparent mode.

As a future work, a joint and efficient DL and UL subchannel planning procedure may increase the accuracy of the model, and improve the computational load for large solution spaces. In addition, studies on the estimated position of interfering users within the cell can help to improve the UL interference characterisation.

Chapter 6: Concluding Remarks

6.1 Conclusions

This work has shed new light on the study of the planning and optimisation of wireless metropolitan area networks, in particular to the access network design of OFDMA-based systems. A practical view for the solution of this problem is presented by means of the development of a novel design framework and the use of multicriteria optimisation. A further consideration of relaying and cooperative communications in the context of the design of this kind of networks has been done.

The thesis is structured in three main blocks covering important gaps in the existing literature in planning (efficient simulation) and optimisation. It first focuses on the study of models and algorithms for classical point-to-multipoint networks, where the optimisation framework is presented. Then, this work analyses the impact of the relay configuration on the network planning process, followed by the analysis of the optimisation procedure with the addition of relay stations and the derived higher complexity of the process.

With the increasing need of supporting new services and applications for different scenarios or multiple objectives, the problem of wireless network architecture design becomes too large in scope to be handled efficiently with a single technique. This thesis presents service-oriented optimization framework that provides a clear and comprehensive description of different options and solutions to achieve an optimal network configuration. The design the problem has been simplified and translated it into an optimization routine with consideration of technology-specific and economic factors. The results of the scenarios analyzed in this article highlight the complexity of the design process, as the operator needs to estimate a number of parameters, but also shows the benefits of the described optimization framework.

A new simple and flexible scheme to integrate multihop communications in the mobile WiMAX frame structure has been proposed and evaluated in a comparison with another basic concept already proposed for the upcoming IEEE 802.16j release of the standard. In addition, efficient capacity calculations that allow intensive system level simulations in a multihop environment have been introduced.

A subchannel restriction of the frequency plan procedure proposed combines the use of the necessary frame fragmentation of in-band relay communications and cooperative procedures. The different size of access zones is considered.

To sum up, this work deals with the complexity of the wireless network design problem and proposes appropriate ways to find close-to-optimal solutions. The consideration of multihop networks involves an increased complexity, which is evaluated in this work.

6.2 Future work

There are some challenges in OFDMA network planning and optimization, in particular in multihop relaying networks that need further investigation. Some work on extending the formulation shown to fully avoid the use of system level simulations would greatly speed up the whole process. This could be achieved by creating a new restriction matrix that may consider intercell relations in terms of CINR and throughput.

This work is focused in the two-hop case since this may be the most common case of this kind of networks. However, the formulation can easily extended to the n-hop case, and the consideration of routing algorithms through the design of several hops together with other parameters may improve final design.

The selection of only one RS per BS corresponds to the focus on the detailed planning and optimisation stage, where parameters with high costs get an initial configuration, like the case of a network expansion need due to an increase of capacity demand or a coverage extension requirement. As a future work, more RSs may be considered per BS besides having more hops in the communication.

The joint consideration of different RRM strategies with new frame structure designs for multihop communications may show interesting results, and this can be evaluated in the framework presented as an extension of this work.

In addition, little work has been done on relay self-optimization procedures, which may be used in an on-line optimization stage once the network has been deployed. In this case, relay stations may need to change their power, frequency settings and other parameters in a dynamic manner, with the only information of their immediate neighbourhood. This may be of big importance in nomadic relay stations.

The data rates considered throughout this work are generally in the modest side (e.g. 12.2 kbps for VoIP and 64 kbps for video); the impact of how the results may look like if higher target data rates are used can be studied.

Other social issues may also be of interest for the network provider, such as health and environmental effects of optimisation. The network designer should find a way to calibrate this objective in the cost function and find the appropriate KPIs.

Appendix A: Publications Derived from this Work

BOOK CHAPTERS

- F. Gordejuela-Sanchez and J. Zhang, "Planning and optimization of multihop relaying networks," accepted in *Evolved Cellular Network Planning and Optimisation for UMTS and LTE*, Auerbach Publications, CRC Press, Taylor & Francis Group, 2010.

JOURNAL PAPERS

- F. Gordejuela-Sanchez, A. Juttner and J. Zhang, "A Multiobjective Optimization Framework for IEEE 802.16e Network Design and Performance Analysis," *IEEE Journal on Selected Areas in Communications (J-SAC) special issue on Broadband Access Networks: Architectures and Protocols*, vol. 27, no. 2, pp. 202-216, February 2009.

REFEREED INTERNATIONAL CONFERENCE PAPERS

- F. Gordejuela-Sanchez and J. Zhang, "LTE Access Network Planning and Optimization: A Service-Oriented and Technology-Specific Perspective," accepted in *IEEE Global Communications Conference (GLOBECOM'09)*, Hawaii, USA, December 2009.
- F. Gordejuela-Sánchez, D. López-Pérez and J. Zhang, "A Two-Step Method for the Optimization of Antenna Azimuth/Tilt and Frequency Planning in OFDMA

Multihop Networks," presented at *ACM International Wireless Communications and Mobile Computing Conference (IWCMC'09)*, Leipzig, Germany, June 2009.

- F. Gordejuela-Sanchez, D. López-Pérez and J. Zhang, "Frequency planning in IEEE 802.16j networks an optimisation framework and performance analysis," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2009 pp.1-6.
- F. Gordejuela-Sanchez and J. Zhang, "Practical design of IEEE 802.16e networks: A mathematical model and algorithms," in *IEEE Global Communications Conference (GLOBECOM)*, 2008, pp.1-5.
- F. Gordejuela-Sanchez, D. Lopez-Perez, L. Q. Zhao and J. Zhang, "Efficient mobile WiMAX capacity estimations in a multihop environment," in *IEEE International Conference on Communication Systems (ICCS)*, 2008, pp. 1640 -1646.

INTERNATIONAL MEETINGS AND ORAL PRESENTATIONS

- D. Lopez-Perez, F. Gordejuela-Sanchez, L. Q. Zhao, W. Huang and J. Zhang, "EMI: Experimental WiMAX simulator for network planning and optimization," presented at *4th MCM COST2100*, Wroclaw, Poland, February 2008.

References

[1] Forsk Atoll Technical Staff, *Technical reference guide: WiMAX BWA Documents.* Forsk Atoll, 2009.

[2] WiMAX Forum, “WiMAX Home Webpage,” June 2009, http://www.wimaxforum.org/home.

[3] F. Ohrtman. *WiMAX handbook:* building 802.16 wireless networks. New York: McGraw-Hill, 2005.

[4] R. Olexa. *Implementing 802.11, 802.16, and 802.20 wireless networks: planning, troubleshooting and operations*. Oxford: Elsevier, 2005.

[5] E. Amaldi, A. Capone, and F. Malucelli, “Planning UMTS Base Station Location: Optimization Models with Power Control and Algorithms,” *IEEE Transactions of Wireless Communications*, vol. 2, no. 5, September 2003.

[6] H. Holma and A. Toskala, *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications,* New York: John Wiley & Sons, 2007.

[7] C. Smith, John Meyer. *3G wireless with WiMAX and Wi-Fi : 802.16 and 802.11.* New York: McGraw-Hill, 2005.

[8] IEEE 802.16 Working Group 16: IEEE Standard for Metropolitan Area Networks, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Systems. IEEE STD 802.16e-2005, approved 280 Feb 2006.

[9] “3GPP; Technical Specification Group Radio Access Network; Requirements for E-UTRA and E-UTRAN”, 3GPP, March. 2008. [On-line]: http://www.3gpp.org/ftp/Specs/html-info/36-series.htm.

[10] H. Lui and G. Li, *OFDM-Based Broadband Wireless Networks Design and Optimization.* New Jersey: John Wiley & Sons, 2005.

[11] J. G. Andrews, A. Ghosh, R. Muhamed, *Fundamentals of WiMax: Understanding Broadband Wireless Networking*. Massachussets: Prentice Hall, 2007.

[12] R. Kwan, C. Leung, and J. Zhang, "Multiuser Scheduling on the Downlink of an LTE Cellular System," *Research Letters in Communications*, vol. 2008, pp.1-4, 2008.

[13] K. Sroka, "GSM-based ran optimisation using measurements data," presented at *COST2100 Training School on Wireless Network Optimisation*, February 2008.

[14] J. Borkowski and J. Lempiainen, "Practical Network-Based Techniques for Mobile Positioning in UMTS" in *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp 1–15, 2006.

[15] J. Yang, M. Aydin, J. Zhang, and C. Maple, "UMTS Base Station Location Planning: a Mathematical Model and Heuristic Optimisation Algorithms," *IET Communications*, vol. 1 (5), pp. 1007–1014, October 2007.

[16] S. Redana, M. Lott "Performance analysis of IEEE 802.16a in mesh operation mode," presented at *IST SUMMIT*, 2004.

[17] Tiraspolsky, S. Rubtsov, A. Maltsev, A. Davydov, A. "Mobile WiMAX - deployment scenarios performance analysis," in *Wireless Communication Systems (ISWCS)*, 2006, pp. 353-357.

[18] ATDI, "Mobile WiMAX from OFDM 256 to S-OFDMA," Kiev, Ukraine, 2007.

[19] T. Theodoros and D. V. Kostantinos, "WiMAX network planning and systems performance evaluation," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2007, pp. 1948-1953.

[20] "The Network Simulator - ns-2," March, 2008. [On-line]: http://www.isi.edu/nsnam/ns/index.html.

[21] "Opnet Technologies," March, 2008. [On-line]: http://www.opnet.com.

[22] F. Gordejuela-Sanchez and J. Zhang, "Practical design of IEEE 802.16e networks: a mathematical model and algorithms," in *IEEE Global Communications Conference (GLOBECOM)*, 2008, pp.1-5.

[23] A. Eisenblätter, H. F. Geerdes, T. Koch, A. Martin, and R. Wessly, "UMTS Radio Network Evaluation and Optimization Beyond Snapshots," *Mathematical Methods of Operations Research*, vol. 63, 2005, pp. 1–29.

[24] O. Alp, Z. Drezner, and E. Erkut, "An Efficient Genetic Algorithm for the p-median Problem," *Annals of Operations Research*, vol. 122, no. RR-5740, 2003, pp. 21–42.

[25] S. Hurley, "Planning Effective Cellular Mobile Radio Networks," *IEEE Transactions in Vehicular Technology*, vol.51, no.2, pp. 243-253, March 2002.

[26] I. Siomina, P. Verbrand, and D. Yuan, "Automated Optimization of Service Coverage and Base Station Antenna Configuration in UMTS Networks," *IEEE Wireless Communications*, vol. 13. pp. 16-25, 2006.

[27] J. Laiho, A. Wacker, T. Novosad, *Radio network planning and optimisation for UMTS*. New York: John Wiley and Sons, 2002.

[28] V. Gunasekaran, D. Fotios and C. Harmantzis, "Affordable infrastructure for deploying WiMAX systems: mesh v. non mesh," in *IEEE Vehicular Technology Conference (VTC Spring)*, 2005, vol. 5, pp. 2979–2983.

[29] R. K. Taplin, D. M. Ryan, S. M. Allen, S. Hurley, and N. J. Thomas, "Algorithms for the automatic design of WiMAX networks," in *The International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST)*, 2007, no. 12-14, pp. 322–327.

[30] S. Modarres Razavi and D. Yuan, "Performance improvement of LTE tracking area design: A re-optimization approach," presented at *ACM InternationalWorkshop on Mobility Management and Wireless Access (MobiWac)*, 2008.

[31] H. Claussen, L.T.W. Ho, L.G, Samuel, "Self-Optimization of coverage for femtocell deployments," in *Wireless Telecommunications Symposium*, 2008, pp. 278 - 285.

[32] L. Farkas, I. Laki, and L. Nagy, "Base Station Position Optimization in Microcells Using Genetic Algorithms," presented at *IEEE International Conference on Telecommunications* (ICT), Bucharest , Romania, 2001.

[33] S. B. Jamaa, Z. Altman, et al, “Manual and Automatic Design for UMTS Networks,” *Mobile Networks and Applications by Springer Inc.*, No. 9, pp. 619-626, 2004.

[34] Kocsis, L. Farkas, and L. Nagy, “3G base station positioning using Simulated Annealing”, in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol.1, 2002, pp. 330-334.

[35] F. Glover and M. Laguna, *Tabu search*. Hingham: MA: Kluwer Academic Publishers, 1997.

[36] D. Lopez-Perez, A. Juttner and J. Zhang, “Optimisation methods for dynamic frequency planning in OFDMA networks,” in *IEEE Networks*, 2008, pp. 1-28.

[37] A. Eisenblätter H. Geerdes, “Wireless Network Design: Solution-Oriented Modeling and Mathematical Optimization”, *IEEE Wireless Communications*, Vol.13, pp. 8-14, December 2006.

[38] K. Jaffres-Runser, J. Gorce, and S. Ubeda, “QoS Constrained Wireless LAN Optimization within a Multiobjective Framework,” *IEEE Wireless Communication*, vol. 13, no. 6, pp. 26–33, December 2006.

[39] Y. Collette and P. Siarry, *Multiobjective Optimisation. Principles and case studies*. Berlin: Springer, 2003.

[40] J. Yang, J. Zhang, M. E. Aydin, and J. Y. Wu, “A novel programming model and optimisation algorithms for WCDMA networks,” in *IEEE Vehicular Technology Conference,(VTC Spring)*, 2007, pp. 1182–1187.

[41] J. Yang, J. Zhang, and M. E. Aydin, “Optimisation of WCDMA radio networks with consideration of link-level performance factors,” *Inernational Journal of Mobile Network Design and Innovation*, vol. 2, no. 1, pp. 26–33, 2007.

[42] IEEE Standard for Metropolitan Area Networks, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Systems - Multihop Relay Specification. IEEE 802.16's Relay Task Group, 2008.

[43] T. Bu, M. C. Chan, and R. Rainjee, “Designing wireless radio access networks for third generation cellular networks,” in *Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2005, vol. 1, pp. 68–78.

[44] S. Mukherjee, H. Viswanathan, “Throughput Range Tradeoff of Wireless Mesh Backhaul Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 24 pp. 593-602, March 2005.

[45] H. Wei, S. Ganguly, R. Izmailov and Z. J. Haas, “Interference-aware IEEE 802.16 WiMAX mesh networks,” in *IEEE Vehicular Technology Conference (VTC)*, 2005, pp. 3102–3106.

[46] N. Esseling, H.S. Vandra, B. Walke, “A Forwarding Concept for HiperLAN/2”, presented at *European Wireless*, 2000.

[47] C. Hoymann, K Klagges and M. Schinnenburg, “Multihop communication in relay enhanced IEEE 802.16 networks,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2006, pp. 1-4.

[48] P. Mach and R. Bestak, “Performance of IEEE 802.16 with relay stations,” presented at *6th Conference on Telecommunications*, Portugal, 2007.

[49] Z. Tao, A. Li, K. H. Teo, J. Zhang, “Frame structure design for IEEE 802.16j mobile multihop relay (MMR) networks,” in *IEEE Global Telecommunications Conference (GLOBECOM)*, 2007, pp. 4301-4306.

[50] P. Lin, H. Ngo, C. Qiao, X. Wang, T. Wang, and D. Y. Qian, “Minimum cost wireless broadband overlay network planning,” in *International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2006, pp. 228–236.

[51] B. Lin, P.H. Ho, L. Xie and X. Shen, “Optimal Relay Station Placement in IEEE 802.16j Networks,” in *ACM International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2007, pp. 25 - 30.

[52] J. Vidal, N. Marina, A. Host-Madsen, “Dimensioning cellular networks with fixed relay stations,” in *International Conference on Telecommunications (ICT)*, 2008, pp. 1-5.

[53] Y. Yu, V. Genc, S. Murphy, L. Murphy, “Network planning for 802.16j networks,” in *WiMAX Network Planning and Optimization*, Auerbach Publications, 2009.

[54] V., Hrishikesh; J., Pradeep K. Revanth, Segu. "Optimal Positions of Relay Stations for Cluster-Based Two-Hop Cellular Network," accepted for *International Journal of Communications Networks and Systems*, 2009.

[55] J. Vazifehdan, H. Shafiee, “Optimal configuration of wireless cooperative networks,” in *Wireless and Optical Communications Networks (WOCN)*, 2005, pp. 102-106.

[56] A. Eisenblätter, “Frequency Assignment in GSM Networks: Models, Heuristics, and Lower Bounds,” Ph.D. dissertation, Technische Universitt Berlin, Berlin, Germany, 2001.

[57] L. Guan, J. Zhang, J. Li, G. Liu, P. Zhang, ”Spectral efficient frequency allocation scheme in multihop cellular network,” in *Vehicular Technology Conference,(IEEE VTC Fall)*, 2007, pp.1446-1450.

[58] Y. Li, Q. Miao, M. Peng and W. Wang, ”Frequency allocation in two-hop cellular relaying networks,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2007, pp.1-5.

[59] T. Liu, M. Rong, H. Shi, D. Yu, Y. Xue, E. Schulz, ”Reuse partitioning in fixed twohop cellular relaying network,” in *Wireless Communications and Networking Conference (WCNC)*, 2006, vol.1, pp.177-182.

[60] L. Chen, N. Martin, C. Cabrera, A. Srinivas, “Joint topology design and frequency allocation for fixed wireless SDR networks”, in *IEEE Workshop on Networking Technologies for Software Defined Radio (SDR) Networks,* 2006, pp. 42 – 51.

[61] M. Werner, P. Moberg, p. Skillermark, M. Naden, W. Warzanskyj, and I. Berberana, "Optimization of cellular network deployments with multiple access point types," presented at *WWRF meeting'20,* Ottawa, Canada, April 2008.

[62] H. Sherali, M. Pendyala, and T. Rappaport, “Optimal Location of Transmitters for Micro-Cellular Radio Communication System Design,” *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 4, pp. 662–673, May 1996.

[63] "Final report on link level and system level channel models," IST-2003-507581 WINNER, D5.4 v.1.4, November 2005.

[64] B. Timus, "Studies on the viability of Cellular Multihop Networks with Fixed Relays," Ph.D. dissertation, University of Stockholm, Sweden, 2009.

[65] L. Galati-Giordano, L. Reggiani, L, Dossi, "Multi-User sub-channel, bit and power allocation in IEEE 802.16 systems," in *Vehicular Technology Conference, (VTC Spring)*, 2007, pp.3120-3124.

[66] "Final report on link level and system level channel models," IST-2003- 507581 WINNER, D5.4 v.1.4, November 2005.

[67] M. E. Aydin and T. C. Fogarty, "A Distributed Evolutionary Simulated Annealing for Combinatorial Optimization Problems," *Journal of Heuristics*, vol. 10, pp. 269-292, No. 3, May 2004.

[68] FlexIble RElay Wireless OFDM-based networks (FIREWORKS) consortium, "Advanced radio resource management algorithms for relay based networks", Deliverable 2D2, IST-27675 STP FIREWORKS, July 2007, available at http://fireworks.intranet.gr/

[69] D. Soldani, S. Dixit, "Wireless Relays For Broadband Access," *IEEE Communications Magazine*, vol.46, no.3, pp.58-66, 2008.

[70] S. W. Peters and R. W. Heath Jr., "The Future of WiMAX: Multi-Hop Relaying with IEEE 802.16j," *IEEE Communications Magazine*, vol. 1, no. 47, 2009.

[71] W. Zirwas, T. Giebel, N. Esseling, E. Schulz, J. Eichinger, "Broadband Multi Hop Networks with reduced Protocol Overhead," in *European Wireless*, 2002.

[72] F. Gordejuela-Sanchez, D. López-Pérez and J. Zhang, "Frequency planning in IEEE 802.16j networks an optimisation framework and performance analysis," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Budapest, Hungary, April 2009 pp.1-6.

[73] D. Lopez-Perez, F. Gordejuela-Sanchez, L. Q. Zhao, W. Huang and J. Zhang, "EMI: Experimental WiMAX simulator for network planning and optimization," presented at *4th MCM COST2100*, Wroclaw, Poland, February 2008.

[74] D. Lopez-Perez, A. Juttner, and J. Zhang, "Dynamic frequency planning versus frequency re-use schemes in OFDMA Networks," in *IEEE Vehicular Technology Conference,(VTC Spring)*, 2009, pp.1-4.