Self-optimization of Antenna Tilt in Mobile Networks

VLAD-IOAN BRATU

KTH Information and
Communication Technology

Master of Science Thesis
Stockholm, Sweden 2012

Self-optimization of Antenna Tilt in Mobile Networks

VLAD-IOAN BRATU

KTH Information and Communication Technology

Master of Science Thesis performed at
the Radio Communication Systems Group, KTH.
June 2012

Examiner: Professor Slimane Ben Slimane
Supervisor: Professor Claes Beckman

KTH School of Information and Communications Technology (ICT)
Radio Communication Systems (RCS)

TRITA-ICT-EX-2012:200

Tryck: Universitetsservice AB

Abstract

The increased complexity of mobile networks, the need to deliver high data rate services and the variation of mobile traffic put a high burden on operation and maintenance in terms of extra workload and additional costs . New approaches to network optimization and management should be taken into account to increase network performance and reduce operational costs. Antenna tilt is a powerful parameter for optimization of a mobile network. It has direct impact on shaping the boundary of the serving cell and hence on the coverage and interference parameters of the network. With the introduction of Remote Electrical Tilt (RET) antennas, tilt optimization can be used in the context of self-optimization. This work discusses how base station antenna tilt can be used as a self-optimization tool for load-balancing and presents a framework for a self-optimization process that can be integrated into existing and future mobile networks. Tests using real traffic data proved that the self-optimization process can be used to correctly identify congested cells. Both link level and system level simulations are performed to determine the impact of tilt adjustments on network performance. The results show that antenna tilt can be an effective tool to achieve load-balance between neighbouring cells and thus increase the Grade of Service (GoS). Furthermore, different tilt adjustment procedures are discussed each with it's advantages and disadvantages. It is concluded that antenna tilt can be successfully used for self-optimization purposes and possible limitations and issues are discussed.

Acknowledgements

This thesis completes my two year Master of Science studies. I would like to thank my supervisor, Prof. Claes Beckman for his guidance and valuable advices during this project. Also I would like to thank my examiner, Prof. Slimane Ben Slimane for his comments and suggestions which helped to improve the work.

The project was supported by Reverb Networks. I would like to thank Arif Ansari and Magnus Friberg for giving me the opportunity to work on this challenging topic and ensuring that I have all the support needed. Also I would like to thank Mustafa Sahin for his assistance in defining the self-optimization framework, his valuable suggestion regarding the simulation process and helpful discussions.

Finally, I would like thank my family. Without their constant support and encouragements, none of this would have been possible.

Contents

List of Tables

List of Figures

List of Abbreviations

3GPP	Third Generation Partnership Project
C/I	Carrier-To-Interference
CCO	Covreage and Capacity Optimization
CDF	Cumulative Distribution Function
CE	Channel Element
CP	Cyclic Prefix
CPICH	Cumulative Density Function
CS	Circuit Switched
DL	Downlink
E-UTRA	Evolved UTRA
GoS	Grade of Service
HPBW	Half Power Beamwidth
HS-DSCH	High-Speed Downlink Shared Channel
HSPA	High Speed Packet Access
HS-SCCH	High-Speed Shared Control Channe
LB	Load Balancing
LTE	Long Term Evolution
MME	Mobility Management Entity
NGMN	Next Generation Mobile Networks
OFDM	Orthogonal Frequency-Division Multiplexing
OVSF	Orthogonal Variable Spreading Factor
PRB	Physical Resource Block
PS	Packet Switched
QoS	Quality of Service
RAN	Radio Access Network
RET	Remote Electrical Tilt
RNC	Radio Network Controller
RSS	Received Signal Strength
SF	Spreading Factor
SINR	Signal to Noise plus Interference Ratio
SIR	Signal to Interference Ratio
SON	Self Organizing Networks
UE	User Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunications System
UTRAN	UMTS Terrestrial Radio Access Network
WCDMA	Wideband Code Division Multiple Access

Chapter 1

Introduction

1.1 Background

The evolution of wireless networks combined with a growing demand for mobile broadband leads to complex heterogeneous deployments with a high density of base stations and a large number of tunable parameters. This type of deployments will require a high level of operation costs and human resources. Furthermore, operators are competing in a new environment, where packet data became the main service that generates traffic in their networks as of 2010. It is estimated that the global mobile data traffic will increase by 100 per cent each year until 2016 [1]. This means that efficient usage of available network resources becomes very important. In this context, new approaches for network operation and management must be considered and network optimization becomes an important tool.

Manual optimization can be time consuming and costly in terms of man power [2]. Also it cannot be performed in real time and thus cannot adapt to fast changing network conditions. Moreover, some tasks are to repetitive or difficult to be performed manually. Self-optimization is considered as an alternative for increasing performance and reducing operational costs [3].

The first high level requirements for the introduction of self-optimization were given in 2006 by the Next Generation Mobile Network Alliance (NGMN) [4]. They were later followed by a detailed list of use cases [5] some of which are now standardized by 3GPP [6]. The use cases are defined as general guidelines and identify possible applications for Self Organizing Networks (SON) which relate to self-configuration, self-optimization and self-healing. Load-balancing was one of the self-optimization use case that was identified and also included in the 3GPP standard.

While the objectives of the use cases are strictly defined, the choice of methods, input measurements and output control parameters are yet to be standardized. This offers flexibility when implementing a SON solution.

1.2 Overview on self-optimization

Self-optimization is a two step process in which first abnormal operation, e.g. congestion, is automatically detected based on the designated inputs and after that, configuration parameters are automatically adjusted to improve performance.

Self-optimization algorithms use Key Performance Indicators (KPI), measurements or counters as inputs. There are different sources for gathering information about the state of a network, depending on the Radio Access Technology, such as the NodeB or eNodeB, the User Equipment (UE), the Radio Network Controler (RNC) in the case of UMTS or Mobility Management Entity (MME) in the case of LTE. While some KPI's can be vendor specific, there are standardized measurements that are also available. For Universal Mobile Telecomunication System (UMTS) the measurement capabilities for both UE and UTRAN are specified in 3GPP TS 25.215 [7], while for Long Term Evolution (LTE) the measurement capabilities for UE and E-UTRA are specified in 3GPP TS 36.214 [8]. A detailed list of LTE measurements that can be used as a support for SON is given in [9].

The outputs are used to control different radio or network configuration parameters, such as antenna tilt, transmit power or handover offset (HO). This work will focus on antenna tilt as an optimization parameter for load balancing, since it has a direct influence on the received signal level and thus on shaping the best server boundaries of a cell.

1.3 Problem Description

Antenna tilt has been proven as an efficient tuning parameter in mobile networks and both manual and automatic optimization techniques have been discussed in literature to some extent. With the ability to remotely control antenna tilt using Remote Electrical Tilt (RET), this parameter can be used in the context of self-optimising networks to address some of the SON use cases. However there are some challenges associated with the introduction of self-optimization of antenna tilt in real-networks which are detailed bellow. An issue that is not discussed and can be seen as suitable for future work is the interaction of self-optimization of antenna tilt with other use cases from the 3GPP specifications.

The self-optimization process for load-balancing using antenna tilt should be able to detect critical cells based on specified performance thresholds. Also it should be able to estimate if tilt changes will have an impact or not, i.e. if the low performance is cause by congestion or other factors. Resources in neighbouring cells should be verified, before making load-balancing adjustments. Furthermore, the load-balancing process should not have a negative impact on the Quality of Service (QoS) and the same services should be supported even after load-balancing is performed.

The traffic load in mobile networks varies based on the user distribution, which in turn changes according to certain patterns. However, the configuration for a cell does not change with the user distribution meaning that at some points, available resources in a cell can be either under-utilized or over-utilized. While antenna tilt can be used to shape the coverage area of a cell and thus the number of served users, the actual distribution of these users in an area will determine if load balancing is efficient or not. In particular, the optimization should become more efficient as the traffic is concentrated towards a single cell in an optimization cluster.

In the case of antenna based optimization, antenna tilt is the main radio configuration setting that will be adjusted for the cells that are involved in the optimization process. There are different tilt change step sizes that can be applied, which can impact the performance and convergence of the optimization algorithm. Also the effect of the tilt adjustment varies with the radiation pattern of the antenna. With a narrower vertical beamwidth, tilting has a greater impact, as shown in the following sections.

1.4 Scope of the Thesis

The work is aimed at studying and simulating an antenna tilt based self-optimization algorithm for load balancing in mobile networks. First, the concept of load-balancing of antenna tilt is discussed and illustrated using a link level simulation in Matlab. Following that, a method for detecting congested cells, selecting neighbour candidates and performing tilt adjustments is presented. Finally, simulations in a radio planning tool will be used to address some of the questions related to self-optimization of antenna tilt, the dependency on the traffic distribution and the possible performance gains that can be achieved.

1.5 Related Work

Automated optimization of radio parameters is discussed in literature both as a tool to help the network design process and as an efficient method to improve coverage, capacity and to address load imbalance situations. Optimization methods are presented both in the context of UMTS and LTE networks. In [10] the author discusses the optimization of three configuration parameters, antenna tilt, antenna azimuth and pilot power in UTMS networks, using a case study from real networks. It is acknowledged that antenna tilt is the most powerful optimization parameter. The results showed a reduction in unserved traffic and a decrease in downlink cell load, although this is not correlated to the uplink cell load and optimization of both UL and DL load is required. The proposed techniques are more suitable for efficient radio planning rather then for SON use cases. Another optimization method for the same three parameters is discussed also in [11]. The authors propose an algorithm based on simulated annealing which allows steps that produce a slight decrease in performance in order to avoid being trapped in a local minimum. This algorithm is suitable for combinatorial optimization problems as in the case of UMTS or LTE radio network optimization. Results from both papers show that antenna tilt has the most powerful impact with respect to the optimization targets considered and accounts for almost half of the gains.

A different algorithm for self-configuration of base station antenna tilt and transmit power is presented in [12] and evaluated with a LTE system level simulation.The optimization target is the $SINR$ measured at the UE for two scenarios, one with base stations placed in a regular hexagonal pattern and one with a irregular set-up. It was shown that terminal measurements can be successfully used as inputs for the algorithm and a centralized architecture is suitable for implementing tilt optimization. Also from the results it can be seen that there is only a small performance increase between tilt optimization and a combination of tilt and power optimization.

The same two parameters, antenna tilt and transmit power are used in [13] as an optimization objective for load balancing in the presence of inhomogeneous traffic. Simulations were performed in a UMTS-FDD environment, with the aim of reducing the uplink load factor for all base stations in the scenario. The work shows that this type of solution can be applied to situations where there are patterns in the traffic distribution which repeat with a certain periodicity. The performance gains where seen both in the reduction of mean UL load factor and also mean downlink transmit power. Observations were made also regarding soft handover areas, which show an increase when only pilot powers are optimized and a decrease when tilt is added.

Using automatic antenna tilt adjustments as a method to adapt to unbalanced traffic distributions is discussed also in [14] and evaluated using a 21 cell network simulation for UMTS-FDD. The results show that antenna tilt optimization provides an average capacity increase of 20 to 30 per cent and is efficient in interference reduction. Also the position of the traffic hotspots relative to the neighbouring base stations has an impact on the capacity gain that can be achieved with tilt adjustments. In particular the largest gains are seen when the traffic hotspots are at unequal distances from the base stations. However, if a particular base station is closer to the traffic hotspot, the performance gain will be reduced, since a large number of UE will be moved at cell borders and power will become a limitation. Also they will be forced to transmit with high power levels, which will increase interference.

Two of the most important optimization targets in mobile networks are capacity and coverage. However, there is a tradeoff between them, which is discussed in [15], with the goal of finding “optimal tradeoff points”. Capacity is dependant on the traffic distribution between cells and antenna tilt can be efficiently used to change cell boundaries in order to achieve load balance. The results show that optimization of antenna tilt for non-uniform traffic distributions gives a 10 per cent capacity decrease for the same coverage, compared to an optimized scenario with a uniform user distribution. Nonetheless, tilt optimization lead to a 16 per cent increase in capacity.

In [16] the discussion of automated optimization for wireless networks is extended to include a framework which covers all stages of the optimization process. Using a case study for a UMTS network to demonstrate their approach, the work is also put in the context of LTE to show how UMTS optimization can be used as a background to develop self-optimization procedures. Furthermore, the authors identify some of the challenges associated with the introduction of SON, which can be used as a guideline for future work In particular they show that both the inputs and outputs of the algorithm should have a reliability index. The optimization procedure must be able to estimate the impact of a parameter change on the network. The SON algorithm has to determine cases were it cannot correct sub-optimal operation using the parameters that it can control.

The impact of electrical and mechanical tilt on the performance of WCDMA networks was studied in [17]. It is concluded that the optimum downtilt angle depends on the network configuration, vertical HPBW and antenna height. In [18] the authors extend the study to LTE and also include the effect of azimuth HPBW angle. In particular the work simulates how the two types of tilt, mechanical and electrical affect coverage and throughput. For LTE downlink performance, it is demonstrated that electrical tilt provides optimal mean throughput, while a combination of mechanical and electrical tilt optimizes the peak rate. The type of tilt or tilt combination has little impact in respect to coverage optimization.

Reasons for introducing SON and a vision on how it could be implemented are discussed in [2], with emphasis on Self Configuration and Self Optimization. Also, an optimization method that uses successive antenna tilt changes to maximise cell edge performance and mean cell throughput is presented. The results showed a 44 per cent improvement of cell edge performance, while the mean cell throughput increased by 7 per cent, with the algorithm converging after approximately 60 iterations.

Chapter 2

Self Organizing Networks

2.1 Introduction to Self Organizing Networks

The need for Self Organization in mobile networks was identified based on the day-to-day operational experience of mobile operators. The requirement was first formulated in 2006 by the Next Generation Mobile Network Alliance (NGMN) [4]. Later, the Self Organizing Networks (SON) concept was introduced together with the Long Term Evolution standard starting from 3GPP Release 8[19] as a new approach to mobile network configuration, maintenance and operation, aimed at reducing operational expenditures and improving the users Quality of Service (QoS).

At a high-level, the introduction of SON should allow for new base stations to be added in a plug-and-play manner, while existing base stations should be able to adapt to network conditions and change their operational parameters. From a functional point of view, SON can be split into Self-Configuration, Self-optimization and Self-Healing.

Self-Configuration is a function that should be performed in the pre-operational state, when the radio interface is not active. This allows newly deployed eNodeB's to automatically configure their operational parameters and download required software, once they have a connection to the core network. The configuration can include radio parameters (operating frequency, transmit power), neighbour relations and cell identification.

Self-optimization is performed in the operational state, where measurements from the UE and eNodeB are used to asses the network performance and automatically tune operational parameters in order to meet defined performance targets. Some examples where self-optimization can be applied are coverage optimization , mobility parameter optimization, interference reduction, load-balancing or energy savings.

Self-Healing will allow for automatic failure detection and response. There can be situations when the failure can be detected and fixed by a single eNodeB or when the process will involve several eNodeB's. An example of the first case can be detecting a software failure and falling back to a previous one, while the more complex second case can be a complete failure of the eNodeB which will create a coverage hole. In this situation neighbouring cells might adjust their radio parameters (antenna tilt, transmit power) to reduce the size of the coverage hole.

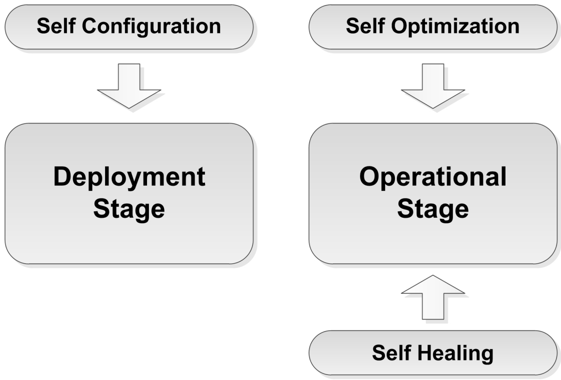

Figure 2.1: Self Organizing Network Concept

2.2 Architectures for Self Organizing Networks

There are different architectures in which a SON solution can be deployed, depending on the network element where the SON process resides. They can either be centralized, distributed or hybrid. The choice of a particular architecture will depend on the use case and the functionalities that need to be achieved.

In a centralized architecture, the SON process will run on a stand-alone server. The inputs will be gathered from the Radio Access Network (RAN) in form of Key Performance Indicators (KPI's) and then analysed. Changes to output parameters can be forwarded back to the RAN at regular intervals or based on triggers. A centralized architecture will be able to maintain global knowledge of the network status. Also it will be able to estimate the impact of parameter changes over multiple cells. However latency will be introduced by forwarding data to and from a central location and in the time interval between receiving, analysing and forwarding changes, the networks status might change again. This makes the centralized approach suitable for use cases that are not delay sensitive.

The distributed architecture will have a SON process running in each network element, e.g. eNodeB. This approach reduces the response times since the data gathering and analysis will be made at the base station. Also it doesn't present a single point of failure, such as the centralized architecture. The distributed architecture will lose the ability to estimate the overall status of the network since the information is limited to the KPI's recorded at a single network element.

The third type, hybrid architecture is a combination of the previous two. In a hybrid architecture, the SON process can be run individually for different use cases, either centralized or distributed, depending on the requirements. For example a time-sensitive or repetitive use case can be run in a distributed manner while one which impacts a larger area of the network or requires a picture of the overall network status might be run centralized.

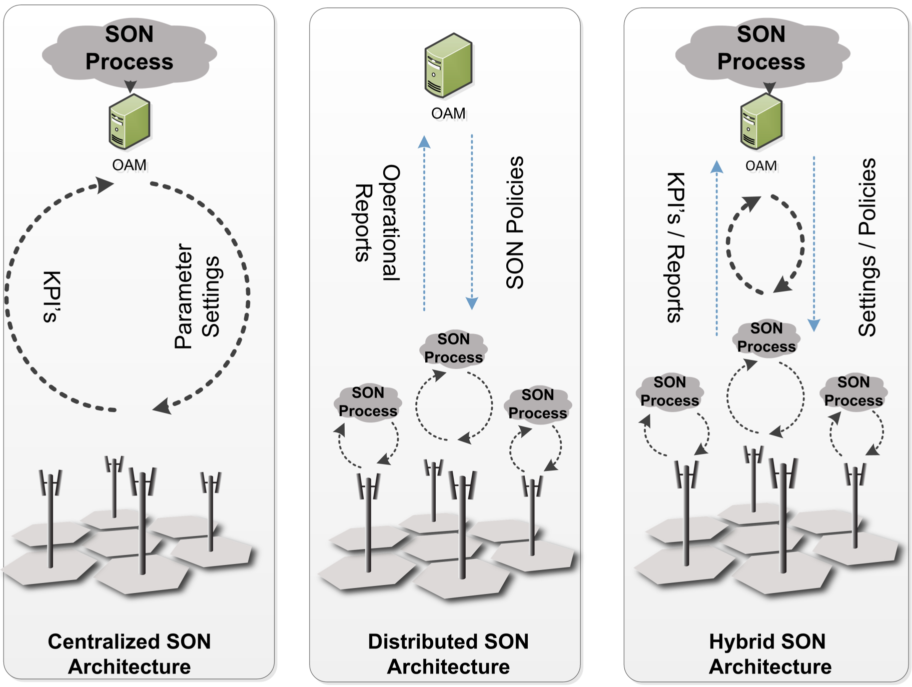

Figure 2.2: Architectures for Self Organizing Networks

2.3 Load Balancing Use Case

Load balancing (LB) is a self-optimization use case related to Covreage and Capacity Optimization (CCO). It address situations where neighbouring cells have significant differences in load and can be defined as a process which ensures that the traffic in an area is shared equally between the serving cells. Load-balancing can be effective when the traffic is unevenly distributed in a region and concentrated mostly in the coverage area of one cell. The estimated effect of load balancing is an increase Grade of Service (GoS), e.g. dropped and blocked calls. It is possible for load balancing to have a negative impact on QoS, e.g. throughput for some users [20].

Load-balance can be achieved by shifting traffic (UE's) from the highly loaded cells to less loaded neighbouring cells. There are several parameters that can be optimized as part of the LB process such as antenna tilt, handover offset or transmit power. Transmit power or handover offset offer more flexibility, since tilt control is conditioned by the availability of Remote Electrical Tilt (RET) controllers. However, adjusting the handover offset might force users to connect to cell from which they receive a lower signal, while adjusting transmit power might not provide the required signal variation. Considering that antenna tilt has a direct influence on the received signal level, load balance might be performed without sacrificing QoS.

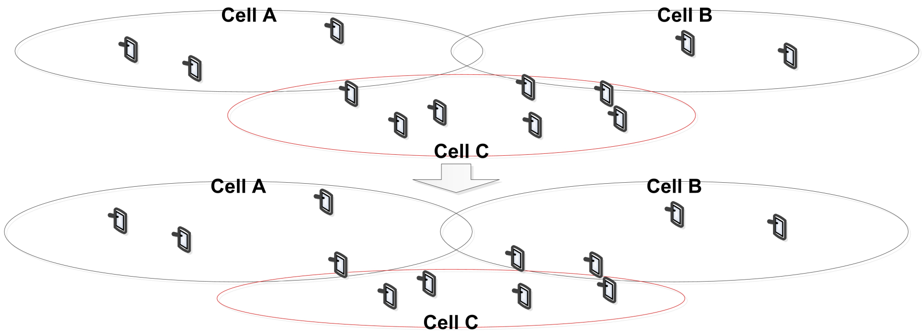

Figure 2.3: Illustration of Load Balancing

2.4 Motivation and Challenges for introducing SON

The motivation behind the introduction of self-organizing functionality in mobile networks is both technical and economical. The technical motivation comes from the need to deliver high data rates and the increased complexity of mobile networks [2]. On the economical side, the main driver is represented by operational costs for configuration and optimization. With a decrease in revenues per-megabit, operators are looking for ways to reduce their costs [3].

There are different challenges that come with SON. Operators might be reluctant to give control of their network to a fully automated process, especially when it will control parameters that have an important impact on network performance. Several SON processes can run at the same time and changes made by one process might impact another process. Also it is possible that several processes have control of the same configuration parameter. For example antenna tilt can be used as an output for both load-balancing and self-healing. This can results in possible configuration conflicts and in this case, coordination between SON processes becomes important.

Chapter 3

Load Balancing using Antenna Tilt

3.1 Antenna Tilt

Antenna tilt is defined as the angle between the main beam of the antenna and the horizontal plane. It is measured in degrees and can have positive and negative values. Positive values mean that the beam is directed downwards, the procedure is called downtilting and the tilt value is refereed to as downtilt. Negative values mean that the beam is directed upwards, the procedure is uptilting and the tilt value becomes uptilt. A tilt value of 0^o shows that the direction of the main beam is parallel to the ground and points towards the horizon.

Figure 3.1: Illustration of uptilting and downtilting

There are two methods by which tilt can be adjusted, either mechanical or electrical. Mechanical tilt implies adjusting the mounting brackets of the antenna in such a way that the whole antenna will be tilted in the desired direction, leaving the radiation pattern unchanged. Electrical tilt is achieved with a phase shifter in the feed network of the individual antenna's elements, which will allow for a uniform modification of the radiation pattern [21].

Mechanical and electrical tilt has a different impact on the antenna footprint, as seen in Fig.3.2 [21]. Mechanical tilt will distort the footprint and create and effect referred to as "pattern blooming", where the 3 dB horizontal HPBW will become wider as the downtilt angle is increased [22] . Electrical tilt will not create these distortions. However, the main drawback with electrical tilt is the limited adjustment range for tilt angle, no more than $10^o - 12^o$. This is why usually, antennas are installed with a combination of both mechanical and electrical tilt.

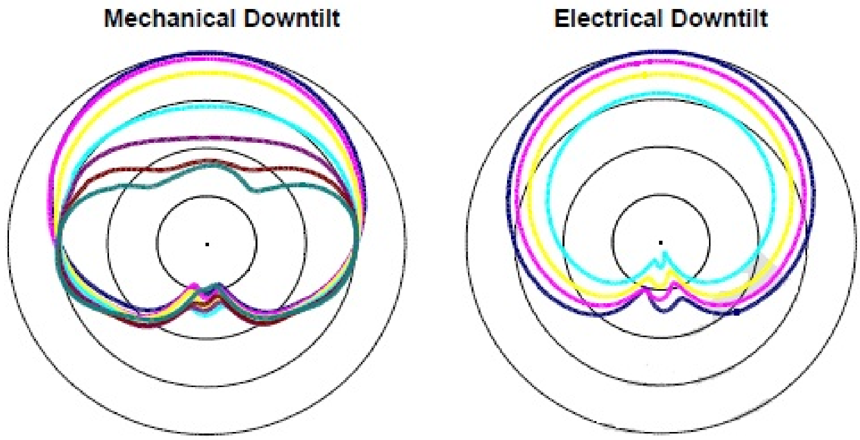

Figure 3.2: Differences in footprint between electrical and mechanical tilt

Initially, base station antennas had fixed electrical tilt, meaning that for each tilt value a different antenna would have to be used. Since this is not a flexible solution, antennas with adjustable electrical tilt were introduced. There are different ways to adjust electrical tilt, however Remote Electrical Tilt (RET) is a method that lends itself to be used in self-optimization methods. RET allows tilt to be adjusted remotely, using electromechanical actuators for phase shifting.

3.2 Link level load-balancing model

The concept of load-balancing using antenna tilt can be illustrate with a link-level model. The model illustrates the impact of antenna tilt on best server boundaries and coverage area of a cell and how this can be used to address situation in which neighbouring cells have significant differences in load. Nine cells with directional antennas, arranged in a hexagonal grid were simulated. The azimuth orientations of the antennas are 90^o, 210^o and 330^o . Users were generated at random positions in the coverage area of the cells, using a uniform distribution . To simulate traffic imbalance, a traffic hotspot with a higher user density was placed on top of the uniform user distribution. The initial scenario is illustrated in Fig. 3.3

The pathloss is modelled as distance dependent and calculated for a frequency of 2GHz, as follows:

$$L = 128.1 + 37.6log(d)$$

where d is the distance in kilometres [23]. With this, the received power at the mobile becomes:

$$P_r = P_t - L + G_h(\varphi) + G_v(\theta)$$

where P_t is the transmitted power, G_h and G_v are the horizontal and vertical gains based on the antenna model described in the next section. The values for the simulation parameters are given in Table 3.1.

Users will connect to cells based on the best downlink Signal to Interference Ration (SIR). The SIR at terminal j which is connected to base station i is:

$$\left(\frac{S}{I}\right)_j = \frac{\frac{P_{t,i} * G_{ij}}{L_{ij}}}{\sum_{k \neq i}^{9} \frac{P_{t,k} * G_{ki}}{L_{ki}}}$$

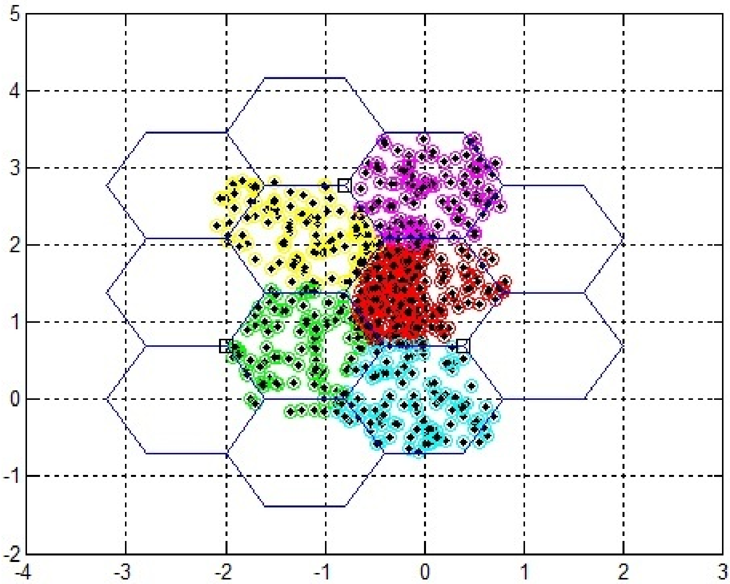

Figure 3.3: Initial scenario with unbalanced traffic and equal tilt

With the traffic hotspot placed in the coverage area of Cell-1-1 (red) a larger number of users will receive a good SIR from Cell-1-1 and connect to it. This will result in a higher load compared to the neighbouring cells, a typical scenario for load-balancing.

Antenna pattern and tilt modelling

For this simulations, the antenna pattern was build according to the model presented in [24] and included a model for adjustable electrical tilt. The gain in horizontal plane is given by:

$$G_h = -min\left(12 * \left(\frac{\varphi}{HPBW_h}\right)^2, FBR\right) + G_{max}$$

where G_{max} is the maximum gain of the antenna, $HPBW$ is the half power beamwidth, FBR is the front to back ratio and φ is the horizontal angle relative to the maximum gain direction. The following values were used:

- G_{max}=21dBi
- $HPBW$=70^o
- FBR=20dB;
- $-180^o \leq \varphi \leq 180^o$

The gain in the vertical plane is given by:

$$G_v = max\left(-12 * \left(\frac{(\theta - \theta_{tilt}}{HPBW_v}\right)^2, SLL\right)$$

where θ is the vertical angle relative to the maximum gain direction, θ_{tilt} is the electrical tilt angle, $HPBW_v$ is the vertical half power beamwidth and SLL is the side lobe level. The following values were used:

- $-90^o \leq \theta \leq 90^o$
- $HPBW_v = 10^o$
- $SLL = -18dB$

The effect of tilt changes on the received signal level using this antenna model is represented in Fig. 3.5. A mobile was placed in a fixed position at 500m from the base station, which was placed at a height of 25m. The vertical angle between the mobile and the base station was calculated as:

$$\theta = \arctan\left(\frac{(h_{BS} - h_m)}{d}\right)$$

where h_{BS}, is the height of the base station's antenna, h_mm is the height of the mobile and d is the distance between the mobile and the base station.

It can be seen how the received signal varies depending on tilt angle for two vertical HPBW values. For a narrower vertical HPBW, tilting has a more powerful impact, with a variation of 11 dBm in received signal compared to only 4.3 dBm in the case of the wider vertical HPBW. Also it can be seen that there is an optimum tilt angle, which provides the highest received signal for this antenna model and that this angle is independent on the vertical HPBW.

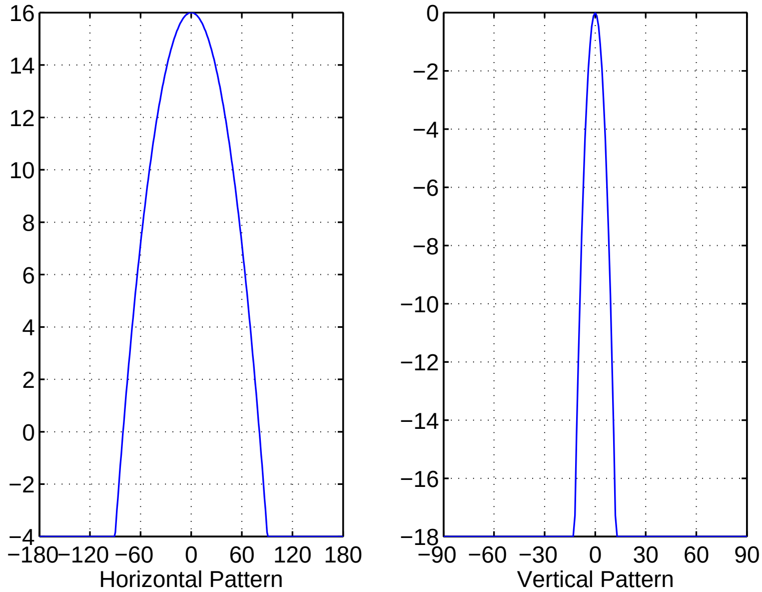

Figure 3.4: Horizontal and vertical radiation patterns for the modelled antenna

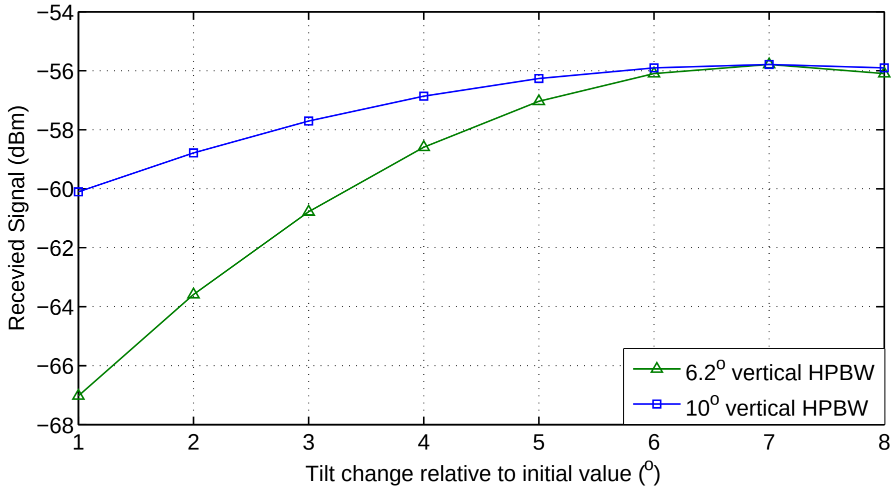

Figure 3.5: Received signal as a function of tilt angle at a fixed position

3.3 Link level load-balancing simulation

Simulations were performed using the model described in the sections above. The goal was to demonstrate how antenna tilt can be used as an efficient tool to achieve load-balance in the presence of irregular traffic distributions. The simulation parameters are given in Tabel 3.1

Starting from the initial tilt angles several tilt changes were performed, that consisted of downtilting for the congested cell and uptilting for the neighbours. After each step, the load of each cell was evaluated. Load was defined as the ratio between the total number of connected users and the maximum connections allowed by the cell.

Base station antenna height	25m
Mobile height	1.5m
Cell radius	800m
Path loss	L = 128.1 + 37.6*log(d)
Base station Tx Power	43 dBm
Antenna Gain	16 dBi
Antenna horizontal HPBW	70^o
Antenna vertical HPBW	10^o
Tilt range	$0^o - 10^o$
Tilt step size	1^o
Type of tilt	Electrical

Table 3.1: Simulation Parameters

The first step was downtilting the antenna of the congested cell, Cell-1-1 with 1^o. The second step was uptilting the antenna of neighbouring cells, Cell-2-3, Cell-3-2 and Cell-3-3 by 1^o. After six iterations load-balance was achieved and no significant improvements could be made. The changes made at each iteration are presented in Table 3.2, while the load of the cell is shown in Fig. 3.4

Iteration	Cell-1-1	Cell-2-3	Cell-3-2	Cell-3-3
1	Initial setting	Initial setting	Initial setting	Initial setting
2	1^o Downtilt	No change	No change	No change
3	1^o No change	1^o Uptilt	1^o Uptilt	1^o No change
4	1^o Downtilt	No change	No change	Uptilt
5	No change	1^o Uptilt	Uptilt	No change
6	Downtilt	No change	1^o No change	No change

Table 3.2: Tilt changes as a result of load-balancing

The load balancing process resulted in a 28 % reduction in load for the congested cell. The load increase in neighbouring cells was between 9 % and 14 % and the maximum load of any cell at the end of the process didn't exceed 50%.

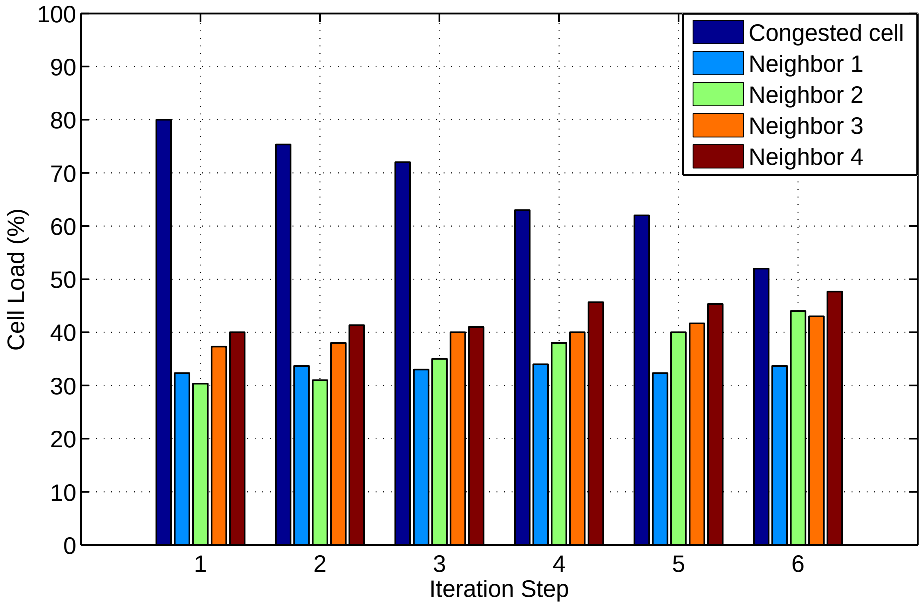

Figure 3.6: Variation of cell load at every iteration step

The impact of antenna tilt on coverage shaping and thus on the best server boundaries for this model can be observed in Fig. 3.5 which shows the coverage area of each cell after the load-balancing process. Compared to the initial settings shown in Fig. 3.3 it can be seen that the coverage area of Cell-1-1 (red) is reduced as a result of downtilting. Also the coverage area of neighbouring cells, Cell-2-3 (green), Cell-3-2 (yellow) and Cell-3-3(magenta) has increased as a result of uptilting. Cell-1-2 (cyan) was not included in the LB process and thus the coverage area remained unchanged. However the load of Cell-1-2 increased slightly as a result of users shifting from Cell-1-1.

The main concern when performing load-balancing is maintaining the same QoS for the users that are shifted to neighbouring cells. The same QoS means that the users can access the same services with the same quality. This mainly concerns data services, where the achieved throughput depends on the radio channel quality and on the received SIR. For the considered model the downlink SIR before and after load-balancing is shown in Fig. 3.6. It can be seen that there is a decrease of around 0.5 dB in SIR for some users, while for others the SIR actually improves. This gives an indication that the same QoS can be theoretically maintained even after load-balancing.

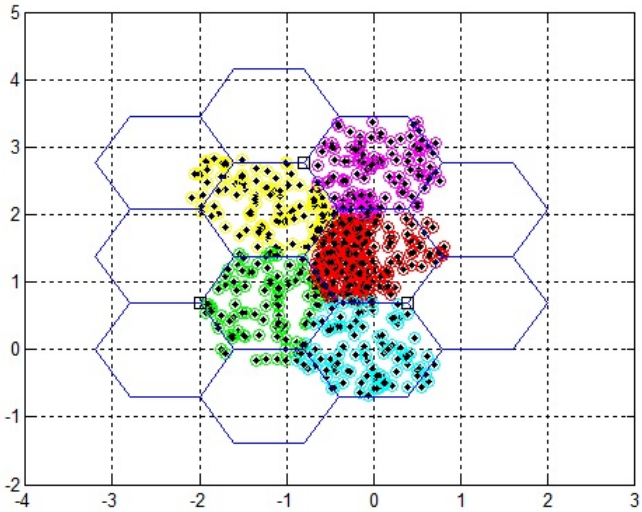

Figure 3.7: Load-balanced configuration with optimized tilt angles

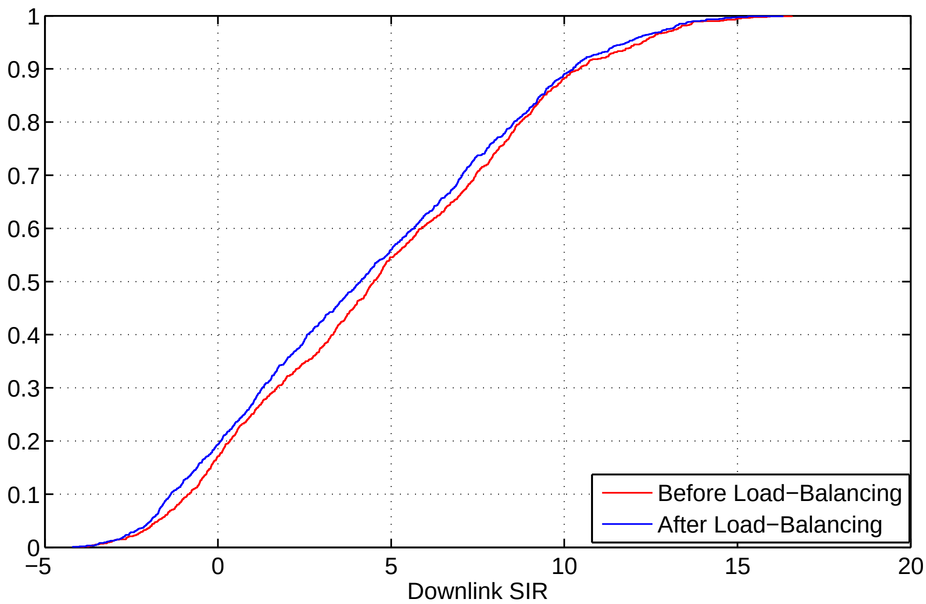

Figure 3.8: CDF plot of downlink SIR before and after load-balancing

Chapter 4

Self-optimization process

4.1 Identification of congested cells

The first requirement for a self-optimization process is to automatically identify cells that experience sub-optimal performance. The identification process relies on Key Performance Indicators (KPI) and in the case of optimization for load-balancing, the KPI's should reflect congestion due to an increased traffic demand.

A congested cell can be identified based on the Grade of Service (GoS) experienced by the users in that cell and the available resources at the base station. The GoS can be used as input regardless of Radio Access Technology (RAT). The definition of resource however will depend on the type of mobile network, e.g WCDMA, LTE.

4.1.1 Grade of Service

GoS is determined based on the call dropping and call blocking ratios. The dropped call ratio is defined as the probability that a call is terminated prematurely, i.e before the user terminates the call. It is calculated as the percentage between the total number of dropped calls and the total number of admitted calls.

$$Dropped\_Call\_Ratio = \frac{N_{drop}}{N_{admitted}}$$

The blocked call ratio is defined as the probability that a call is not admitted in the network. It is calculated as the percentage between the total number of blocked calls and the total number of attempts.

$$Blocked\_Call\_Ratio = \frac{N_{block}}{N_{attempt}}$$

In order to combine both metrics described above, the Grade of Service (GoS) is defined as:

$$GoS = 1 - Blocked\_Call\_Ratio + Dropped\_Call\_Ratio$$

There are different services offered in mobile networks, e.g. voice and data. Each of these services will have an associated blocking and dropping ratio and thus an associated GoS. In this way, the GoS can be used as a first indicator for congestion, per service. The second indicator is represented by the availability of resources in the cell. This is required to determine if the value of the GoS is reduced by congestion and if load-balancing using antenna tilt can be employed.

4.1.2 Resource Utilization

Base stations for mobile communications have a limited number of resources that can be allocated to users. Depending on the mobile technology considered, these resources differ. However, the level of resource utilization at the base station can be used to identify congested cells regardless of technology.

Resources in WCDMA networks

The main downlink resources for a WCDMA cell are transmit power and channelization codes [25]. Also the concept of Channel Elements (CE) is used to quantify the hardware and processing capabilities of a base station. In the uplink, the load is given by the noise rise.

The total transmit power of a NodeB is shared between common channels and traffic channels. The common channels are transmitted at fixed power, and thus their contribution to the downlink load is constant. However, the power required to support traffic channels will be dependent on the user location and supported service. For the purpose of the cell identification method the downlink power load can be defined as:

$$\eta_{DL} = \frac{P_{used}}{P_{total}}$$

where P_{used} represents the amount of utilized power and P_{total} represents the total available power at the cell. However, this formula does not give an estimate for the interference on the air interface [26].

Channelization codes are used for multiple access and they are based on Orthogonal Variable Spreading Factor (OVSF) codes. They are arranged in a tree structure based on the Spreading Factor (SF), with a total of 256 codes with SF 256. Depending on the service delivered to the user a specific SF is allocated, with higher data rates requiring a lower spreading factor. For example, a 12.2 kbps voice bearer requires SF 128 while a 384 kbps data bearer requires SF 8. The codes allocation is governed by two rules. A code can be allocated only if there are no other codes in use on the path to the root of the tree and also if there are no codes in use in the underlying branches. For the example above, this would result in 34 codes being used out of the 256 available (two SF 256 for the voice service and 32 SF 256 for the data service) . Furthermore, part of the code tree will be reserved for High Speed Packet Access (HSPA) if only one carrier is used. HSPA uses fixed SF 16 codes, and depending on the configuration, several SF 16 codes can be reserved Assuming five SF 16 codes (80 SF 256 codes) used for HSPA, this results in 176 SF 256 codes that are available for other services. An example of code tree allocation is given in Fig. 4.1, with green representing available codes, orange, used codes and grey, codes that are not available.

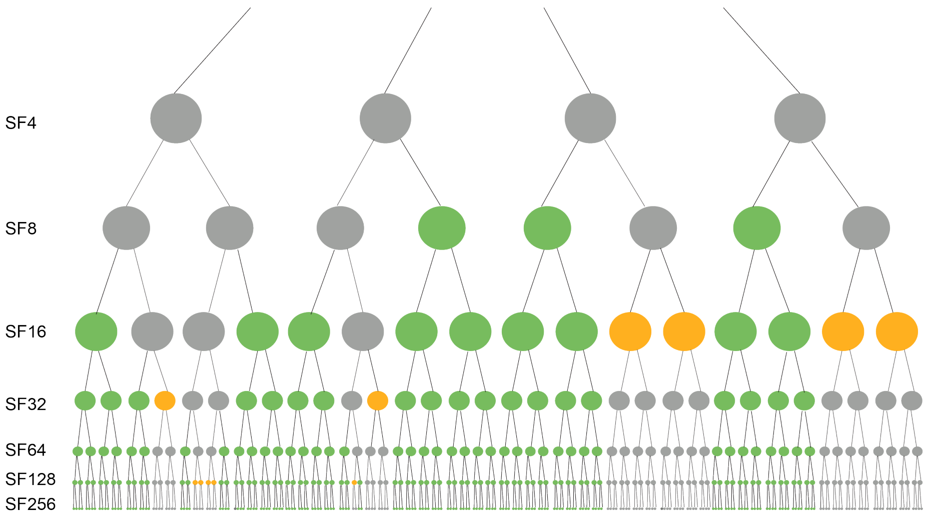

Figure 4.1: Example of OVSF code tree allocation

To quantify the channelization code usage for the purpose of identifying congested cells, the following formula is used:

$$\eta_{CC} = \frac{N_{used}}{N_{total} - N_{reserved}}$$

where N_{used} represents the number of codes in use, N_{total} represents the total number of available codes and $N_{reserved}$ represents the number of reserved codes for common channels and HSPA.

The required processing and hardware for one speech bearer (12.2 kbps) represents one channel element (CE). In this way, higher bit rate services will consume more CE. The exact definition and calculation method for CE is not standardized, and thus it varies based on vendor. However, each base station will have a limited number of CE. The channel element utilization is calculated as:

$$\eta_{CE} = \frac{N_{used}}{N_{total}}$$

where N_{used} represents the number of CE that are in use and N_{total} represents the total number of available CE.

Resources in LTE networks

The downlink transmission scheme in Long Term Evolution (LTE) is based on Orthogonal Frequency Division Multiplex (OFDM). Using OFDM, a wideband carrier can be split into multiple narrowband subcarriers. In this way the downlink physical resource in LTE is allocated in both frequency and time. A basic resource element is represented by one OFDM subcarrier during one OFDM symbol interval [25].

Multiple basic resource elements are grouped into Packet Resource Blocks (PRB), as shown in Fig. 4.2 . The PRB is the smallest resource unit that can be allocated to a user. One PRB consists of 12 consecutive subcarriers transmitted in one time slot. The timeslot has a duration of 0.5ms and consists of 6 or 7 OFDM symbols, depending on the length of the cyclic prefix (CP). This results in a PRB containing either 72 or 84 resource elements.

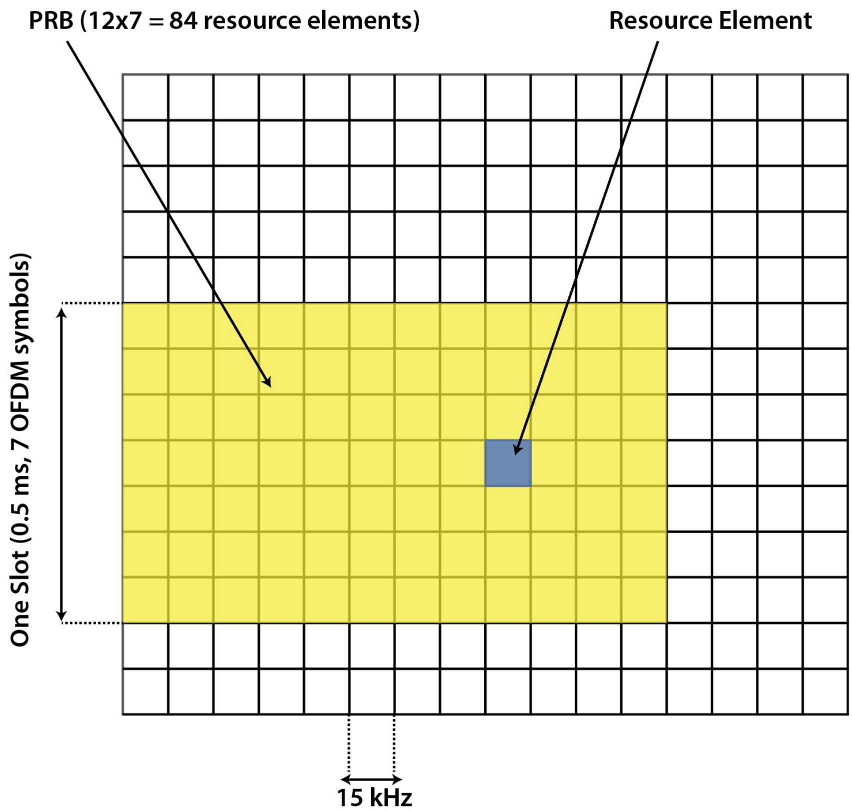

Figure 4.2: Downlink time-frequency resource blocks in LTE

The frequency spacing between two subcarriers in LTE is $\Delta f = 15$kHz. A PRB using 12 subcariers will have a total bandwidth of 180 kHz. Depending on the maximum system bandwidth, the number of available PRB's will vary.

System Bandwidth (MHz)	1.4	3	5	10	15	20
Number of PRB's	6	15	25	50	75	100

Table 4.1: Number of available downlink PRB's depending on system bandwidth

The ratio of PRB's which are in use to the number of total available PRB's will give an indication of the resource utilization level at the eNodeB:

$$\eta_{PRB} = \frac{N_{used}}{N_{total}}$$

4.1.3 Identification algorithm

The identification algorithm should use KPI's, related to GoS and resource utilization, collected over a period of several days. This time interval is required to establish if a cell constantly experiences sub-optimal performance, i.e congestion and filter out isolated events. A longer analysis period will give more accurate results.

Depending on the interval at which KPI's are collected and averaged, the number of available values will vary. For example, if the KPI's are collected each hour, this will result in 24 values per day and a total of $24 * Analysis\_Period$ values per cell. A detailed implementation based on the KPI's described in the sections above is presented next.

For every cell, at each collection interval, e.g. one hour, the GoS for each service is compared against the threshold for that service. In this way each service can have different threshold values. If the GoS value is bellow the set threshold, one event is marked at that time for that service. The events are weighted, so that priority can be given to certain services. Each cell will have a number of events for each time interval the GoS was collected. In the last step, the events are averaged over the analysis period, giving a total number of averaged events equal to the GoS collection interval for each cell. For example, if GoS values are collected every hour and the analysis period is 30 days, each cell will have 24 events averaged over 30 days, each representing one hour of the day. The events are calculated as:

$$Cell\_Events = \sum_{i=1}^{N_s} \alpha_i$$

where N_s is the number of services in the cell and α_i $0 \leq \alpha_i \leq 1$ is the weight for service i. A flowchart describing the identification algorithm is presented in Fig. 4.3.

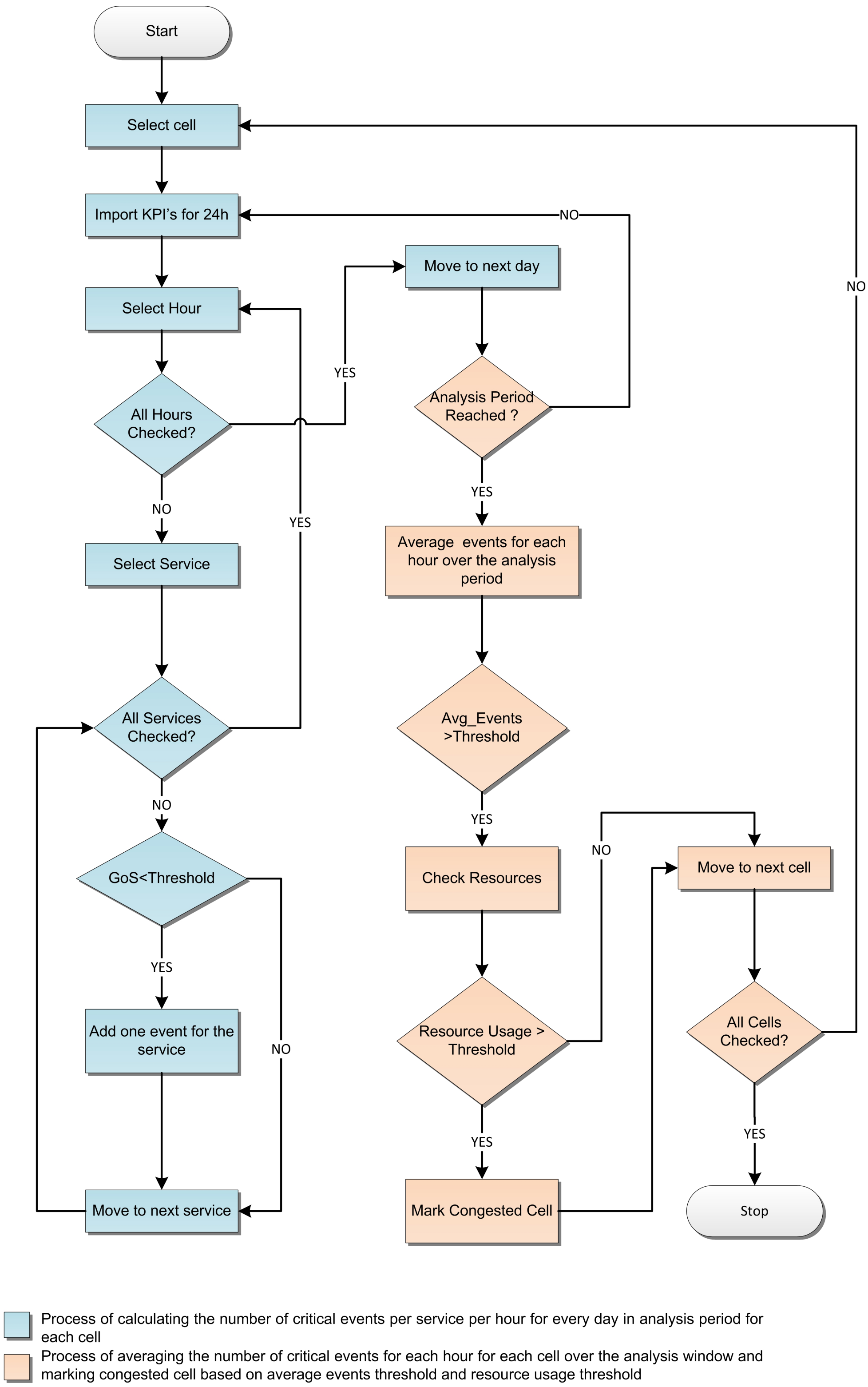

Figure 4.3: Flowchart for cell identification process

Cells that constantly experience congestion at a given time of day can be identified by comparing the average number of events to a threshold. For example, if a cell with one service experiences congestion in the same time interval, for half of the days in the analysis period, the average number of events for that time interval, will be equal to 0.5. The threshold value should be set according to the number of analysed services and the value of the weights for each service .

When the average number of events in a cell, for a given time interval is higher that the set threshold, the last step in the algorithm is to check the resource utilization in that cell. This is done to ensure that the sub-optimal performance in that cell is caused by congestion and that antenna tilt optimization can be employed. There might be cases when a cell has a low GoS, and thus creates a large number of events, although the traffic in that cell is relatively low

4.2 Example of congested cell identification based on GoS

The congested cell identification process was implemented in Matlab and applied to traffic data from a 3G network. The provided data was gathered from a number of 143 cells in an urban environment and averaged each hour for a period of 8 days. Using the blocking and dropping rates, the GoS value was calculated for each service at each hour. Since the available data did not contain resource utilization information, a different condition was used to determine if a low GoS is caused by congestion . The set condition was a minimum number of connection attempts for each service per hour. This will eliminate cells that have lower GoS, but relatively low traffic.

The analysed services were voice and data. The GoS threshold for each service was set at 99 %, which represents a combined call dropping and blocking ratios of at least 1 %. Equals weights where used for both services, voice and data, which resulted in a maximum number of two events per hour, when both services have an GoS which is lower that the threshold. The threshold for marking an hour that has a low GoS in consecutive days was set at 0.375. That means that cells which show traffic congestion for at least one service, at the same hour for 3 or more days out of 8 are marked.

The results, shown in Fig. 4.4, indicate that 15 cells out of 143 experienced congestion at the same hour for 3 or more days. Furthermore, some cells showed congestion for multiple hours of the day, with the highest value being 15 hours out of 24. This cell could be a candidate for load-balancing using antenna tilt.

Fig. 4.5 shows the number of hours with registered events, separate by services. It can be observed that voice, Circuit Switched (CS) generated more events compared to data, Packet Switched (PS). This also indicates that voice services had a lower GoS over the available analysis period.

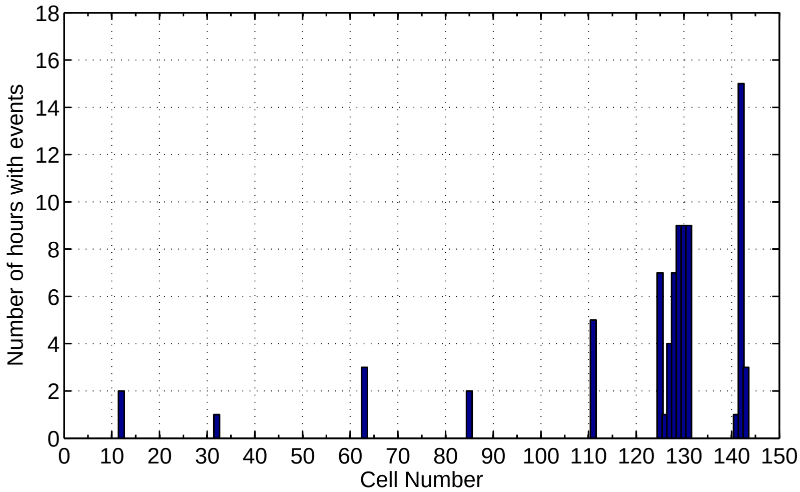

Figure 4.4: Number of average events above the set threshold

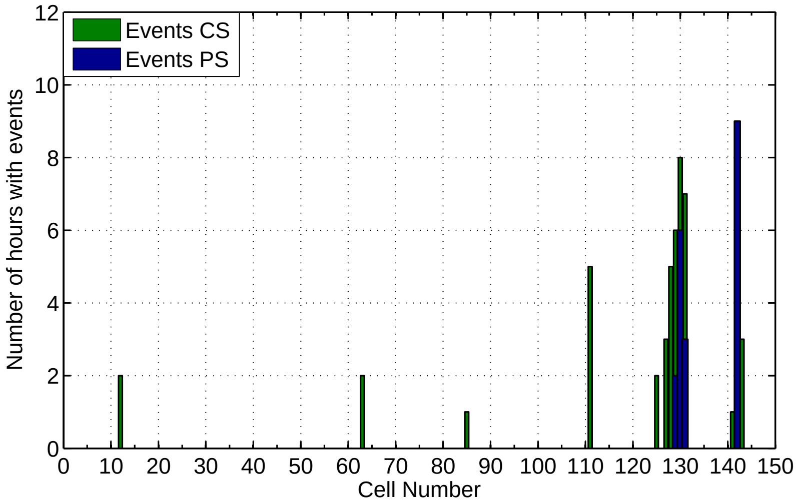

Figure 4.5: Number of average events per service above the set threshold

Traffic in mobile networks is periodic and usually varies with a period of 24 hours and within the same geographical region [27]. Fig. 4.6 shows that the traffic in the cells that experience congestion over several days is concentrated between 10:00 AM and 17:00 PM. This coincides with normal business hours. The size of the bubble is proportional to the number of events averaged over the 8 day analysis period. The smallest bubble represents only one event at that hour during 8 days. The antenna tilt self-optimization procedure can be triggered by considering the distribution of hours when congestion is presented.

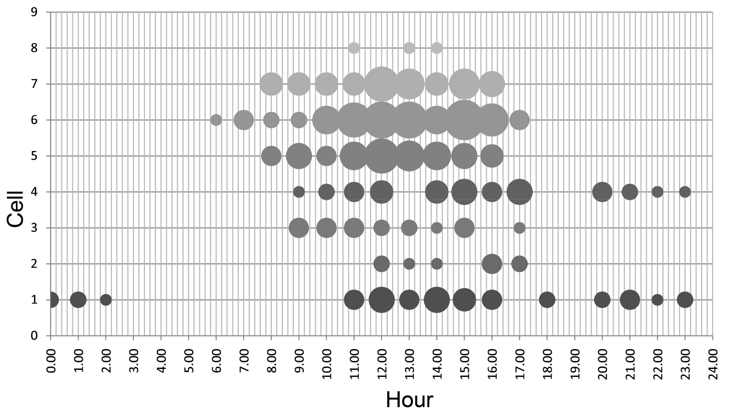

Figure 4.6: Distribution of averaged critical events for each hour

4.3 Neighbour cell selection

The next step after identifying a congested cell as a candidate for tilt optimization is to select the best neighbour cells to be included in the optimization process. The neighbours will be used to offload some of the traffic from the congested cell. The following selection criteria are used:

- Number of handovers (HO) between the congested cell and possible neighbour
- Site location
- Antenna azimuth orientation
- Resource availability

At a high level, handover can be defined as the process of changing the serving base station, while a call is in progress. This means that the mobile crosses the boarder between two cells. The number of HO to the congested cell gives a first indication on which cells can be possible neighbours. There are different decision methods for handover and these methods use either on Received Signal Strength (RSS) measurements or Carrier to Interference (C/I) measurements. Also, there are different types of handover, e.g. hard handover, soft handover or softer handover. Nonetheless, this metric can be used in the neighbouring cell selection process.

The antenna tilt load-balancing process is based on modifying the best server coverage area of the congested cell and of the neighbours. Site location and antenna azimuth, which is the horizontal angle between the main lobe of the antenna and the north, measured clockwise in degrees are used to determine if the neighbouring cell is spatially located in such a way that its coverage area can be expanded towards the congested cell.

Finally, resource availability should be taken into account, to determine if the neighbour cell can support additional users. The same resource definition as described in previous sections can be used, with respect to the considered RAT. The neighbour cell will be selected only if its resource utilization is bellow a defined threshold.

4.4 Tilt adjustment process

Tilt adjustments are the last phase in the self-optimization process. The adjustments consist in steps of downtilting for the congested cell and uptilting for the neighbours. The first step is downtilting the congested cell, followed by uptilting the neighbours. Since adjustments should be done automatically, Remote Electrical Tilt (RET) will be used. This implies that each base station antennas has a RET module.

The required number of tilt changes will vary depending on the optimization target, e.g. FR threshold, user location inside the congested cell and the granularity of tilt variations. Also, a tradeoff must be made between the duration of the optimization process and it's accuracy. A lower granularity in tilt variations, e.g. using one degree increments instead of two, will give better accuracy, although it will increase the optimization time. Using higher granularities will speed up the process, but some tilt configurations will not be checked.

In the case of multiple neighbours, the uptilt adjustments can be done either individually, each cell one after the other, or simultaneously, by uptilting all neighbour cells at the same time. As in the case of the granularity of tilt changes a tradeoff must be made. The tradeoff is between individual or simultaneous uptiliting. Simultaneous uptilting will reduce the optimization time at the cost of increasing the probability for missing an optimum configuration. This probability will increase even more with a higher number of neighbours.

The time interval between two tilt changes should be proportional to the variation of the traffic which is causing congestion. Considering the example in the section above, where the traffic varied with a period of 24 hours, one tilt change should be made after every 24 hours, until the optimization target is reached or no significant improvements can be seen.

The main limitation for the tilt adjustment process is the electrical tilt adjustment range of the antenna. A typical range can be between $0^o - 10^o$. Also the antennas will already be downtilted with a certain value in the initial configuration and this will reduce the available range even more, specially for downtilting.

4.5 Architecture

Self-optimization of antenna tilt is suited for implementation in a centralized SON architecture, where the SON process resides on a dedicated server. This architecture has the advantage of providing global knowledge of the network, i.e. KPI's from all cells in the network will be gathered and analysed in one location. This is required both for the congested cell identification process and for neighbour cell selection. Once a congested cell is identified, the SON process will send uptilt and downtilt commands to the RET controllers of the selected cells. After the first commands are sent, the process will wait a defined period of time to analyse the impact of tilt changes and decide if further actions are needed or not.

Centralized self-optimization can be introduced in a mobile network without any modifications to the existing infrastructure or base stations. The only requirement is the availability of antennas with adjustable electrical tilt and RET controllers. In the case of multi-vendor networks, where KPI's can be vendor specific, their values should be normalized to a common indicator.

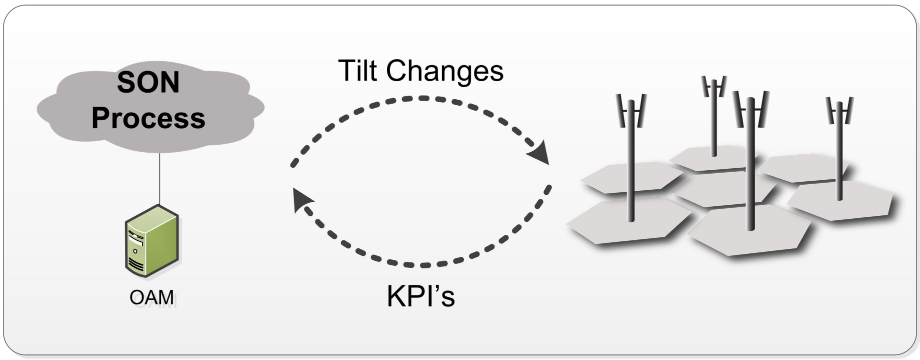

Figure 4.7: Centralized SON architecture for self-optimization of antenna tilt

Chapter 5

System level simulation

5.1 System Model and Parameters

Simulations of load-balancing using antenna tilt were performed for a WCDMA with HSPA network model. The model was created using the Mentum Cellplanner radio planning tool [28]. Both network and traffic parameters were included in the model. The map used was for a sub-urban environment and it included terrain profiles and clutter data, at a resolution of 20m.

A number of 10 sites, each with 3 cells where distributed over the simulations area, so that sufficient coverage in terms of E_c/I_o was achieved for both voice and data (HSPA) services. In order to have a cell that experiences congestion and thus can become a candidate for load-balancing, two traffic demands were generated. The first demand was a uniform distribution over the entire analysis area, using a scale factor for each service. The scale factor will determine the number of active users per square km. The second traffic demand was defined as a traffic hotspot, using a polygon, in which a certain number of active users where generated. By placing this polygon in the coverage area of one of the cells, the extra traffic was added to the uniform distribution and thus, congestion occurred in that cell.

The transmit power of the cell was set at 43 dBm, while the power for the common pilot channel (CPICH) was set at 33 dBm. The CPICH is used by the mobile (UE) for cell selection, handover and channel estimations. Also, the power of the CPICH is used as a reference for setting the power of other common channels.

The total available power is shared between the common channels and the traffic channels. The downlink (DL) power admission threshold represents the percentage of the total available power that can be allocated to traffic channels. The threshold level has an impact on the blocking probability for voice users. When deploying HSPA on the same carrier with voice, the power that is not allocated to voice can be used for HSPA. It can be seen that in this situation the DL power admission threshold should be set in such a way that it can accommodate both services. For this simulation a threshold of 75 % was used.

The noise rise gives an indication of the uplink (UL) of the cell and it is used in admission control. It is defined as the ration between the total received wideband power to the noise power [26]. Each user that connects to the cell will increase the value of the noise rise. In the case that the value will be increased above the threshold, that user will be blocked by admission control. A value of 6 dB, which corresponds to 75% uplink load was set.

The main characteristic of HSPA is the use of shared channel transmission [25]. This means that the resource of the cell, channelization codes and transmit power are shared between users in the time domain. The decision of which user gains access to the resources at a given time is made using a scheduler. There are different scheduling methods and the one used in this simulation is proportional fair. Proportional fair scheduling assigns resources to users based on current channel quality and previous received data rates.

The downlink channels used in HSPA are the High Speed Downlink Shared Channel (HS-DSCH) and the High Speed Shared Control Channel (HS-SCCH). When one carrier is shared between voice and HSPA, channelization codes have to be reserved for these channels. The HS-DSCH is transmitted using SF 16 codes, while the HS-SCCH is transmitted using SF 128 codes. The number of channelization codes reserved for the HS-DSCH determine also the maximum bit rate that can be achieved on this channel. For this simulation 5 x SF 16 codes where reserved for the HS-DSCH and 1 x SF 128 code for the HS-SCCH, which gives a total of 82 SF 256 codes. This leaves 174 SF 256 codes for voice.

Two radio bearers were defined, to support the voice and HSPA services. The required DL E_b/I_o for voice was set at 7.4 dB, and at 5 dB for data . The average bit rate for data users was set at 500 kbps in the downlink and 128 kbps in the uplink. The proportion of users between the two services was 30% voice and 70% data users, which resulted in higher blocking rates for data users.

Carrier frequency	2100 MHz
No. of carriers	1
Antenna model	Kathrein 742 215
Antenna gain	18 dBi
Antenna horizontal HPBW	65^o
Electrical tilt	Adjustable: $0^o - 10^o$
Tx Power	43 dBm
CPICH Power	33 dBm
Power admission DL	75 %
Max UL noise rise	6 dB
No. of HS-SCCH	1
Reserved SF 16 codes	5
Scheduler	Proportional fair
Radio bearers	Voice and HSPA
Traffic proportion	30 % Voice 70 % HSPA

Table 5.1: System Parameters

Figure 5.1: Simulation scenario with congested cell and two neighbours

5.2 Simulation Method

The aim of the simulations process is to determine the effect of load-balancing using antenna tilt on the GoS and QoS experienced by users in the congested cell and the selected neighbouring cells. In this case the GoS is determined by the blocking probability, while the QoS, is related to cell throughput. A two step approach is taken. First an electrical tilt change was made on one of the selected cell. Following each tilt change, a Monte Carlo simulation was run and the results analysed and compared to the previous ones. The process is terminated when no further improvements can be seen in the observed metrics or the tilt adjustment limit has been reached.

A number of three cells were included in the optimization process, as shown in Fig. 5.1, one congested cell and two neighbours. Due to the proportion of users between the two services, HSPA showed a higher blocking probability compared to voice. The average HSPA blocking probability for these cells, was selected as the optimization target. This combined metric is used in order to determine the impact of the tilt changes, both on the congested cells and on the neighbours. In the same way, the average throughput of the cells was used to determine the effect of tilt changes on QoS.

Three scenarios were simulated and are described in detail in the following sections. In the firs one, the impact of the proposed self-optimization procedure is evaluated considering a uniform traffic distribution inside the congested cell. The second scenario analyses the effect of user location inside the congested cell by placing a traffic hotspot at different positions in the area of the congested cell and comparing the results. Finally, in the third scenario, simulations are performed to determine the effect of using simultaneous uptilt adjustments in the neighbouring cells.

Monte Carlo Simulations

Monte Carlo simulations were used in order to provide significant statistical results. This method simulates random processes, by taking snapshots of random variables with a predefined probability density function and then averaging the results. When simulating mobile networks, the random variable represents the distribution of users and their activity factor. At each snapshot users attempt to connect to the network based on their specified service. The simulator measures the achieved C/I for each user, calculates the required UL and DL power and applies admission and congestion control based on the base station settings. In case that one or several conditions are not met, users are disconnected. There are several parameters to be configured for the Monte Carlo analysis, which are summarized in Table 5.2 and detailed bellow.

For each analysis a number of trials is specified, representing the number of distinctive snapshots, thus user distributions which are analysed. A higher number of trials will provide more accurate results and for this simulation 50 trials were used.

The number o iterations during each trial represents the number of allowed changes to be made to the system, in order to achieve convergence. At each iteration, power changes are made in both uplink (UL) and downlink (DL) . The value, in dB, of the power change is given by the maximum step size. The convergence limit is the allowed variation in E_b/I_o for a radio bearer to be considered connected.

The cell ranking criteria gives the order in which cells are tested for connection by the UE. There are different ranking types that can be selected, e.g. strongest signal, lowest pathloss, minimum UE power or best E_c/N_0. For this simulation the ranking was done using best E_c/N_0. In this way, the simulator will first try to connect the UE to the cell which gives the best CPICH quality.

With the Monte Carlo method the user distribution is changed after each trial. However, for the purpose of simulating load-balancing using antenna tilt, the user distribution must be fixed. This condition ensures that any modification in the observed metrics can be attributed only to tilt changes. Using a fixed seed for the simulator's random number generator will allow the same user distribution to be maintained, over multiple trials.

Number of trials	50
Number of iterations	120
Maximum step size	1 dB
Convergence limit	0.5 dB
Convergence criteria	User based (C/I)
Cell ranking	Best E_c/N_0
Fixed seed	Yes

Table 5.2: Monte Carlo Analysis Parameters

5.3 Uniform traffic distribution

For this first scenario, the traffic distribution in the congested cell is uniform. The additional traffic demand leads to an initial average blocking probability of 5.57 % between the three cells, which translates into a GoS of 94.43 %. Since most of the traffic is concentrated in the coverage area of only one of the three cells, this becomes a typical situation where load balancing using antenna tilt can be used.

The tilt adjustment procedure was applied as described in the sections above, with individual uptilt changes for the two neighbours. While this method increases the duration of the optimization process, it also increases the accuracy, by allowing more uptilt/downtilt combinations to be checked. The results are presented in Fig. 5.2, where one iteration step corresponds to an uptilt or downtilt change. It can be seen that the lowest value for average blocking probability was achieved after seven iterations, with a value 0.43 %. This results in a 5.44 % improvement in GoS. Following that, the value of the blocking probability starts to increase, as a result of congestion created in the neighbouring cells by the users that were shifted through the load-balancing process. The optimization process was terminated at this point. The tilt settings for the load-balanced configuration were obtained at step 7, with 2^o downtilt for the congested cell and 2^o uptilt for each of the neighbours.

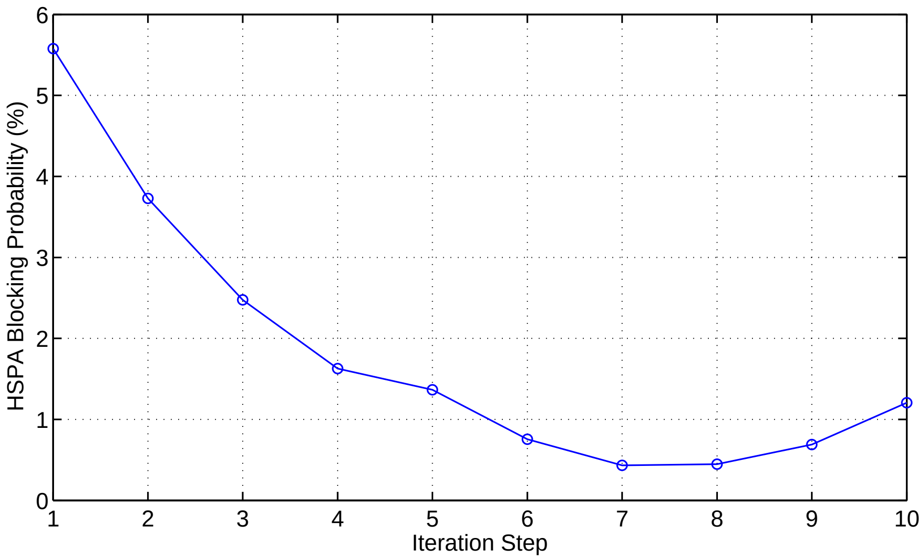

Figure 5.2: Variation of average HSPA blocking probability

The second metric was the average cell throughput, which relates to the QoS experienced by the users. The cell throughput is calculated by multiplying the total number of data service users with the average bit rate per user. It can be assumed that by increasing the GoS and thus the number of users which can connect to the cell, the cell throughput also will increase. However, for this simulation scenario, the results show that this was not the case. Throughput values oscillated at every iteration step with a clear decreasing trend. The highest decrease in throughput was 15.57 % compared to the initial value and was measured at step 9. At step 7, the point where the highest GoS was achieved, the throughput decreased by 11.25 %.

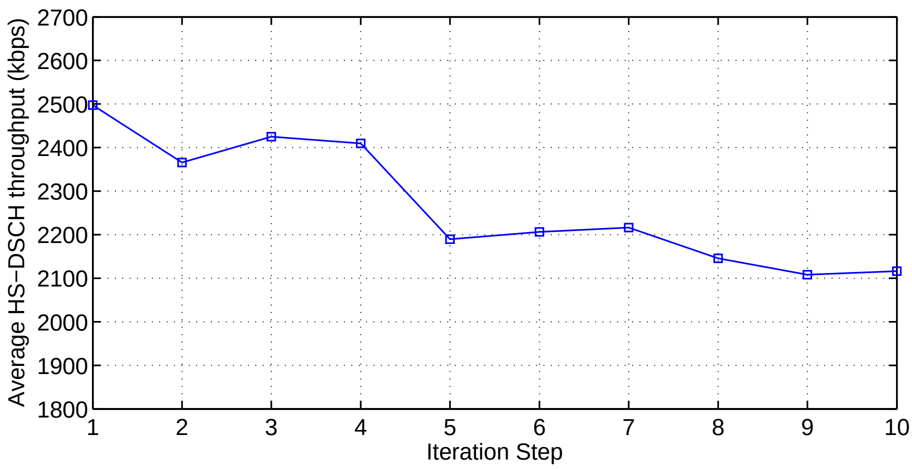

Figure 5.3: Variation of average cell throughput

Fig. 5.4 shows a comparison between the average number of served users and the average downlink data rate per user, at each step of the optimization process. It can be seen that while the number of users increases, the average data rate per users decreases. This result helps to explain the average cell throughput variation observed, variation that is not correlated to the increase in GoS.

Besides it's effects on GoS and QoS, load-balancing using antenna tilt has also an influence on the resource utilization at the base station. Fig. 5.5 shows the variation of Channel Element (CE) utilization during the optimization process. This indicator was used because, in this scenario, HSPA and voice were deployed on the same carrier and thus the transmit power will not vary with the number of users, since the remaining power after serving voice users will be allocated to data users. The number of utilized CE decreases for the congested cell and increases for the neighbours. At step 7, where the highest GoS was measured the CE utilization decreased by 26.34 % in the congested cell. The CE utilization increased in the neighbouring cells. However, the increase was asymmetrical meaning that users were shifted predominately to one neighbour.

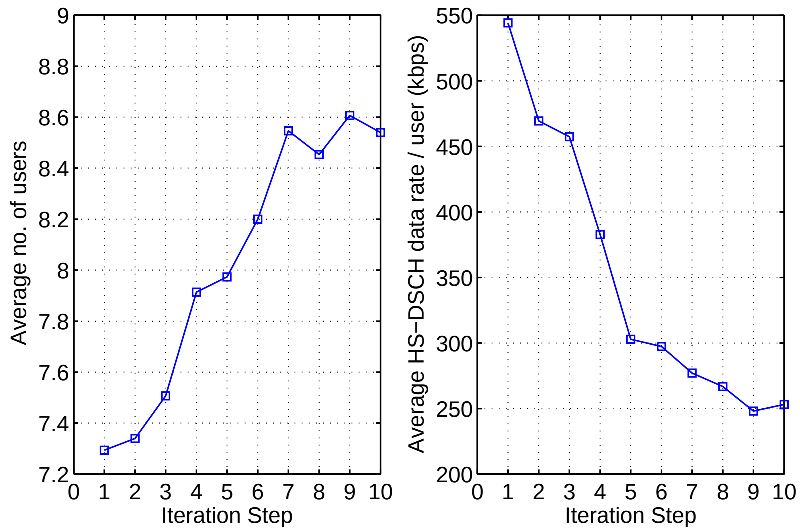

Figure 5.4: Comparison between the average number of users and average data rate per user

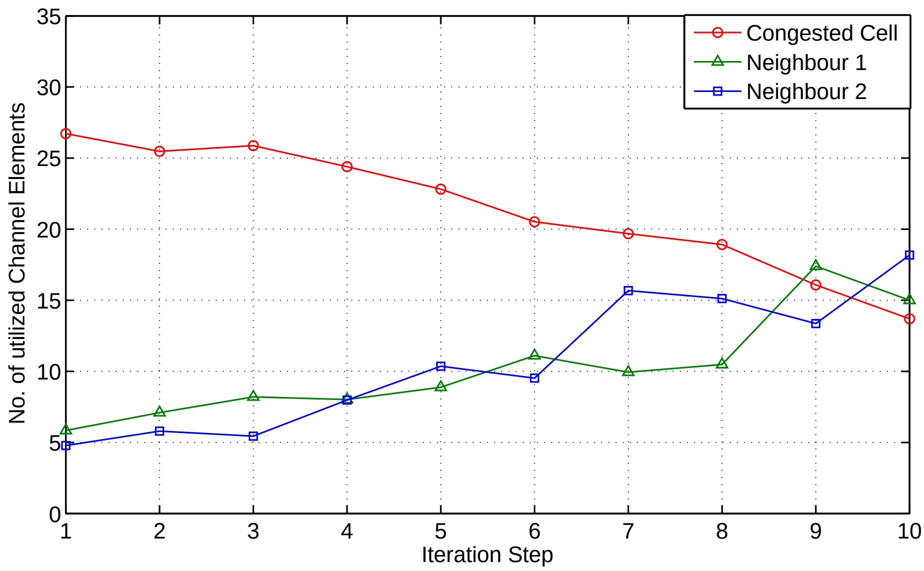

Figure 5.5: Variation of Channel Element Utilization

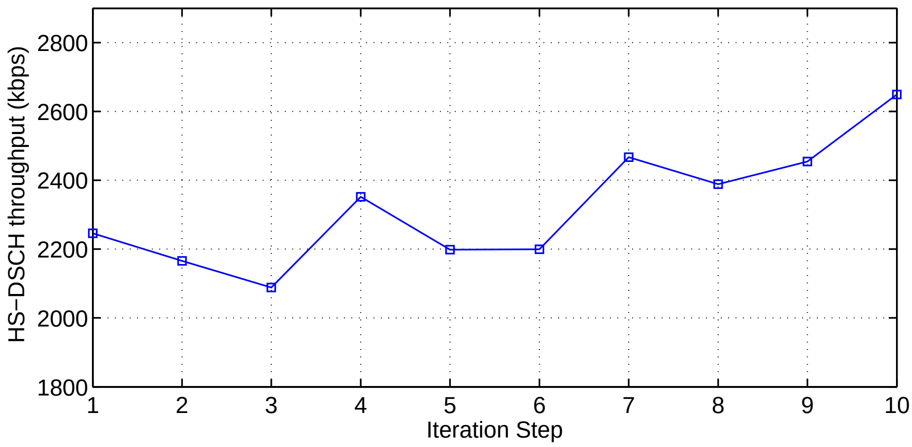

Figure 5.6: Variation of throughput in the congested cell

Although there was an overall decrease in average throughput between the three cells in the optimization cluster, the congested cell alone showed an increase in throughput, as it can be seen in Fig. 5.6. A maximum increase of throughput of 17.9 % was measured, while the throughput at the step were the highest GoS value was registered increased by 9.8 %.

5.4 Effect of inter-cell traffic distribution

The efficiency of load-balancing is also related to the traffic distribution inside the congested cell. Traffic distributions might exhibit users concentrated in a certain region of the cell, e.g. closer to the cell edge. Fig. 5.6 illustrates the effect of load-balancing using antenna tilt on blocking probability value and thus GoS, for two different position of the traffic hotspot. The results show that, when the hotspot is closer to the cell edge, the lowest average value of blocking probability is reached at an earlier optimization step compared to the case when the traffic hotspot is closer to the base station. Also the value of the average blocking probability at the optimum point is lower with the traffic hotspot towards the cell edge, having a value of 0.3 % compared to 0.65%.

The variations in throughput during the optimization process follow slightly different trends depending on the position of the traffic hotspot. The average cell throughput is higher when the traffic hotspot is closer to the base station, due to the fact that users experience a better E_c/N_0. Also, in this case, a small increase of 3.5 % in throughput is seen at step 4. However, at step 8, when the highest GoS was obtained, the average throughput decreased by 6.24 %

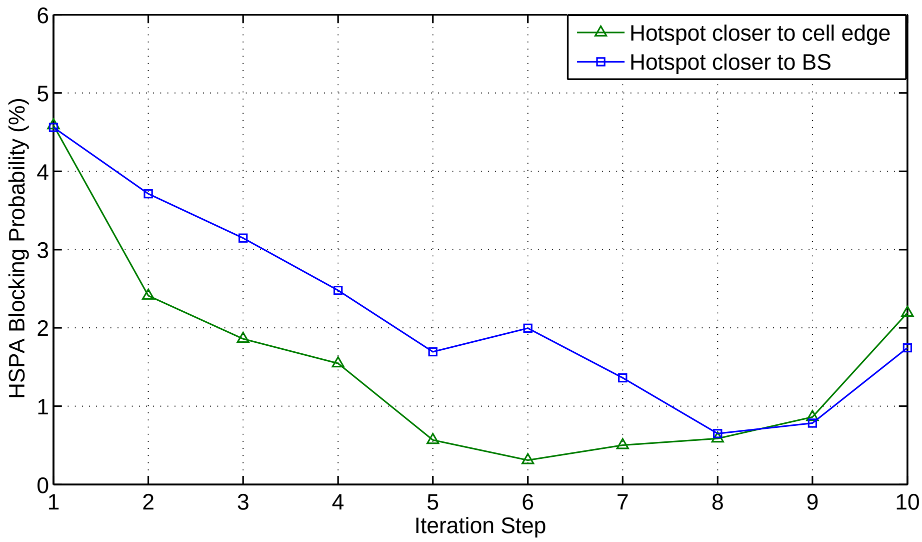

Figure 5.7: Variation of average HSPA blocking probability for different traffic hotspot positions

For the case when the hotspot was positioned towards the cell edge, the variation in cell throughput followed a similar pattern to the case when the traffic hotspot was uniform distributed, although the throughput values were lower. The highest decrease was 17.91 % and was observed at step 6 of the optimization process, which is also the step where the highest GoS was measured.

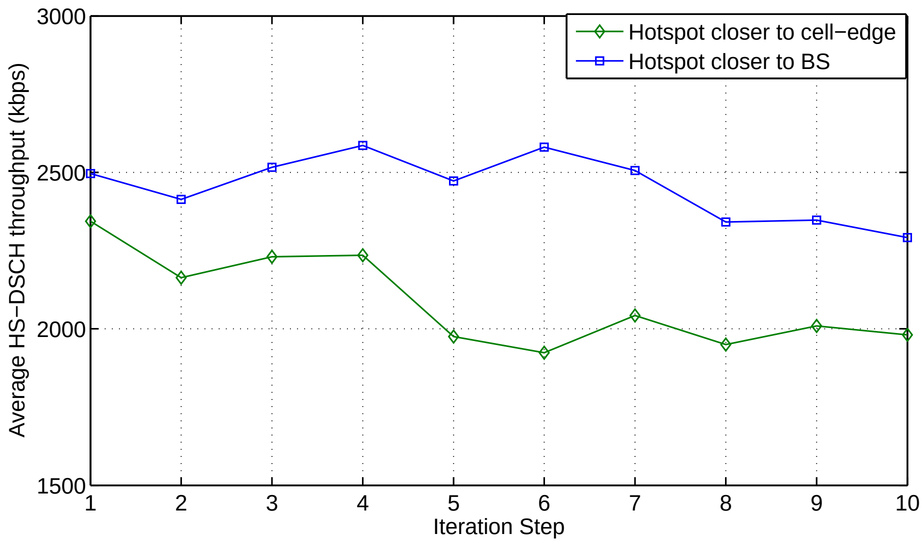

Figure 5.8: Variation of average cell throughput for different traffic hotspot positions

5.5 Simultaneous uptilt for neighbour cells

As discussed in previous sections, the duration of the optimization process can be reduced by performing simultaneous uptilt adjustments of neighbouring cells antennas. However, with this method accuracy might be reduced since some tilt configuration will not be checked. Simulation were performed to compare the optimization results for individual and simultaneous uptilting of neighbouring cells antennas. The same three scenarios presented in the sections above were used, uniform distribution, traffic hotspot closer to the cell edge and traffic hotspot closer to the base station.

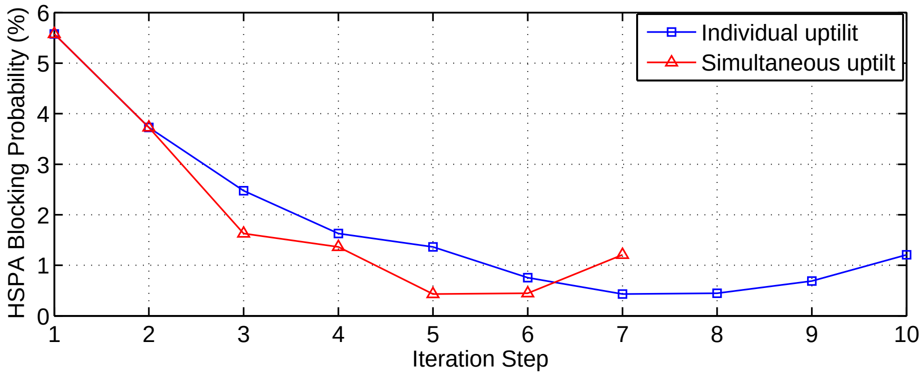

Figure 5.9: Effect of applying simultaneous uptilt for neighbour cells with uniform traffic distribution

The results confirm that with simultaneous uptilt, the optimization process is faster and the lowest blocking probability value is found at an earlier step. They also show that in some cases, the lowest blocking probability value found with simultaneous uptilt is not equal to the lowest blocking probability value found with individual upitlt. This means that a tradeoff between the speed and accuracy of the optimization process has to be made.

In the scenario with uniform user distribution, the lowest blocking probability value found during the optimization process is the same with both methods, as seen in Fig. 5.9. This is due to the fact that this value is found at a step where both neighbours are uptilted with 2^o. It can be seen that accuracy is not lost in this case.

Fig. 5.10 shows a case where the lowest blocking probability value differs depending on the uptilt method. With individual uptiliting, the lowest value is found at step 6 in the optimization process, which corresponds to a 1^o uptilt of one neighbour and 2^o uptilt of the other. This configuration will not be checked in case of simultaneous uptilt and thus the difference in the lowest value for blocking probability, of 0.3 % compared to 0.56 %.

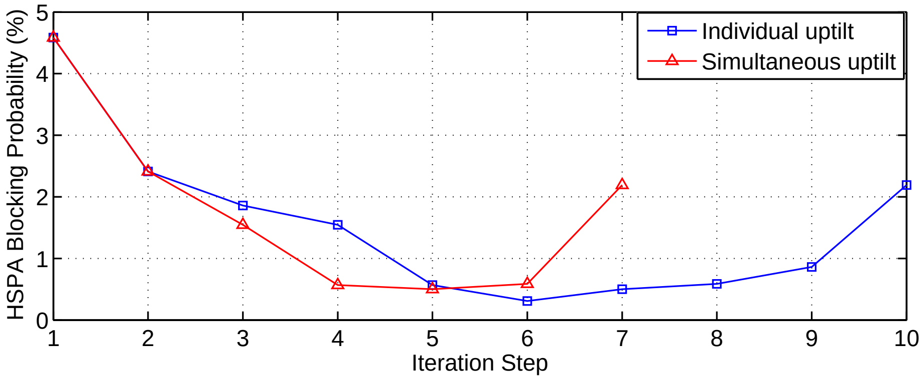

Figure 5.10: Effect of applying simultaneous uptilt for neighbour cells with the traffic hotspot closer to cell edge

In the last scenario, the value for the lowest blocking probability was the same with both methods. As in the previous case in which the blocking probability values where equal with both methods, at the optimization step where the value was obtained, the two neighbours had been uptilted with 2^o. This means that for this case there would be no performance difference between the two methods, and the faster optimization time provides an advantage for simultaneous uptilting.

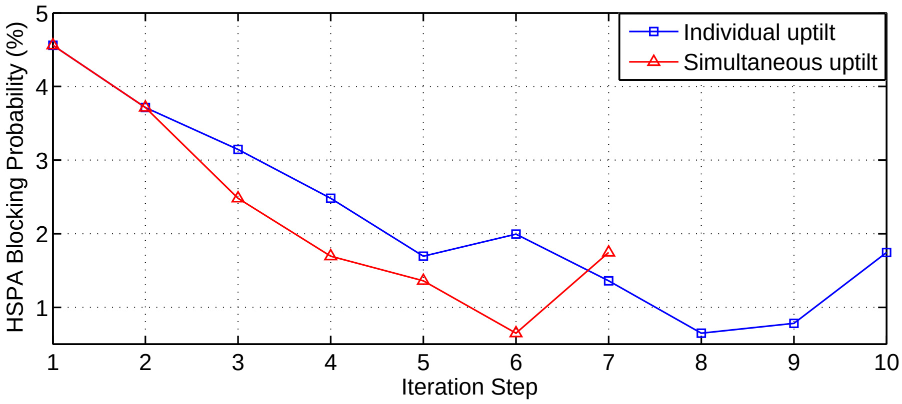

Figure 5.11: Effect of applying simultaneous uptilt for neighbour cells with the traffic hotspot closer to the base station

Chapter 6

Discussion and Conclusions

This work presented the concept of self-optimization using antenna tilt in mobile networks with emphasis on its application for load-balancing. First, an introduction to Self Organizing Network (SON) was given, including a brief history, desired functionality and possible implementation architectures. Following that, a link-level model was created starting from previous work, which acknowledges antenna tilt as an efficient optimization parameter. This model showed how antenna tilt can be successfully used to shift users between neighbouring cells and thus address situations when these neighbouring cells have a significant difference in load. Also, it was shown that the SIR values before and after the tilt optimization process have similar values, which means that theoretically, users will benefit from the same QoS.

The next step was to propose a framework on which a self-optimization process can be build. The process should be able to automatically and independently respond to situations where sub-optimal performance is detected, by identifying congested cells, selecting neighbours to be included in the optimization process and making tilt adjustments based on defined performance targets. The congested cell identification method was tested using real traffic data from a 3G network and the results showed that this method is efficient in identifying cells which can be candidates for the optimization process. Also traffic patterns were identified from the available data, where the congestion predominantly occurred during business hours. During these hours self-optimization could be employed to improve the GoS, without adding extra capacity to the base stations. Finally, a centralized SON architecture was proposed for introducing the self-optimization method in existing and future mobile networks.

System level simulations for a WCDMA network were performed, to illustrate the effect of load-balancing using antenna tilt. The results showed that this method can achieve a reduction in the average GoS of an area where one cell experiences a higher blocking probability compared to its neighbours. Furthermore, when the traffic hotspot is placed near the cell-edge, load-balancing becomes more effective, since users will be shifted to the neighbouring cells at an earlier stage.

The optimization process produced variations in average throughput for the whole optimization cluster , with an overall decrease in throughput values. Although the number of served users increased, their individual bit rate and thus cell throughput decreases. This trend was observed in all simulation scenarios and the results can be attributed to a reduction in the E_c/N_0. The users which have been connected to neighbouring cells as a result of tilt adjustments will be located at the cell-edge. Also, downlink interference might increase as a result of uptilting. This means that users will receive a lower signal and experience higher interference. Considering only the congested cell, the throughput values increase as a result of load-balancing. However, this method is aimed at improving the throughput in the whole optimization cluster and not only the congested cell. In a realistic scenario this outcome will not be accepted since increasing the throughout in the congested cell should not come with the cost of decreasing the throughput in the optimization cluster . The throughput variations are specific for this scenario and at this point cannot be generalized. Also the reduction in throughput is not the same for the three scenarios, with the lowest reduction being measured when the traffic hotspot was placed near the base station and the highest reduction being measured when the traffic hotspot was placed at the cell edge.

Simultaneous uptilt of neighbouring cells can improve the optimization time, but some tilt configuration will not be checked. The results showed that simultaneous and individual uptilting achieves the same results when the lowest blocking probability value is reached at a step where both neighbours were uptilted with equal amounts. In the case when the best configuration is found at a step where the neighbours don't have the same uptilt adjustments, individual uptilitng provides better results. Also the probability of finding the optimum value at a step where both neighbours have equal uptitl adjustments decreases with increasing the number of neighbours.

Overall, the results proved that antenna tilt self-optimization has the potential to improve performance in congested areas. This method will provide benefits both from user perspective, by improving the GoS and also from network perspectives by reducing the load on the base stations and achieving a more efficient resource utilization. The resource utilization becomes more efficient because users will be shared between multiple base stations in an area. In this way it is possible to take advantage of resources from base stations which have low traffic and increase the performance in that area.

Possible issues and limitations

Using this method presents some limitations and introduces possible issues that might have a negative impact on network performance. First of all, self-optimization of antenna tilt requires that antennas have RET capabilities. Another limitation could be the electrical tilt adjustment range, which will have an impact on the number of optimization steps that can be made. Also, antennas would already be downtilted with a certain value before starting the optimization process. For example an antenna which is already electrically downtilted with 8^o and has a maximum range of 10^o will allow for only 2^o of downitlting for optimization purposes.

Since antenna tilt has a direct influence on the coverage area of the cell, one concern could be creating coverage holes during the optimization process. This might occur mostly as a result of downtilting and will depend also on the coverage overlap between the neighbouring cells. Another issue is increasing the downlink interference by uptilting the neighbouring cells. This might reduce the C/I level and thus influence the services available to the users, e.g. decrease their maximum achievable data rate.

Future work

The simulations results showed that this method produces a small degradation in cell throughput. However, at this point this statement cannot be generalized and this is why the effect on cell throughput is suitable for future work. More specifically, the variation in downlink interference during the optimization process can be studied. Also the work can be extended to include LTE system level simulations.

Determining the average number of tilt changes which are required to achieve load-balance for different scenarios could provide insight on average optimization times. Also this would determine if the tilt adjustment range is actually a limitation or not.

The influence of cell size and propagation environment can be also a subject for future work. The simulations in this work were performed in a sub-urban environment model with medium cell sizes. Also only one traffic hotspot was introduced in the network. Simulations which include the presence of multiple traffic hotspots can give more insight on the performance of this method.

The SON concept includes different types of self-configuration, self-optimization and self-healing use cases which might be addressed by separate SON processes. The interaction and coordination between antenna based self-optimization and other SON use cases can be seen as future work. For example, a reduction in GoS caused by load-balancing might trigger an energy savings use case, which considers the cells have low traffic, although that is not the case.

References

[1] Cisco. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011 to 2016. Whie paper, , 2012.

[2] U. Barth and E. Kuehn. Self-organization in 4g mobile networks: Motivation and vision. In *Wireless Communication Systems (ISWCS), 2010 7th International Symposium on*, pages 731–735. IEEE, 2010.

[3] 4G Americas. Self-Optimizing Networks - The Benefits of SON in LTE. White paper, 4G Americas, 2011.

[4] Next Generation Mobile Networks Alliance. Next Generation Mobile Networks Beyond HSPA and EVDO. White paper, , 2006.

[5] Next Generation Mobile Networks (NGMN) Alliance. NGMN Use Cases related to Self Organizng Network, Overall Description. Technical report.

[6] 3GPP Technical Report 36.902. LTE Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Self Configuring and Self Optimizing Network (SON) use cases and soultions. TR 36.902,version 9.3.1 Release 9 , 3rd Generation Partnership Project (3GPP).

[7] 3GPP Technical Specification 25.515. Universal Mobile Telecommunications System (UMTS)Physical layer Measurements (FDD). TS 25.515,version 10.0.0 Release 10 , 3rd Generation Partnership Project (3GPP).

[8] 3GPP Technical Specification 36.214. Evoloved Universal Terrestrial Radio Access (E-UTRA) Physical layer Measurements. TS 36.214,version 10.1.0 Release 10 , 3rd Generation Partnership Project (3GPP).

[9] N. Scully, K. Zetterberg, S. Stefanski, and L. Christoph. INFSO-ICT-216284 SOCRATES d5. 10. *Measurements*, 5:10.

[10] U. Turke and M. Koonert. Advanced site configuration techniques for automatic UMTS radio network design. In *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, volume 3, pages 1960–1964. IEEE, 2005.

[11] I. Siomina, P. Varbrand, and D. Yuan. Automated optimization of service coverage and base station antenna configuration in UMTS networks. *Wireless Communications, IEEE*, 13(6):16–25, 2006.

[12] A. Temesvary. Self-configuration of antenna tilt and power for plug and play deployed cellular networks. In *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE*, pages 1 –6, april 2009.

[13] M. García-Lozano, S. Ruiz, and J.J. Olmos. UMTS optimum cell load balancing for inhomogeneous traffic patterns. In *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, volume 2, pages 909–913. IEEE, 2004.

[14] M. Pettersen, L.E. Braten, and A.G. Spilling. Automatic antenna tilt control for capacity enhancement in UMTS FDD. In *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, volume 1, pages 280 – 284 Vol. 1, sept. 2004.

[15] G. Hampel, K.L. Clarkson, J.D. Hobby, and P.A. Polakos. The tradeoff between coverage and capacity in dynamic optimization of 3g cellular networks. In *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, volume 2, pages 927–932. IEEE, 2003.

[16] O. Sallent, J. Pérez-Romero, J. Sanchez-Gonzalez, R. Agustí, M.A. Diazguerra, D. Henche, and D. Paul. A roadmap from UMTS optimization to LTE self-optimization. *IEEE Communications Magazine*, 49(6):172–182, 2011.

[17] J. Niemelä, T. Isotalo, and J. Lempiäinen. Optimum antenna downtilt angles for macrocellular wcdma network. *EURASIP Journal on Wireless Communications and Networking*, 2005(5):816–827, 2005.

[18] F. Athley and M.N. Johansson. Impact of Electrical and Mechanical Antenna Tilt on LTE Downlink System Performance. In *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*, pages 1 –5, may 2010.

[19] 3GPP TS 32.500. Universal Mobile Telecommunications System (UMTS);LTE;Telecommunication management;Self-Organizing Networks (SON);Concepts and requirements. TS 32.500,version 8.0.0 Release 8 , 3rd Generation Partnership Project (3GPP).

[20] Neil Scully, Stefan Thiel, et. al. INFSO-ICT-216284 SOCRATES D2.1 Use Cases for Self-Organising Networks. Technical report.

[21] C. Beckman and B. Lindmark. The evolution of base station antennas for mobile communications. In *Electromagnetics in Advanced Applications, 2007. ICEAA 2007. International Conference on*, pages 85 –92, sept. 2007.

[22] Louis J. Meyer. Electrical and Mechanical Downtilt and their Effects on Horizontal Pattern Performance. White paper, Comm Scope.

[23] 3GPP Technical Report 25.814. Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA). TR 25.814,version 7.1.0 Release 7 , 3rd Generation Partnership Project (3GPP).

[24] F. Gunnarsson, M.N. Johansson, A. Furuskar, M. Lundevall, A. Simonsson, C. Tidestav, and M. Blomgren. Downtilted base station antennas-a simulation model proposal and impact on HSPA and LTE performance. In *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*, pages 1–5. IEEE, 2008.

[25] E. Dahlman. *3G evolution: HSPA and LTE for mobile broadband*. Academic Press, 2008.

[26] H. Holma and A. Toskala. *WCDMA for UMTS: HSPA Evolution and LTE*. 2007.

[27] U. Paul, A.P. Subramanian, M.M. Buddhikot, and S.R. Das. Understanding traffic dynamics in cellular data networks. In *INFOCOM, 2011 Proceedings IEEE*, pages 882–890. IEEE, 2011.

[28] Mentum Cellplanner v.10.1. `http://www.mentum.com/index.php?page=cellplanner&hl=en_US`.