Analysis of Wireless Cloud Implementation in LTE-Advanced

Pedro Miguel Pardal Martins

Thesis to obtain the Master of Science Degree in

Electrical and Computer Engineering

Examination Committee

Chairperson: Prof. Fernando Nunes

Supervisor: Prof. Luís Manuel de Jesus Sousa Correia

Member of Committee: Prof. Mário S. Nunes

Member of Committee: Engineer Lúcio Ferreira

November 2013

To Inês

Acknowledgements

First I would like to offer my deep gratitude to Prof. Luís M. Correia for the opportunity of doing this M.Sc. Thesis, in cooperation with GROW. All his supervision and guidance were extremely important for this thesis, as well as all the help with technical problems, and the great general advices given, which will certainly help me on this transition to a professional career.

It was a pleasure to have been a member of GROW and to have an insight on the current research topics on Wireless Communications. Thus I want to thank all the personal of GROW, which always tried to help whenever some doubts or problems happened, and also for the development of my presentation skills. A special thank you to Lúcio Ferreira, for all the help and guidance.

To my roommates, Gonçalo Fernandes, Ricardo Santana and Sofia Teles, which followed me throughout most of my academic years, I want to give my special gratitude for their friendship and support and thank them for all the moments that we shared together.

To my best friends, André Alberto, Pompeu Santos, Jaime Pereira, Joana Falcão, Gonçalo Brás, Daniel Ferreira, Rodrigo Batista, Gonçalo Brito, Gonçalo Tomé, Tiago Pereirinha, Guilherme Santos, João Féria, and to all the rest of my friends that were not mentioned, a very special thanks for being there throughout this journey.

Last, but not least, I would like to thank my family, specially my Parents and Sister, for all their support, comprehension and advices that they gave me over these last years. Also I want to thank the Inês, the love of my life, for helping me get through these years; without her, it would not have been possible to get through these years.

Abstract

Mobile communications suffered a great increase in traffic with the introduction of LTE-Advanced, leading to the development of new centralised architectures that split Base Stations into two main components, one responsible for all the processing functions and the other for the radio signal transmissions. This work consists of the introduction of datacentres and development of an algorithm for traffic balance among them in a LTE-A network with a very high load, followed by an analysis of network capacity and datacentre load. This was achieved by upgrading an existing simulator for LTE DL implemented on a real urban network with appropriate user densities, according to the heterogeneity of the city of Lisbon. The obtained results show that, depending on the network configurations and average throughput per radio units, the number of required datacentres to cover the whole city will differ between one and three, and also proves that no latency problems will occur. Moreover, it was possible to see different datacentres load percentage, depending on their locations being closer to the centre area where more users are active. Furthermore, a theoretical analysis was made considering a larger region, which reduced the network capacity to 50%, leaving around 2 400 radio units without connectivity.

Keywords

LTE-Advanced, mobile clouds, load balancing, capacity, urban scenarios, Lisbon.

Resumo

As comunicações móveis sofreram um grande aumento de tráfego com a introdução do LTE-A, dando origem ao desenvolvimento de novas arquiteturas que separam as estações base em duas componentes, uma responsável por todas as funções de processamento e o outra pelas transmissões do sinal de rádio. Este trabalho consiste em introduzir centros de dados e desenvolver um algoritmo que faz uma distribuição equilibrada do tráfego numa rede muito carregada, baseada em LTE-A, de modo a analisar a capacidade da rede e carga dos centros de dados. De modo a atingir este objetivo, foram feitos melhoramentos num simulador já existente de LTE DL, implementado numa rede urbana real, com diferentes densidades de utilizadores de acordo com a heterogeneidade da cidade de Lisboa. Os resultados obtidos mostram que, dependendo da configuração da rede e do ritmo de dados obtidos por unidade de rádio, o número de centros de dados necessários para cobrir a cidade toda variam entre 1 e 3, provando também que não existiriam problemas de latência. Foi também possível verificar diferentes percentagens na carga dos centros de dados, dependendo das suas distâncias ao centro da cidade, onde se encontram mais utilizadores activos. Para finalizar, um estudo teórico foi feito, abrangendo uma área bastante maior, levando a uma redução na capacidade total da rede de 50%, deixando aproximadamente 2400 unidades de rádios sem conectividade.

Palavras-chave

LTE-Advanced, nuvens móveis, balanceamento da carga, capacidade, cenários urbanos, Lisboa.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AD/DA	Analogue to Digital/Digital to Analogue
AMBR	Aggregate Maximum Bit Rate
ARP	Allocation and Retention Priority
BBU	Baseband Unit
BS	Base Station
CA	Carrier Aggregation
CAPEX	Capital Expenditure
CoMP	Coordinated Multipoint Transmission/Reception
CP	Cyclic Prefix
CPRI	Common Public Radio Interface
C-RAN	Cloud – Radio Access Network
DL	Downlink
D-RAN	Decentralised – Radio Access Network
D-RoF	Digital Radio over Fibre
eNodeB	evolved Nodes B
EPC	Evolved Packet Core
EPS	Evolved Packet System
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
FDD	Frequency Division Duplex
FFT	Fast Fourier Transform
FST1	Frame Structure Type 1
FST2	Frame Structure Type 2
FTP	File Transfer Protocol
GBR	Guaranteed Bit Rate
GOPS	Giga Operations per Second
GPP	General Purpose Processors
GTP	GPRS Tunnelling Protocol
HARQ	Hybrid Automatic Repeat Request
HSS	Home Subscription Server
ICIC	Inter-Cell Interference Coordination
IFFT	Inverse Fast Fourier Transform

IP	Internet Protocol
ISI	Inter-Symbol Interference
LoS	Line of Sight
LTE-A	LTE-Advanced
MAC	Medium Access Control
MBR	Maximum Bit Rate
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity
MSS/BBU	Multi-site/standard BBU
NLoS	Non Line of Sight
O&M	Operation and Management
OBSAI	Open Base Station Architecture Initiative
OFDMA	Orthogonal Frequency Division Multiple Access
OPEX	Operational Expenditure
PAR	Peak to Average Ratio
PCRF	Policy and Charging Resource Function
PDCP	Packet Data Convergence Protocol
PDNG	Packet Data Network Gateway
PMIP	Proxy Mobile Internet Protocol
QAM	Quadrature Amplitude Modulation
QCI	QoS Class Identifier
QoS	Quality of Service
RAT	Radio Access Technologies
RB	Resource Block
RF	Radio Frequency
RLC	Radio Link Control
RRH	Remote Radio Head
RRM	Radio Resource Management
SAE	System Architecture Evolution
SC-FDMA	Single Carrier – Frequency Division Multiple Access
S-GW	Serving Gateway
SMS	Short Text Message
SNR	Signal to Noise Ratio
TA	Timing Advance
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
UE	User Equipment
UL	Uplink
UP	User Plane
VBS	Virtual Base Station

VO	Virtual Operators
VoIP	Voice over IP
WDM	Wavelength-division Multiplexing

List of Symbols

$\overline{N}_{RB}$	Average number of RBs per RRH
$\overline{N}_{users/RRH}$	Average number of users per RRH
$\bar{R}_{b,RB}$	Average throughput per RB
h_b	BS height
h_m	Height of mobile terminal
A_x	Complexity associated to each processing power function
B_{RB}	Bandwidth per RB
$C_{network}$	Network capacity
H_B	Buildings heights
L_0	Free space propagation loss
L_{bsh}	Loss between different heights of antennas and rooftops
L_{ori}	Street orientation loss
L_p	Path loss
L_{rm}	Loss between BSs and rooftops
L_{rt}	Loss between rooftop and mobile terminal
N_A	Number of antennas
$N_{bits/symb}$	Number of bits per symbol
N_{dc}	Number of datacentres
$N_{sub/RB}$	Number of subcarriers per RB
$N_{symb/sub}$	Number of symbols per subcarrier
$P_{BBU,N}$	Processing power per BBU
P_{CPRI}	Processing power due to CPRI functions
P_{DPD}	Digital pre-distortion processing power
$P_{FD,linear}$	Processing power for frequency domain functions, with linear dependency
$P_{FD,nonlinear}$	Processing power for frequency domain functions, with non-linear dependency
P_{Filter}	Up-down filtering processing power
P_{OFDM}	Processing power required by FTT and OFDM functions
$P_{PDCP-ciphering}$	Processing power due to PDCP ciphering functions
$P_{PDCP-compression}$	Processing power due to PDCP compression functions
$P_{PHYuser,DL}$	Processing power of the physical functions, for downlink

$P_{RLC-MAC}$	Processing power of RLC, MAC functions
$P_{S1-termination}$	Processing power required on functions related to the S1 interface
P_{fixed}	Constant processing power of a BBU-pool
$P_{pool,\%}$	Processing power of a BBU-pool in percentage
P_{pool}	Processing power of a BBU-pool
P_{pool}	Processing power of the BBU-pool
$P_{scheduler}$	Processing power due to scheduling functions
$RRH_{connected}$	Number of RRHs connected to datacentres
RRH_{total}	Number of total RRHs
$R_{b,RB}$	Throughput per RB
$R_{b,limit}$	Throughput limit for the datacentres
T_{RB}	Time interval of each RB
d_{max}	Maximum distance between RRHs and datacentres
k_a	Increase in path loss for BS antennas below rooftops
k_d	Dependence of multi-screen diffraction loss versus distance
k_f	Dependence of multi-screen diffraction loss versus frequency
w_B	Buildings separation distance
w_s	Width of the streets
δ_{CoMP}	Delay due to CoMP processing
δ_{max}	Maximum delay budget
$\delta_{optical}$	Delay in the optical link
$\delta_{processing}$	Delay due to processing
$\delta_{propagation}$	Delay due to propagation of the radio signal
δ_{switch}	Delay in switches devices
B	Bandwidth
C	Coding rate
M	Modulation used
d	Distance
df	Frequency domain duty-cycling
dt	Time domain duty-cycling
f	Frequency
ϕ	Angle of incidence of the signal into the buildings

List of Software

Microsoft Word	Text Editor software
Microsoft Excel	Calculation and Graphing software
Microsoft Visio	Graphics and diagrams editor
MapInfo	Geographic Information System software
Map Basic	Programming software and language to create additional functionalities for MapInfo
C++ Builder 6	C++ Compiler and Development software
Notepad++	C++ Development Tool software
Google Maps Area Calculator	Tool to calculate areas based on Google Maps
Google Earth	Tool used to analyse the file containing all the BSs in Portugal

Chapter 1

Introduction

This chapter introduces the topic of this thesis, by giving a brief overview on the evolution of telecommunications networks and by presenting the motivation that lead to this work, as well as the work structure.

1.1 Overview

Mobile communications brought a whole new perspective on the way that people communicate between each other. It started by working with very low functionalities and it has suffered an exponential growth, leading to the present world of mobile communication, where users have access to a wide range of services. This evolution brought a huge increase in the network traffic, shown in Figure 1.1, forcing operators to frequently expand and upgrade their networks, maintaining multiple-standards interoperability.

Standards-developing organisations, such as Third Generation Partnership Project (3GPP) played a very important role in the development of the systems, leading to LTE-Advanced (LTE-A), which is the latest generation of mobile communications systems. LTE-Advanced is standardised in Release 10 of 3GPP, which specifies peak data rates of 3 Gbit/s for Downlink (DL) and 1 Gbit/s for Uplink (UL) and introduces additional functionalities such as Uplink (UL) MIMO, Carrier Aggregation (CA), enhanced Inter-Cell Interference Coordination (ICIC) and relaying.

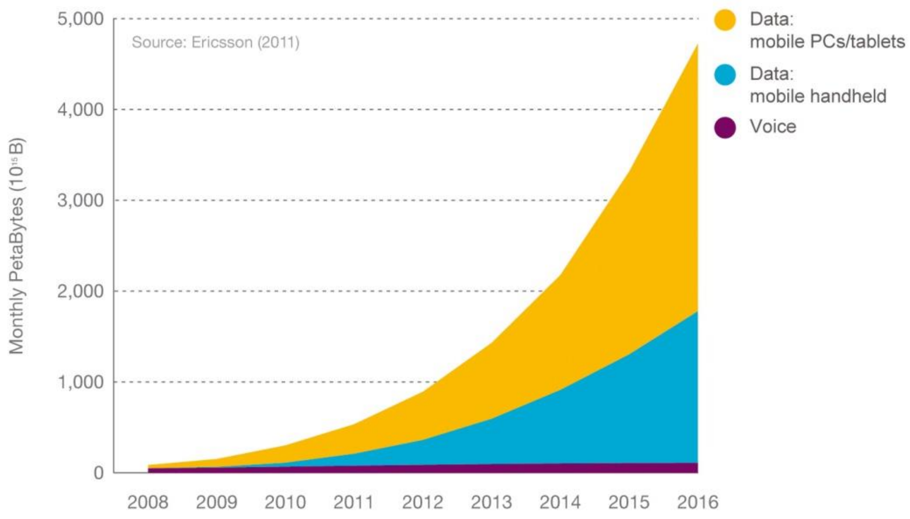


Figure 1.1: Traffic growth evolution (extracted from [Eric11]).

Mobile network architectures usually are split between a Radio Access Network (RAN) and a core network. RAN is the main asset for mobile operators to provide high data rates, good quality of service and to guarantee a 24×7 service, and thus it is where most of the upgrades have been made. In LTE-A, all the processing functions are concentrated on the eNodeBs, which are the base stations of the RAN. Therefore, costs for operators to build, upgrade and maintain their infrastructures increased significantly. On the other hand, the increase in revenue did not increase at the same rate, which is why operators have to find solutions in order to decrease their expenses without jeopardising the services provided.

The Total Cost of Ownership (TCO) is an estimation that determines the direct and indirect costs of any specific system. In this case, considering the TCO of mobile operators, regarding cell sites, it consists of the Capital Expenditure (CAPEX), which is associated with the network infrastructure build, and the Operational Expenditure (OPEX) which is associated with the network operation and management. Figure 1.2 and Figure 1.3 show an example of the CAPEX and OPEX per year of a cell site. As it can be seen, the cost of the site acquisition, civil works, etc., are more than 50% of the CAPEX, while more than 70% of the OPEX is spent on electricity and site rent.

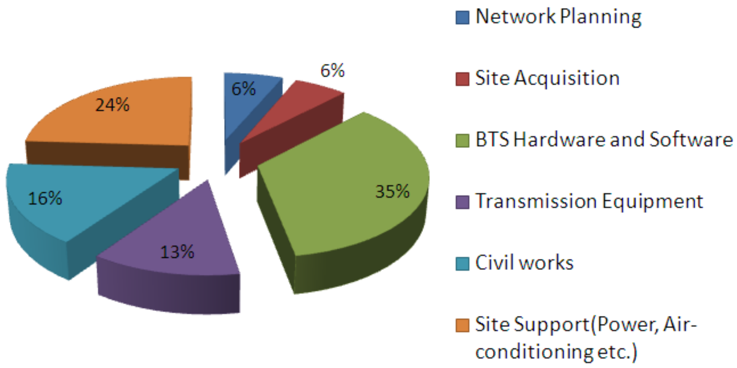


Figure 1.2: CAPEX of a Cell Site (extracted from [CMRI10]).

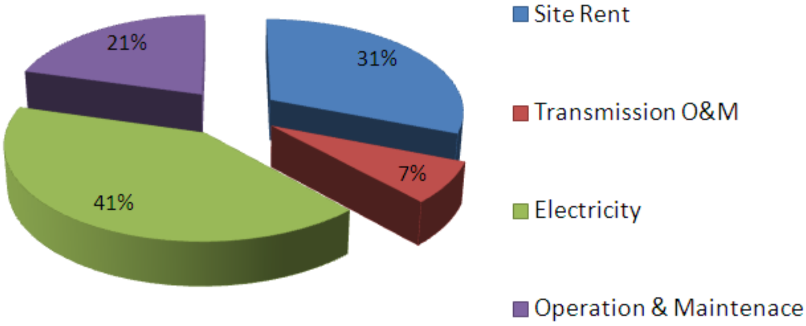


Figure 1.3: OPEX per year of a Cell Site (extracted from [CMRI10]).

In addition to the problems with the increased costs, the mobile network load is very dependent on time. This effect is shown on Figure 1.4, being called “tidal effect”. At the beginning of the day, most of the people move from the residential areas to the office ones, and at the end they go back to their residential areas. Consequently, there is a great amount of resources and power that is wasted at Base Stations (BSs) that are not busy at all. Operators have to guarantee a 24×7 service and so these idle BSs consume almost the same energy as if they were busy.

This lead to the concept of Mobile Cloud Networks, which transfer all the processing functions of the BSs to centralised locations, called datacentres or central offices. This would enable the implementation of new functions, allowing joint scheduling and processing in order to obtain better spectral efficiency, and thus coverage. Datacentres would run with open-platforms, supporting multi standards and allowing a smooth introduction of network upgrades and enhancements. Also, a significantly reduction in CAPEX and OPEX could be achieved, since mobile clouds could also provide a solution to balance the traffic load, according to the traffic expectation along the day.

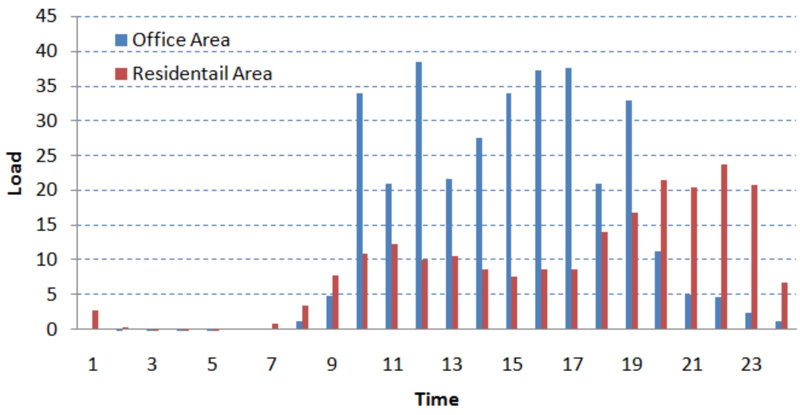


Figure 1.4: Mobile Network Load during one day (extracted from [CMRI10]).

Furthermore, mobile clouds could also be used to obtain additional revenues, since they can provide the RAN as a service. This means that operators can easily share the available infrastructures, and each one would be responsible for the appropriate processing at the datacentres. Consequently, new operators could join the market without huge investments, because they could simply pay for the rental of the RAN as a service and use the already deployed network. These are called Mobile Virtual Network Operators (MVNO). These MVNOs are the focus of [YrRu11] and are characterised into different types, as shown on Figure 1.5:

- True MVNO, which owns some network elements, such as the Home Local Register, switching and Intelligent Network platforms, and also provides billing and services;
- Weak MVNO, which does not own any of the network elements mentioned, and basically controls service marketing, branding and services;
- Reseller, which basically provides services by implementing marketing and branding.

This centralised architecture still needs further studies and field trials in order to correctly obtain the gains in terms of coverage, user experience and to quantify the reduction in expenses for the operators. Nonetheless, as more operators begin to notice the need for a cloud architecture, more studies are focusing on this area, bringing huge developments related to virtualisation techniques and different solutions for the current limitations.

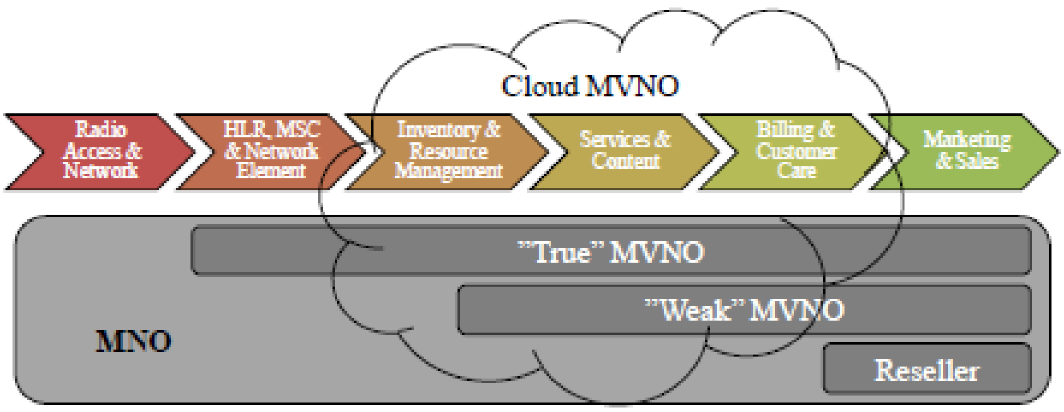


Figure 1.5: Types of MVNOs (extracted from [YrRu11]).

1.2 Motivation and Contents

The increase in network costs and in the total data traffic will be quite a limitation for Mobile Operators in the future, especially with the introduction of LTE-A. Consequently, it is necessary to innovate and upgrade network architectures, focusing mainly on the RAN, since it is the biggest expense. Green and centralised RANs are the solution, since they will be constituted of datacentres that use virtualisation techniques in order to simulate BSs, thus decreasing maintenance costs and improving some main aspects on the flow of signalling and controlling information.

The main scope of this thesis is to analyse the impact of introducing a similar centralised architecture in the region of Lisbon, by analysing some key performances, such as network capacity and connectivity problems, and to provide a load balancing algorithm that distributes traffic among the datacentres. In the end, the required number of datacentres required to cover the region under study is estimated. To accomplish these objectives, a model was made and implemented in a previously created simulator. The main results obtained were LTE parameters, although applied on a higher layer; in order to obtain results on the performance of the datacentres, the used scenarios consisted mainly of high load networks.

The principal output of this work is an upgrade of the previous model of the LTE simulator, with changes that allow the introduction of datacentres in the city, and their respective analysis in terms of capacity and connectivity problems.

This thesis consists of 5 chapters, including the present one, and a group of appendixes. Chapter 2 starts by introducing the LTE-A network architecture and radio interface, followed by a similar approach on the centralised architecture. It concludes with the state of the art, which presents the latest work developments on the subject of this thesis, and the characterisation of the different services and applications used in LTE-A.

Chapter 3 starts by presenting information regarding the developed models and the main parameters to be analysed. It is followed by an explanation of the simulator, and the changes made to implement the previous models. It ends with the results assessment in order to check the validity of the model and the simulator.

Chapter 4 consists of the analysis of the obtained results. It begins with the analysis of the reference scenario, being followed by a high load analysis, where some key parameters are changed in order to see their fluctuations and impact on the network. To conclude, a theoretical analysis is made using an area larger than the city of Lisbon. Along the results discussion, conclusions are presented and discussed.

Chapter 5 is a summary of this thesis, displaying the most important conclusions and results, also addressing some suggestions for future work on this topic.

At the end of the work, a group of appendixes are presented, in order to give auxiliary information. Appendix A provides the auction results from ANACOM, regarding the frequency bands in use. It is followed by Appendix B, which provides tables with the required complexities to calculate the processing power of a BS. Appendix C presents the equations used for the throughput calculations used in the simulator. Appendix D contains the user manual, explaining how to perform the simulations done in this thesis. Appendix E illustrates the datacentres locations, by showing the appropriate maps. Appendix F has an explanation of the COST 231 Walfisch-Ikegami path loss model. Appendix G presents the antennas radiation patterns, both for vertical and horizontal orientations. Finally, Appendix H has some maps that were used for the theoretical analysis performed in Section 4.3.

.

Chapter 2

Fundamental Concepts and State of the Art

This chapter provides an overview on LTE-A and Mobile Clouds, describing the main aspects relevant to this thesis. Sections 2.1 and 2.2 describe the radio interface and architecture for the LTE-A systems. It is followed by Section 2.3, which explains the basic concepts of C-RAN. To conclude, Sections 2.4 and 2.5 present the state of the art relative to Mobile Cloud Networks, and the services and applications characterisation, respectively.

2.1 Network Architecture

In this section, the basic network architecture for LTE-Advanced (LTE-A) is addressed, based on [HoTo09] and [3GPP12a].

The architecture in LTE-A is divided into four high level domains: User Equipment (UE), Evolved UMTS Terrestrial Radio Access Network (E-UTRAN), Evolved Packet Core (EPC) and the Services domain. This architecture can be seen in Figure 2.1.

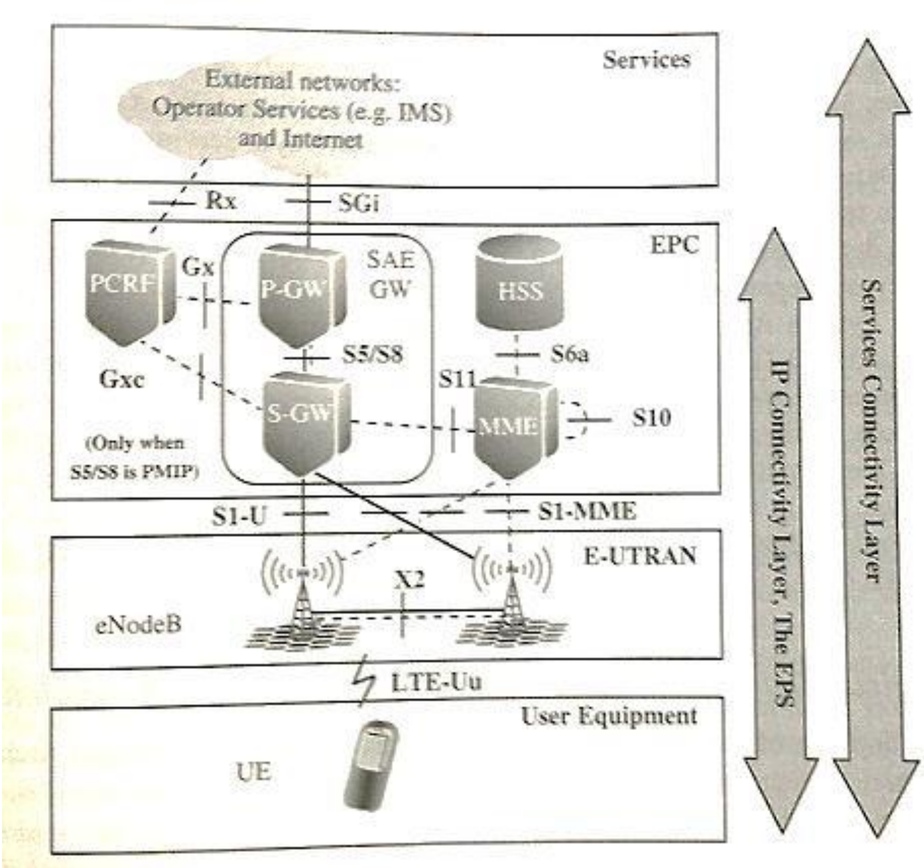


Figure 2.1. EPS architecture including connection to external networks (extracted from [HoTo09]).

The Evolved Packet System (EPS) was the result of the standardisation work by the 3GPP System Architecture Evolution (SAE), and it is completely optimised to provide Internet Protocol (IP) connectivity. All services are provided on top of IP, which eliminates the use of circuit switched nodes and interfaces. The EPS consists of the three first layers defined above: UE, E-UTRAN and EPC.

The E-UTRAN is formed by a mesh of evolved Nodes B (eNodeBs). The nodes are connected between themselves through the X2 interface, and with the EPC through the interfaces S1-U and S1-MME. The last interfaces refer to the connection between eNodeBs and the Serving Gateway (S-GW) and Mobility Management Entity (MME) respectively. The main functions of eNodeBs are:

- Radio Resource Management (RRM), which is responsible for functions such as radio bearer control, radio admission control, connection mobility control, and dynamic allocation of resources to UEs.
- Mobility Management, which is responsible for measurements and signalling information for handovers.
- To act as a bridge between the UE and the EPC, protection of the User Plane (UP) data and compression of IP headers.

The EPC is organised into five distinct nodes and respective functions:

- MME, which is the main control element and does not deal with the UP. It is responsible for authentication and security, mobility management and managing subscription profile and service connectivity. The MME uses the interfaces S10, S6a and S11 to connect with others MMEs, Home Subscription Server (HSS) and S-GW respectively.
- S-GW has a minor role in the control functions. It acts as a local mobility anchor during mobility, it relays the data between eNodeBs and the Packet Data Network Gateway (P-GW) and it is responsible for the allocation of its own resources based on requests from other nodes. The S-GW uses the interface S5/S8 based on GPRS Tunnelling Protocol (GTP) or Proxy Mobile Internet Protocol (PMIP) to connect with the P-GW, and in case PMIP is used it also has to connect with the Policy and Charging Resource Function (PCRF) through the Gxc interface.
- P-GW is the edge between the EPS and external data networks. Its main functionalities are IP allocation for the UE and filtering, charging and rate enforcement. The P-GW is connected to the PCRF through the Gx interface and to the external networks through the SGi interface.
- The PCRF is responsible for Policy and Charging Control (PCC) and decides how to handle services in terms of Quality of Service (QoS).
- HSS is a database that contains all permanent user data and subscriber profiles.

2.2 Radio Interface

This section addresses the radio interface for LTE-Advanced, and is based on [HoTo09] and [3GPP12b].

3GPP has specified 11 working bands for Time Division Duplex (TDD) and 25 ones for Frequency Division Duplex (FDD) for the E-UTRAN. The list of the available bands auctioned in Portugal is shown in Appendix A. In Europe, the available spectrum only includes the bands of 800, 900, 1800, 2100 and 2600 MHz. In Portugal, ANACOM, which is the regulatory authority for the communications sector, auctioned one extra band at 450 MHz.

LTE-A uses 2 different multiple access schemes: for Downlink (DL) it is Orthogonal Frequency Division Multiple Access (OFDMA) and for Uplink (UL) it is Single Carrier – Frequency Division

Multiple Access (SC-FDMA). Some of the main aspects that motivated the use of OFDMA were: good performance in frequency selective fading channels, low complexity for base-band receiver, good spectral properties, frequency domain scheduling, handling of multiple bandwidths, compatibility with advanced receiver and antenna technologies. OFDMA is implemented using the Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (IFFT) so that it is possible to process and analyse the signal in both the time and frequency domains. In OFDM, the centre frequency of a certain sub-carrier is chosen so that at the sampling instant all neighbouring sub-carriers have zero value. The spacing between sub-carriers in LTE-A is 15 kHz disregarding the total transmission bandwidth. LTE-A uses Cyclic Prefix (CP) in order to avoid Inter-Symbol Interference (ISI). It consists of copying the last part of the symbol to the beginning in order for the symbol to appear as periodic, and usually it is large enough to ensure that it exceeds the delay spread.

Due to the fact that LTE-A can work with both FDD and TDD, there are 2 type of frames used. Frame Structure Type 1 (FST1) is used for FDD while Frame Structure Type 2 (FST2) is used for TDD. This thesis focuses on FST1, since it is used for FDD, which is the most implemented method. FST1, shown on Figure 2.2, defines a maximum frame length of 10 ms, which is divided into 20 slots of 0.5 ms each. The scheduling is based on sub-frames, which are two consecutive slots. LTE-A benefits from the frequency domain scheduling, which is the dynamic allocation of resources in the frequency domain, based on Resource Blocks (RB). An RB consists of a group of 12 sub-carriers, which occupy 180 kHz in the frequency domain, and 7 or 6 OFDM symbols in the time domain, depending on if the normal CP or the extended one is being used, respectively. The minimum resource unit is the resource element that consists of one sub-carrier during one OFDM symbol, which implies that an RB has 84 (12 sub-carriers × 7 OFDM symbols) or 72 (12 sub-carriers × 6 OFDM symbols) resource elements per slot in the time domain. The channel bandwidth defined in LTE-A ranges from 1.4 MHz to 20 MHz, the corresponding number of resource blocks ranging from 6 to 100. The allocation of resources is shown in Figure 2.3. Still, the bandwidth can be extended by using Carrier Aggregation (CA).

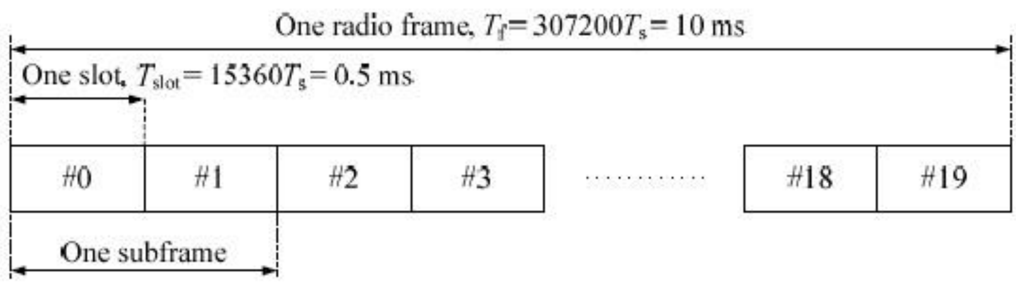


Figure 2.2 - Frame Structure Type 1 (extracted from [3GPP12b])

CA is one of the major improvements in LTE-A, in comparison with the previous versions of LTE. It can extend the channel bandwidth up to 100 MHz by aggregating up to five carriers, which are basically LTE Release 8 carriers that have a maximum bandwidth of 20 MHz. This aggregation may happen with carriers from the same band (Intra-band CA) or between different bands (Inter-band CA) as shown in Figure 2.4.

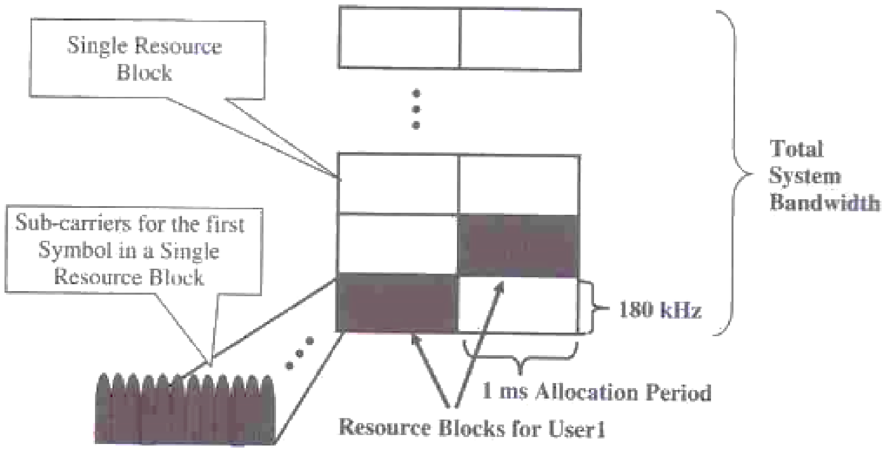


Figure 2.3. Resource allocation in OFDMA (extracted from [HoTo09]).

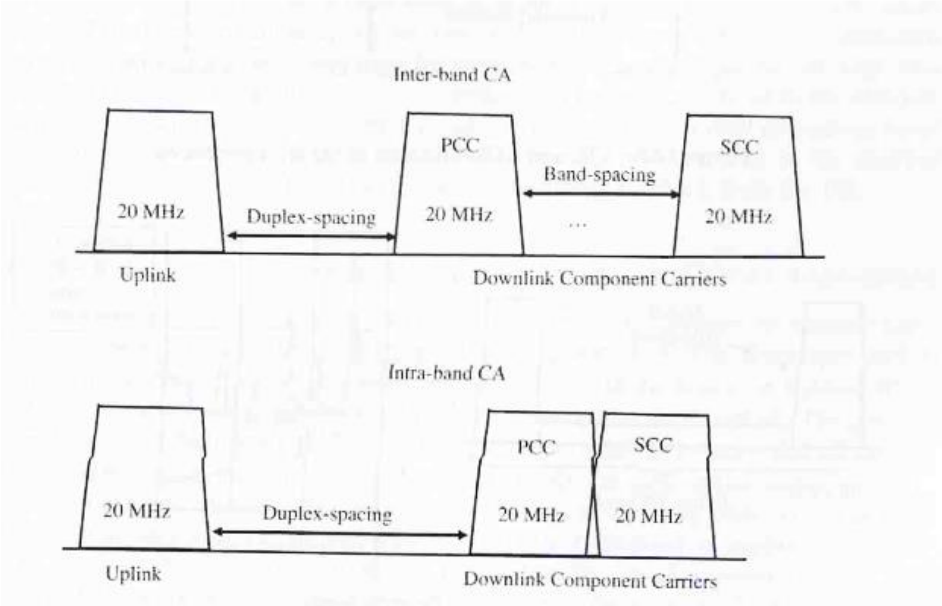


Figure 2.4. Inter-band CA and Intra-band CA (extracted from [HoTo09]).

Despite the advantages mentioned before, OFDMA has some challenges. One of the most important constrains is the high Peak to Average Ratio (PAR), because the signal is composed of several parallel sub-carriers, which correspond to multiple sinusoidal waves in the time domain, making the signal envelope to vary strongly. This strong variation creates some difficulties to the amplifier design, because it requires additional back-off, compared to a regular single carrier signal. Consequently, extra power is required, thus a faster consumption of the battery's energy exists. This is one of the reasons why the multiple access scheme used for the UL in LTE is SC-FDMA, which is more efficient in terms of battery energy.

SC-FDMA sends one symbol at a time, similarly to Time Division Multiple Access (TDMA). But, in contrast with a normal Quadrature Amplitude Modulation (QAM) modulator, the signal is generated in the frequency domain, adding the good OFDMA spectral proprieties. For SC-FDMA, the CP is added periodically instead of being added after each symbol, due to the increase in the symbol rate. In UL, the resource allocation is done similarly to OFDMA but the RBs are allocated to one user consecutively in the frequency domain, as shown in Figure 2.5. The maximum bandwidth that can be allocated is 20 MHz, but there has to be some margin for the guard bands, and so the useful channel bandwidth is smaller. For example, if the channel bandwidth is 20 MHz the occupied bandwidth is 19 MHz, which corresponds to 100 RBs.

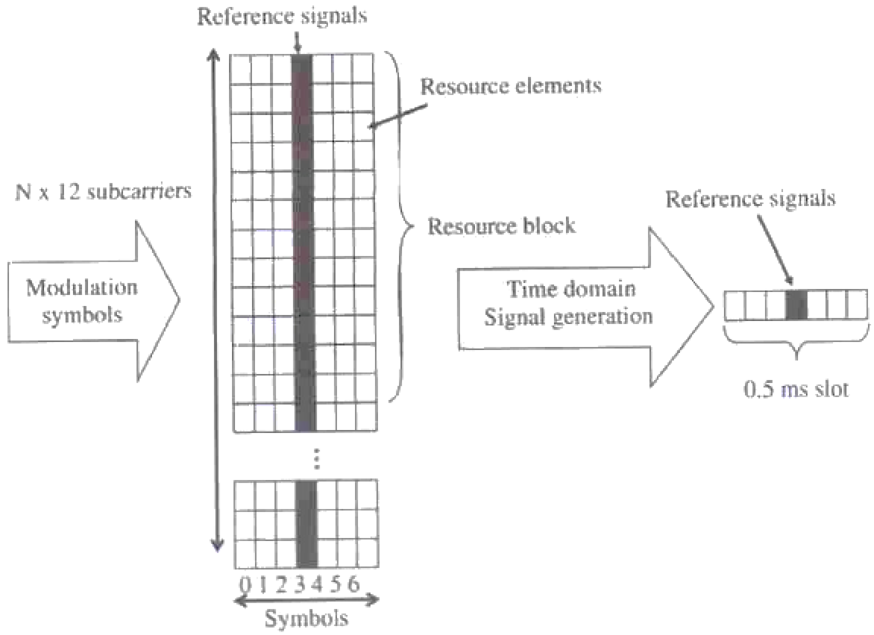


Figure 2.5. Resource allocation in SC-FDMA (extracted from [HoTo09]).

An important enhancement of LTE is Multiple Input Multiple Output (MIMO) operation, which is illustrated in Figure 2.6. MIMO introduces some important functions:

- Spatial multiplexing, where two or more antennas transmit signals, with different data streams, separated at the receiver and then processed, increasing the peak data rate by a factor of two or more, depending on the configuration. In LTE-A the available configurations for MIMO increased to 8×8 in DL and 4×4 in UL.
- Pre-coding of the signals transmitted from the different antennas in order to maximise the Signal to Noise Ratio (SNR) in the receiver.
- Transmit diversity, which corresponds to a transmission of the same signal through multiple antennas in order to exploit the gains from independent fading between the antennas.

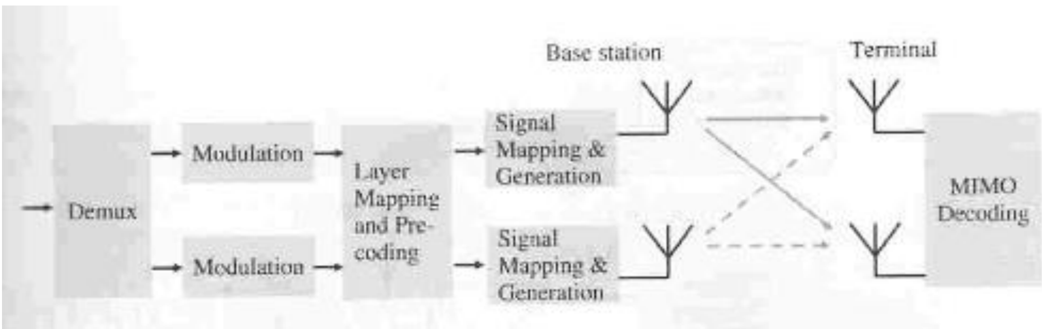


Figure 2.6. MIMO with a 2×2 configuration (extracted from [HoTo09]).

Another important feature of LTE-A is Coordinated Multipoint Transmission/Reception (CoMP), even though it is not standardised in Release 10. CoMP consists of coordinating geographically separated eNodeBs in order to provide joint scheduling and transmissions. This way, it can greatly improve signal reception/transmission and increase throughput at the cell edges. LTE-A also brings some improvements to Inter-Cell Interference Coordination (ICIC), which is also quite important due to the introduction of multi-layer Heterogeneous Networks.

2.3 Cloud – Radio Access Network Architecture

In order to improve capacity and reduce operator's expenses, a new centralised architecture has been proposed, Cloud – Radio Access Network (C-RAN), which is presented in Figure 2.7. Note that this architecture is just a prototype and there could be some variations in the future implementation.

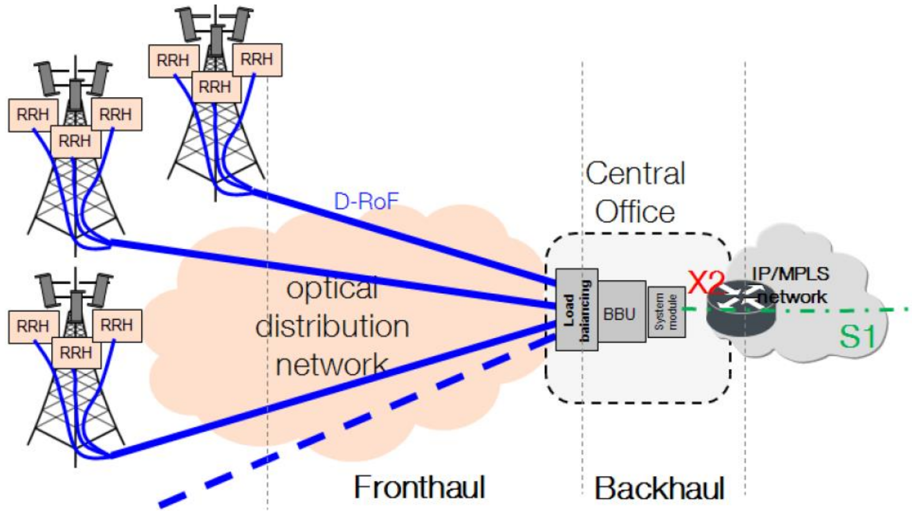


Figure 2.7: C-RAN architecture with resource pooling (extracted from [PiChCl]).

The main idea of C-RAN consists of the division of BSs into 2 different components, Remote Radio Heads (RRHs) and Baseband Units (BBUs). Depending on the architecture, these units may have different functions. In [CMRI10], 2 different architectures are defined, shown in Figure 2.8:

- Fully centralisation: BBUs are responsible for baseband processing as well as Layer 2 and Layer 3 functions (solution 1).
- Partial centralisation: BBUs are only responsible for Layer 2 and Layer 3 functions (solution 2).

Fully centralisation has higher requirements in the transport of the signal between the BBU and RRH, due to the high bandwidth of the baseband radio signal. Nonetheless, it makes it easier to upgrade the network, improves usage of joint processing, and has better capabilities to support multi-standards, so it will be the architecture upon which this thesis will focus on.

The RRHs, explained more in-depth in [PMJS12], deal with amplification of the Radio Frequency (RF) signal, up/down conversion, filtering, Analogue to Digital/Digital to Analogue (AD/DA) conversion, and interface adaptation, as shown in Figure 2.9.

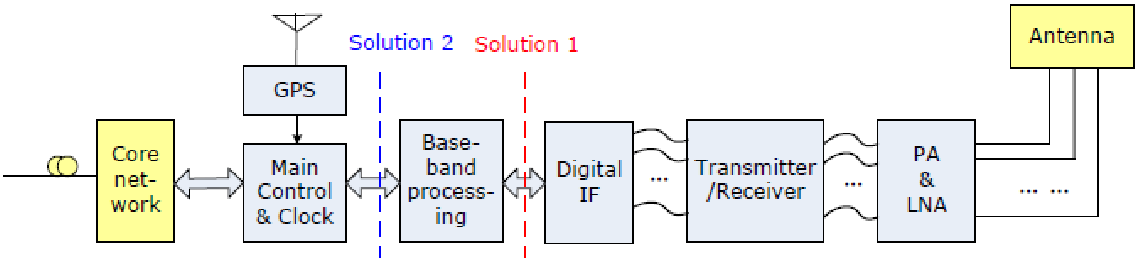


Figure 2.8: Split of functions between RRH and BBU for both architectures (extracted from [CMRI10]).

The connection between RRHs and BBUs is done through a low latency optical network, and the transmission can be done using 2 different standards, Common Public Radio Interface (CPRI) and Open Base Station Architecture Initiative (OBSAI), and they both use radio signal digitalisation, or Digital Radio over Fibre (D-RoF). This work focuses on CPRI and some of its parameters, such as latency and bit rate, that could create some constrains for the transmissions in C-RAN.

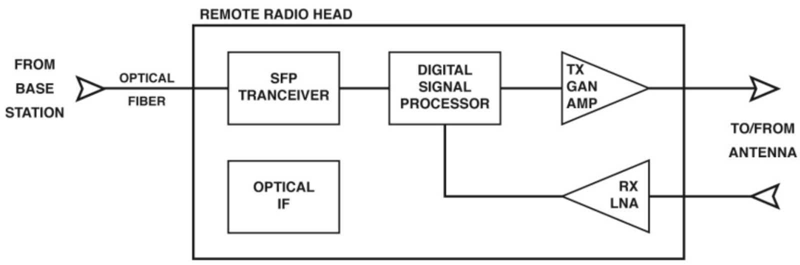


Figure 2.9 Components of a RRH (extracted from [PMJS12]).

The BBU consists of high performance programmable processors, which allow the use of real-time virtualisation techniques, and it is capable of supporting different Radio Access Technologies (RATs). It is responsible for cell broadcast, control, scheduling and Operation and Management (O&M) functions. Gioven the virtualisation of resources, a BBU pool can be created, which consists of many identical BBUs interconnected between them, centralised in one physical location. This enables the use of resource aggregation and pooling. The BBU pool requires a controller that is responsible for the instantiation of the virtual BBUs, and it should be able to control more than one BBU pool in order to improve joint processing, resource usage and spectral efficiency.

The only element left from the C-RAN architecture is the Load Balancer, explained in-depth in [Habe12], which acts as a router between the RRHs and the BBU pools, assigning the radio bearers to a selected BBU, depending on parameters such as latency and resource usage of the BBU. One example of how it can be used is shown on Figure 2.10.

The main advantages and objectives obtained by the C-RAN are:

- Energy efficient and friendly for the environment because, in comparison with the traditional architecture where BS's hardware is located in a cabinet close to the antenna, it will only be required to install the RRHs (which have a lower CAPEX and OPEX), while the processing equipment will be located in a Central Office, and probably in much less quantity, given the virtualisation. C-RAN will also allow for a reduction of the power consumption and the use of support equipment such as air conditioning.

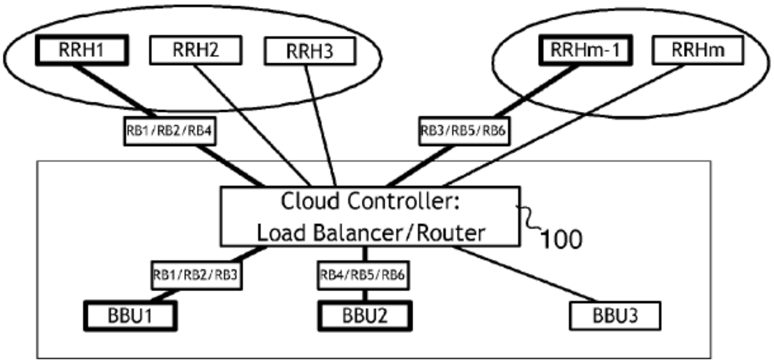


Figure 2.10: Load Balancer example (extracted from [Habe12]).

- Cost savings in terms of CAPEX and OPEX, mainly due to the reduction in the Operation & Management costs.
- Capacity improvement due to the virtualisation of BSs at the BBU pool. This also provides sharing of signalling, traffic data and channel information about the active UEs, thus, it is easier to implement joint processing and scheduling, improving the spectral efficiency and ICI.
- Adaptive to non-uniform traffic due to load-balancing capability in the distributed BBU pool.
- Multi-RAT support, which includes other legacy systems, i.e., GSM, UMTS and LTE.
- Enabling the use of some of LTE-A techniques like CoMP, ICIC and Heterogeneous Networks, and even other techniques that are not defined by 3GPP, like joint scheduling.

Although C-RAN has a lot of advantages in comparison with the traditional architecture, it also has some important constrains, such as:

- High bit rates needed at the fronthaul.
- Latency constrains at the fronthaul.
- Jitter and synchronisation.
- Baseband pool interconnection.
- BS virtualisation technology.

2.4 State of the Art

In [NGMN13], a more in-depth analysis of the architecture of the C-RAN can be found. In this paper, different deployment cases for C-RAN are analysed, such as Large/Medium Scale C-RAN, Small Scale C-RAN using legacy site room, Indoor Coverage, Super Hot Spot Coverage and Railway/Subway coverage. It also presents the key functionalities and possible solutions for the C-RAN implementation, for both the fronthaul and the architecture of the network elements.

As mentioned before, the bandwidth required for transmission at the fronthaul is quite big, thus it is important to address the possibility of data compression. In [GuCT12] an algorithm is suggested to compress data to be transported through the CPRI standard. This is done through spectral redundancy removal, block scaling and quantisation. It is mentioned that there has to be a trade-off between performance and consumption.

In [ZLZT11], the advantages of virtualisation of LTE air interface are investigated, with the use of realistic simulation models and mixed services of File Transfer Protocol (FTP) and Voice over IP (VoIP). The paper also proposes a generalised multi-party model to enable spectrum sharing between different Virtual Operators (VOs) and a mechanism for spectrum budget estimation based on real time services. The results obtained show a large multiplexing gain, as expected.

In [ZGWK11], two different ways to structure a Virtual Base Station (VBS) pool are proposed, as depicted in Figure 2.11 (Scenario 1 and 2). Both structures rely on multi-core General Purpose Processors (GPP). In Scenario 1, all VBS instances are deployed on top of the same GPP node, while in Scenario 2, 2 or more GPP nodes can be combined in order to process one single VBS instance (as an example, Figure 2.11 shows that the Physical layer is treated in one GPP node while the Medium Access Control (MAC) layer is treated on another GPP node). The latter scenario will provide better processing efficiency.

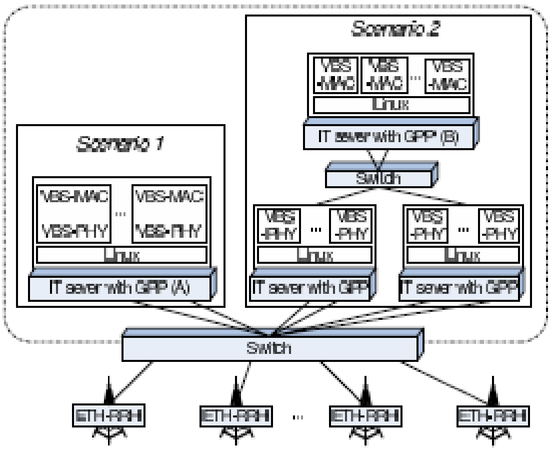


Figure 2.11: Two scenarios for a VBS pool (extracted from [ZGWK11]).

As mentioned above, cloud BSs facilitate the use of features such as CoMP. In [PRPK12], the mobility performance of applying CoMP is compared to non-CoMP systems based on three mobility metrics. The results show that a combination of the cloud architecture and CoMP joint processing scheme can greatly improve the mobility performance. Another study of CoMP in C-RAN is shown in [LZYH12], where the authors demonstrated the gains of CoMP in terms of throughput at the cell edge, in comparison to a normal RAN architecture.

In [LYWS12], the handover performance of the C-RAN is compared over a Decentralised–RAN (D-RAN) for all systems, i.e., GSM, UMTS and LTE. For LTE, X2-based inter-eNB handovers in C-RAN demonstrate a delay reduction, and to a large extent eliminate the risk of connection lost for the UE. On the other hand, S1-based handovers can rarely occur, since BBUs are co-located and in most cases there is a direct link between BBUs or BBU pools. The comparison between C-RAN and D-RAN is shown in Table 2.1.

Table 2.1: Comparison between handover performance between C-RAN and D-RAN (extracted from [LYWS12]).

Figures of merits	Performance evaluation
Handover delay	Decrease
Handover interrupt time	For the LTE-A UE with the good RACH procedure, the handover interrupt time may decrease
Handover failure rate	No influence
Number of handovers	The number of the handovers will be reduced to some extent if more RRHs share one cell.

In [Habe13], different solutions on where to split BSs are shown, presented in Figure 2.12 (each number represents where the partition should be made), and the advantages and constrains for each solution. The authors show a possible architecture for the Multi-site/standard BBU (MSS-BBU) and for the overall Mobile Cloud network. Some metrics are defined for the evaluation of the resource pooling and in the end simulations show a significant gain.

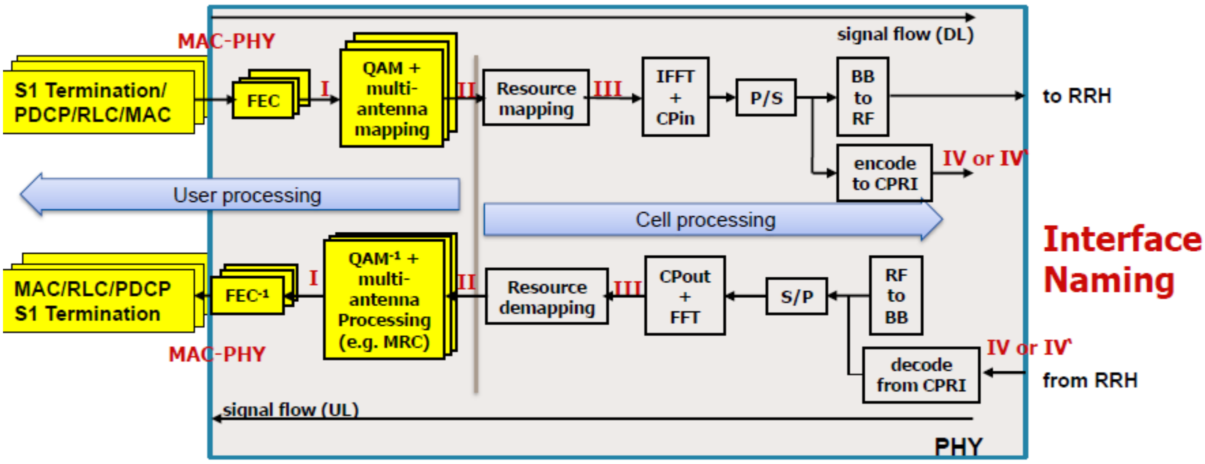


Figure 2.12: Different locations to split functionalities in a BS (extracted from [Habe13]).

RRHs must be associated with the BBUs depending on the traffic at a certain period of time. Thus, a good algorithm to describe the switching schemes between RRHs and BBUs is required. In [NaWK12], an example of an algorithm is made that explains how this switching is done, according to metrics such as processing power and RRH localisation. With the help of some simulations, it is concluded that the number of required BBUs could be reduced between 26% and 47%.

In [DDGF12], a model is proposed to quantify the required processing power of a BS. The total processing power could be calculated by summing the contributions from four different components: analogue part (RF signal), baseband processing, power amplifier functions and overhead, which consist of power systems and cooling. The contributions of each component to two different types of BSs are exemplified in Figure 2.13.

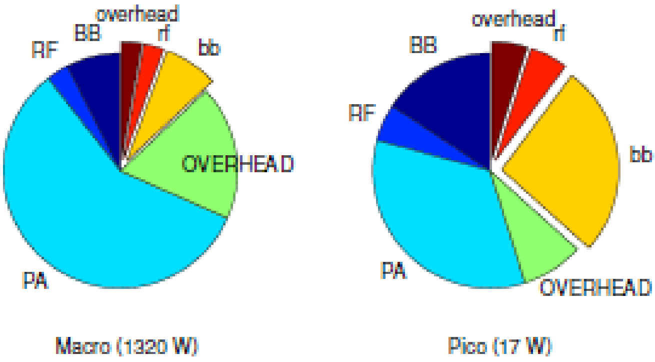


Figure 2.13: Contributions of each component for a macro- and pico-cell BS (extracted from [DDGF12]).

2.5 Services and Applications

The most popular service is voice, which is available for all UEs. Voice is quite predictable on its demanding performance from the network because the bit rate for voice is known and is quite constant along time. Therefore, operators focus on the capacity of providing access to the voice service with reasonably good quality, during a certain time interval. For this type of service, delay is the main factor that could jeopardise the good Quality of Service (QoS).

Nonetheless, data services have become much more important, with services such as Short Text Message (SMS), FTP, video streaming, online gaming, video conferences, web browsing, email, real time and interactive applications, etc. These services have very different requirements and specifications, such as bit rate, delay and duration, and involve a tremendous amount of traffic. To guarantee a good QoS, many quality requirements have to be monitored according to the specific service. Thus operators evolve in order to meet the requirements for the growth in the data services while guaranteeing a good QoS, while mobile equipment manufacturers have to improve their devices in order to provide the best experience for users, i.e., smartphones that are very data oriented.

Therefore, it is required to classify and distinguish the requirements for each type of service. 3GPP defined 4 classes of services for the previous system, UMTS. Conversational and Streaming were obviously the most important because they depend a lot on the delay. These include for example, telephony, VoIP, teleconference and music/video streaming, respectively. The other 2 classes are Interactive and Background with applications such as web browsing or download of e-mails on the background. These classes and their fundamental characteristics are shown in Table 2.2.

Table 2.2 UMTS QoS Classes (extracted from [3GPP12c]).

Traffic class	Conversational class conversational RT	Streaming class streaming RT	Interactive class Interactive best effort	Background Background best effort
Fundamental characteristics	- Preserve time relation (variation) between information entities of the stream -Conversational pattern (stringent and low delay)	- Preserve time relation (variation) between information entities of the stream	- Request response pattern - Preserve payload content	- Destination is not expecting the data within a certain time - Preserve payload content
Example of the application	- voice	- streaming video	- Web browsing	- background download of emails

The evolution of the network to a more data oriented one forced 3GPP to redefine the QoS parameters, which are:

- QoS Class Identifier (QCI): It is an index that identifies some attributes that are locally configured for each class. These attributes are resource type, delay budget, loss rate and priority. There are 9 standardised classes, shown in Table 2.3.
- Allocation and Retention Priority (ARP): Indicates the priority of a bearer in comparison to the others bearers.
- Maximum Bit Rate (MBR): Indicates the maximum bit rate for the bearer.
- Guaranteed Bit Rate (GBR): Indicates the bit rate that is guaranteed to the bearer.
- Aggregate Maximum Bit Rate (AMBR): Indicates the total maximum bit rate that a UE may have for all bearers in the same connections.

Table 2.3 QoS parameters for QCI (extracted from [3GPP12d]).

QCI	Resource Type	Priority	Delay Budget [ms]	Loss Rate	Example Application
1	GBR	2	100	1e-2	Conversational Voice
2		4	150	1e-3	Conversational Video (Live Streaming)
3		3	50	1e-3	Real Time Gaming
4		5	300	1e-6	Non-Conversational Video (Buffered Streaming)
5	Non-GBR	1	100	1e-6	IMS Signalling
6		6	300	1e-6	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, FTP, P2P file sharing, progressive video, etc.)
7		7	100	1e-3	Voice, Video (Live Streaming), Interactive Gaming
8		8	300	1e-6	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, FTP, P2P file sharing, progressive video, etc.)
9		9			

Chapter 3

Algorithm development

This chapter presents the developed model and gives an overview of the simulator, in order to assess the introduction of datacentres and to analyse the load balancing between them. Section 3.1 begins with a detailed explanation about the key parameters used to create the model. It is followed by Section 3.2, which presents the developed model using the mentioned parameters. Section 3.3 explains how the simulator works and how the model was implemented. To conclude, Section 3.4 gives the results assessment in order to check the validity of the implemented models and results.

3.1 Performance Parameters

3.1.1 Latency

Latency is one of the main constrains for the fronthaul in the centralised architecture. This parameter determines what can be the maximum distance for the optical link between the RRH and the BBU. Figure 3.1 shows the different latency requirements for most of the procedures in LTE-A. These procedures are independent, so it is only required to look at the most critical one, which is the UL synchronisation Hybrid Automatic Repeat Request (HARQ), as shown in Figure 3.1. This procedure is used to correct transmission errors, having a target delay of 3 ms. If this value is exceeded, it may have negative impacts, such as performance loss and reduced peak data rates per user.

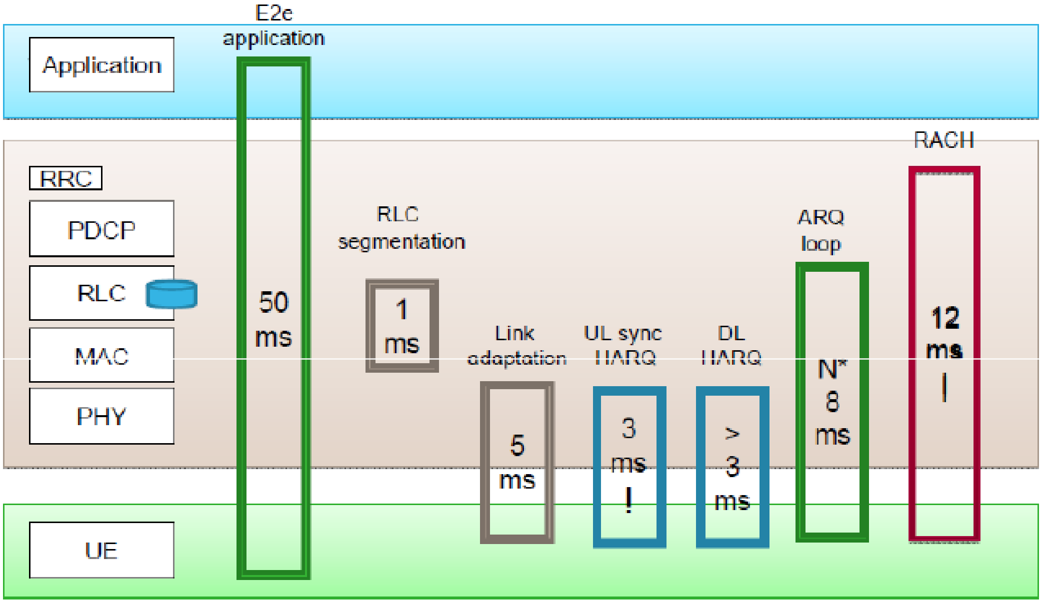


Figure 3.1 Delay constrains for the different procedures of LTE-A (extracted from [Pizz12]).

Another parameter that has to be taken into account is Timing Advance (TA), which is a small time interval that compensates for the propagation time between the UE and the BS, in order to guarantee that signals from users at the cell border arrive at the RRH at the same time as signals from users in the centre of the cell. The maximum cell coverage defined by 3GPP is 100 km, which corresponds to a TA of 667μs. Even though most cells have a radius much lower than the 100 km, the TA is always used assuming that value, in order to have an increased margin for the other delays. This implies that the maximum delay budget for the round trip is:

$$\delta_{max} = 3.667\text{ms} \tag{3.1}$$

The round trip processing delay for the BBU+RRH is assumed to be (this value was estimated by Alcatel-Lucent and Ericsson and is included in [Pizz12], although it is not experimentally verified):

$$\delta_{processing} \approx 2.8\text{ms} \quad (3.2)$$

The round trip propagation time between the UE and the RRH is given by:

$$\delta_{propagation} = \frac{d}{c} \times 2 \times 10^3\text{ms} \quad (3.3)$$

It is still required to consider CoMP processing time and switch delay, which are given by the following equations:

$$\delta_{CoMP} \approx 0.5ms \quad (3.4)$$

$$\delta_{switch} \approx 0.02 \times 2\text{ms} \quad (3.5)$$

Thus, the time left for the propagation in the optical link and its respective length are given by (3.6) and (3.7), respectively:

$$\delta_{optical} \approx \delta_{max} - (\delta_{processing} + \delta_{propagation} + \delta_{CoMP} + \delta_{switch})[\text{ms}] \quad (3.6)$$

$$d_{max} = \frac{\delta_{optical}[\text{ms}] \times 10^{-3} \times v_{fiber}[\text{km/s}]}{2}[\text{km}] \quad (3.7)$$

These parameters are susceptible to variations, and so it is better to assume that the maximum distance is 10% or 15% lower than the value given by the last equation. For a UE located 1 km away from the RRH, and assuming that the velocity in the fibre is 200 000 km/s, the available fibre length is around 25 km. Nevertheless, it is important to mention that some delays were left out of the equations, which could significantly reduce this value. Some examples of extra delays not considered are the delays caused by active equipments for transmission, such as the ones used for Wavelength Division Multiplexing (WDM) systems and equipment for CPRI compression.

3.1.2 Processing Power

The BBU pool is implemented on physical machines, which enable the use of virtualisation techniques in order to create multiple instances of a VBS in a single hardware machine. In order to achieve an efficient load balancing of the traffic, the processing power has to be quantified.

The total processing power that is used by the BBU-pool is measured in Giga operations per second (Gops) and can be expressed as:

$$P_{pool}[\text{Gops}] = \sum_{i=1}^{N} P_{BBU,N} + P_{fixed} \quad (3.8)$$

where:

- $P_{BBU,N}$ represents the required processing power per BBU,

- P_{fixed} represents a fixed processing power required for scheduling and signalling functions, which does not depend on the number of BBUs.

In terms of percentage it can be written as:

$$P_{pool,\%} = \frac{P_{pool} \times 100}{P_{pool,max}} \tag{3.9}$$

In [DDGF12], a method to calculate the required processing power of a BS is proposed, for both DL and UL. The total processing power required by a BBU is shown in (3.10), and it involves the sum of all the processing effort required for the BBU to compute, before transmitting, the signal to the core network. This includes the processing power for the Packet Data Convergence Protocol (PDCP), which has the functions of compression/decompression and ciphering/deciphering, the Radio Link Control (RLC) and MAC functions, the scheduler functions, the S1 termination functions and the physical layer functions.

$$P_{total} = P_{scheduler} + P_{RLC-MAC} + P_{PDCP-ciphering} + P_{S1-termination} + P_{PDCP-compression} + P_{PHYuser} \tag{3.10}$$

where

- $P_{scheduler}$ is a function of the number of radio bearers and the resource allocation algorithm.
- $P_{RLC-MAC}$ is a function of the S1 data rate and the RLC mode used to perform the data transfer.
- $P_{PDCP-ciphering}$ is a function of the S1 data rate and the type of security configured by the upper layers.
- $P_{PDCP-compression}$ is a function of the S1 data rate and the type of header compression used.
- $P_{S1-termination}$ is a function of the S1 data rate.

This thesis focuses on the physical functions, because they can be quantified using [DDGF12]. It is important to note that it will be assumed that the BBU will have all the baseband functions, while the RRH is left with the radio functions. Also, the processing power required for power amplification and overhead was not considered, since it is not relevant for the purpose of this thesis.

The processing power required by the physical functions for the DL case is given by the following expression (note that for the UL case some functions may not exist):

$$P_{PHYuser,DL}[\text{Gops}] = A_1 P_{DPD} + A_2 P_{Filter} + A_3 P_{CPRI} + A_4 P_{OFDM} + A_5 P_{FD,lin} + A_6 P_{FD,nlin} + A_7 P_{FEC} \tag{3.11}$$

where:

- P_{DPD} – Digital Pre-Distortion.
- P_{Filter} – Up/down sampling and filtering.
- P_{CPRI} – Link to backbone network.
- P_{OFDM} – FTT and OFDM specific functions.
- $P_{FD,linear}$ – Frequency domain functions with a linear dependency.

- $P_{FD,nonlinear}$ – Frequency domain functions that have a non-linear dependency.
- P_{FEC} – Forward Error Correction functions.
- A_x- Corresponds to the complexity associated with each function, given in Giga Operations per Second (GOPS). The value for each complexity depends on the type of BS cell (macro, micro, pico, femto) and on the flow of information (DL or UL). The tables with all the values can be found in Appendix B.

Each processing power can be calculated using the following equations:

$$P_{DPD} = P_{Filter} = P_{OFDM} = \frac{B}{B_{ref}} \times \frac{N_A}{N_{A,ref}} \times \frac{dt}{dt_{ref}} \tag{3.12}$$

$$P_{CPRI} = P_{FEC} = \frac{M}{M_{ref}} \times \frac{B}{B_{ref}} \times \frac{C}{C_{ref}} \times \frac{N_A}{N_{A,ref}} \times \frac{dt}{dt_{ref}} \times \frac{df}{df_{ref}} \tag{3.13}$$

$$P_{FD,linear} = \frac{B}{B_{ref}} \times \frac{N_A}{N_{A,ref}} \times \frac{dt}{dt_{ref}} \times \frac{df}{df_{ref}} \tag{3.14}$$

$$P_{FD,nonlinear} = \frac{B}{B_{ref}} \times (\frac{N_A}{N_{A,ref}})^2 \times \frac{dt}{dt_{ref}} \times \frac{df}{df_{ref}} \tag{3.15}$$

where:

- B – Bandwidth [MHz],
- N_A – Number of antennas,
- C– Coding Rate,
- M – Modulation used [bits per symbol],
- dt – time-domain duty cycling [%],
- df – frequency-domain duty cycling [%].

The time-domain duty cycling is basically the fraction of time on which the RRH is active. In this thesis, it is always considered as 100%, unless the RRH does not have any RBs allocated to users, and in that case it is considered 0% since it is not transmitting. On the other hand, the frequency-domain duty cycling is the fraction of RBs used in comparison to the maximum allowed by the RRH bandwidth. Also note that the modulation and code rate used are not defined parameters for the simulation but they depend on the channel quality and are chosen appropriately.

The reference scenario used for the equations above correspond to a scenario where the network is fully loaded (dt = 100% and df= 100%), using a single antenna, a modulation scheme of 64-QAM, coding rate of 1, and a bandwidth of 20 MHz. As an example, using 2×2 MIMO, with a frequency duty cycle of 50%, code rate of 1/3, 60 MHz bandwidth and 64-QAM modulation, the obtained processing power for a macro BS is around 3200 Gops.

3.1.3 RRH and Datacentre Positioning

The localisation of the RRHs is also quite significant for the load balancing. It is much more efficient for neighbour RRHs to be processed in the same BBU-pool, since an external communication between different BBU-pools will not be required, greatly improving handovers and joint scheduling and processing. For the operators this is quite easy to implement, since the RRHs locations are well known, and thus it is possible to create neighbours lists for each RRH, containing specific information regarding which RRHs are closest.

For the purpose of this thesis, it would be too demanding, in terms of computational resources, to calculate each RRH neighbours, and create those previously mentioned neighbour lists. Instead, when a RRH wants to connect to a datacentre one considers a circular area, with a radius of 500 m, and an analysis is made in order to count how many RRHs are connected to which datacentre, thus discovering which datacentre contains more neighbours. This value for the distance was chosen because the area under study has a very high density of RRHs and a maximum radius in the order of 5 km, thus meaning that this value represents an accurate average distance between RRHs.

The datacentres locations are also important for the distribution of the traffic among them. Since latency is the main constrain affecting the connectivity between RRHs and the datacentres, it is expectable that the RRHs will try to connect to the nearest available datacentre. Consequently, one expects a higher load in datacentres located nearest the densest areas, where the data traffic and number of neighbours for each RRH is higher. This implies that the datacentres locations have to be carefully planned in order to serve all the RRHs in a given area.

3.1.4 Capacity

For the purpose of this thesis, network capacity is measured at the RRH level instead of the user one. Basically, capacity is the sum of all active RRHs that are correctly connected to a BBU-pool. It is expressed as:

$$C_{network} = \frac{RRH_{connected}}{RRH_{total}} \tag{3.16}$$

where:

- $RRH_{connected}$ - Number of RRHs that are associated to a BBU-pool.
- RRH_{total} - Number of total RRHs in the network.

The term capacity can also be related to the required number of datacentres that are necessary to serve a specific region. Consequently, it is required to obtain an estimation of that value, which can be calculated using (3.18) or (3.19), depending on throughput and processing power, respectively:

$$N_{dc} = \frac{\overline{N}_{users/RRH} \times \overline{R}_{b,RB} \times \overline{N}_{RB} \times RRH_{total}}{R_{b,limit}} \tag{3.18}$$

$$N_{dc} = \frac{RRH_{total} \times \bar{P}_{power,RRH}}{P_{power,limit}}$$

where:

- N_{dc} – number of required datacentres,
- $\bar{N}_{users/RRH}$ – average number of users per RRH,
- $\bar{R}_{b,RB}$ – average throughput per RB,
- $\bar{N}_{RB}$ – average number of RBs used per RRH,
- $R_{b,limit}$ – throughput limit for each datacentre,
- $\bar{P}_{power,RRH}$ – average processing power per RRH,
- $P_{power,limit}$ – processing power limit for each datacentre.

It is important to refer that there is no information available on the datacentres processors limitations. Thus, it was not possible to take limitations relative to each datacentre maximum processing power, meaning that the limitation had to be made based on the throughput. The value for the limitation was based on some theoretical calculations, where it was assumed an average of 600 active RRHs, each one with an average throughput of 300 Mbit/s. The resulting value might be higher than expected, but since LTE-A already defines bandwidths of 100 MHz and 8×8 MIMO, it is considered an appropriate approximation. Since the chosen reference scenario consists of 3 deployed datacentres, the reference value taken for the limitation was 60 Gbit/s. Note that, in this work, it is not considered that a datacentre can have more than one BBU-pool, therefore each datacentre is considered as one big BBU-pool.

3.1.5 Throughput

Throughput represents the data rate transmitted to a user and can be theoretically obtained using the following expressions:

$$R_{b,RB}[\text{bit/s}] = \frac{B_{RB} \times N_{sub/RB} \times N_{symb/sub} \times N_{bits/symb}}{T_{RB}} \tag{3.19}$$

where:

- B_{RB} – Bandwidth per RB,
- $N_{sub/RB}$ – Number of sub-carriers per RB,
- $N_{symb/sub}$ – Number of symbols per RB,
- $N_{bits/symb}$ – Number of bits per symbol,
- T_{RB} – Time interval of the RBs.

This result will only represents the throughput per RB, thus it is required to sum all the contributions, from each RB, to get the final throughput per user:

$$R_{b,user} = \sum_{i=1}^{N_{RB,user}} R_{b,RB} \tag{3.20}$$

Note that this method of calculating the throughput is not very accurate, because it does not take into account the received power by the user nor the interference. Furthermore, gains due to MIMO or antennas orientation are also not taken into account. Basically it would mean that users at difference distances and scenarios, which correspond to different received powers, would obtain the same throughput. That is why this rough approximation is not used in this thesis. The equations used are derived from graphs that correlate the SNR with the obtained throughput, based on measures done by manufacturers. These equations are further explained in [Alme13] and are presented in Appendix C.

3.2 Model Development

The proposed model has the objective of efficiently balancing the load of each RRH throughout the available BBU-pools, in order to address problems such as the imbalance of traffic along the day.

The algorithm starts by randomly selecting a target RRH that needs to be aggregated to a specific BBU-pool in order to create a VBS. Using the previous latency (3.7), and assuming that a user is located at the cell edge so that it corresponds to the worst case scenario in terms of latency, calculations are done to find the maximum distance available for the optical link. If the maximum distance obtained is lower than the distance of the available fibre between the RRH and any available datacentre (where all the available BBU-pools are located) there will be problems due to the latency. This process is presented in the flowchart illustrated in Figure 3.2. To avoid the latency problems, the following equation has to be verified:

$$d_{max} > d_{fibre\_RRH-BBU} \tag{3.11}$$

After this first step, the target RRH has to be allocated to one specific BBU-Pool. If there is only one active BBU-Pool, there will be an analysis of the processing power currently in use by that BBU-Pool. Furthermore, it will be observed if the addition of the processing power required by the new RRH makes the overall processing power exceed the maximum available value.

$$P_{pool} + P_{RRH} < P_{pool,max} \tag{3.12}$$

On the other hand, if there is already more than one BBU-Pool activated, both parameters have to be taken into account, the RRH positioning and the processing power of the BBU-Pools. In this case, the target RRH will check if any of the active BBU-Pools have associated some of its neighbours, which, as mentioned in Section 3.1.3, consists of checking if the distances between the RRHs are below 500 m.

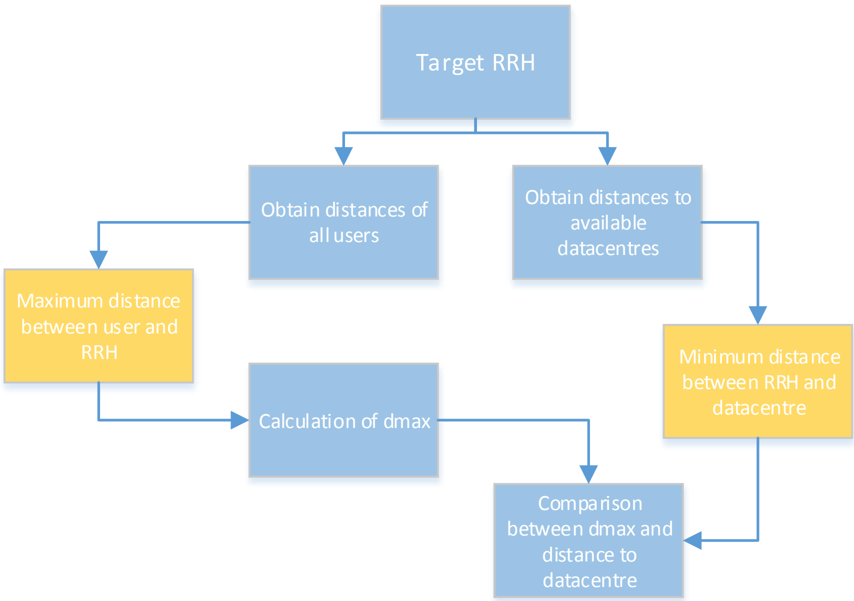


Figure 3.2: Flowchart with the method used to check for latency constrains.

The used method is illustrated in Figure 3.3. If there are neighbours on the BBU-Pool, the above condition has to be verified again to guarantee that the processing power does not exceed the limit. In case there are no neighbours the BBU-Pool is chosen in order to balance the processing powers of the active BBU-Pools.

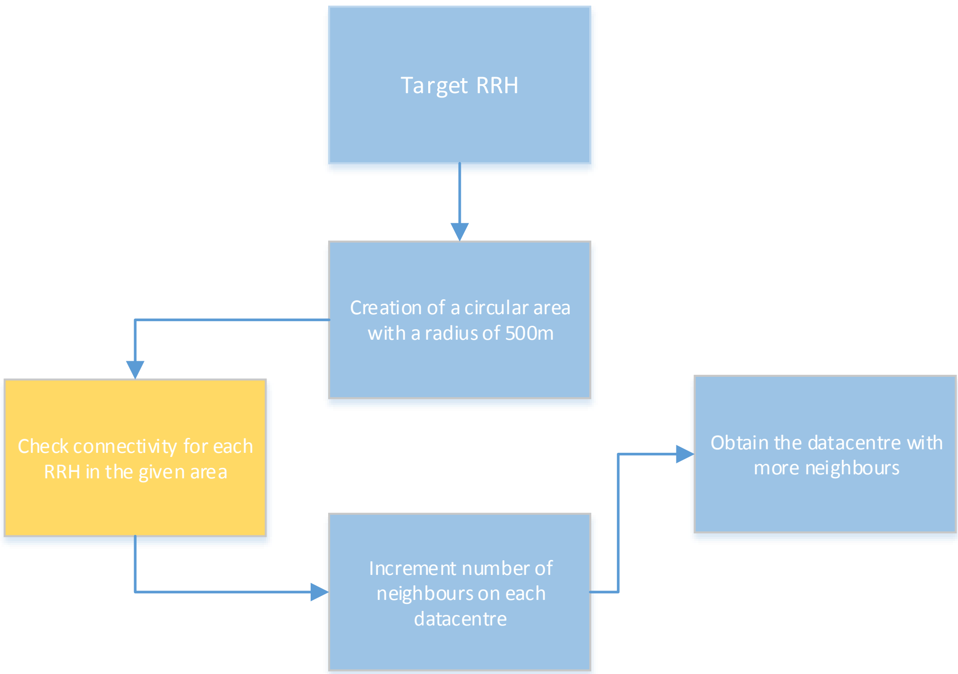


Figure 3.3: Flowchart illustrating the neighbour calculation.

Note that since the used simulator has no temporal variation, it was not possible to analyse the changes happening in the connections between the datacentres and RRHs along time. In order to analyse this factor, a first aggregation between the RRHs and datacentres is made, where RRHs connect to the closest datacentre, making it possible for changes to occur if the network traffic is very high, forcing some RRHs to change their connection to datacentres with lower loads.

A flowchart is presented in Figure 3.4 explaining the overall structure of the algorithm.

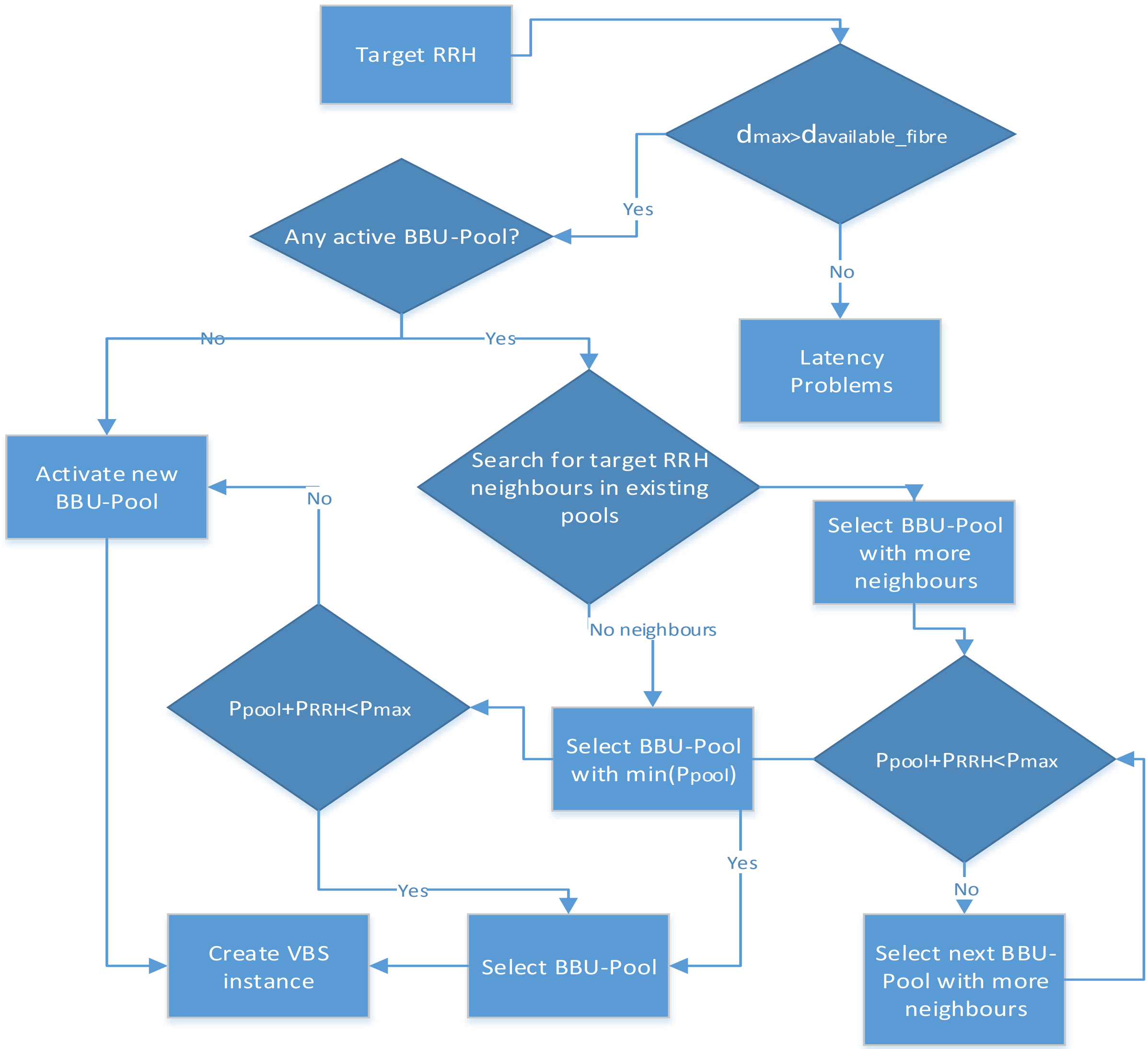


Figure 3.4: Algorithm Flowchart

3.3 LTE Simulator

In this section a brief explanation of the simulator is presented, followed by the introduction of this thesis algorithm on the simulator.

3.3.1 Overview

The simulator has the objective of analysing the performance of an LTE network, deployed over a dense urban area cellular network, in this case the city of Lisbon. It is the result of several previous studies, namely the master thesis from [Card11], [Duar08], [LaCo06], [Lope08], [Pires12], [Salv08] and [SeCa04], and it is composed of 3 main modules, as shown in Figure 3.5:

* User’s generation (SIM3).
* Network Deployment (UMTS_Simul&Net_Opt).
* LTE DL ICI Analysis (LTEDL_Stat).

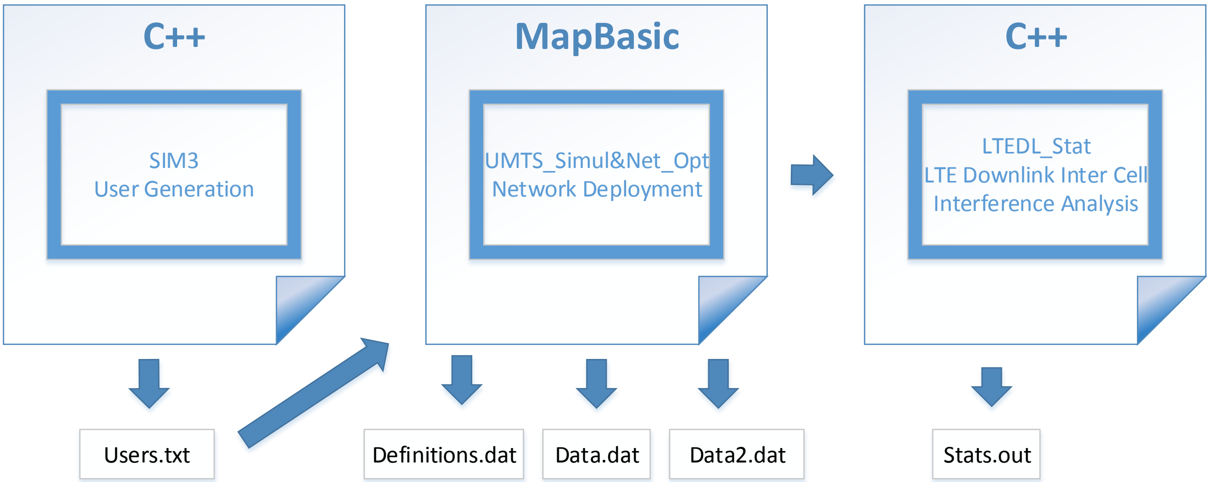


Figure 3.5: Simulator Overview

The user’s generation module, described in detail in [Lope08] and [Salv08], is responsible for the creation of a file with information about the users: user’s id, latitude and longitude, attenuation, service and type of user. These users are created randomly, both for their location as well as the type of service, according to a given service penetration percentage. This module was used to obtain the user files used throughout the simulations, as described in Section 4.1.

The network deployment module, described in detail in [LaCo06], receives the file with the information of the users and places them in the corresponding locations, followed by the positioning of the BSs. During the deployment, calculations are made in order to check the coverage of each BS and to check which BSs can serve each user. Note that the users that are not covered will not be analysed in the next module and do not appear in the output files. Some modifications were made to this module, using the Map Basic programming tool, in order to allow the configuration of additional network parameters and introduction of datacentres on the Map Info executable, called UMTS_Simul. This module concludes with the creation of 3 files, Definitions.dat, Data.dat and Data2.dat. The first two were already implemented, but the third file was created for this thesis, in order to write the information relative to the available datacentres, which consists basically of their id, latitude and longitude.

Finally, the last module, developed by [Pires12], receives the previous output files from the Network Deployment module, being responsible for the analysis of all LTE-A functionalities. It retrieves the information about the users and the BSs, and simulates their access to the network, based on a snapshot. It is responsible for the RBs distribution along the users, the calculation of their respective throughputs, and the ICI calculations in order to get the real throughputs for each user. In the end, outputs files are created in order to show different stats about the network. This module is also responsible for the implementation of this thesis’ models, which analyse the deployment of datacentres and their aggregation with the BSs.

The details and instructions on how to perform the simulation are described in Appendix D.

3.3.2 Algorithm Implementation

As mentioned above, the LTE DL ICI Analysis module described in [Pires12], has the objective of analysing the performance of an LTE network, using a snapshot approach, by obtaining results with information regarding capacity, user's throughput, interference and some other key parameters. In this section, the changes made to the module are explained in order to understand how the models in this thesis are implemented. Note that it is assumed that each sector of a BS corresponds to an RRH, so when an RRH is mentioned, in the simulator it corresponds to a sector.

The analysis begins by associating users to the closest RRH and distributing RBs among them, according to the requested services, taking into account the previously defined QoS priorities and requirements for each service. In this work, the only service used was FTP, meaning that the distributing of RBs was very similar to all users, depending only on the load of the RRH. Since no interference was considered, after the RBs distribution, the aggregation between RRHs and datacentres occurs. The analysis concludes with a calculation of the appropriate statistics and generation of the output files. The overall workflow of the simulator is presented in Figure 3.6, including the used algorithms along the process. Note that the scheduling and reduction algorithms were already implemented previously and are not within the scope of this thesis.

There are many aspects of radio and strategy parameters that can be used to tune settings for the simulations. Some of these parameters are:

- Bandwidth,
- MIMO configuration,
- Frequency band,
- Transmission Power,
- BS and MT gains,
- User and cable losses,
- Noise figure,
- Frequency Reuse Schemes,
- Strategy for RB reduction,
- Scheduling Algorithm,
- Antennas down tilt,
- Interference Analysis switch,
- Scenario parameters (Street widths, BS height, etc.),
- Datacentres switch,
- Latency parameters.
- Datacentre throughput limit.

The number of users, their services and the penetration percentage for each service are defined by the Users Generation module. For this thesis, the most important parameters are the datacentre switch, the latency parameters, bandwidth, and MIMO configuration.

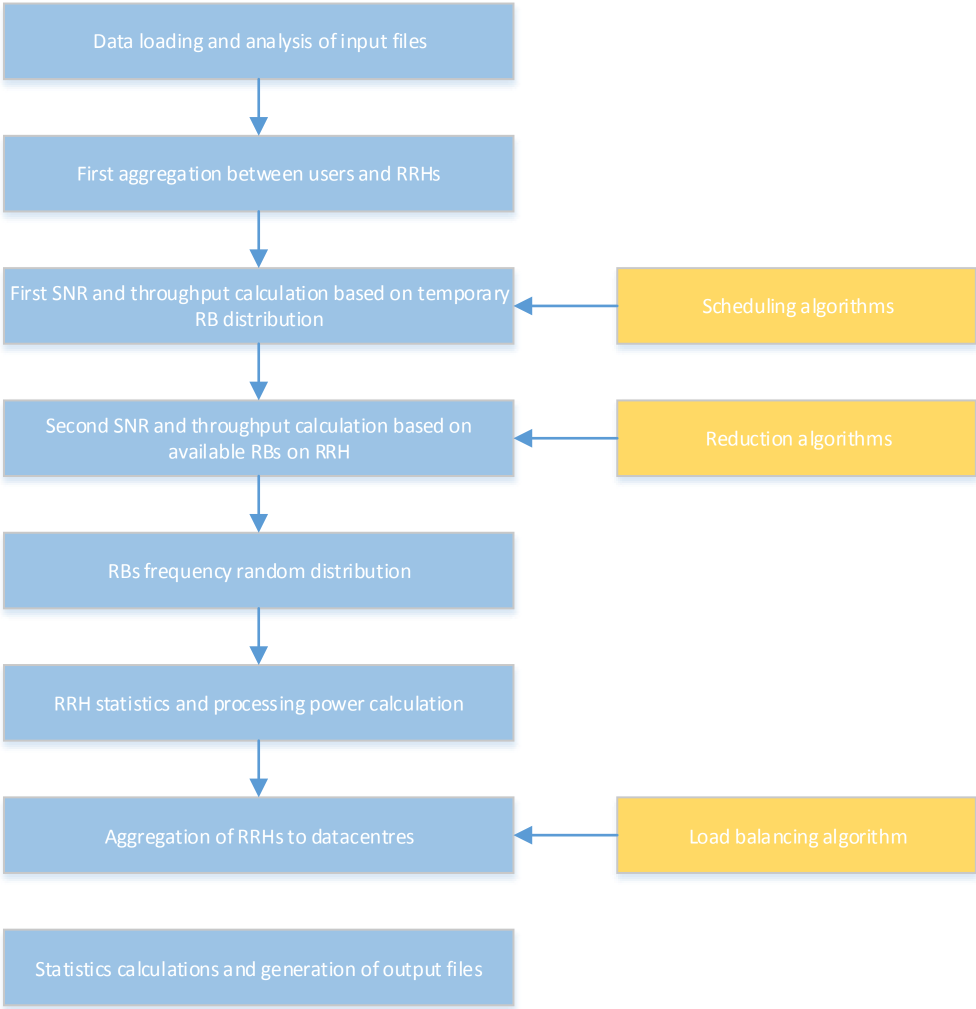


Figure 3.6: General workflow of the simulator.

The Bandwidth and MIMO configuration have a direct impact on the required processing power of an RRH, as explained in Section 3.1.2. On top of that, bandwidth also determines the number of available RBs per RRH, and thus it also has impact on the frequency usage percentage, which is also one of the parameters for the processing power calculation. The datacentre switch is simply used to indicate if there will be datacentres deployed or not. If the switch is on, the simulator will request a table containing the datacentres information, which were created using the Map Info tool. All the maps created with the datacentres locations are shown in Appendix E. Latency parameters are used for the latency calculations and since those parameters are not completely established and can vary greatly, it was decided to make them as an input so that it would be more dynamic and easy to adjust. These parameters include: CoMP delay, processing delay, fibre velocity, switches delay and latency margin. Lastly, the datacentres throughput limit determines the maximum throughput that one datacentre can handle, which is fundamental for the network capacity calculations and aggregation of RRHs to datacentres.

To adapt the simulator to the analysis of RRHs association with datacentres, a new class representing datacentres had to be created, and some of the previously made classes had to suffer some changes. The main changes were the following:

- The new class datacentre consists of a list with all the deployed datacentres, and basically, every object (datacentre) contains information about its current processing power, total throughput, latitude and longitude.
- The sector class (RRHs) was the class that suffered most changes. It now includes a function to calculate the maximum radius, based on the user that is furthest away from it, in order to have the worst case possible in terms of latency. And it also includes some new variables to have information regarding:
  - Datacentre on which it is connected,
  - Latency problems vector for all the available datacentres (it is basically a flag that indicates if there are latency problems for a certain datacentre),
  - Distance to the connected datacentre,
  - Total throughput in use,
  - Total processing power.
- A new class was created with all the required functions and algorithms to associate RRHs to the datacentres, as well as some additional required functions, such as the ones used to calculate the processing power.

As mentioned before, the Network Deployment module generates an extra output file called Data2.dat. This file contains the information relative to the datacentres locations, and it is used during the data loading and network configuration. During the normal course of the simulation, the only key aspect changed was the throughput calculations, since more realistic equations were available. These equations were derived by an extrapolation of different functions that relate the SNR with the respective throughput per RB. The way that these equations were derived are better explained in [Alme13], and the respective graphs and formulas are shown in Appendix C. Also, a rough implementation of Non Line of Sight (NLoS) was made, which randomly defines if a user is in Line of Sight (LoS) or NLoS, thus resulting in different ways to calculate the SNR. Some small changes had to be made in order to enable the introduction of MIMO 8×8 and higher bandwidths values; although they were roughly implemented, since it was not within the scope of this work.

When the RBs are properly allocated to users, the datacentre aggregation algorithms start to run. It initiates by obtaining crucial information about each RRH such as:

- Total number of RBs in use;
- Total throughput;
- Average modulation used (it is made a sum of the modulation used by each user that has allocated RBs, and then divided by the respective number of users);
- Code rate, based on the obtained average modulation;
- Number of Micro BSs.

These variables are then used to calculate the processing power for each RRH. When the calculations are concluded, some statistics are shown to provide information about the average, maximum and minimum processing power, as well as information about the total network throughput and processing power. This is important because, since there is no information that limits the maximum processing power for each datacentre, this limitation is done based on the total required throughput. That way, it is possible to evaluate the network in terms of its capacity, as explained in Section 3.1.4.

It is important to note that there are different processing powers depending on the class of the BS cell, i.e., macro- or micro-cell. Since most of the micro-cells are located at the centre of Lisbon, one BS is located at the limit of this centre area, and then one calculated the distance between this reference BS and all the others. If that distance is lower to 4 km, which is the distance from the reference BS to the outer limit of the centre area, it is a micro-cell, otherwise it is a macro-cell, as shown in Figure 3.7.

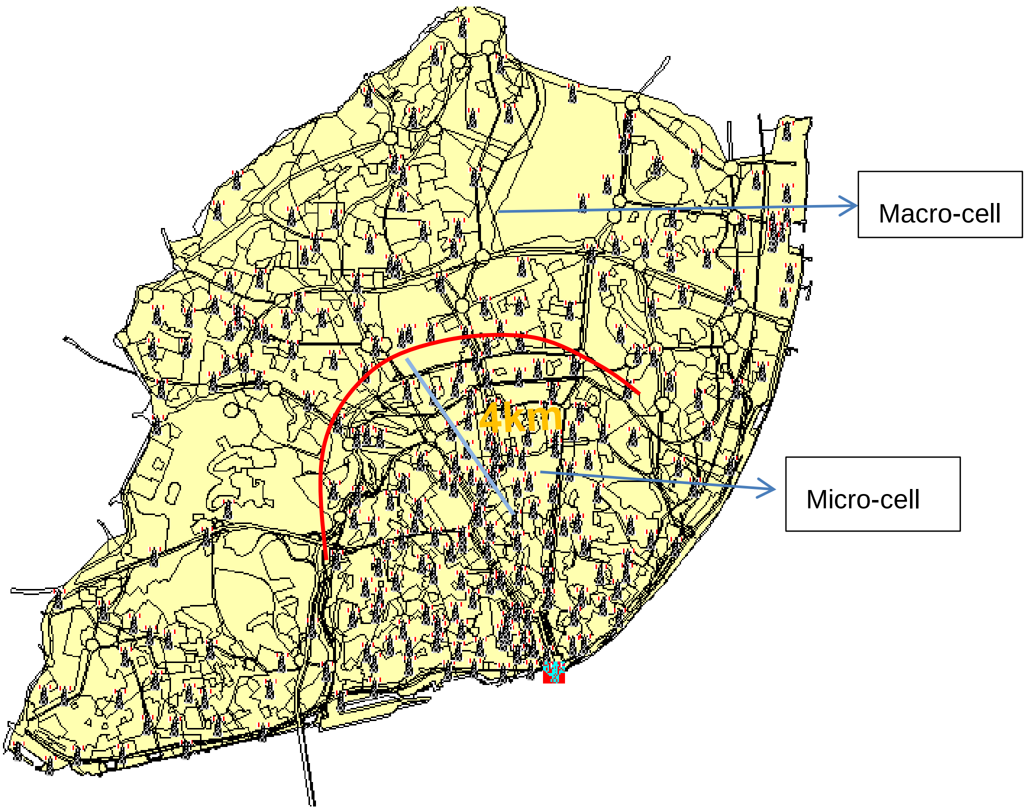


Figure 3.7: Micro- and macro-cells regions.

Subsequently, there is a first aggregation between the RRHs and the datacentres, based on their distance and latency. For every pair of RRH and datacentre, the latency is calculated in order to obtain the maximum distance for no latency problems. This distance is then compared with the physical distance between both components, in order to get feedback on the latency constrains. If there is latency problems, the RRH latency vector on the position equal to the datacentre id is changed to 1, otherwise it stays at 0. This way, in the end of this algorithm, every RRH knows for which datacentres there will be latency problems, thus it will not allow any connections between them. Nevertheless, if there are no latency problems, the RRH connects to the datacentre closer to it, which is the one with lower latency. This initial connection is necessary due to the lack of timing variation of the simulator. It enables to see changes in the distribution of RRHs throughout the datacentres, because otherwise, in the initial phase of the simulation, there would be no neighbours for each RRH, thus producing a very static distribution, meaning that the neighbour's parameter would have a very small contribution on the distribution of RRHs.

After this initial association, the main algorithm is responsible for the proper aggregation between the RRHs and the datacentres. It begins by calculating each RRH neighbour's and storing those values in a vector. This is called the priority vector, as it indicates which datacentres have more neighbours for the target RRH and thus that RRH will try to connect to datacentres using that order. Using this order, the RRH tries to connect to a datacentre, and two outcomes are possible, as illustrated in Figure 3.8:

- Priority vector value is 0: this means that for that specific datacentre there are no neighbours allocated. It starts by verifying if the RRHs current datacentre (allocated from the first algorithm) exceeds its throughput limit. In case it does not exceed, it maintains connected to the current datacentre. Otherwise, the RRH tries to connect to any other datacentre that does not have latency problems with this RRH and does not exceed the throughput limit.
- Priority vector value different than 0: this means that there are neighbours on that datacentre. Firstly, it is verified whether the current datacentre of the RRH corresponds to the one with most neighbours. Then there are verifications over latency and throughput limitations, and finally the RRH connects to the datacentre.

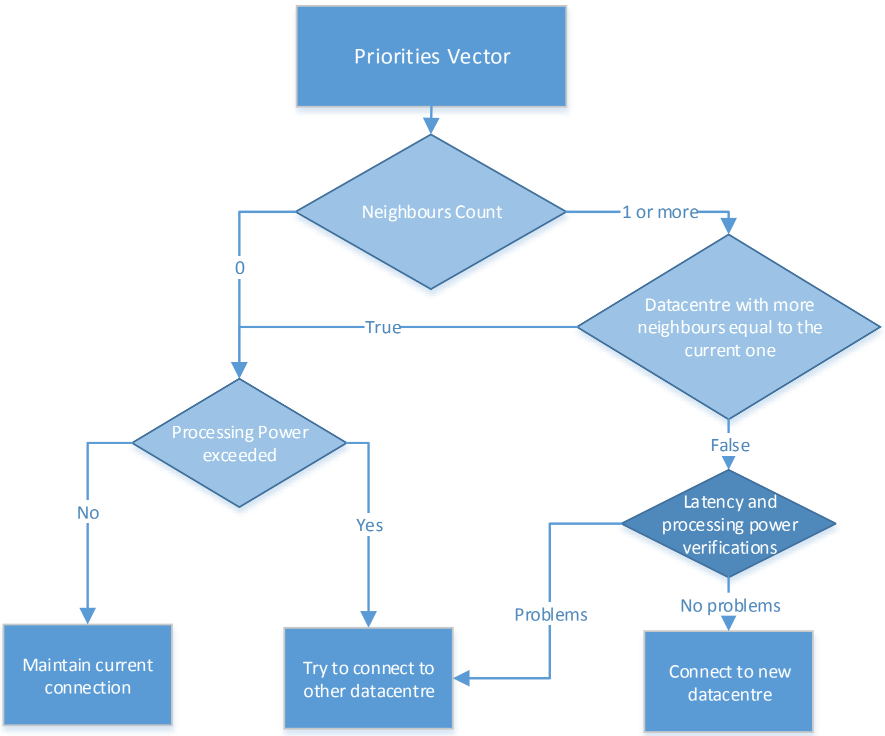


Figure 3.8: Flowchart explaining the final aggregation between RRHs and datacentres.

In the end, there is an analysis of the network capacity and output files are generated to demonstrate which RRH are connected to which datacentres. A general workflow is presented in Figure 3.9 to demonstrate the whole structure of the algorithm.

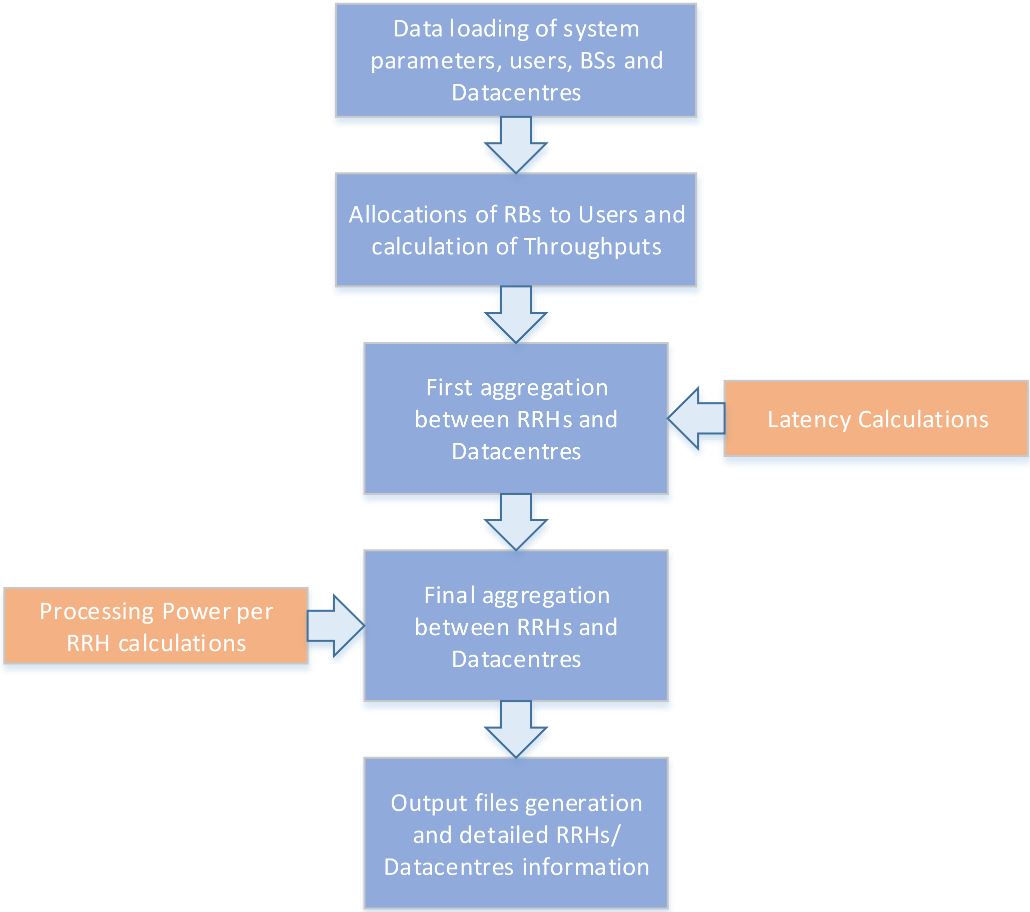


Figure 3.9: Load Balancing Algorithm Workflow

3.3.3 Input and Output Files

To start the simulator, in the Network Deployment module, some files are required:

- Ant65deg.TAB, which contains information about the radiation pattern for the BSs.
- Kathrein Horizontal.txt, Kathrein Vertical.txt, which also contains information about the radiation pattern for the antennas, one for the horizontal gains and another for the vertical, accordingly.
- DADOS.TAB, which contains information about the different areas in Lisbon, their population, distribution of population by age for the different regions, average income and scholarship level.
- ZONAS.TAB, which is used to characterise the different regions,
- users.txt, that contains all the users, their id, localisation, service and type of user,
- Vodafone Network.TAB, which has the position of all the BSs,
- datacentre.TAB, which has the position of all the datacentres.

After the required processing, 3 output files are generated, to be used in the LTE DL ICI Analysis module:

- definitions.dat, which has all the parameters required for the simulation, such as Bandwidth, MIMO configuration, etc.;

- data.dat, which has information about each user, the BSs and RRH that are covering him, as well as its coordinates, his service, type of user, distance to the RRH and his localisation,
- data2.dat, which has all the required information about the datacentres.

At the end of the simulation, in the LTE DL ICI Analysis module, some output files are generated to present the results of the implemented algorithms directly on the Map Info program:

- stats.out, which provides the simulation results, in terms of network analysis and services statistics;

Lastly, some text files are also generated to provide detailed information regarding the different steps of the algorithm, in order to check if it is working correctly. The most important ones are:

- processing_power.txt, which contains the values for the different parts required for the processing power calculations;
- resourceblocks.txt, which contains the number of RBs per user and per RRH, as well as the corresponding throughput. It also contains information about the frequency usage percentage per RRH, average modulation used, the processing power;
- dc_capacity.txt, which is run after the algorithms, in order to check which RRHs are connected to datacentres, and the corresponding network capacity;
- model.txt, which gives detailed information about the association between the RRHs and the datacentres;
- results.txt, which contains the average parameters, standard deviations and other relevant values;
- latency.txt, which contains information about the distances and latency calculations made during the first association algorithm.

3.4 Results Assessment

To validate the latency, processing power and throughput of each RRH, calculations had to be made in Excel, to verify that the obtained results were correct and in accordance to the expected values given by the equations in Sections 3.1.1 and 3.1.2. For the throughput, since new equations were implemented, calculations were also made to verify that they had been correctly calculated.

In order to test the algorithm, a scenario of 286 BSs, 3 datacentres and using 9 000 active users, was considered, which corresponds to the reference scenario explained in Section 4.1. The latency equation was temporarily modified so that the given value would be lower, in order to force a situation where latency problems would occur. Similarly, the datacentre throughput limit was also given a lower value, forcing a situation where this limit would be achieved by every datacentre, thus making some RRHs unconnected. By analysing some of the output files mentioned in Section 3.3.3, it was verified that the algorithm was properly running, and that the aggregation between the RRHs and datacentres was being made correctly. It is important to mention that the ICI switch was turned off, since the aim of this thesis will be to analyse the network with the maximum possible load. If the ICI was on, there would be a big decrease in the number of RBs used per RRH, thus decreasing the throughput and processing power.

To find the number of simulations required in order to ensure statistical relevance, it was required to calculate the average and standard deviation of some of the main output parameters. These were done using the following equations, correspondingly:

$$\mu = \frac{\sum_{i=1}^{N_z} Z_i}{N_z} \tag{3.13}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N_z} (Z_i - \mu)^2}{N_z}} \tag{3.14}$$

where:

- Z_i is the value of event i,
- N_z corresponds to the total number of events.

The parameters analysed were throughput per RRH, processing power per RRH, and latency, and they were tested for 30 simulations. In Figure 3.10, one shows the throughput and its respective standard deviation. The results show that there are no considerable changes between each simulation. This approach was repeated for the other parameters, obtaining the same results. Each simulation took between 40 minutes and 2 hours, and they were performed in a Genuine Intel®CPU U7300 1.30GHz. In Figure 3.11, one shows the standard deviation divided by the average for each of the parameters mentioned before in relation to the number of simulations.

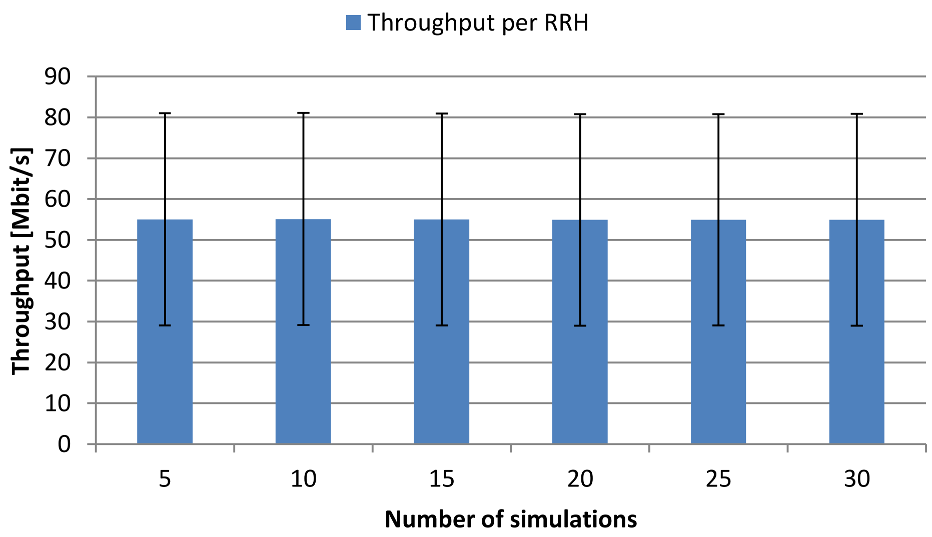


Figure 3.10: Average throughput per RRH along the simulations.

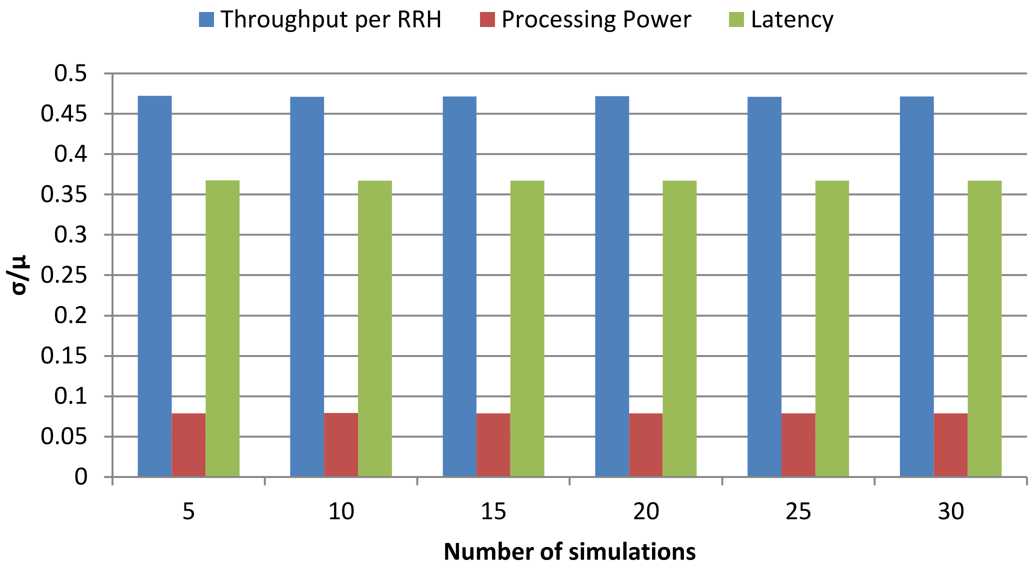


Figure 3.11: Standard deviation over average along the simulations.

These results are as expected, because the number of users is already big enough that statistical relevancy is acquired using only five or less simulations.

Chapter 4

Results Analysis

This Chapter presents the scenarios description and results analysis. It starts by defining, in Section 4.1, a reference scenario and network configurations. It is followed by a high load analysis in Section 4.2, divided into different simulations where variations in key parameters are made. To conclude, a theoretical analysis was made in Section 4.3, in order to address a bigger area than the one considered on the reference scenario. Conclusions relative to the results obtained are made throughout the whole chapter.

4.1 Scenarios Description

This section presents the two main scenarios used during the analysis. First, one explains the main scenario, used in the simulator, as well as the main characteristics and network configurations. It is followed by an analysis of the second scenario, where a theoretical analysis was performed.

4.1.1 Lisbon Area

The main scenario for this thesis is the city of Lisbon, illustrated in Figure 4.1. Even though there is no specific distinction between the centre and the outer region of the city, the analysis being made considering the town as a whole, it is important to refer that by observing the maps with the BSs positions, these two regions are quite different in terms of BS density. This fact could have an impact on the datacentres positioning, since the requested throughput in the centre region might be much higher than in the outer region. Also, as mentioned above, the BS density in the centre region is much higher, thus it was assumed that the BSs in this area are characterised as micro-cell BSs instead of macro-cell ones.

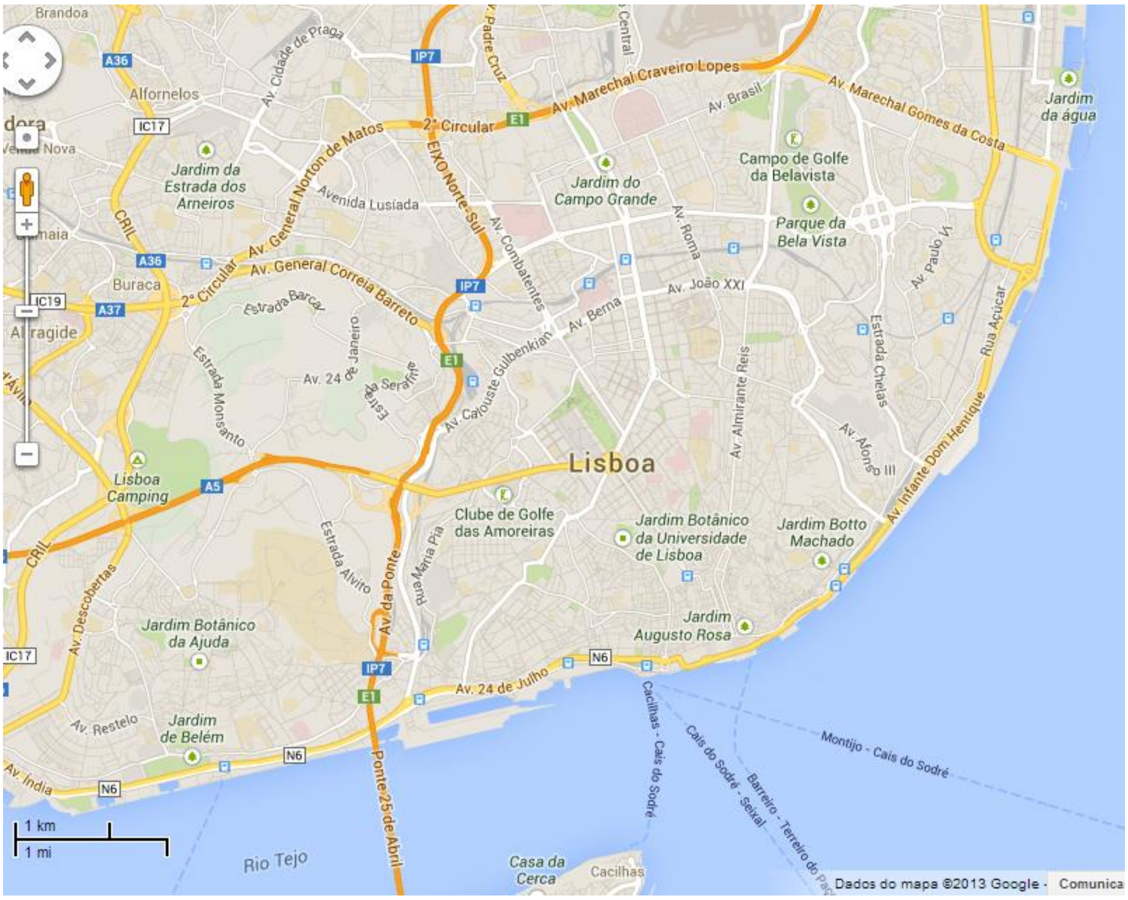


Figure 4.1: Map of Lisbon

Since Lisbon is mainly an urban environment, the propagation model used to calculate the path loss was the COST 231 Walfisch-Ikegami, with the parameters shown in Table 4.1. This model is detailed in Appendix F. Users were characterised by 4 different environments: outdoors, vehicular and indoor, high and low loss. Each of these environments gives additional loss, with the values illustrated in

Table 4.2, decreasing the received SNR. The distribution of users along the city is presented in Appendix G.

Table 4.1: Propagation Model Values.

Propagation model parameter	Value
BS height (h_b) [m]	26
Buildings height (H_B) [m]	24
Streets width (w_s) [m]	24
Width between buildings' centres (w_B) [m]	48
Departing angle from the closest building (ϕ) [º]	90
MT height (h_m) [m]	1.80

Table 4.2: Users additional losses depending on different environments.

User Scenario	Additional Loss [dB]
Pedestrian	0
Vehicular	11
Low Indoor	11
High Indoor	21

The reference scenario uses the frequency band of 2600 MHz with a transmission power of 46 dBm, a bandwidth of 20 MHz and MIMO configuration of 2×2. The down tilt angle used for the antennas is 6º and the gains for both vertical and horizontal planes are shown in Appendix H. It is relevant for this work that the network is highly loaded, thus one did not used any specific Frequency Reuse Scheme and the Interference Switch was turned off, since it would decrease the allocated number of RBs. Typical values were used for cable and user losses as well as for the noise figure. These values are presented in Table 4.3.

Regarding datacentres, the default number was 3 and their throughput limit was 60 Gbit/s. This value was chosen because there are 858 RRH in Lisbon, but on average only 600 have active users, as shown in Section 4.2.1, with a value of 15 users per RRH. Assuming an average throughput of 300 Mbit/s per RRH, the resulting network throughput would be 180 Gbit/s and thus a limitation of 60 Gbit/s would be sufficient to cover the entire city. This value of throughput per RRH is very high but it was chosen because the use of higher bandwidths combined with higher MIMO configurations should be able to achieve such throughputs, as proved in Section 4.2.4. The position of the datacentres was chosen in order to be approximately at the same distance to all BS, the first two datacentres being closer to the centre, while the third is located in the north of the city, closer to the border. Their location and id number is illustrated in Figure 4.2.

Table 4.3: LTE DL reference parameters.

Parameters	Values
DL Transmission Power [dBm]	46
Maximum Antenna Gain [dBi]	17
User Losses [dB]	1
Cable Losses [dB]	3
Noise Factor [dB]	7
Bandwidth [MHz]	20
Frequency Band [MHz]	2600
Down Tilt Angle [º]	6
MIMO Configuration	2×2
Frequency Reuse Scheme	None
Scheduling Algorithm	Proportional Fair
Interference Analysis	Off
Datacentre Limit [Mbit/s]	60000

For latency calculations, the used values were the ones explained in Section 3.1.1 and shown in Table 4.4. The results obtained show that there are no any latency problems for the region of Lisbon, and thus a theoretical study was made with a bigger area to check where problems would start to appear.

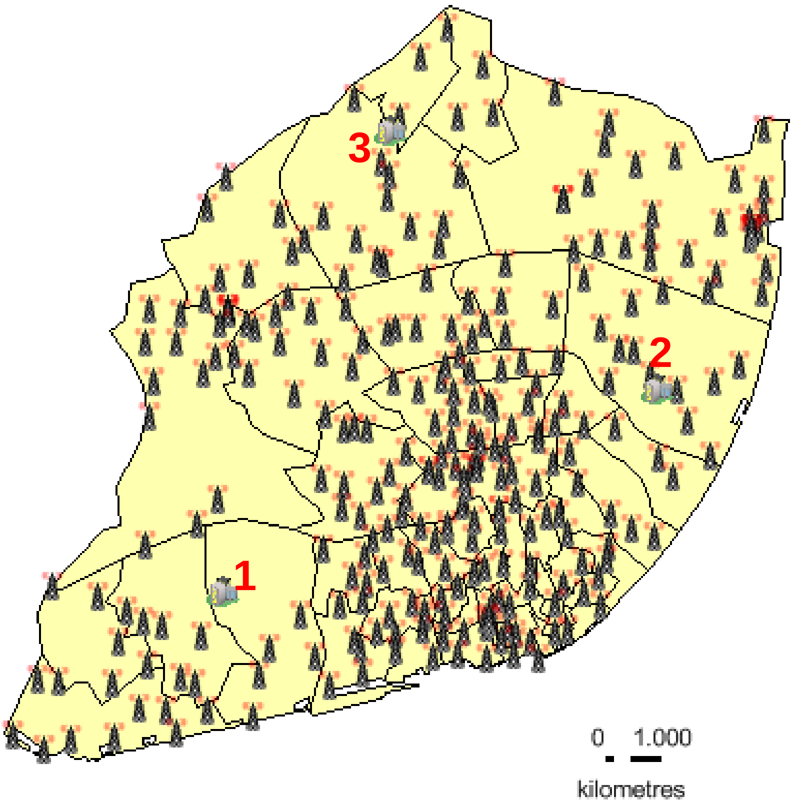


Figure 4.2: Datacentres locations and BS locations (corresponding to 3 RRHs per site).

In order to analyse the implementation of the datacentres and their total capacity, it is required to have a full loaded network, as it represents the worst case scenario in terms of required throughput and processing power. The number of active users is around 9 000 and the only service used is FTP, since it is the one that has higher requests in terms of throughput. It is important to note that for each type of modulation, the respective maximum coding rates were used, for QPSK 0.588, 16QAM 0.602; 64QAM 0.926, which were the values obtained in [Alme13] that better simulate the values obtained on real networks.

Table 4.4: Reference Latency Parameters.

Latency parameters	**Value**
Processing Delay ($\delta_{processing}$) [ms]	2.8
CoMP Delay (δ_{CoMP}) [ms]	0.5
Switch Delay (δ_{switch}) [ms]	0.02
Fibre Velocity (v_{fibre}) [km/s]	200 000
Delay Margin [%]	10%

4.1.2 Peripheral Area

The peripheral area is used for the theoretical analysis done in Section 4.3. In Figure 4.3, the total area taken into consideration is shown, corresponding to around 1 970 km^2 and by analysing a Google Earth file with the placements of BSs throughout Portugal, two main regions are defined: urban and rural, depending on the density of BSs. Nonetheless, 90 km^2 were subtracted from that value, which were the values obtained for the area of Lisbon. This was made because the theoretical analysis was made excluding Lisbon. The values for BS density considered for each region as well as the respective area are shown in Table 4.5. These values are a rough estimation, since the exact values of BS per area were unknown. Using the values mentioned, it is possible to do a similar analysis as in Section 4.2, in order to check whether it would be possible to increment the area covered by the datacentres. The values for the areas were obtained by using the Google Maps Area Calculation tool, [Daft09].

Figure 4.3: Total area to analyse.

Table 4.5: Characterisation of the different environments.

	Area [km^2]	BS Density [BS/km^2]
Urban	370	10
Rural	1510	3

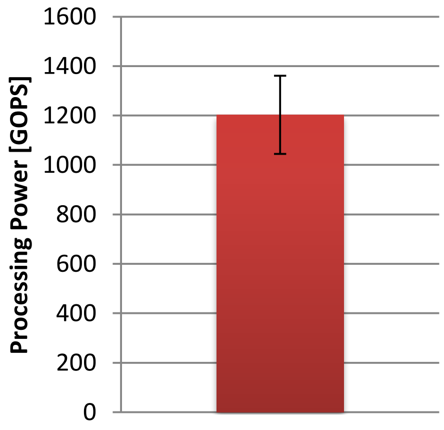
To obtain the values of both areas, the total area was split into main regions, explained better in Appendix I, which are characterised by different densities of BSs per squared kilometre. Note that in this scenario there is no heterogeneity in terms of users distribution. Also, regarding the network configurations, they are the same as in the reference scenario explained in Section 4.1.1.

4.2 High Load Scenario Result Analysis

This section presents the results analysis concerning the first reference scenario, which is a high loaded network since the number of users exceeds 12 000. First, the reference scenario is evaluated according to a set of key parameters. Secondly, an analysis of the sensitivity to the variation of these parameters is performed. The key parameters that will be analysed are the Throughput, Processing Power, Latency, number of RRHs connected per datacentre, number of Users and Network Capacity.

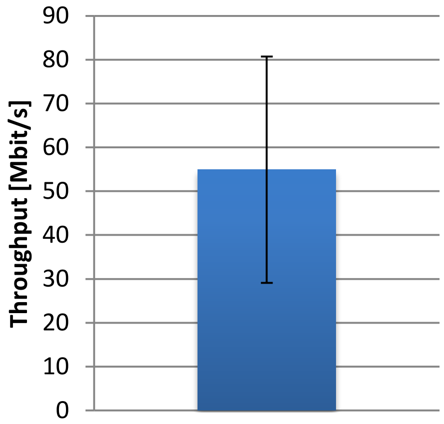
4.2.1 Reference Scenario

To initiate the analysis of the reference scenario explained in Section 4.1, one can observe, in

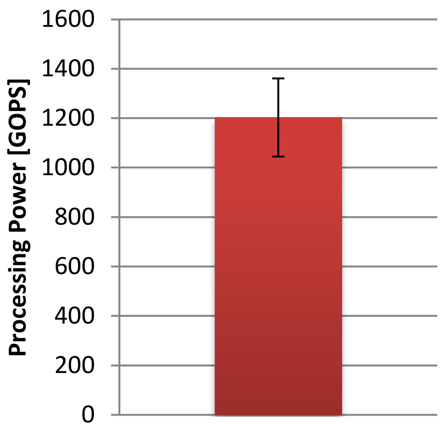


(b) Processing power per RRH

Figure 4.4, the average throughput and processing power per RRH, while in Figure 4.5 the average number of active users.



(a) Throughput per RRH



(b) Processing power per RRH

Figure 4.4: Average throughput and processing power per RRH.

Even though there is no interference, it is possible to observe that the average throughput per RRH, 55 Mbit/s, is quite lower than the maximum values of 150 Mbit/s. This is due to the user distribution, which creates a lot of divergence in the load of each RRH, and that is why this parameter is susceptible to big variations, as shown by the standard deviation. Nonetheless, this value is higher than the ones obtained in [Pires12], mainly because interference was not taken into account. The processing power on the other hand has a more static value, around 1 200 Gops, which is good because it could help quantify the number of BBUs supported by each BBU-pool processor, based on the processing effort limitations. Still, this value is a rough estimation, since the real processing effort is not easy to quantify as it depends on a great number of parameters, as well as on the architecture of the processors used.

Concerning the number of active users, since no interference is considered most of them are able to be served, obtaining an average of served users around 90%, which is a high value, but expected, because this value considers the users served with a minimum required throughput which can easily be achieved. On the other hand, the percentage of satisfied users might lower than expected, since most of the users can only be served using the minimum throughputs. This happens mainly because the RB distribution among users is made in order to guarantee that they have the minimum throughput, and after all users guarantee that, the remaining RBs are equally distributed among them. Consequently, the wanted throughputs for FTP are hard to achieve since it is the most demanding service.

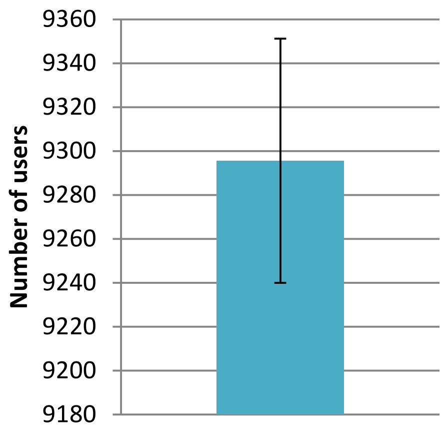


Figure 4.5: Number of active users on the network.

As pictured in Figure 4.6 it is possible to see that the variations in the maximum distance to the connected datacentres are very small, which means that the latency problems are mainly due to the processing time and delay in the equipment, thus not very depended on the distance of the users, which is why there was no need to perform simulations using other frequency bands, which would create some fluctuations on this average distance. The values obtained were around 25 km which means that there are not any latency problems in the Lisbon area, and thus a theoretical study is shown in Section 4.3 which covers a larger area. The average distance between the RRHs and the datacentre on which it is connected is approximately 2.5 km, which is to be expected since the diameter of Lisbon is around 10 km.

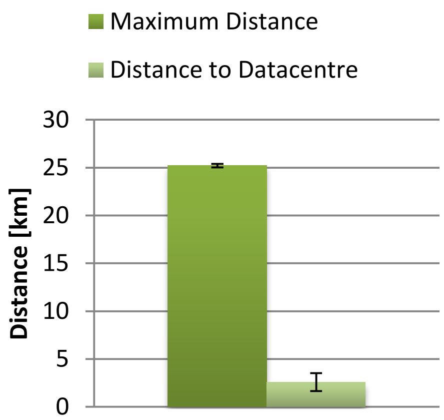
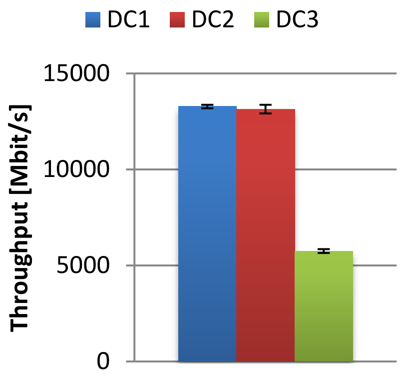
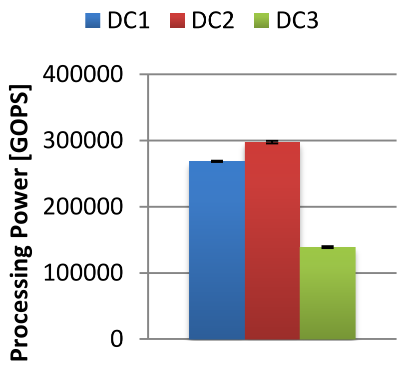


Figure 4.6: Maximum possible distance and distance to connected datacentre.

In Figure 4.2, it is possible to see that datacentres 1 and 2 are closer to central Lisbon and thus more RRHs will be connected to those datacentres, due to the first aggregation explained in Section 3.3.2. This will only happen if the throughput limits are not exceeded, because otherwise some of the central RRHs will try to connect to datacentre 3. In Figure 4.7, one illustrates the throughput and processing power served by each datacentre, where values range from 6 and 13 Gbit/s and between 140 000 and 300 000 GOPS, respectively.



(a) Throughput per Datacentre.



(b) Processing Power per Datacentre.

Figure 4.7: Average throughput and processing power served per datacentre.

The reason why the processing power in datacentre 2 is higher than in datacentre 1 can be seen in Figure 4.8, by analysing the number of RRHs allocated to each datacentre. As it is shown, datacentre 1 has around 220 RRH while datacentre 2 has almost 250. This implies that there is no linear relationship between throughput and processing power since, and so one cannot say that a datacentre with a higher throughput also has a higher processing power. Furthermore, Figure 4.8 shows that datacentre 3 has a lower number of aggregated RRHs, around 120, due to its location. Consequently the algorithm could be enhanced by taking into account the number of RRHs allocated on each datacentre, and not only the datacentre throughput/processing power limitation.

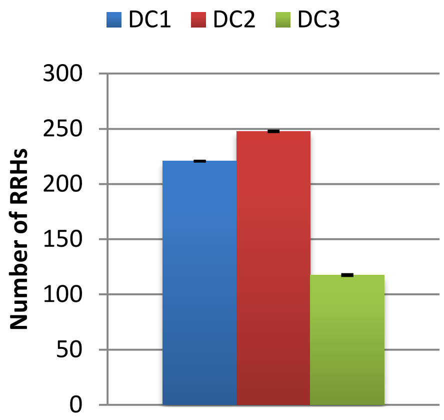
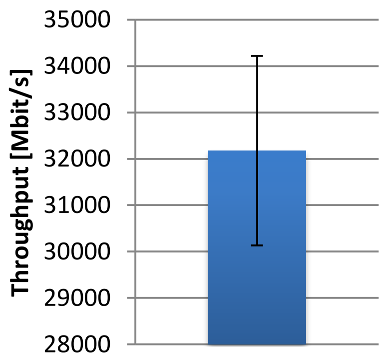
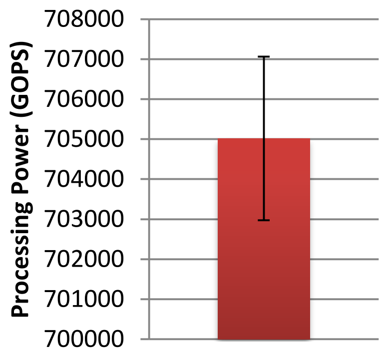


Figure 4.8: Number of aggregated RRHs per datacentre.

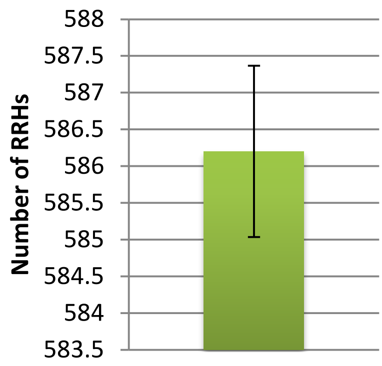
Observing the values of throughput per datacentre, it is possible to say that there will not be any problems due to throughput limitations, since none of the datacentres reached the maximum throughput limit. Figure 4.9 shows that no problems occur, as the obtained values are consistent with the sum of each individual value per datacentre. The total network throughput is in the order of the 32 Gbit/s, the processing power 705 000 Gops and the number of active RRHs is around 586.



(a) Total network throughput.



(b) Total network processing power.



(c) Total number of active RRHs.

Figure 4.9: Average network throughput, processing power and active RRHs.

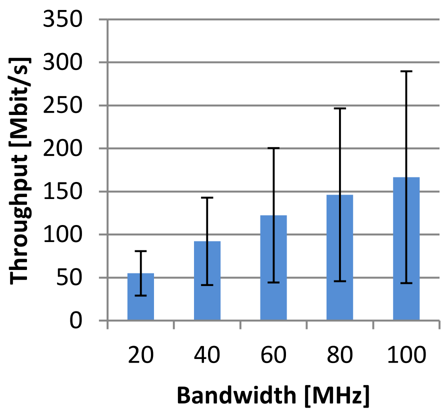
In conclusion, for the reference scenario all existing RRHs are capable of connecting to a datacentre, meaning that there are no problems due to throughput or latency. Thus, the resulting network capacity for these conditions would be 100%. Even if the number of users increased drastically, it would be difficult to achieve degradation in the capacity because of the high datacentre limits. By analysing some of the previous values, it is possible to see that 9 000 users generate a network traffic of approximately 32 Gbit/s, which means that it would require an increase of almost 6 times in the number of users to achieve the full datacentres capacity.

4.2.2 Bandwidth

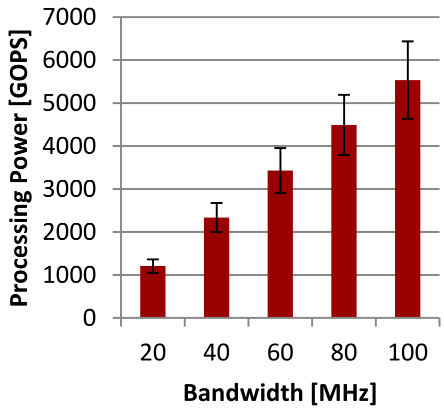
In terms of bandwidth, simulations were made using bands of 20, 40, 60, 80 and 100 MHz, corresponding to 100, 200, 300, 400 and 500 RBs. Note that these bands are an optimistic expectation since techniques, like CA, are required, which were not implemented since it is not within the scope of this thesis.

Nonetheless, Figure 4.10 shows the results obtained for the throughput and processing power per RRH for all the different bandwidths. Since simulations were made without any CA and interference, it is normal that both the throughput and the processing power increase linearly with the bandwidth. On the other hand, it is possible to see an increase in the standard deviation along the bandwidths, meaning that throughput at the higher bandwidths is susceptible to larger fluctuations than at lower bandwidths. The processing power follows the same evolution as the bandwidth but its standard deviation does not increase as much as the one from the throughput, thus confirming the more strict behaviour of the processing power.

In Figure 4.11, one can observe that the number of served users does not follow the same behaviour as the other key parameters. It starts with an aggressive increase, but after 40MHz it arrives at a saturation point. This is due to the fact that the files with the users contained around 12 000 users, and up from 60 MHz forward, practically all the users can be served with a minimum throughput, due to the increase in the number of available RBs, thus the number of served users does not increase.



(a) Throughput variation along different bandwidths.



(b) Processing power variation along different bandwidths.

Figure 4.10: Throughput and Processing Power per RRH along different bandwidths.

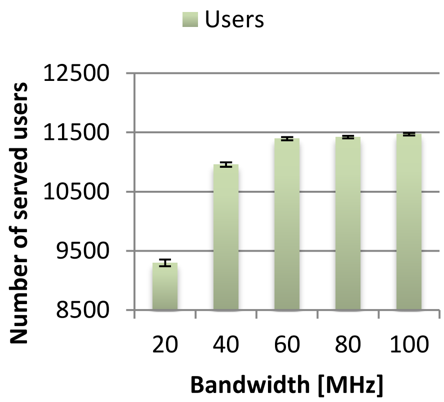
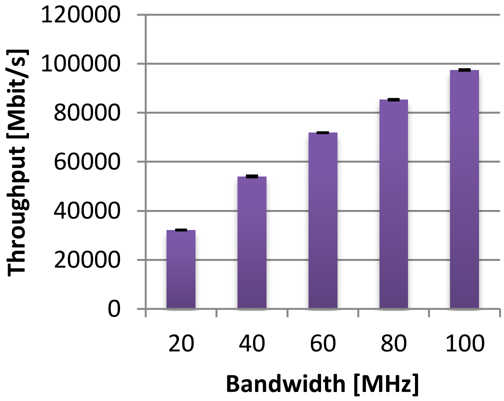
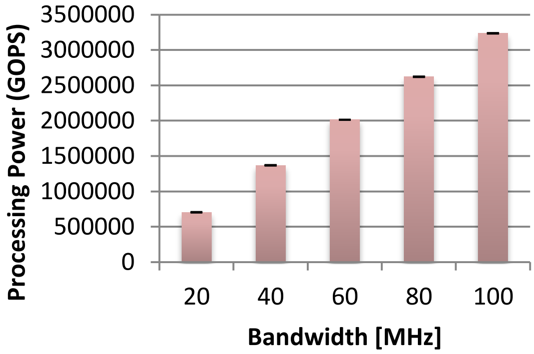


Figure 4.11: Number of served users on the network, along different bandwidths.

The total network throughput and processing power is shown on Figure 4.12. Since one has defined a throughput limitation of 60Gbit/s per datacentre, there can be a maximum throughput of 180 Gbit/s in the network. As one can see, even on simulations with 100 MHz bandwidth this value is not reached, the average network throughput being obtained in the order of 100 Gbit/s, corresponding to a processing power around 3×10^{6} GOPS. Consequently, it is predictable that the throughput distribution along the datacentres follows the same increase, since all RRHs keep their current connection as long there are no throughput limitations. This fact occurs because the RRHs are always located at the same positions, and thus, their first aggregation to the datacentres, and later its neighbours count, will be similar along different simulations. Figure 4.13 illustrates this fact.

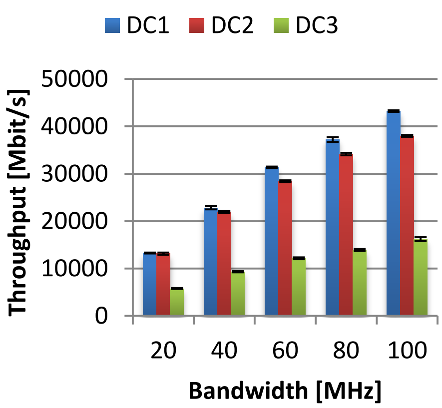


(a) Total network throughput.

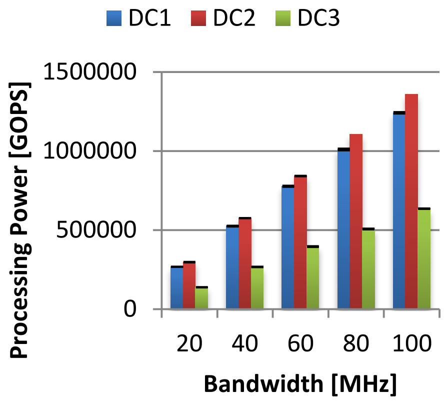


(b) Total network processing power.

Figure 4.12: Total network throughput and processing power along different bandwidths.



(a) Throughput per datacentre.



(b) Processing power per datacentre.

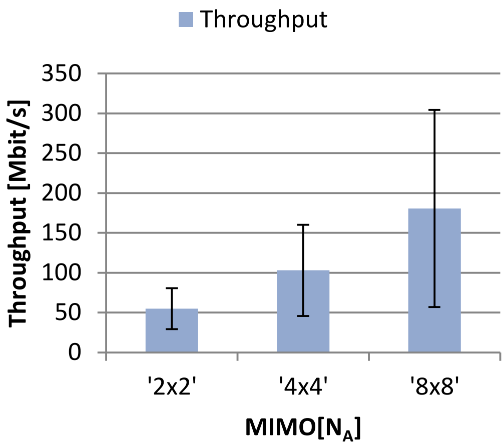
Figure 4.13: Throughput and processing power per datacentre along different bandwidths.

4.2.3 MIMO

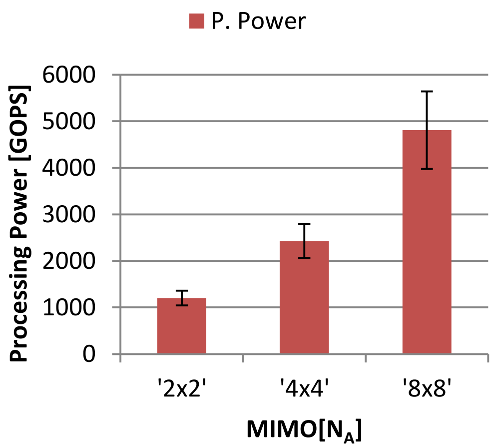
The different configurations of MIMO used were 2×2, 4×4 and 8×8. This factor was also roughly implemented since it was not within the scope of this thesis, and the gains used for each configuration was 0, 2 and 4, because the throughput equations were already made specifically to MIMO 2×2. Thus, it is expectable to see these gains in the throughput per RRH and in other key parameters.

In Figure 4.14, one represents the throughput and processing power per RRH for the different configurations. Comparing the values obtained with the ones from Figure 4.10, one can see that with MIMO 8×8 the values are quite close to the ones obtained with 100 MHz bandwidth. But on the contrary, as Figure 4.15 shows, the number of active users does not reach the saturation point. This is as expected, since MIMO only adds a gain to the throughput, but the number of RB available does not increase. The number of users increases a little bit, but this is because the minimum throughput can be given to users using less RBs.

Regarding the processing power, the maximum value obtained is also quite similar to the ones obtained for the higher bandwidths. This fact could be misleading since the MIMO configuration values generate a higher increase in comparison to the reference values than the bandwidth. The reference bandwidth is 20 MHz, while the maximum bandwidth taken into consideration was 100 MHz, meaning that the processing power would be five times higher for the maximum bandwidth. On the other hand, MIMO reference number of antennas is one, while for 8×8 MIMO it is increased to 8, thus it would be expected to see a higher increase. But, even for the reference scenario of 20 MHz, the MIMO configuration used was 2×2, meaning that there is only an increase by a factor of 4, instead of 8. Thus proving the reliability of the results obtained for the MIMO average processing powers, which are a bit lower than the ones for the bandwidth.



(a) Throughput per RRH.



(b) Processing power per RRH.

Figure 4.14: Throughput and processing power along different MIMO configurations.

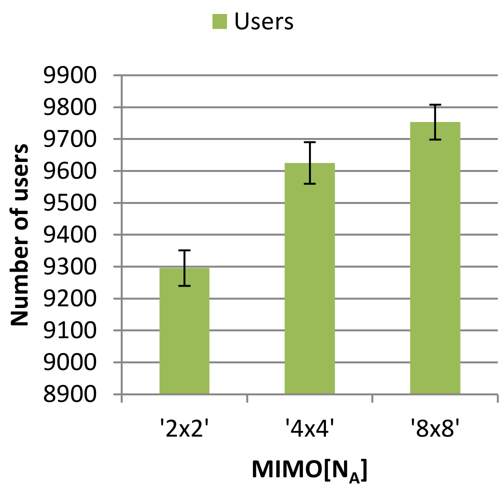
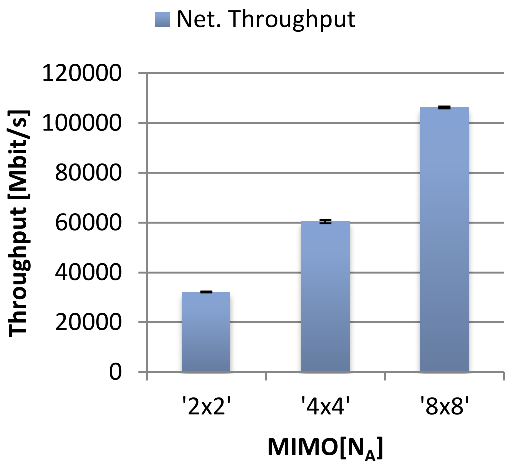
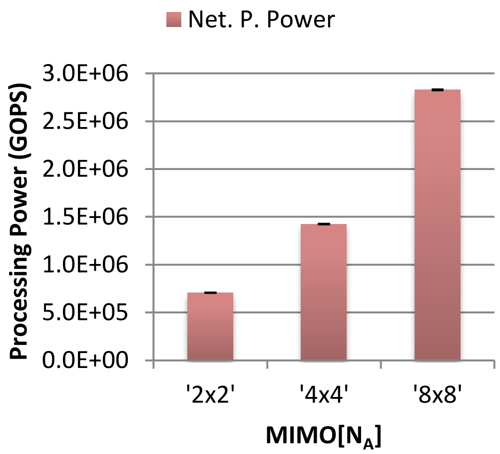


Figure 4.15: Number of active users along different MIMO configurations.

The total network throughput and processing power in 8×8 MIMO is also quite similar to the one obtained with bandwidth of 100MHz, which is expected since the throughput and processing power per RRH follows the same behaviour. Thus it is also predictable to see the same distribution of RRHs along the datacentres. Both parameters are displayed in Figure 4.16 and Figure 4.17.

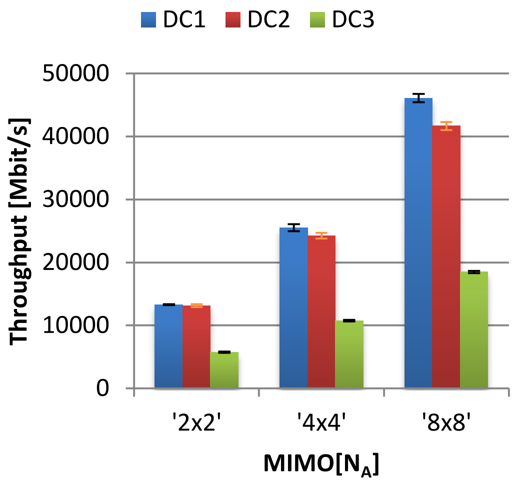


(a) Total network throughput.

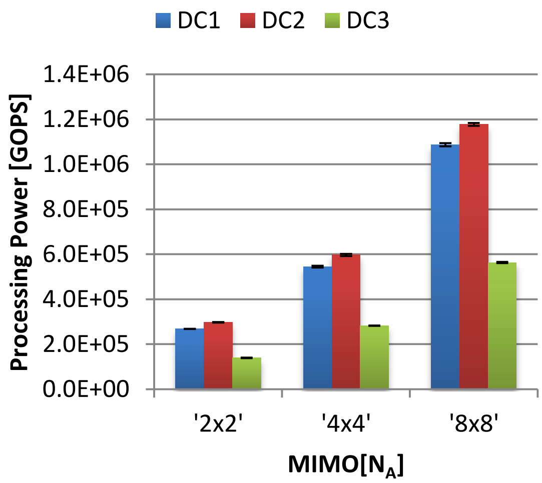


(b) Total network processing power

Figure 4.16: Total network throughput and processing power along different MIMO configurations.



(a) Throughput per datacentre.



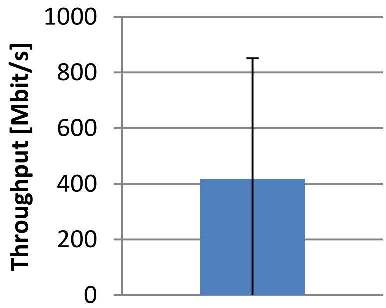
(b) Processing power per datacentre.

Figure 4.17: Distribution of throughput and processing power per datacentre, along different MIMO configurations.

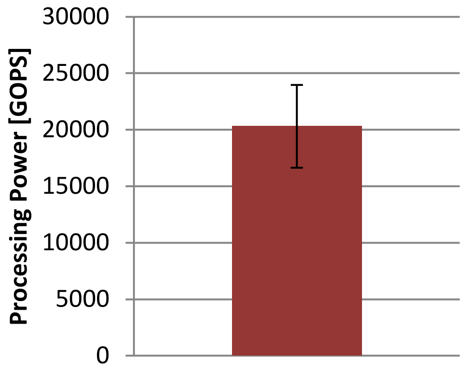
4.2.4 Datacentre Throughput Limit

In this section, the datacentre throughput limit was changed between 50Gbit/s and 100Gbit/s in order to analyse capacity problems and the RRH distribution along the datacentres. Note that for these simulations the used bandwidth was 100 MHz, with a MIMO configuration of 8×8, in order to obtain the maximum network throughput possible.

For these configurations, the throughput and processing power per RRH is shown in Figure 4.18. As one can see, there is quite a big increase of both parameters in comparison with the values obtained for both parameters analysed independently. This increase is due to the amount of RBs available together with the increase in throughput per RB obtained due to MIMO. The number of served users also increases by a small amount compared to the ones obtained with the bandwidth of 100 MHz because of the better conditions provided by MIMO, as it is shown in Figure 4.19.



(a) Throughput per RRH.



(b) Processing power per RRH.

Figure 4.18: Throughput and processing power per RRH for 100MHz with 8×8 MIMO.

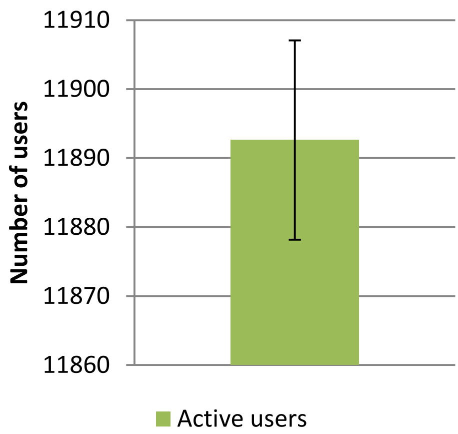
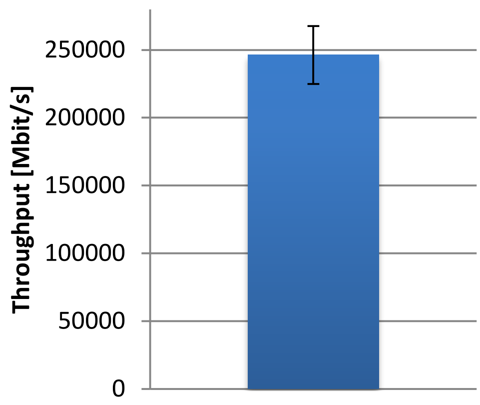
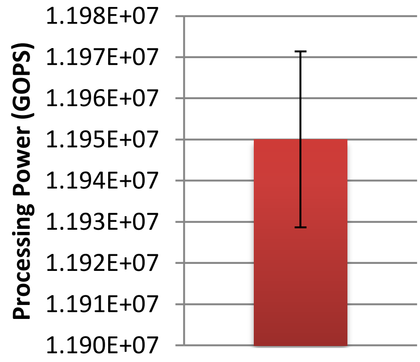


Figure 4.19: Number of active users for 100MHz bandwidth and 8×8 MIMO.

The network requested throughput and processing power is pictured in Figure 4.20. These values are the ones that increase more significantly which might create an impact on capacity and RRH distribution along the datacentres. The last parameters mentioned are equal throughout the simulations for different datacentres throughput limits, thus the values are only presented once.



(a) Network throughput.



(b) Network processing power

Figure 4.20: Network requested throughput and processing power.

In Figure 4.21, one can see the variations in throughput per datacentre along the different throughput limitations. The maximum limitation of 100 Gbit/s demonstrates that the saturation point of both datacentre 1 and datacentre 2 is reached. Both those datacentres have throughputs very close to its limit, while datacentre 3 has a much lower value. For the 90 Gbit/s limit, there is less 20 Gbit/s available for both datacentre 1 and datacentre 2 and thus it is expectable to see an increase in the throughput of datacentre 3 around the same value. For all the other values, the limit is achieved by all datacentres, which creates capacity problems, meaning that some RRHs will not be able to obtain allocated processing capacity. This results in the disruption of the RRH. One solution would be to lower the modulation used on each user or to accept only a percentage of users on the RRH, in order to decrease the total RRH throughput and processing power.

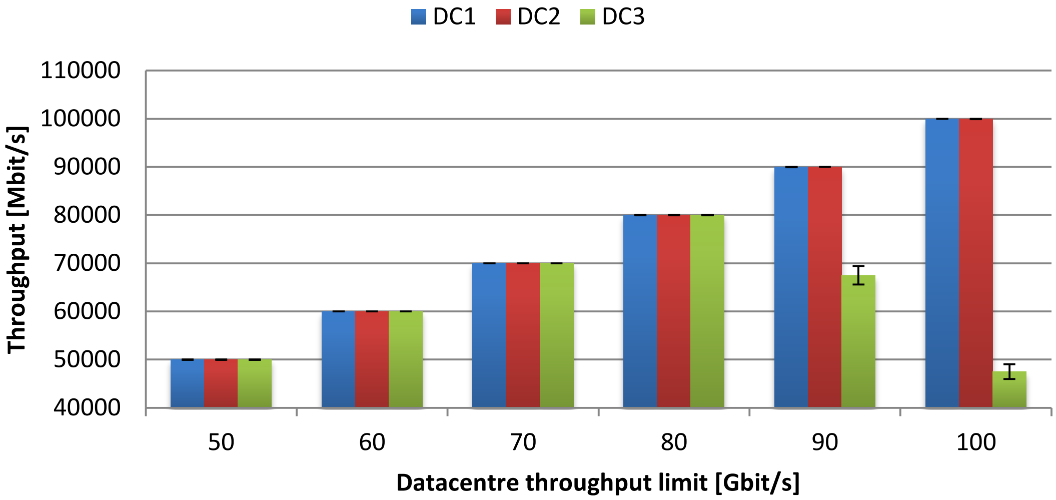


Figure 4.21: Throughput per datacentre along different throughput limits.

The processing power on the other hand is not limited and has some fluctuations, but as it depends on the number of RRHs allocated, the general flow is the same as the throughput. The reason why it is not limited is because there are no reference values regarding this parameter, thus it would probably result in value far from reality. As it is demonstrated in Figure 4.22, the processing power increases for all datacentres until reaching the 80 Gbit/s limitation, where the capacity problems disappear. Afterwards, the processing power of datacentre 3 decreases until it reaches a minimum at 100 Gbit/s. This can be explained since the two first datacentres are closer to the problematic area of centre Lisbon, and thus they both reach their limitations very early. Consequently, when the throughput limit is achieved by datacentre 1 and datacentre 2, all the RRHs that are still unconnected will be associated to datacentre 3. On the other hand, it is important to note that for the lower limitations, datacentre 3 has a bigger processing power. This occurs because the area with more traffic and users is at the centre, and thus the most loaded RRHs, in terms of throughput, will try to connect to the first two datacentres, achieving the corresponding datacentres limits much faster. Thus, datacentres 1 and 2 will have less RRHs, but with higher throughputs than datacentre 3. On the other hand, since datacentre 3 has more RRHs, the processing power is higher than in the other datacentres.

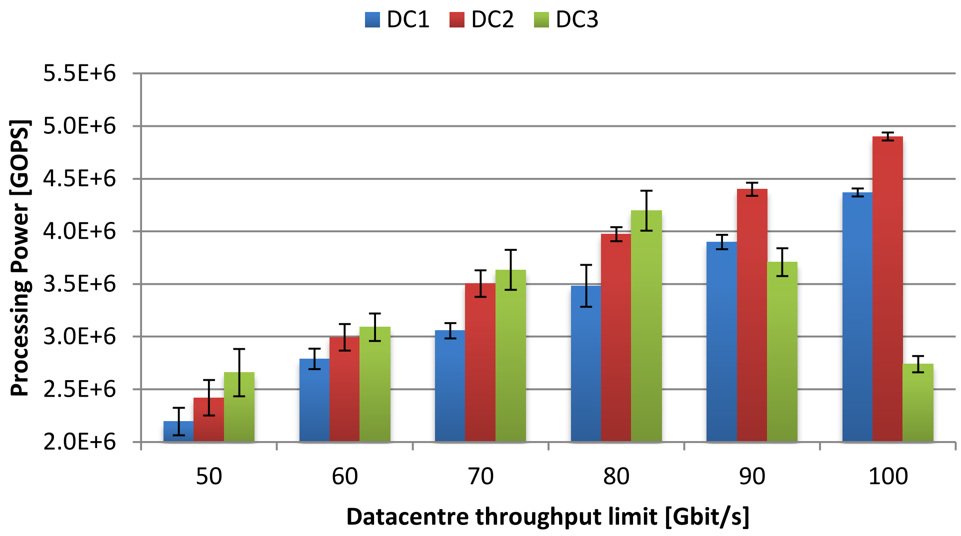


Figure 4.22: Processing power per DC along different throughput limits.

To support the conclusions above, Figure 4.23 shows the changes in RRH distribution along the datacentres. As pointed out before, the number of allocated RRHs per datacentre will follow a very close behaviour as the one from the processing power. All datacentres have an increase in their number of RRHs up to the 80 Gbit/s limit. Afterwards, there is a decrease in the number of RRHs of datacentre 3, because most of the RRHs are already able to connect to the first two datacentres, since the limitation is higher. Since more RRHs are being connected to datacentre 3 at the lower limit levels, there is an increase in the average distance at which RRHs are connected to datacentres, in comparison to the values obtained for the reference scenario. This is shown in Figure 4.24, with values in the order of the 3 km, which is still an acceptable value, and does not create any latency constrains. At the 100 Gbit/s limit, it is already possible to see a decrease in the average distance, converging to the 2.5 km obtained for the reference scenario.

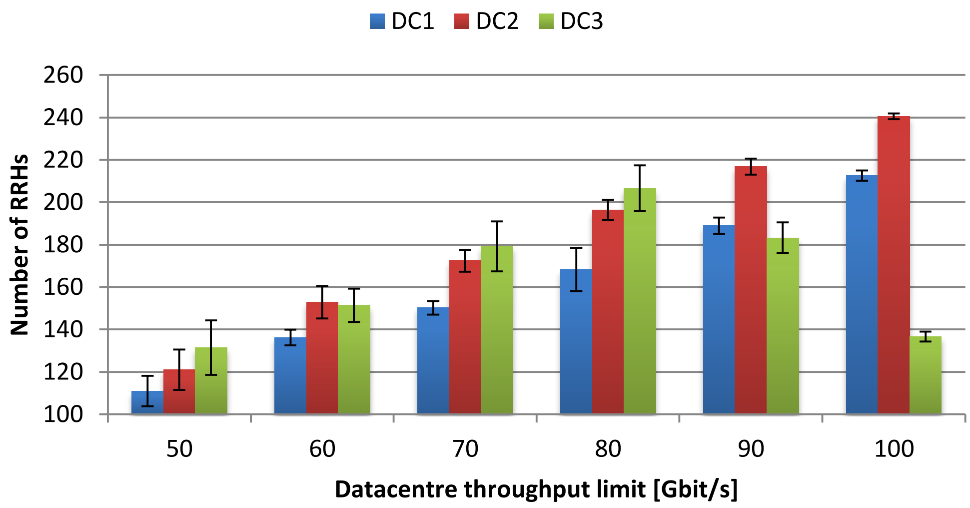


Figure 4.23: Number of RRHs allocated to each DC along different throughput limits.

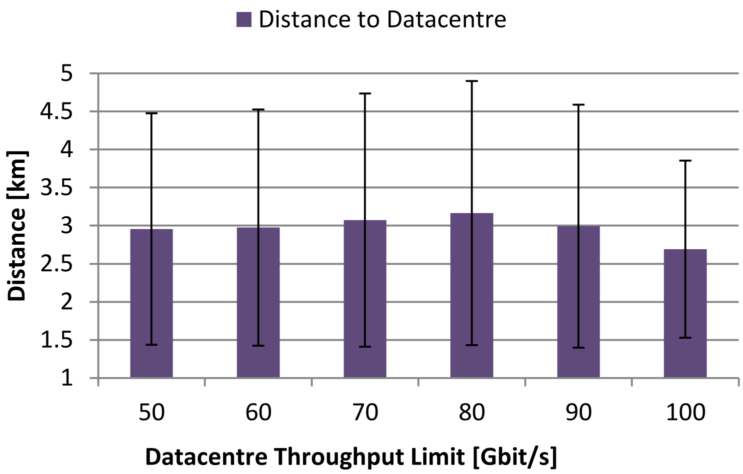


Figure 4.24: Distance between RRHs and DCs along different throughput limits.

In conclusion, Figure 4.25 and Figure 4.26 represent the network capacity and the number of RRHs not connected due to throughput, respectively. As mentioned before, the limitations values below and including the 80 Gbit/s, introduce capacity problems. For every 10 Gbit/s increase in the throughput limitation, the capacity also increases by a factor close to 10%. The number of unconnected RRHs for the minimum throughput limit is around 225, and for each 10 Gbit/s limitation increase, it decreases by a factor close to 10% of the total number of active RRHs, which are around 60. These unconnected RRHs are disrupted and the traffic going through them is not processed. As mentioned before, the solution could be to reduce the throughput on each RRH, by lowering the modulation used or the percentage of users on each RRH, or by allowing different MIMO configurations on each RRH.

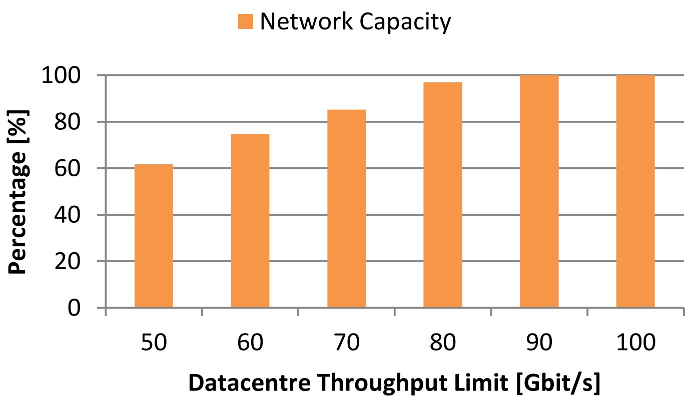


Figure 4.25: Network capacity for different throughput limits.

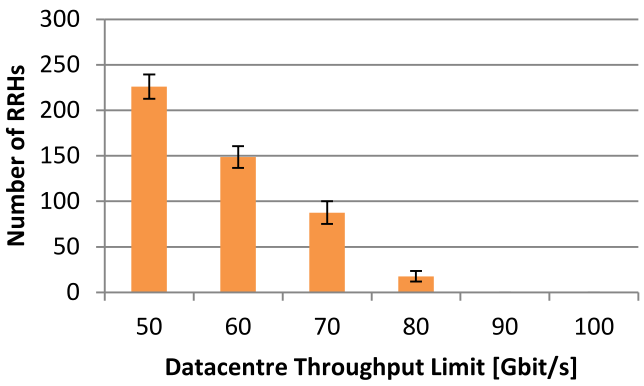


Figure 4.26: Number of RRHs without connectivity due to throughput limitations.

4.2.5 Number of Datacentres

In this section, simulations were made using a different number of deployed datacentres. Using the reference scenario, which does not have any capacity problems, one can observe in Figure 4.27 the fluctuations in the datacentre usage ratios for each of the datacentres deployed. Datacentre 1 is the one that has more throughput usage, reaching a maximum of around 50% in the case of only 1 datacentre. Since in all deployments datacentre 1 is the closest one to the centre of Lisbon, it is expected that its usage will be higher throughout every simulation. The only exception happens with the use of 3 datacentres, mainly due to their positioning, which puts both the datacentres at relatively the same distance between themselves and the centre of Lisbon. Further information regarding the positioning of the datacentres is explained in Appendix E.

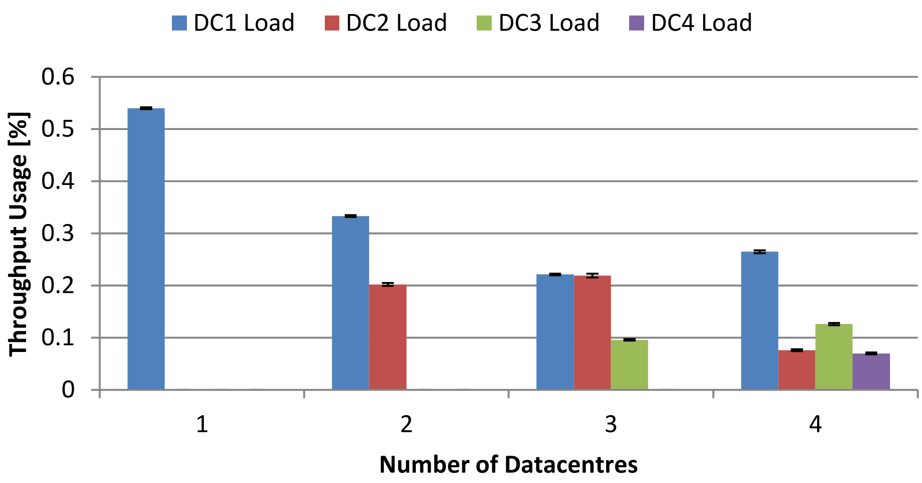


Figure 4.27: Usage ratio of each datacentre along different number of deployed datacentres.

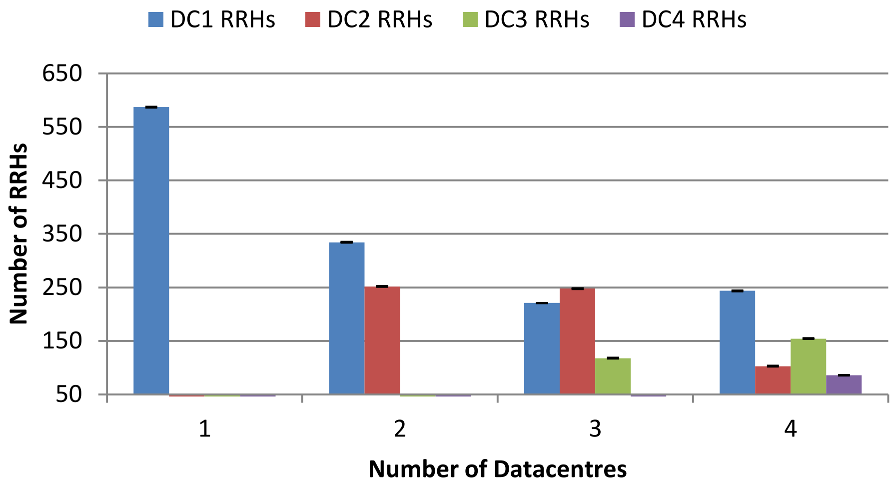


Figure 4.28: Number of RRHs associated to each datacentre for different number of deployed datacentres.

In Figure 4.29, the average distance between the RRHs and the datacentres is represented, which as expected, shows decreases of around 1.5 km between the maximum and minimum value, obtained for 1 and 4 datacentres respectively. This is easily explained since all the RRHs try to connect to the closest datacentre, and as the number of datacentres increase, more options are available for the RRHs, thus reducing their average distance between themselves and the datacentres. This factor might not be of very importance for this case, but in a bigger area, where problems due to latency exists it could have a great influence, as it is explained in Section 4.3.

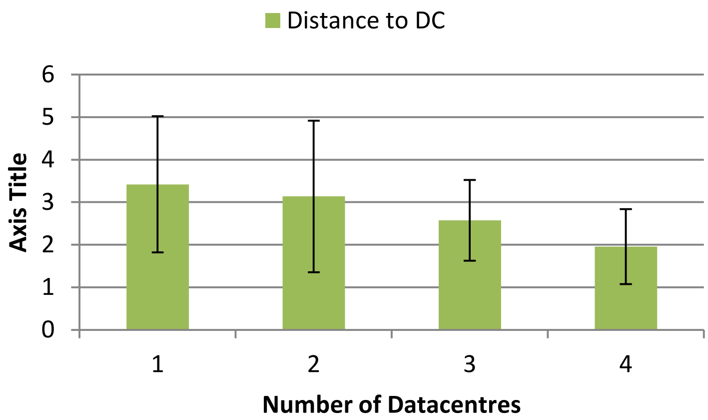


Figure 4.29: Distance between RRHs and datacentres for different number of deployed datacentres.

These last results imply that for the region of Lisbon one single datacentre would be enough to serve all of the existing RRHs with a relatively low load, in the order of 50%. Nonetheless, it is important to mention that with the implementation of 8×8 MIMO, the use of high bandwidths and an increase in the number of users, it would be necessary to deploy at least 3 datacentres. Note that all these values would depend on the limitations of the datacentres, which is still an unknown factor since it is required more testing on the virtualisation techniques and equipment.

4.3 Theoretical Analysis

Since there are no latency problems for the region of Lisbon analysed in the previous simulations, this section takes into account a larger area, surrounding the main cities and places outside of Lisbon, as mentioned in Section 4.1.2.

From Section 4.2.1, it is possible to conclude that the maximum distance allowed between the RRHs and the datacentres was around 25 km. Consequently, the total area considered can be cut into a smaller area with a maximum distance of 25 km. This reduction is shown in Figure 4.30, and by observation, it is assumed that 15% of the total area suffers from latency problems, which refers to the area outside of the circular region. Thus, most of the RRHs located in those regions are not able to connect to the datacentres in Lisbon. This value is a rough estimation, and, even though some of the outer regions are further away than 25 km, they could still be served by datacentres located at the borders of Lisbon, since there is approximately a 10 km diameter between the centre and border zones of the city. Nevertheless, these values allowed to obtain an estimation of latency problems occurring in this area.

Figure 4.30: Area outside of the circle might have latency problems.

In a first approach, the total number of extra RRHs is calculated, resulting in around 5 100 extra RRHs, as illustrated in Figure 4.31. It is only considered the extra RRHs because the results obtained on the previous analysis were added up to the new values. This increase is acceptable since the area considered is 20 times bigger than the area studied before, but there are much less populated areas and some big areas that include mountains and forests that have a very low number of BSs. In order to obtain this value, the number of additional BSs is calculated using the values from Table 4.5 and multiplied by 3 in order to get a total number of RRHs. Since some RRHs might not be active, this total number was reduced by a factor of 25 %, which is close to the value obtained in the other simulations. Figure 4.31 also shows the number of active RRHs, which was obtained after calculating the number of RRHs that would have latency problems, which is around 760.

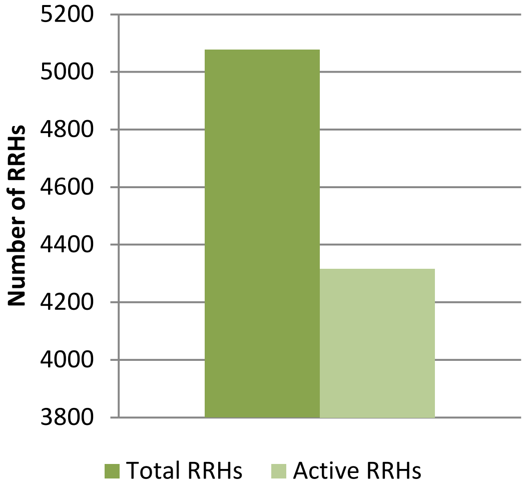
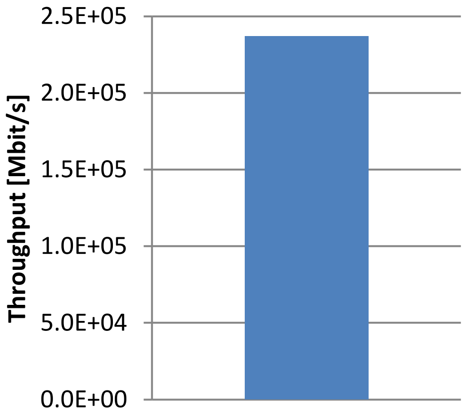
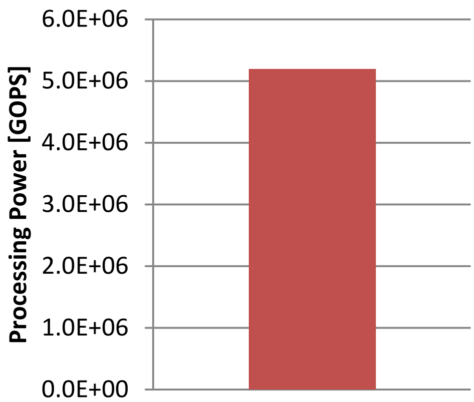


Figure 4.31: Number of extra RRHs.

Taking the values of average throughput and processing power per RRH obtained for the reference scenario, the total network throughput and processing power were calculated, as illustrated in Figure 4.32. The obtained value for the throughput of the extra RRHs is around 240 Gbit/s, and considering the same limitations as in the reference scenario, it is enough to achieve the maximum load at each datacentre. The maximum value for all the datacentres is 180 Gbit/s and considering the required network throughput for the Lisbon area (around 30 Gbit/s), around 90 Gbit/s are left without being able to connect to any datacentre. Therefore, the number of RRHs without connectivity to datacentres would be around 1 640, as demonstrated in Figure 4.33.



(a) Total network throughput



(b) Total network processing power

Figure 4.32: Total network throughput and processing power for the whole area.

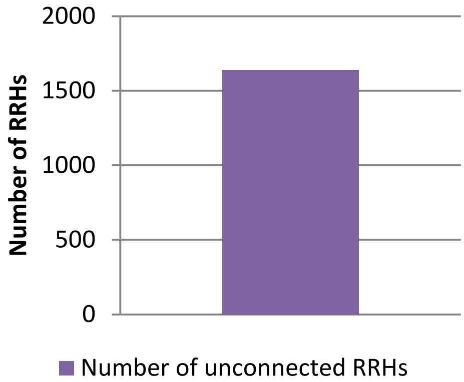


Figure 4.33: Number of RRHs without connection due to throughput limitations.

Including the RRHs with latency problems, for the region in study, the resulting capacity is around 50%, as pictured in Figure 4.34. It is possible to conclude that using a larger area, latency problems would arise, and quite a lot of RRHs would have problems to connect since the datacentres would be full. Nonetheless, if more datacentres would exist and their location chosen carefully, it could be enough to cover the whole region. This is quite a good result, despite all the limitations, because if well planned, a few datacentres would be able to cover huge regions, proving one of the main objectives of this work.

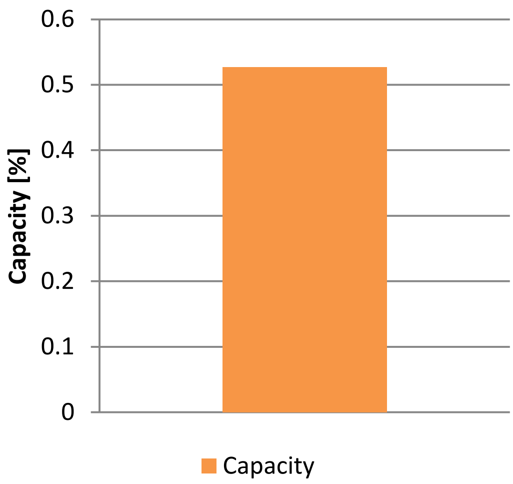


Figure 4.34: Total network capacity.

Chapter 5

Conclusion

This chapter concludes the developed work by summarising the main conclusions obtained, and gives an overview of some aspects to be developed in future work.

As mentioned before, the main scope of this thesis was to analyse the introduction of datacentres in a network, coming from the separation of BSs into RRHs and BBU-pools, in order to simulate a centralised architecture, and to implement a load balancing algorithm that associates RRHs to the datacentres, depending on some key parameters of LTE-A. The implemented algorithm was tested through intensive simulations with a very high loaded network, the number of users being above 12 000 and the only service being FTP, which is the most demanding in terms of throughput.

Chapter 2 begins with an explanation of the LTE-A network architecture by identifying the main components and their functionalities. It is followed by the Radio Interface, which addresses the main aspects of spectrum distribution, multiple access schemes used, resource blocks organisation, modulation used and some of the key aspects and new functionalities of LTE-A. The third section introduces the C-RAN architecture, explaining the main differences between a centralised and decentralised one, its main components and their functionalities (RRHs and BBUs), different ways to split functions between the components, and concludes with the advantages and constrains of these type of architecture. The fourth section is the state of the art, which describes the latest developments on C-RAN, focusing on aspects relative to the load balancing, virtualisation techniques, and some performance enhancements relative to the normal architecture. To conclude, the fifth section describes the different services and applications used on LTE-A, enumerating their priority, required data rates and associated errors and delay tolerances.

Chapter 3 presents the developed model in order to access the performance on the datacentres and the distribution of the traffic load among them. The first section is divided into five sub-sections, each one addressing a specific performance parameter that is of great importance for the developed load balance algorithm. These parameters are latency, processing power, RRH positioning, capacity and throughput. The second section takes into account these parameters and gives a detailed explanation on how the algorithm works in order to distribute the RRHs throughout the datacentres, based on the key parameters explained in the first section. Section 3.3 gives a full explanation of the simulator. It starts with the overview, which explains how the simulator is organised, and gives a brief explanation on the main functions of each component. It is followed by the algorithm implementation, which provides detailed information on how the algorithm was implemented on the simulator, stating the main changes and new functionalities provided. Finally, it concludes with the description of the Input and Output files and the information given in each of the files. To conclude Chapter 3, there is an assessment of the results obtained, in order check the validity of algorithm.

Chapter 4 starts by explaining the reference scenarios, which consist of the city of Lisbon and in a peripheral area that includes the city of Setubal. This was considered a very heterogeneous network since the distribution of BSs and users along the different areas of the city are not uniform. It is important to note that this scenario does not take into account variations such as the different terrain tilt or some high buildings that could provide higher propagations losses. Rough approximations were made regarding the connections between the RRHs and the datacentres, as well as the number of active RRHs. Since the centralised architecture requires an optical link between each RRH, or each group of co-located RRHs, to the datacentres, in order to obtain accurate values it would be required to have a map of the optical links already deployed in the area, thus obtaining the true distance between both components. Regarding the number of RRHs, not all the deployed BSs are tri-sectorised, and thus the value obtained for the number of active RRHs, which correspond to the number of active sectors, could be higher than reality. It is important to mention that no interference was taken into account, as it would decrease quite significantly the number of used RBs and thus the total network throughput and processing power. Nonetheless, since this study focuses on a worst case scenario in terms of network load, this assumption is not very relevant.

A very high number of simulations were made using the reference scenario, as well as simulations with successive variations of one parameter at a time, thus observing its influence on the total network analysis. The parameters tested were bandwidths, MIMO configurations, datacentre throughput limits and number of deployed datacentres. Some approximations were also made regarding the tested bandwidths and MIMO configurations, since the values used correspond to the standard values from 3GPP for LTE-A. The simulator was not configured to support these values, and thus a very rough implementation was made since it was not within the scope of this thesis. The network was composed of almost 300 BS and with a number of users between 9 000 and 12 000, which could make some parameters suffer big variations making them harder to analyse due to the randomness and heterogeneity of the network. But, on the other hand, since the performed analysis was made on a network layer, instead of the user one, the values obtained are very conclusive and easy to analyse.

The results obtained for the reference scenario are quite optimistic, which makes sense since there was no interference taken into account. The average throughput per RRH is around 55 Mbit/s while the processing power is 1200 GOPS. This results in a network throughput of around 32 Gbit/s and a total processing power of 75×10^4 GOPS. Another important parameter taken into consideration was the average distance between the RRHs and the datacentres which was around 2.5 km, an expected result since Lisbon has a radius of around 5 km. Considering the results obtained and the fact that they were acquired for a worst case scenario, it is possible to say that the city of Lisbon can be covered with a single or a small number of datacentres. But note that this value can be susceptible to changes since the assumed value for the datacentres limitation, in terms of throughput or processing power, was a rough estimation and the real values are not known. This is due to the fact that most of virtualisation technology is still in development, meaning that the exact number for the limitation in the processing power of a BBU-pool, as well as the number of RRHs that can be processed on each one, are unknown. By changing the values of bandwidth and MIMO, the maximum values obtained for the throughput per RRH, and consequently the total network throughput, result in an increase of more than 200 %, meaning that the throughput per RRH is over 150 Mbit/s and the network throughput over 100 Gbit/s. This value is still low considering the limitations chosen, proving that, even with the implementation of new functionalities that increase quite a lot the throughput and resource usage, all the city of Lisbon can still be covered by a few datacentres. As for the processing power, the values obtained in relation to the bandwidth and MIMO configuration follow a quadratic evolution. As mentioned before, there is no information relative to the limitations on the BBU-pools, thus simulations were made where different values for this datacentre limitation were tested. It is important to note that on these simulations, the network was configured with the values that enable the maximum possible throughput, which are the 100 MHz bandwidth with an 8×8 MIMO configuration. For this case, the RRHs average throughput is around 400 Mbit/s, while the total network throughput and processing power are 250 Gbit/s and 1.2×10^7 GOPS, respectively. The limitations values used range from 50 000 GOPS until 100 000 GOPS, which proves that, by using the reference number of datacentres, problems with capacity would appear for the lower limitation values. Nevertheless, up from 90 000 GOPS it would be possible to cover the whole city without any constrains. On these last simulations the position of the datacentres and RRHs was kept equal, meaning that there are no significant fluctuations on the average distance connecting the two components. Therefore, simulations were made using the reference scenario and network configurations, but with different number and locations of datacentres, in order to check how it would affect the datacentres load and the network capacity. The obtained results were as expect, showing that the datacentres located nearest the city centre have a much higher load than the others (in the case of more than one datacentre). Also, it was possible to show that even with one datacentre the city of Lisbon can be completely covered, the load of the datacentre being below 60 %.

Since there are no problems due to latency, to conclude Chapter 4, an extra analysis was made considering a much larger area than the city of Lisbon, which includes the main cities and regions to the south and north, in a radius of approximately 30 km from the centre of Lisbon. In this analysis, some rough estimations were made by analysing the region maps and BSs locations, in order to obtain values for density of BS per km^2 and to categorise the main regions into urban and rural, according to that parameter. It was considered that 25 % of the RRHs had no users, thus obtaining a total increase of almost 5100 RRHs. Since the maximum possible distance between the RRHs and datacentres is around 25 km, a margin of 15 % was assumed for RRHs that would have latency problems, which means that around 730 RRHs would have connectivity problems. Using the previous obtained values for average throughput and processing power per RRH, the resulting total network throughput is almost 250 Gbit/s, while the processing power is 5×10^6 GOPS. This concludes that in this case there would be a decrease in the network capacity, obtaining a value close to 50 %, which means that, besides the RRHs with latency problems, there would be around 1640 RRHs without connectivity due to throughput limitations, giving a total number of unconnected RRHs around 2400. Nevertheless, the obtained results are quite positive, since the considered number of datacentres is very low, and thus it is possible to conclude that by increasing the number of deployed datacentres, a much larger area than the city of Lisbon can be covered without problems.

The simulator proves to be a reliable tool since it offers a general picture of how the load balancing of traffic would happen in a LTE-A network, in spite of all the approximations made. Nevertheless, it is based on a snapshot, and thus there is no temporal variation, which affects the scheduling of RBs to users, and the way that RRHs connect to datacentres. This last point has a high relevance for the purpose of this work because the algorithm developed is supposed to occur over some timing reference, meaning that the connections between RRHs and datacentres would change over the course of the day, based on the required throughput and processing power requested at a specific time. Even though this main point could not be analysed using this simulator, the obtained results are able to give an overview of how the distribution of traffic among the datacentres could occur.

The main observations and results obtained throughout this thesis are all based on the datacentre limitations in terms of processing power/throughput, and thus they are very susceptible to changes. But nonetheless, the results confirm that a few datacentres can be used to cover the city of Lisbon and some of its whereabouts, performing traffic balancing between them and thus proving that a centralised architecture can be a very effective and efficient upgrade to the network actual infrastructure, proving one of the main objectives of this thesis.

Regarding future work, some improvements have to be made in order to properly implement the new features of LTE-A, such as CA and higher MIMO configurations as it would significantly affect the throughputs and processing power of the RRHs. Another important enhancement would be the introduction of a temporal analysis where users and active RRHs would change over a specific timing interval, in order to check the fluctuations on the network traffic. Relaying and small cells introduction could also be implemented as their introduction would increase the network traffic. Besides these general LTE-A aspects, it would be important to add a map with all the available optical links in the considered area, in order to obtain a value for the distance between RRHs and datacentres closer to reality since there is almost no direct connections between them. As mentioned before, it is also necessary to correctly quantify the datacentres limitations, since it is one of the main constrains over all the simulations. This might be possible in the nearest future since the centralised architecture and virtualisation techniques are currently under development.

Appendix A

Frequency bands assignment

This appendix presents the available frequency bands for LTE-A, their respective bandwidths and which bands were assigned to which Telecommunication Operator.

ANACOM is the Portuguese authority that is responsible for the spectrum distribution among the Mobile Operators. The auction took place at November 2011 are the results are presented in TABLE. The blocks that lack operator are the ones that did not receive any bid. The total value of the available blocks sold was €372.000.000.

Table A.1: ANACOM auction results.

Block	Frequency Band	Block Size	Price	Operator assigned
A1	450	2×1.25	-	-
B1	800	2×5	45.000.000	TMN
B2	800	2×5	45.000.000	TMN
B3	800	2×5	45.000.000	Vodafone
B4	800	2×5	45.000.000	Vodafone
B5	800	2×5	45.000.000	Optimus
B6	800	2×5	45.000.000	Optimus
C1	900	2×5	-	
C2	900	2×5	30.000.000	Vodafone
D1	1800	2×5	4.000.000	TMN
D2	1800	2×5	4.000.000	TMN
D3	1800	2×5	4.000.000	Vodafone
D4	1800	2×5	4.000.000	Vodafone
D5	1800	2×5	4.000.000	Optimus
D6	1800	2×5	4.000.000	Optimus
D7	1800	2×5	-	-
D8	1800	2×5	-	-
D9	1800	2×5	-	-

E1	1800	2×4	3.000.000	TMN
E2	1800	2×4	3.000.000	Vodafone
E3	1800	2×4	3.000.000	Optimus
F1	2100	5	-	-
F2	2100	5	-	-
G1	2600	2×5	3.000.000	TMN
G2	2600	2×5	3.000.000	TMN
G3	2600	2×5	3.000.000	TMN
G4	2600	2×5	3.000.000	TMN
G5	2600	2×5	3.000.000	Vodafone
G6	2600	2×5	3.000.000	Vodafone
G7	2600	2×5	3.000.000	Vodafone
G8	2600	2×5	3.000.000	Vodafone
G9	2600	2×5	3.000.000	Optimus
G10	2600	2×5	3.000.000	Optimus
G11	2600	2×5	3.000.000	Optimus
G12	2600	2×5	3.000.000	Optimus
G13	2600	2×5	-	-
G14	2600	2×5	-	-
H1	2600	25	3.000.000	Vodafone
I1	2600	25	-	-

Appendix B

Processing Power Complexity Tables

In this appendix, it is shown some auxiliary values required for the calculation of the Processing Power per RRH.

As explained in chapter 3.1.2 the calculations for the processing power depend on the flow of information, whether it is UL or DL, and on the type of BS. The following tables show the different values of complexity for each function required for the processing power formula.

Table B.0.1: Complexity of baseband operations for DL.

GOPS per function	Macro	Micro	Pico	Femto
DPD	160	160	0	0
Filter	200	160	120	100
CPRI	360	300	0	0
OFDM	80	80	70	60
FD (linear)	30	30	20	20
FD (non-linear)	10	10	5	5
FEC	20	20	20	20

Table B.0.2: Complexity of baseband operation for UL.

GOPS per function	Macro	Micro	Pico	Femto
Filter	200	160	160	150
CPRI	360	300	0	0
OFDM	80	80	80	60
FD (linear)	60	60	40	30
FD (non-linear)	20	20	10	10
FEC	120	120	120	110

Appendix C

Throughput Equations

In this section the equations used to calculate the throughput, relative to a given SNR, are presented.

The throughput equation used were derived in [Alme13], based on measurements performed by Mobile Operators, presented in Figure C.1.

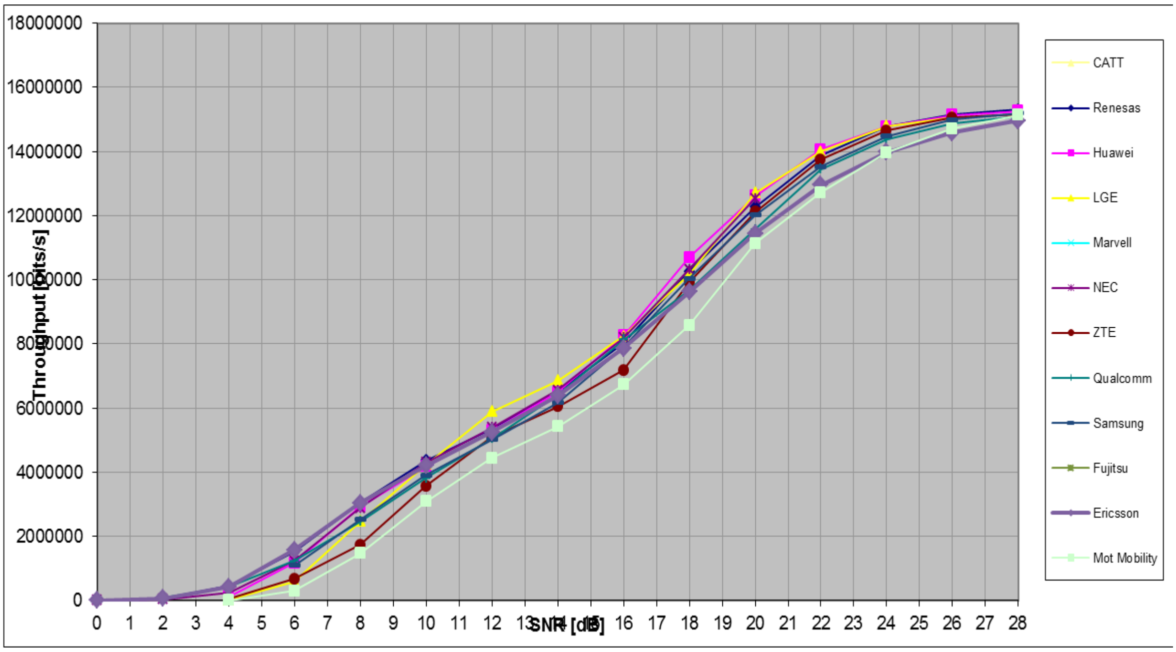


Figure C.1: Measurements from different operators, relating the throughput with the obtained SNR.

To obtain the equations, an extrapolation of those functions was made, resulting in:

$$R_{b1}[Mbit/s] = \frac{1.027 \times 3 \times C_{max,1} \times 2280441.2727217874}{(14.005140037510571 \ + \ e^{(-0.5778969006043214 \times SNR)}) \times 10^6} \tag{C.1}$$

$$R_{b2}[Mbit/s] = \frac{1.027 \times 2 \times C_{max,1} \times 47613.05094787189}{(0.0926274904521111 \ + \ e^{(-0.295838412098985 \times SNR)}) \times 10^6} \tag{C.2}$$

$$R_{b3}[Mbit/s] = \frac{1.027 \times 2 \times C_{max,1} \times 17603.857570084674}{(0.022018582206702525 \ + \ e^{(-0.2449102441508849 \times SNR)}) \times 10^6} \tag{C.3}$$

where:

- R_{bx}, is the obtained throughput for each modulation,
- $C_{max,x}$, is the maximum coding rate for each modulation.

On Table C.1 it is shown the considered maximum coding rates and the modulations that each throughput equation represents:

Table C.1: Maximum coding rates and modulation used.

Throughput Equation	Modulation	Coding Rate
1	QPSK	0.588
2	16QAM	0.602
3	64QAM	0.926

Appendix D

User’s Manual

This Appendix gives detailed instructions on how to configure the parameters and run a simulation

To run the simulator, one must start by running the file UMTS_Simul.MBX which will ask for 4 inputs files, as exemplified in Figure D.1:

- "Ant65deg.TAB", that contains the BS antennas gains;
- "Eb_No.TAB", not relevant for this thesis;
- "DADOS_Lisboa.TAB", with information regarding Lisbon different districts;
- "ZONAS_Lisboa.TAB", relative to the different areas characterisations.

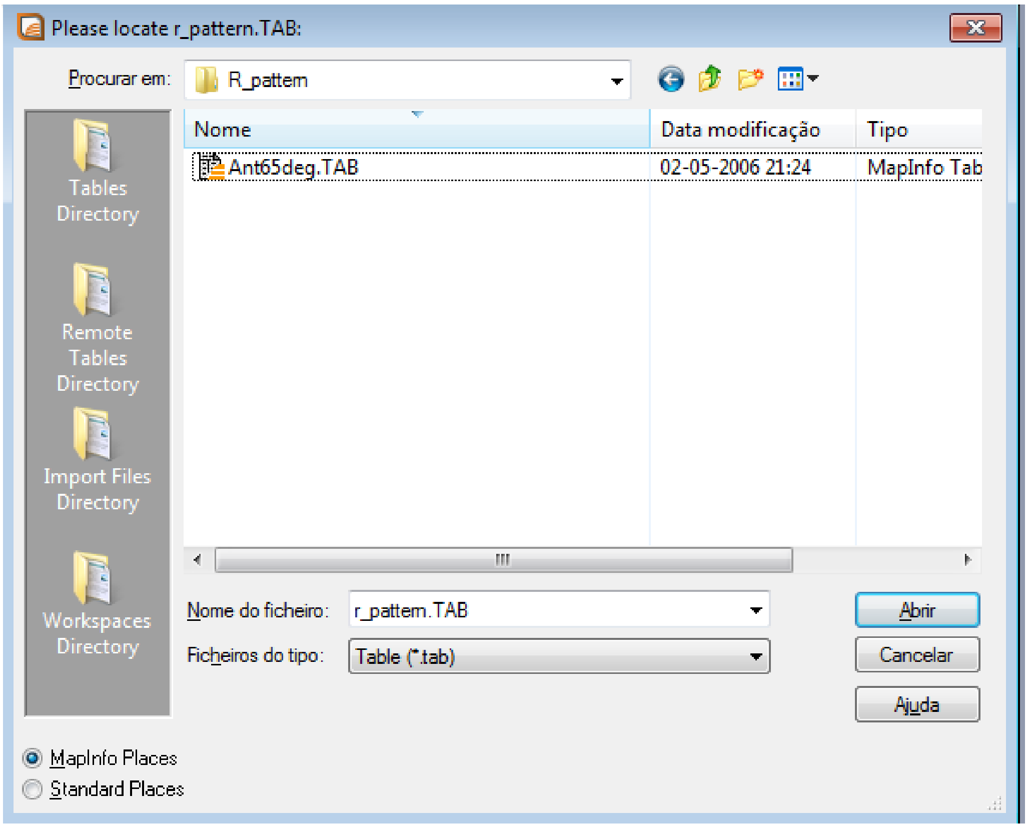


Figure D.1: Window for selection of the radiation pattern file.

This will give a picture of the city of Lisbon, divided into the districts, and enables the use of a new net tab called System, where it is possible to choose the LTE DL option, as shown in Figure D.2. Inside this last option one could choose from 5 different possibilities, being the most important ones:

- Propagation Model, where it is possible to define the parameters used for the propagation models used, such as building height, etc.;
- Latency Model, which enables the definition of delay parameters and margins, for example processing delay and fibre velocity, illustrated in Figure D.3;
- Network Settings, where most network configurations and algorithms are chosen, as illustrated in Figure D.5;

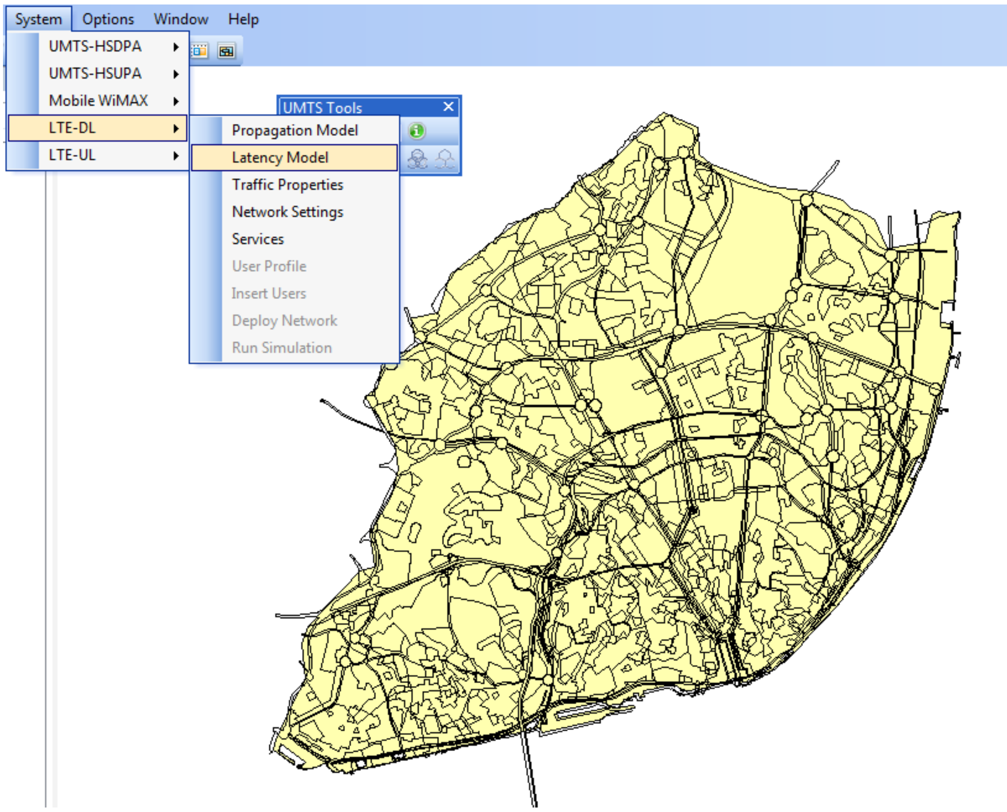


Figure D.2: System net tab and map of Lisbon.

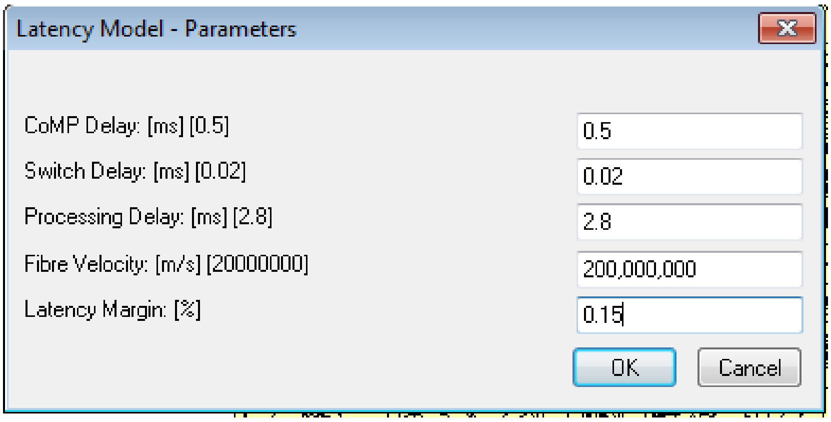


Figure D.3: Latency parameters.

After defining the Network Settings, the User Profile tab becomes enabled. This window defines the request throughput, and the minimum required for each service, as can be seen in Figure D.5. After hitting the OK button, the Insert Users tabs unlocks itself. This allows to choose the user input files as showed in Figure D.6. For big user files, this process could take a while, and the resulting output is shown in Figure D.7, as well as the new option available.

The new available option is the Deploy Network tab, which allows to chose the tables containing the BSs and datacentres locations (if the datacentre swith is on). By clicking on it, a window appears requesting the BSs table, followed by another request, this time for the datacentres table, as exemplified in Figure D.8.

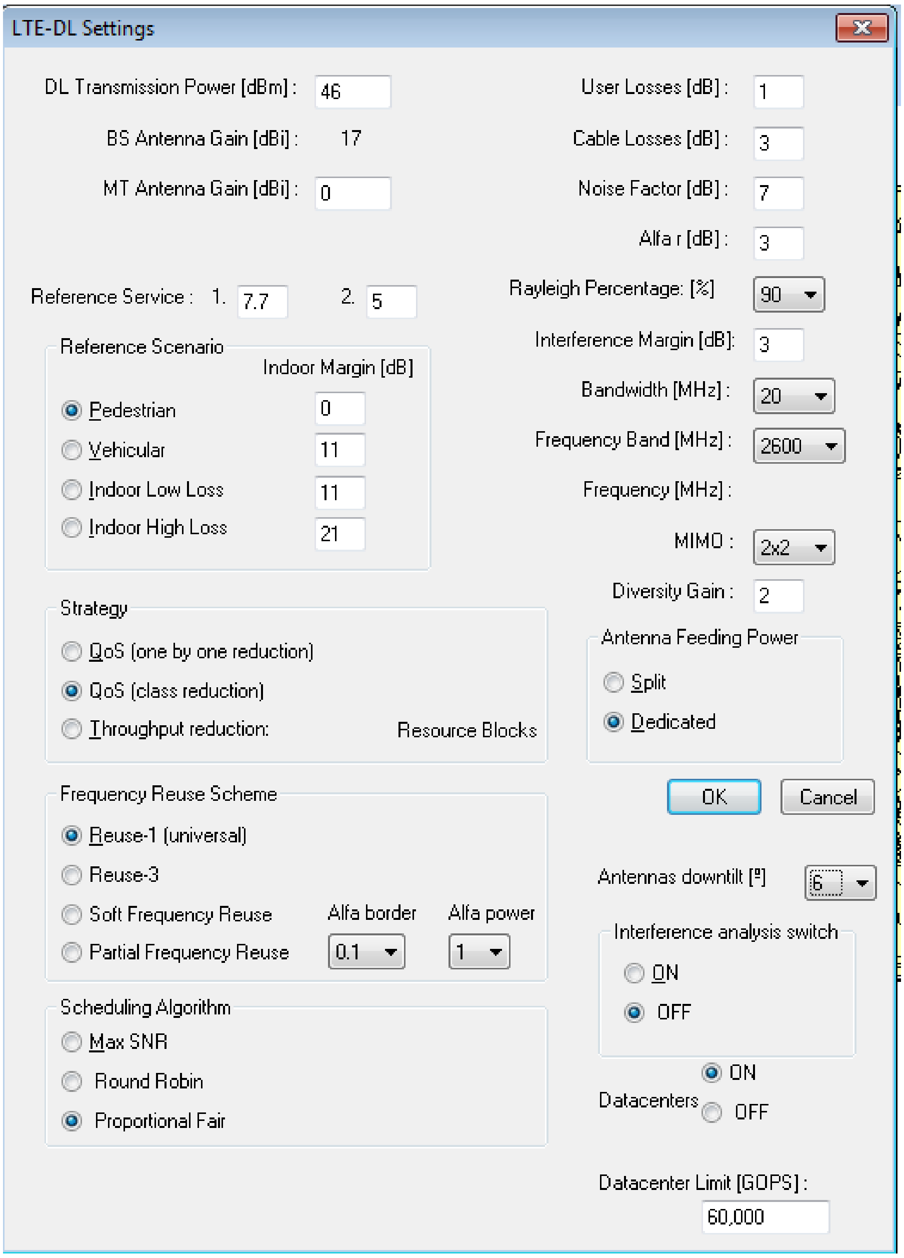


Figure D.4: Network Settings

The deployment of the BSs might also take a while to complete, and the obtained output is presented in Figure D.9. After it is complete, the option Run Simulation unlocks, which will run the implemented models and algorithms. To conclude, some statistics are shown presenting relevant information regarding the network. Although, most of these results were not relevant for this thesis, as those results are written in some other output files, as explained in Section 3.3.3.

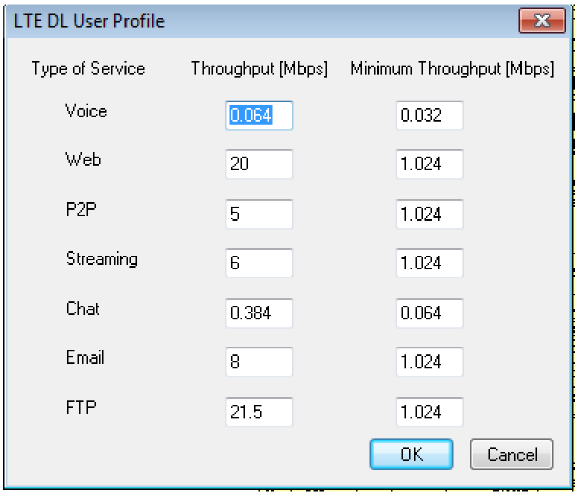


Figure D.5: User Profile regarding throughputs for each service.

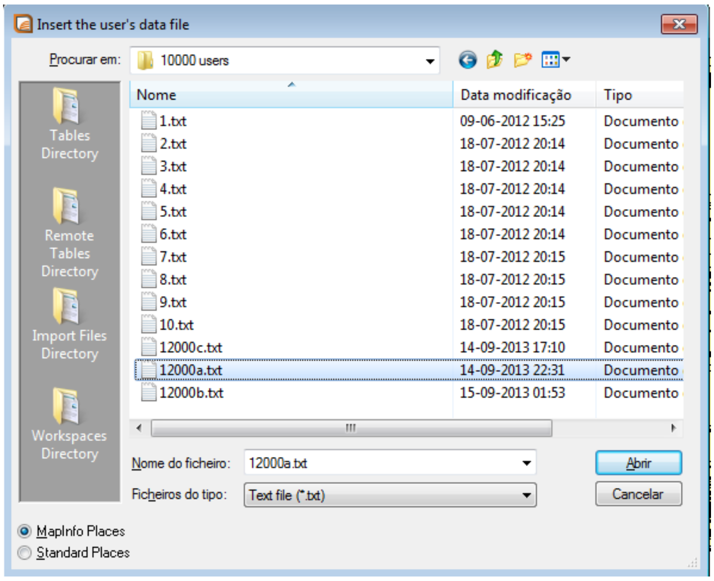


Figure D.6: Window to select the users input file.

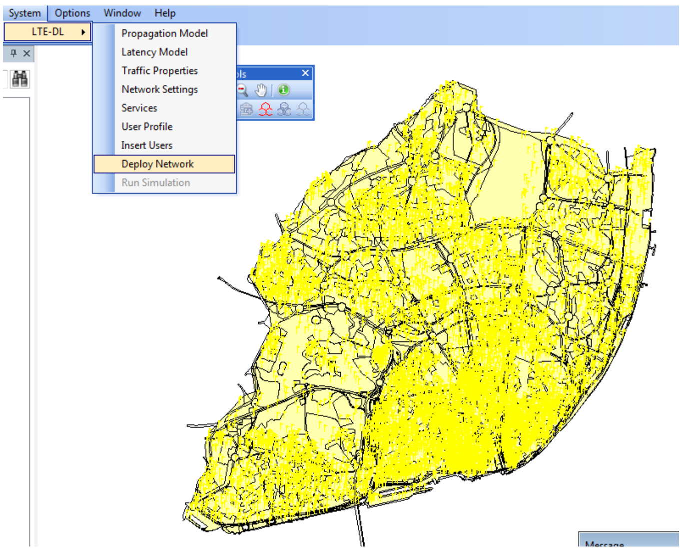


Figure D.7: Map with users deployed and net unlocked option, Deploy Network.

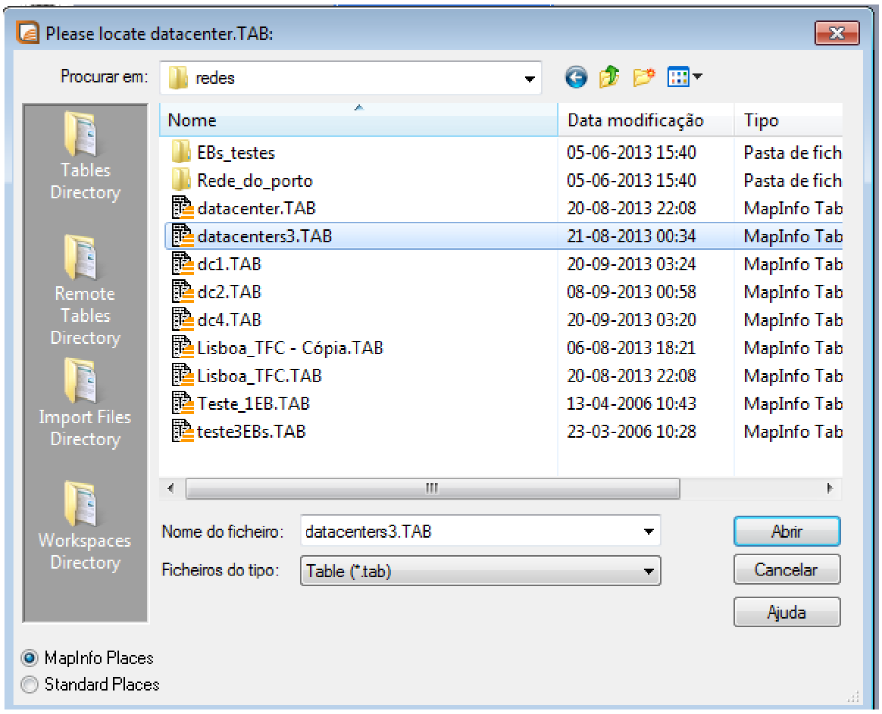


Figure D.8: Window requesting the datacentre tables.

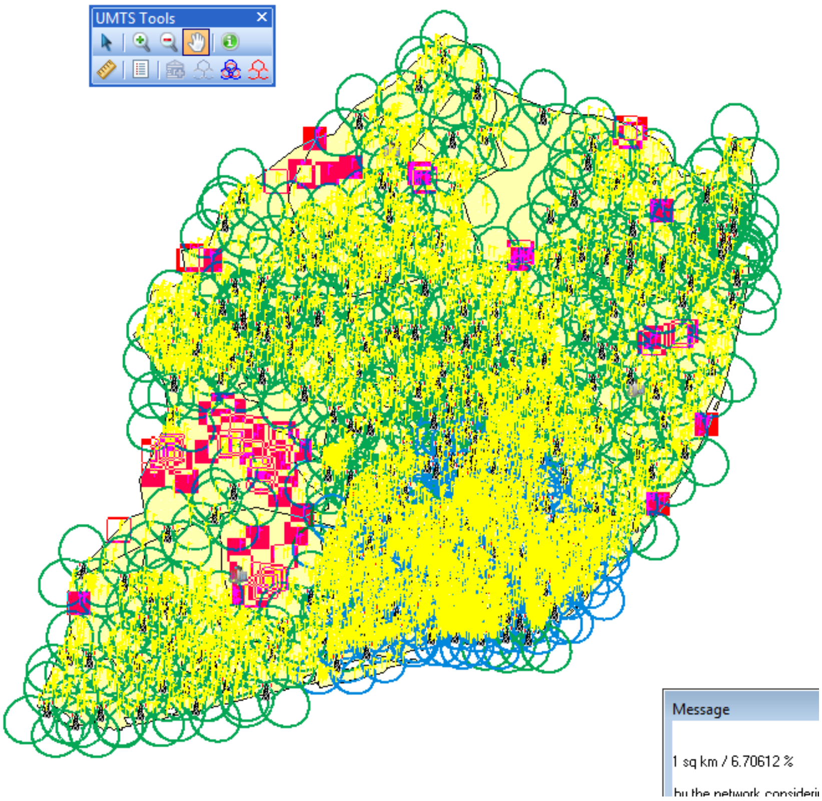


Figure D.9: Network with the deployed users and BSs.

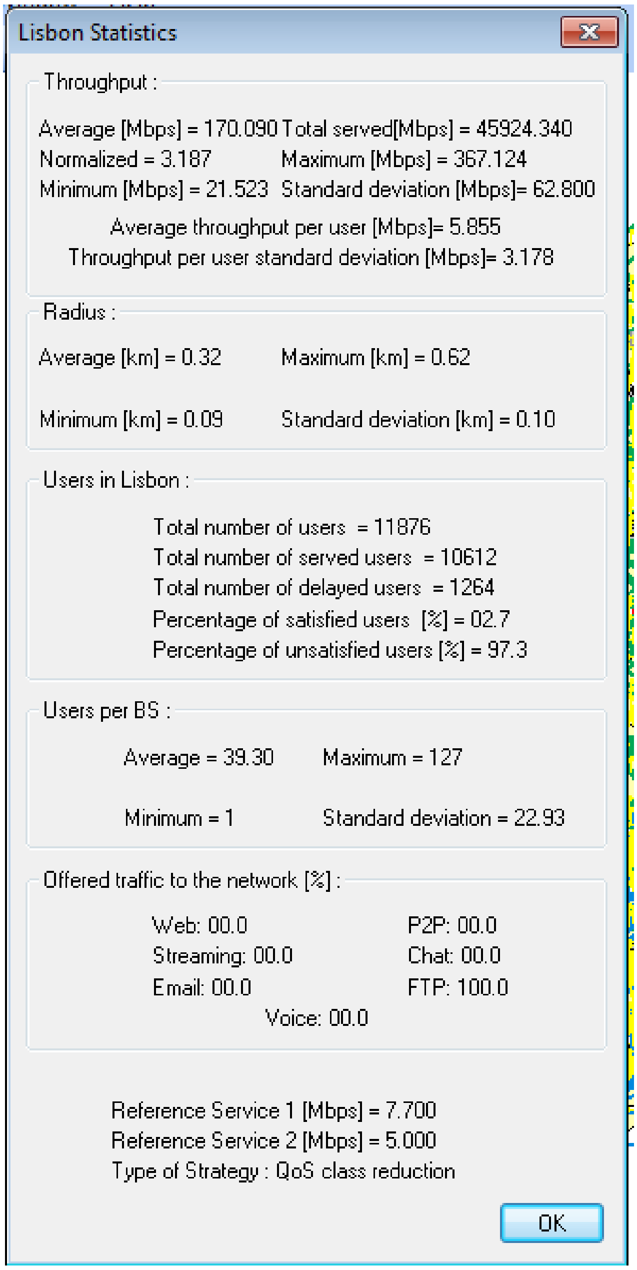


Figure D.10: Statistics obtained in the Map Info program.

Appendix E

Datacentres Maps

This Appendix provides information regarding the location of the datacentres used throughout the simulations

The following figures are representative of the different datacentres locations used in Chapter 4 during the simulations

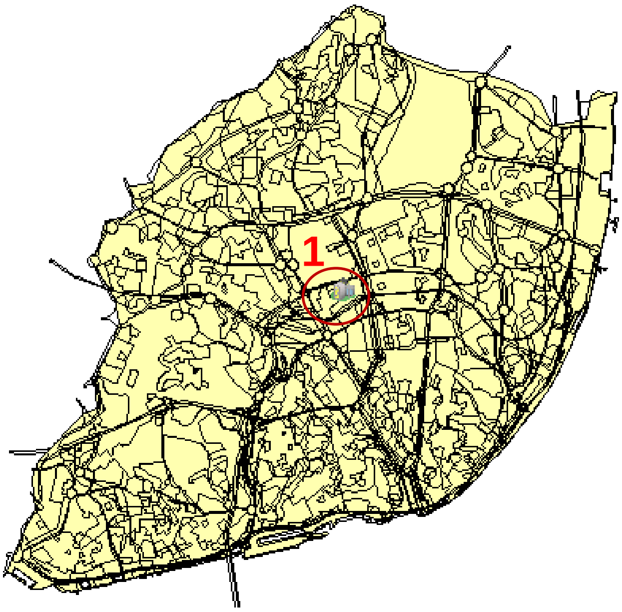


Figure E.1: Lisbon with one datacentre deployed.

Figure E.2: Lisbon with two datacentres deployed.

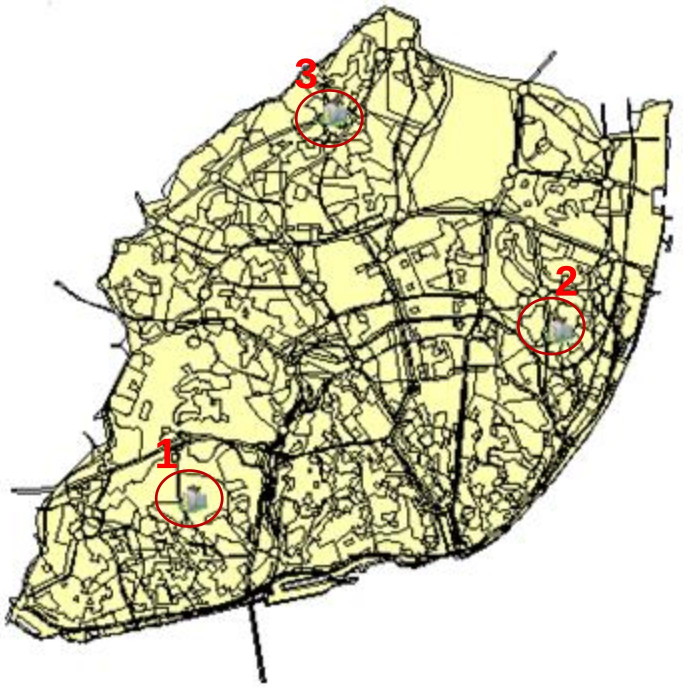


Figure E.3: Lisbon with three datacentres deployed.

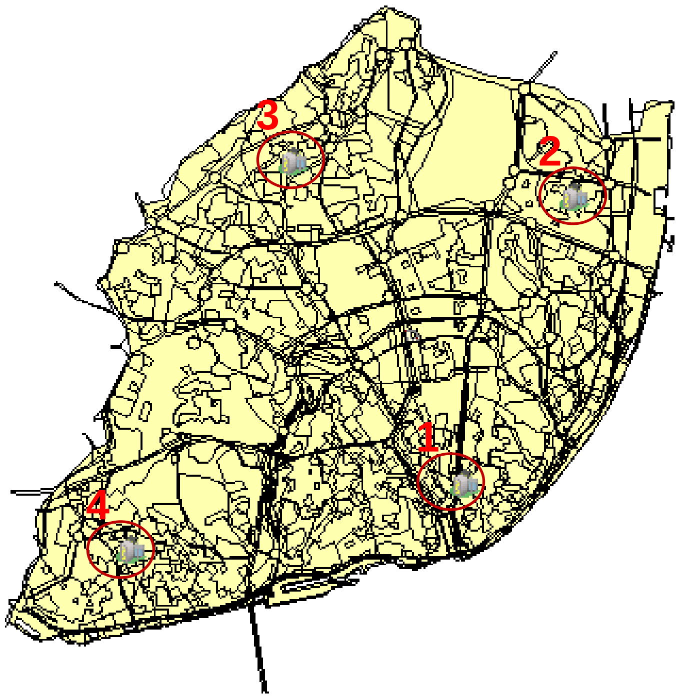


Figure E.4: Lisbon with four datacentres deployed.

Appendix F

COST 231 Walfisch-Ikegami

This appendix presents information regarding the propagation model used to calculate the path loss

To calculate the path loss in urban scenarios, the model chosen was the COST231 Walfisch-Ikegami, which consists of the parameters defined in Figure F.1.

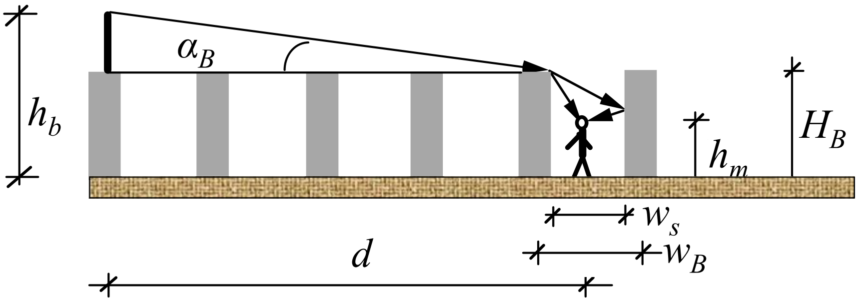


Figure F.1: Walfisch-Ikegami parameters (taken from [Corr09]).

In case of LoS propagation, one has:

$$L_{p[\text{dB}]} = 42.6 + 26\log d_{[\text{km}]} + 20\log f_{[\text{MHz}]} \text{ , for } d > 0.02 \text{ km} \tag{F.1}$$

where:

- d, is the horizontal distance between the user and the BS,
- f, is the frequency of the signal.

On the other hand, for NLoS the equations are:

$$L_{p[\text{dB}]} = \begin{cases} L_{0[\text{dB}]} + L_{rt[\text{dB}]} + L_{rm[\text{dB}]} & \text{, for } L_{rt} + L_{rm} > 0 \\ L_{0[\text{dB}]} & \text{, for } L_{rt} + L_{rm} \leq 0 \end{cases} \tag{C.2}$$

where:

- L_0, corresponds to the free space propagation loss,
- L_{rt}, is the loss between the last rooftop and the MT,
- L_{rm}, is the loss between the BS and the last rooftop.

The L_0 is calculated using equation C.3:

$$L_{0[\text{dB}]} = 32.44 + 20\log d_{[\text{km}]} + 20\log f_{[\text{MHz}]} \tag{C.3}$$

The loss between the rooftop and the MT is given by:

$$L_{rt[\text{dB}]} = L_{bsh[\text{dB}]} + k_a + k_d \log d_{[\text{km}]} + k_f \log f_{[\text{MHz}]} - 9\log w_{B[\text{m}]} \tag{C.4}$$

where:

- L_{bsh}, loss between the difference in heights between the rooftops and the antennas,
- k_a, increase in path loss for BSs antennas below the adjacent buildings,
- k_d, dependence of multi-screen diffraction loss versus distance;
- k_f, equal to k_d but relative to the frequency;
- w_B, building separation distance.

The loss between the differences in heights is given by:

$$L_{bsh[\text{dB}]} = \begin{cases} -18\log\left(h_{b[\text{m}]} - H_{B[\text{m}]} + 1\right), & \text{for } h_b > H_B \\ 0, & \text{for } h_b \leq H_B \end{cases} \tag{C.5}$$

where:

- h_b, is the BS height;
- H_B, is the buildings height.

The way to calculate the other parameters are shown in equations C.6 to C.8:

$$k_a = \begin{cases} 54, & \text{for } h_b > H_B \\ 54 - 0.8\left(h_{b[\text{m}]} - H_{B[\text{m}]}\right), & \text{for } h_b \leq H_B \text{ and } d \geq 0.5 \text{ km} \\ 54 - 1.6\left(h_{b[\text{m}]} - H_{B[\text{m}]}\right)d_{[\text{km}]}, & \text{for } h_b \leq H_B \text{ and } d \geq 0.5 \text{ km} \end{cases} \tag{C.6}$$

$$k_d = \begin{cases} 18, & \text{for } h_b > H_B \\ 18 - 15\frac{h_b - H_B}{H_B}, & \text{for } h_b \leq H_B \end{cases} \tag{C.7}$$

$$k_f = \begin{cases} -4 + 0.7\left(\frac{f_{[\text{MHz}]}}{925} - 1\right), & \text{for urban and suburban scenarios} \\ -4 + 1.5\left(\frac{f_{[\text{MHz}]}}{925} - 1\right), & \text{for dense urban scenarios} \end{cases} \tag{C.8}$$

Lastly, the loss between the BS and the rooftops is obtained using equations C.9:

$$L_{rm[\text{dB}]} = -16.9 - 10\log w_{s[\text{m}]} + 10\log f_{[\text{MHz}]} + 20\log\left(H_{B[\text{m}]} - h_{m[\text{m}]}\right) + L_{ori[\text{dB}]} \tag{C.9}$$

where:

- w_s, is the width of the street,
- h_m, is the height of the MT;
- L_{ori}, is the street orientation loss.

The street orientation loss is given by:

$$L_{ori[\text{dB}]} = \begin{cases} -10.0 + 0.354\phi_{[°]}, & \text{for } 0° < \phi < 35° \\ 2.5 + 0.075\left(\phi_{[°]} - 35\right), & \text{for } 35° \leq \phi < 55° \\ 4.0 - 0.114\left(\phi_{[°]} - 55\right), & \text{for } 55° \leq \phi \leq 90° \end{cases} \tag{C.10}$$

where:

- ϕ is the angle of incidence of the signal into the buildings.

To conclude, this model is valid under the following conditions:

- $f \in [800, 2000]$ MHz;
- $d \in [0.02, 5]$ km;
- $h_b \in [4, 50]$ m;
- $h_m \in [1, 3]$ m.

It takes standard deviations ranging from 4 to 7 dB.

Appendix G

User Distribution

This Appendix describes the user distribution in the city of Lisbon.

As one can see from Figure G.1, most of the users are located in the centre area of Lisbon. The outer zones has a more heterogeneity in the user density. There are two distinguished places where there is almost no users, which are the Airport of Portela at north and the Forest Park of Monsanto at west. The other locations have a high density of users but still much inferior compared to the centre region. It is important to note that this user diversification coincides with the BSs distribution.

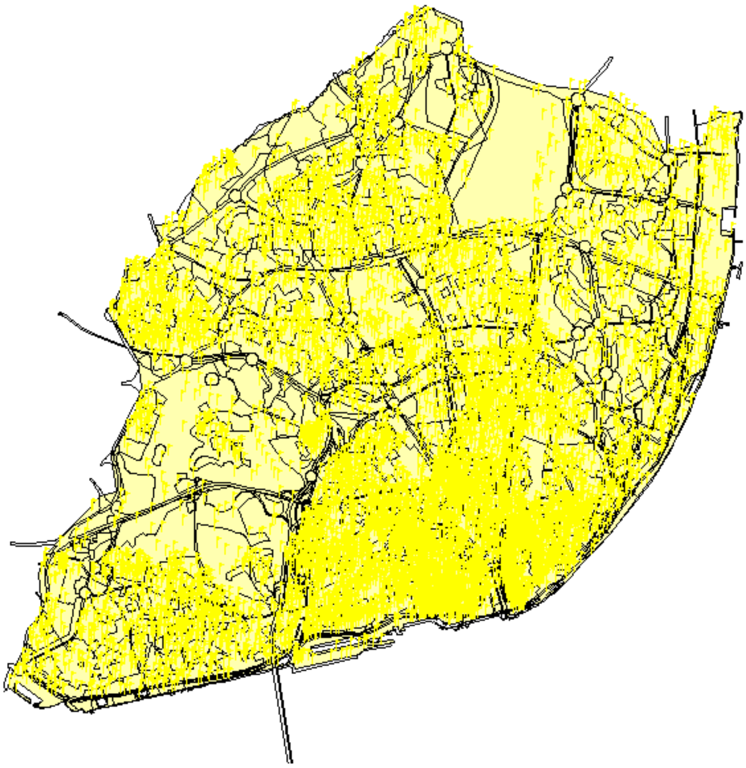


Figure G.1: User distribution in the Lisbon area.

Appendix H

Antennas' Radiation Patterns

This Appendix describes the antennas used throughout this simulator.

Figure H.1 illustrates both horizontal and vertical radiation patterns, used for the antennas in the simulator. In the figure, Frequenz 1 represents the horizontal pattern and Frequenz 2 the vertical one. Note that the vertical pattern has a 90º offset.

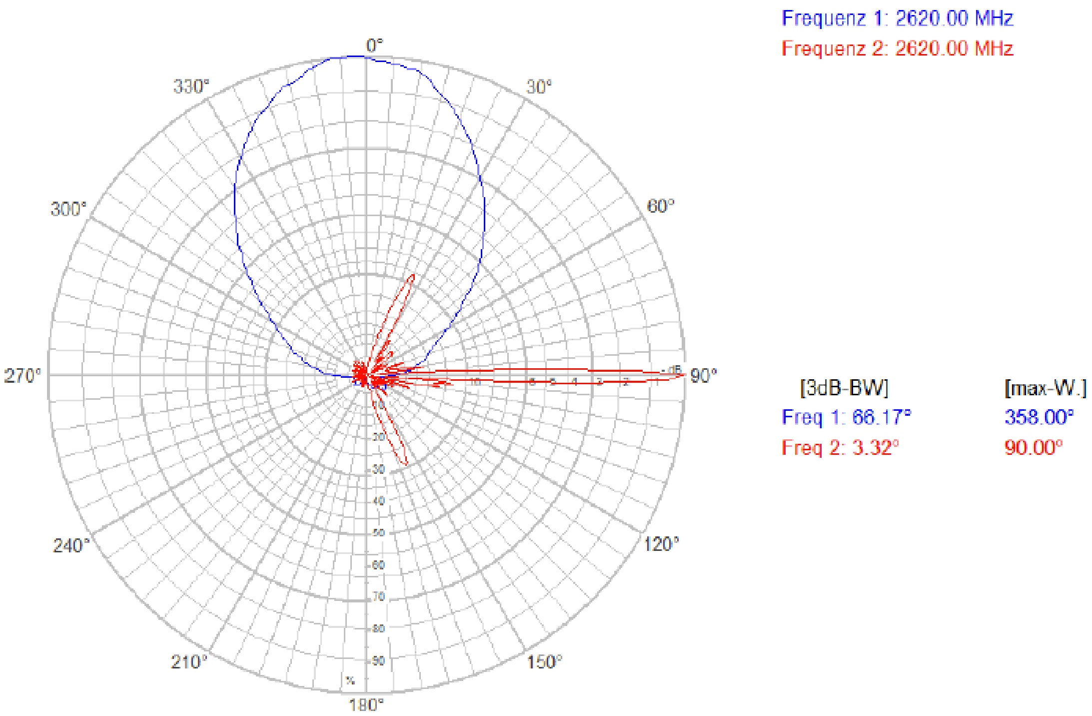


Figure H.1: Radiation pattern of the antennas used.

Appendix I

Maps and Information on the outskirts of Lisbon

This Appendix shows the maps used for the theoretical analysis, in Section 4.3.

Firstly, in Figure I.1: Total area analysed in Section 4.3.Figure I.1 one can observe the total area considered for the theoretical analysis.

Figure I.1: Total area analysed in Section 4.3.

By comparing the map files with the ones used on the Map Info program, the total area already studied was calculated, as shown in Figure I.2.

Figure I.2: Area already implemented in the Map Info analysis.

Lastly, Figure I.3 and Figure I.4 shows which areas were considered to be urban areas, thus having a higher BS density per square km. To conclude, Table H.1 shows the results obtained for the mentioned areas.

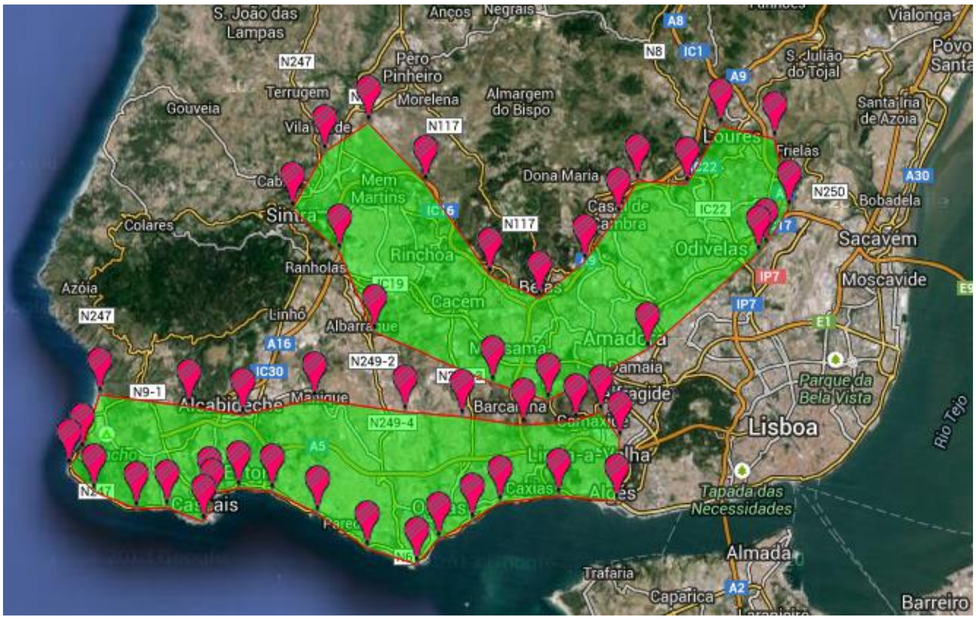


Figure I.3: Northern area of Lisbon considered urban.

Figure I.4: Southern area of Lisbon considered urban.

Table H.1: Considered areas for all the different regions.

Region	Area [km^2]
Total Area	1967
City of Lisbon previously studied	87
Northern Urban Area	172
Southern Urban Area	200

References

[3GPP12d] 3GGP, *Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control architecture*, ETSI TS, No. 23.203, Ver. 10.8.1, November 2012 (http://www.3gpp.org).

[3GPP12c] 3GPP, *Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Quality of Service (QoS) concept and architecture*, ETSI TS, No. 23.107, Ver. 10.2.0, 2012 (http://www.3gpp.org).

[3GPP12a] 3GPP, *LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2*, ETSI TS, No. 36.300, Ver 10.8.9, July 2012 (http://www.3gpp.org).

[3GPP12b] 3GPP, *LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation*, ETSI TS, No. 36.211, Ver. 10.5.0, July 2012.

[Alme13] Almeida, D., *Inter-Cell Interference Impact on LTE Performance in Urban Scenarios*, M.Sc. Thesis, Technical University of Lisbon, Lisbon, Portugal, 2013.

[Card11] Cardeiro, J., *Optimisation of Base Station Location in UMTS-FDD for Realistic Traffic Distribution*, M.Sc. Thesis, Instituto Superior Técnico, Lisbon, Portugal, Oct. 2011.

[CMRI10] China Mobile Research Institute White Paper, "C-RAN The Road Towards Green Radio RAN", 2010, (http://labs.chinamobile.com/report/view_34516).

[Corr09] Correia, L.M., *Mobile Communications Systems, Lecture Notes*, Instituto Superior Técnico, Lisbon, Portugal, 2009.

[DDGF12] Desset, C., Debaillie, B., Giannini, V. and Fehske, A., "Flexible power modeling of LTE base stations," in *IEEE Wireless Communications and Networking Conference: Mobile and Wireless Networks*, 2012.

[Duar08] Duarte, S., *Analysis of Technologies for Long Term Evolution in UMTS*, M.Sc. Thesis, Instituto Superior Técnico, Lisbon, Portugal, Sep. 2008.

[GuCT12] Guo, B., Cao, W. and Tao, A., "CPRI compression transport for LTE and LTE-A ," in *2012 7th International ICST Conference on Communications and Networking in China (CHINACOM)*, 2012.

[Habe12] Haberland, B., "Concept for load balancing in a radio access network," , 2012.

[Habe13] Haberland, B., "Smart Mobile Cloud," in *Conference on Networked Systems - KiVS 2013*, 2013.

[HoTo09] Holma, H. and Toskala, A., *LTE for UMTS - OFDMA and SC-FDMA Based Radio Access*, Wiley, Chichester, United Kingdom, 2009.

[Daft09] http://www.daftlogic.com/projects-google-maps-area-calculator-tool.htm, *Google Maps Area Calculator Tool*, 2009.

[Eric11] http://www.ericsson.com/news/1561267, , Nov. 2011. [Online]. http://www.ericsson.com/news/1561267

[LaCo06] Ladeira, D. and Costa, P., *Optimal planning of UMTS cellular networks for HSDPA data-based services (in Portuguese)*, M.Sc. Thesis, Instituto Superior Técnico, Lisbon, Portugal, Mar. 2006.

[LYWS12] Liu, L. et al., "Analysis of Handover Performance Improvement in Cloud-RAN Architecture," in *7th International ICST Conference on Communications and Networking in China (CHINACOM)*, 2012.

[Lope08] Lopes, J., *Performance of UMTS/HSDPA/HSUPA at the Cellular Level*, M.Sc. Thesis, Instituto Superior Técnico, Lisbon, Portugal, Mar. 2008.

[LZYH12] Luo, Y. et al., "Performance assessment of C-RAN: methodology, analysis and preliminary results," in *2012 7th International ICST Conference on Communications and Networking in China (CHINACOM)*, 2012.

[PMJS12] McClusky, P. and Schroeder, J., "Fiber-to-the-antenna: Benefits and protection requirements," in *Telecommunications Energy Conference (INTELEC) IEEE 34th International*, Scottsdale, 2012.

[NaWK12] Namba, S., Warabino, T. and Kaneko, S., "BBU-RRH Switching Schemes for Centralized RAN," in *7th International ICST Conference on Communications and Networking in China (CHINACOM)*, 2012.

[NGMN13] NGMN Alliance, "Suggestions On Potencial Solutions to C-RAN", vol. Ver. 4.0, January 2013.

[PRPK12] Park, S., Ryu, B., Park, N. and Kim, D., "The Impact of Cloud Base Station's Coordinated Multi-point Schemes on Mobility Performance," in *ICT Convergence (ICTC), 2012 International Conference*, Jeju Island, 2012.

[Pires12] Pires, R., *Coverage and Efficiency Performance Evaluation of LTE in Urban Scenarios*, M.Sc. Thesis, Technical University of Lisbon, Lisbon, 2012.

[Pizz12] Pizzinat, A., *Fronthaul segment requirements and solutions*, 2012.

[PiChCl] Pizzinat, A. et al., "Optical fiber solution for mobile fronthaul to achieve Cloud Radio Access Network,".

[Salv08] Salvado, L., *UMTS/HSDPA comparison with WiMAX/IEEE 802.16e in mobility scenarios*, M.Sc. Thesis, Instituto Superior Técnico, Lisbon, Portugal, Fev. 2008.

[SeCa04] Sebastião, D. and Cardeiro, J., *Modelation and Traffic Dimensioning in the UMTS Radio Interface (in Portuguese)*, Graduation Thesis, Instituto Superior Técnico, Lisbon, Portugal, Oct. 2004.

[YrRu11] Yrjo, R. and Rushil, D., "Cloud Computing in Mobile Networks – Case MVNO," in *15th International Conference on Intelligence in Next Generation Networks*, 2011.

[ZLZT11] Zhao, L., Li, M., Zaki, Y., Timm-Giel, A. and Gorg, C., "LTE virtualization: From theoretical gain to pratical solution," in *Teletraffic Congress (ITC) 23rd Internacional*, San Francisco, 2011.

[ZGWK11] Zhu, Z. et al., "Virtual base station pool: towards a wireless network cloud for radio access networks," in *Proceedings of the 8th ACM International Conference on Computing Frontiers*, New York, 2011.