TÉCNICO LISBOA

Leveraging Capacity with Energy Consumption in 4G and Beyond Refarming and Reconfiguration Scenarios

Sara Raposo Henriques

Thesis to obtain the Master of Science Degree in

Electrical and Computer Engineering

Supervisors: Prof. António José Castelo Branco Rodrigues
Prof. Pedro Manuel de Almeida Carvalho Vieira

Examination Committee

Chairperson: Prof. José Eduardo Charters Ribeiro da Cunha Sanguino
Supervisor: Prof. António José Castelo Branco Rodrigues
Member of the Committee: Prof. Pedro Joaquim Amaro Sebastião

February 2021

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Acknowledgments

First of all, I would like to thank my supervisor, Professor António Rodrigues, and my co-supervisors, Professor Pedro Vieira and Professor Maria Paula Queluz, for all the support, insights and knowledge share given throughout this Thesis.

I would also like to thank CELFINET for the unique opportunity to develop the Thesis within a great company environment, while providing the support I needed, and all my collegues from the department, specially Eng. David Duarte, whose help and guidance were crucial during the thesis development. I would also like to express my gratitude to Eng. Thaína Saraiva and Eng. Diogo Parracho, for having shared the work developed in their thesis providing all the necessary support, and to André Alves for his patience and advice. I also thank my first internship mentor Eng. Bruno Paixão, who has never stopped supporting and advising me. In addition, I would like to thank Instituto de Telecomunicações (IT) for providing me the means for the completion of this dissertation.

I would like to thank to all my friends who accompanied me throughout this academic journey, especially to my friend Raquel Reis. Finally, to my parents and brother for providing me this opportunity, without whom all of this would not be possible.

Abstract

The evolution of mobile communication technologies, associated with the emergence of 5th Generation (5G) systems, have increased the variety and quantity of functionalities provided by the wireless network, thus being one of the main contributors to global mobile traffic growth. This evolution has a major impact on Base Stations (BSs) energy consumption, along with the gradual increase of subscribers. However, traffic growth and radio capabilities are not easily predictable, requiring operators to constantly revise their planning forecasts to fulfill all Quality of Service (QoS) and Quality of Experience (QoE) requirements.

The aim of this dissertation is to present capacity models for both 4th Generation (4G) and 5G technologies, each one for a distinct vendor, and provide energy saving scenarios in order to increase the BSs energy efficiency, using existing energy consumption models. The proposed capacity models are both based on supervised Machine Learning (ML) techniques using data collected from two real mobile network operators. This approach provides the detection of capacity saturation problems and the prediction of the maximum cell capacity, under realistic conditions. The considered scenarios are based in frequency bands and technologies switch-off techniques, combined with further analysis of both energy and capacity impact in the concerned BS.

Keywords

Mobile Networks, Capacity Modeling, Energy Consumption, 4G, 5G, Machine Learning.

Resumo

A evolução das tecnologias de comunicação móvel, associada ao surgimento dos sistemas de 5ª Geração (5G), aumentou a variedade e a quantidade de funcionalidades disponibilizadas pela rede sem fios, sendo assim um dos principais contribuintes para o crescimento do tráfego móvel global. Esta evolução tem um impacto importante no consumo de energia das Estações Base (BS), juntamente com o aumento gradual de utilizadores. No entanto, o crescimento do tráfego e os recursos rádio não são facilmente previsíveis, exigindo que as operadoras revisem constantemente as suas previsões de planeamento para cumprir todos os requisitos de Qualidade de Serviço (QoS) e Qualidade de Experiência (QoE).

O objetivo desta dissertação é apresentar modelos de capacidade para as tecnologias de 4ª Geração (4G) e 5G, cada uma para um fornecedor distinto, e fornecer cenários de poupança energética a fim de aumentar a eficiência energética das BSs, utilizando modelos de consumo energético existentes. Os modelos de capacidade propostos são baseados em técnicas de Aprendizagem Automática (ML) supervisionada utilizando dados recolhidos de duas operadoras de rede móvel reais. Esta abordagem fornece a detecção de problemas de saturação de capacidade e a previsão da capacidade máxima da célula, em condições realistas. Os cenários considerados são baseados em técnicas de *switch-off* de bandas de frequência e de tecnologias, combinadas com uma análise mais aprofundada do impacto energético e capacidade na BS em questão.

Palavras Chave

Redes Móveis, Modelação de Capacidade, Consumo Energético, 4G, 5G, Aprendizagem Automática.

Contents

List of Figures

List of Tables

Acronyms

2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
5GC	5G Core
ACK	Acknowledgement
AI	Artificial Intelligence
AMR	Adaptive Multi-Rate
AMR-WB	Adaptive Multi-Rate Wideband
AR	Augmented Reality
ARQ	Automatic Repeat Request
AuC	Authentication Centre
BER	Bit Error Rate
BBU	Baseband Unit
BCCH	Broadcast Control Channel
BS	Base Station
BSC	Base Station Controller
BSS	Base Station Subsystem
BTS	Base Transceiver Station

CCCH Common Control Channel

CCE Control Channel Element

CDMA Code Division Multiple Access

CE Channel Element

CEPT Conference of European Posts and Telegraphs

CM Configuration Management

CN Core Network

CQI Channel Quality Indicator

CS Circuit Switched

DC Dual Connectivity

DCH Dedicated Transport Channel

DS-CDMA Direct-Sequence Code Division Multiple Access

DSCH Downlink Shared Channel

EDGE Enhanced Data GSM Environment

EIR Equipment Identity Register

EIRP Equivalent Isotropic Radiated Power

EM Energy Management

eMBB enhanced Mobile BroadBand

EN-DC E-UTRAN New Radio – Dual Connectivity

eNB evolved Node B

ETSI European Telecommunications Standards Institute

E-UTRAN Evolved UMTS Terrestrial Radio Access Network

EPC Evolved Packet Core

EPS Evolved Packet System

E-SMLC Evolved Serving Mobile Location Centre

FACH Forward Access Channel

FCCH Frequency Correction Channel

FDD Frequency Division Duplex

FDMA Frequency Division Multiple Access

FR Frequency Range

FTP File Transfer Protocol

G900 GSM 900 MHz

GERAN GSM Edge Radio Access Network

GGSN Gateway GPRS Support Node

GMLC Gateway Mobile Location Centre

GMSC Gateway MSC

gNB next generation Node B

GPRS General Packet Radio Service

GPS Global Positioning System

GRAN Generic Radio Access Network

GSM Global System for Mobile Communications

GU900 GSM UMTS 900 MHz

HARQ Hybrid Automatic Repeat Request

HLR Home Location Register

HS High Speed

HSCSD High Speed Circuit-Switched Data

HSDPA High Speed Downlink Packet Access

HS-DPCCH High Speed Dedicated Physical Control Channel

HS-DSCH High Speed Downlink Shared Channel

HS-PDSCH High Speed Physical Downlink Shared Channel

HS-SCCH High Speed Shared Control Channel

HSPA High Speed Packet Access

HSPA+ Evolved High Speed Packet Access

HSS Home Subscriber Server

HSUPA High Speed Uplink Packet Access

IETF Internet Engineering Task Force

IMEI International Mobile Equipment Identity

IMSI International Mobile Subscriber Identity

IMT International Mobile Telecommunications

IoT Internet of Things

IP Internet Protocol

ISDN Integrated Services Digital Network

ITU International Telecommunication Union

KPI Key Performance Indicator

L800 LTE 800 MHz

L1800 LTE 1800 MHz

L2600 LTE 2600 MHz

LAN Local Area Network

LCS LoCation Services

LTE Long Term Evolution

M2M Machine-to-Machine

MAPE Mean Absolute Percentage Error

ME Mobile Equipment

MIMO Multiple-Input Multiple-Output

ML Machine Learning

MLR Multiple Linear Regression

MME Mobility Management Entity

mMTC mobile Machine-Type Communications

mmWave millimeter Wave

MNO Mobile Network Operator

MS Mobile Station

MSC Mobile Services Switching Centre

MT Mobile Termination

NACK Negative Acknowledgement

NE Network Element

NG-RAN Next Generation – Radio Access Network

NR New Radio

NSA Non-Standalone

NSS Network and Switching Subsystem

OFDM Orthogonal Frequency Division Multiplexing

OFDMA Orthogonal Frequency Division Multiple Access

OMC Operations and Maintenance Center

OS Operating System

OSS Operation and Support Subsystem

OVSF Orthogonal Variable Spreading Factor

PCRF Policy Control and Charging Rules Function

PDCCH Physical Downlink Control Channel

PDN Packet Data Network

PDU Protocol Data Unit

P-GW Packet Data Network Gateway

PM Performance Management

PN Pseudo-Noise

PRB Physical Resource Block

PS Packet Switched

PSCell Primary Secondary Cell

PSTN Public Switched Telephone Network

QAM Quadrature Amplitude Modulation

QoE Quality of Experience

QoS Quality of Service

QPSK Quadrature Phase Shift Keying

Q-Q Quantile-Quantile

R99 Release 99

RAN Radio Access Network

RE Resource Element

RF Radio Frequency

RI Rank Indicator

RLC Radio Link Control

RMSE Root Mean Squared Error

RNC Radio Network Controller

RNS Radio Network Subsystems

RRU Remote Radio Unit

SA Standalone

SAE System Architecture Evolution

SC-FDMA Single-Carrier Frequency Division Multiple Access

SCH Synchronization Channel

SDCCH Stand-alone Dedicated Control Channel

SF Spreading Factor

SGSN Serving GPRS Support Node

S-GW Serving Gateway

SIM Subscriber Identity Module

SINR Signal to Interference plus Noise Ratio

SM Spatial Multiplexing

SMS Short Message Service

SNMP Simple Network Management Protocol

SSE Sum of Squared Error

SWIPT Simultaneous Wireless Information and Power Transfer

TBS Transport Block Size

TCH Traffic Channel

TDD Time Division Duplex

TDMA Time Division Multiple Access

TE Terminal Equipment

TMN Telecommunication Management Network

TTI Transmission Time Interval

U900 UMTS 900 MHz

U2100 UMTS 2100 MHz

UE User Equipment

UICC Universal Integrated Circuit Card

UMTS Universal Mobile Telecommunications System

URLLC Ultra-Reliable Low Latency Communications

USIM UMTS Subscriber Identity Module

UTRAN UMTS Terrestrial Radio Access Network

VLR Visitor Location Register

VoIP Voice over IP

VR Virtual Reality

WCDMA Wideband Code Division Multiple Access

1

Introduction

Contents

This chapter provides the motivation and some objectives of the work developed within the Thesis, as well as a summary of the document structure.

1.1 Motivation

Nowadays, mobile traffic is escalating with the growing number of wireless users along with the increased traffic volume per subscriber, derived from advanced applications such as high-resolution video streaming, remote monitoring and real-time control applications. Mobile network data traffic grew by 56% between Q1 2019 and Q1 2020 [1]. Figure 1.1 presents the total global monthly network data and voice traffic from Q1 2014 to Q1 2020, along with the year-on-year percentage change for mobile network data traffic.

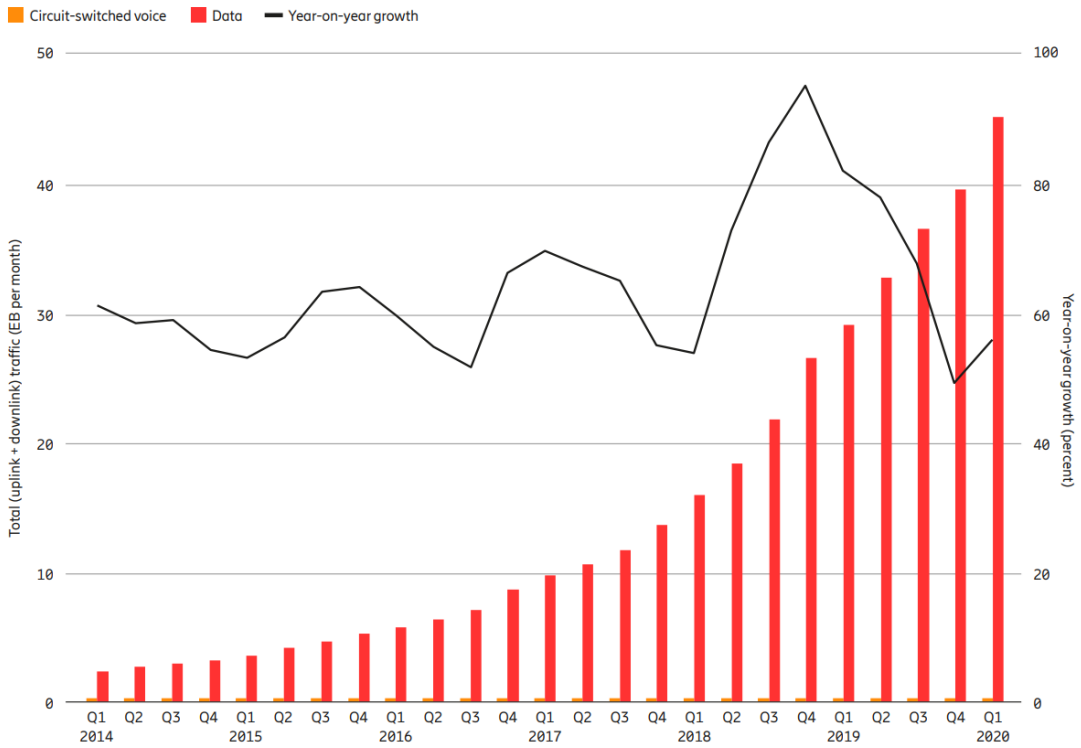


Figure 1.1: Global mobile network data traffic and year-on-year growth [1].

In addition to the evolution of the current mobile communication technologies, the gradual emergence of Fifth Generation (5G) devices will have a strong impact on this growth. By 2025, it is estimated that 45% of total mobile data traffic will be carried by 5G networks [1].

The exponential growth of mobile data traffic requires operators to estimate the capacity of the cells, in order to guarantee an adequate Quality of Service (QoS) and Quality of Experience (QoE) for end users. Once the capacity of a cell is exceeded, these parameters reach unacceptable levels. Thus, an

accurate cell capacity estimate is necessary to fulfill QoS requirements with minimal network investment [2]. The downlink throughput performance is often considered to be the significant metric that best defines the end user satisfaction on the mobile network performance, since it directly affects the data download speed of several packet applications, such as video streaming and web browsing, especially in Fourth Generation (4G) systems [3].

This rapid evolution of mobile data consumption has led to an increase in the number of Base Stations (BSs). However, as shown in Figure 1.2, the BS is the main energy consumer in a cellular network and can reach 57% of all the mobile operator energy consumption [4]. Thus, BS energy consumption monitoring helps operators to become aware of their energy consumption, in order to assess the impact of their subsequent decisions.

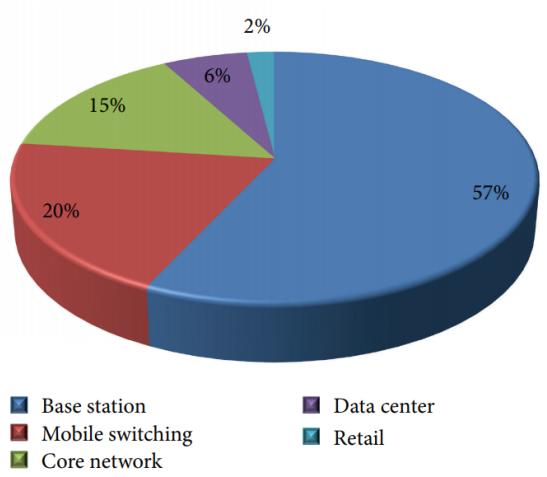


Figure 1.2: Energy consumption composition of a mobile operator [4].

1.2 Objectives

This Thesis aims to develop models that predict the cell capacity in both 4G and 5G mobile networks, along with energy efficiency scenarios of traffic migration. The capacity models are based on real measurements, using Multiple Linear Regression (MLR) algorithms. The main advantage of these measurement-based approaches is to consider the peculiarities of each cell, such as propagation conditions, channel quality and latency. Subsequently, two traffic migration scenarios, associated with Universal Mobile Telecommunications System (UMTS) technology, are developed in order to provide energy efficiency solutions of the BSs, by using existing power consumption models for the radio equipment. Finally, a simple 5G data traffic migration scenario is added to predict the impact of 4G traffic on the resources that 5G provides.

During the development of the work, R and Python languages were used.

1.3 Thesis Outline

The work developed within the Thesis scope is divided into five main chapters: Chapter 2 provides an overview of the various subjects studied and considered relevant, such as Radio Access Network (RAN), from Second Generation (2G) to 5G and also energy consumption monitoring; Chapter 3 addresses Machine Learning (ML) concepts including a introduction to ML models, in which the MLR approach is detailed, along with an explanation of the metrics used to evaluate the models developed; Chapter 4 presents a summary of the procedures and the power consumption models used in order to estimate the Remote Radio Units (RRUs) power consumption; Chapter 5 provides a description of the several steps of the 4G and 5G models development, presenting the algorithms used as well as their results; Chapter 6 discusses both the approaches and results of the traffic volume migration scenarios. Finally, in Chapter 7 conclusions are drawn and future work is suggested.

1.4 Publications

In the context of this Thesis, the following scientific paper was presented:

- S. Henriques, D. Duarte, P. Vieira, M. P. Queluz and A. Rodrigues, ”Leveraging Capacity with Energy Consumption in 4G and Beyond Refarming and Reconfiguration Scenarios”, 23rd International Symposium on Wireless Personal Multimedia Communications (WPMC), September 2020.

2

State of the art

Contents

This chapter provides an overview of the various subjects studied and considered relevant for the development of the work being done within the scope of the Thesis, aiming for its better understanding. Thus, an introductory section about RANs is presented, as well as a section about each mobile technology (2G, Third Generation (3G) and 4G). This chapter also provides a section about energy consumption monitoring and then a final section about 5G is presented, which also mentions the concept of green networks.

2.1 Radio Access Networks

This section presents an introduction to the RANs, such as a brief overview and its planning. The content of this section is mainly based on the following references: [5] in Subsection 2.1.1; [6] in Subsection 2.1.2.

2.1.1 RANs Overview

The RAN started to be used at the beginning of cellular technology and has evolved through the development of the various mobile communications generations. In a RAN, radio sites provide radio access and coordinate resource management. An individual device is connected (wireless) to the RAN, which transmits the device signal to numerous wireless endpoints and so this signal travels through other networks' traffic.

Some types of radio access networks include Generic Radio Access Network (GRAN), which utilizes base transmission stations and base transmission controllers to manage radio links for both Circuit Switched (CS) and Packet Switched (PS) core networks, and GSM Edge Radio Access Network (GERAN), which supports real-time packet data. Those types of RANs include as well UMTS Terrestrial Radio Access Network (UTRAN), which supports CS and also PS services, and Evolved UMTS Terrestrial Radio Access Network (E-UTRAN), for Long Term Evolution (LTE), which is focused only on PS services and provides high data rates and low latency.

RAN components include a BS, which is a fixed communication location for customer cellular phones on a carrier network, and antennas, connected to the BS, which receive and transmit the signals in the cellular network to customer devices. A BS is typically located in a position far above the grounded area providing coverage. Different types of BSs are set up according to the coverage needed, while the antennas cover a specific region depending on their capacity.

Figure 2.1 shows a current BS architecture. A typical wireless BS consists of the baseband processing unit, Baseband Unit (BBU), placed in the equipment room, and the Radio Frequency (RF) processing unit, RRU, generally installed in towers near the antenna, which reduces RF signal insertion loss and improves efficiency. The BBU, connected with the RRU via optical fiber, has the advantage of modular design, small size, high integration, low power consumption and easy deployment. The RRU are controlled by a controller located nearby the tower inside a closed shelter.

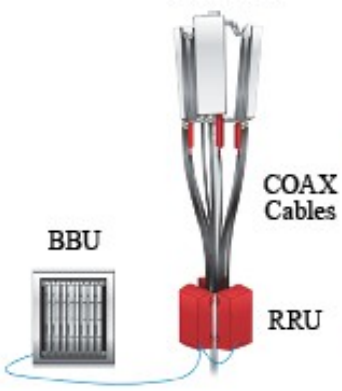


Figure 2.1: Distributed base station (site) architecture (adapted from [5]).

2.1.2 RAN Planning

The process of RAN planning has several ambitions:

- Achieve sufficient coverage in the target area ensuring an adequate QoS/QoE and low Bit Error Rate (BER).
- Guarantee the required network capacity taking into account a low service blocking, decent user throughput and also low dropped call rates.
- Obtain an economic and efficient network infrastructure avoiding unnecessary sites, maintaining the demanded capacity, coverage and quality.

These purposes can be accomplished selecting favorable locations for the sites and carefully setting cell parameters.

2.2 GSM Background

This section provides an overview of the technology Global System for Mobile Communications (GSM), presenting a brief introduction to this technology, the services it provides and an architectural overview of GSM. There is also a subsection about radio transmission aspects of this system and, finally, another one about capacity in GSM, which introduces briefly both General Packet Radio Service (GPRS) and Enhanced Data GSM Environment (EDGE) systems.

The content of this section is mainly based on the following references: [7, 8] in Subsection 2.2.1; [9] in Subsection 2.2.2; [7–9] in Subsection 2.2.3; [7, 8] in Subsection 2.2.4; [10–13] in Subsection 2.2.5.

2.2.1 Introduction to GSM

GSM is a 2G mobile wireless standard, developed by the European Telecommunications Standards Institute (ETSI), that uses digital radio transmission to provide voice, data and multimedia communication services. In general, it operates on the 900 MHz and 1800 MHz mobile communication bands. The GSM system has the capacity to carry 64 kbps to 120 Mbps of data rates and it allows up to 8 to 16 voice users to share each radio channel, which is 200 kHz wide and it is divided into 8 time-slots with 25 kHz.

GSM started to be developed in 1979, at the World Administrative Radio Conference, with the reservation of the 900 MHz band. In 1982, the Conference of European Posts and Telegraphs (CEPT) formed the Groupe Spécial Mobile (the initial meaning of GSM), to implement a common mobile phone service in Europe in the 900 MHz range (spectrum previously allocated). By 1987 the basic design of the system was configured establishing properties such as the multiple access scheme: Time Division Multiple Access (TDMA) and the modulation technique. The responsibility for GSM was assigned to the ETSI later in 1989 and the Phase I specifications were published in 1990 (and the Phase II recommendations in 1995) while there were requested specifications for higher user densities with low-power mobile stations and operating at 1800 MHz, which were published in 1991. In the same year started the GSM networks commercial operation in Europe. In 2014, the system was globally standardized with over the 90% market share, operating in over 193 countries, making ”Global System for Mobile Communications” the current term for acronym GSM.

The system of GSM is part of the evolution of wireless mobile telecommunications along with High Speed Circuit-Switched Data (HSCSD), GPRS, EDGE and UMTS.

2.2.2 Services

Initially the services provided by GSM were focused on CS voice service, but then they were evolved starting to provide data, messaging, multicast and multimedia services.

- **Voice service** was a full rate voice service in the beginning allowing 8 user per radio channel. However, decreasing the audio quality, the original design allowed the use of only half rate voice service increasing the quantity of simultaneous users to 16 per radio channel.
- **Short Message Service (SMS)** is a text messaging service to exchange short text messages (140 characters), which has evolved into executable messages that enable advanced two-way messaging features.
- **Data services** started as low speed CS data with user bit rates up to 9.6 kbps. Posteriorly it was evolved allowing the combination of multiple CS data connections providing HSCSD services.
- **Packet Data** (GPRS) is an efficient approach to upgrade the existing GSM system to a packet radio

service allowing users to share packet data resources dynamically for services such as Internet browsing.

- **Multicast** is a point-to-multipoint interface specification design to provide efficient delivery of multicast and broadcast services such as group call (dispatch type services) and voice broadcast (traffic alerts).

2.2.3 Architecture

The functional architecture of a GSM network is a simple and effective way to provide the services needed for a 2G cellular system. The GSM specifications define several functional entities in terms of functions and interfaces but it does not specify hardware, in order to not limit so much the designers and, at the same time, allowing the operators to buy equipment from different suppliers. The structure of a GSM system, as shown in Figure 2.2, involves four main subsystems: the Mobile Station (MS), Base Station Subsystem (BSS), Network and Switching Subsystem (NSS) and Operation and Support Subsystem (OSS). Each subsystem contains functional units which communicate through the various standard interfaces using specified protocols.

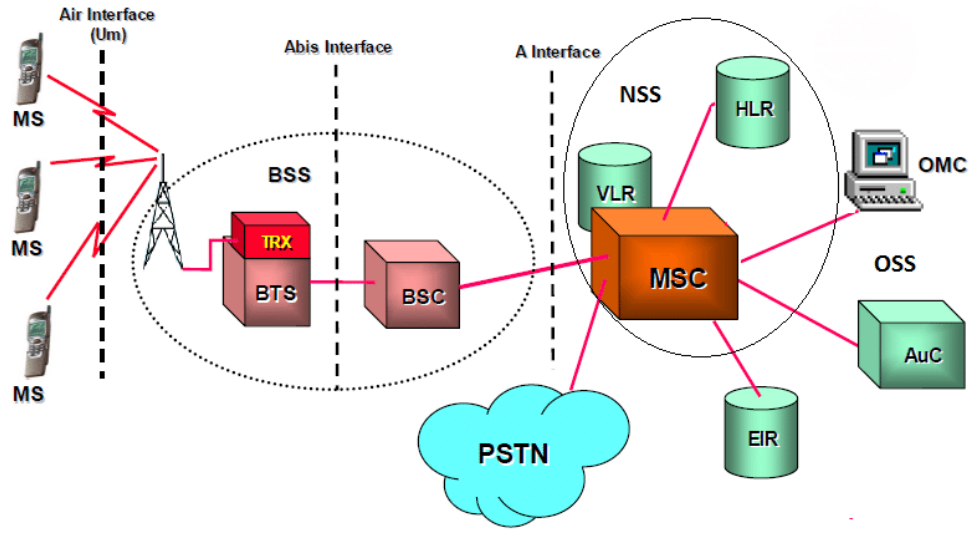


Figure 2.2: General architecture of a GSM network (adapted from [9]).

The **MS** is composed of the following components:

- **Mobile Equipment (ME)** is the actual hardware and it is identified with an International Mobile Equipment Identity (IMEI), which is installed in the phone at manufacture and it cannot be changed. During the registration of the ME in the network, this access the IMEI to check if the equipment has been reported as stolen.

- **Subscriber Identity Module (SIM) card** stores the subscriber information, which includes a unique International Mobile Subscriber Identity (IMSI). The insertion of the SIM card into any GSM mobile equipment allows the user to make and receive calls on that terminal and to use the other subscribed services. The IMEI and IMSI are independent in order to provide personal mobility to the users.

The **BSS** consists of two elements:

- **Base Transceiver Station (BTS)** comprises the radio transceivers that are responsible for the radio transmissions with the MS, which occur across the radio (Um) interface. The BTS requirements are ruggedness, reliability, portability and minimum cost, due to the usual large amount of BTSs.
- **Base Station Controller (BSC)** manages the radio resources for a group of BTSs and it controls functions such as the allocation/release of radio interface channels, power control algorithms, frequency hopping, encryption algorithm, radio link monitoring and handover management. It communicates with a BTS through the Abis interface.

The **NSS** provides the main control and interfacing for the whole cellular network and it is composed of several components:

- **Mobile Services Switching Centre (MSC)**, which is the central component of the NSS, is responsible for the switching of calls between mobile users and also between mobile and fixed network users, managing the incoming/outgoing calls from several types of networks including Public Switched Telephone Network (PSTN), Integrated Services Digital Network (ISDN) and Packet Data Network (PDN). In addition provides the functionality needed to handle a mobile subscriber, such as registration/authentication of a user, location updating, inter-MSC handovers and also call routing to a roaming subscriber. The MSC communicates with the BSS through the A interface.
- **Home Location Register (HLR)** is one of the NSS databases, associated with the MSC, providing certain capabilities such as call-routing and roaming. It contains all the administrative information that belongs to each of the registered subscribers within the corresponding GSM network, including the IMSI and the list of services the user subscribes, along with the current location of the subscriber.
- **Visitor Location Register (VLR)** is the other database, containing selected administrative information from the HLR for call control and to provide the subscribed services for each subscriber currently located in the geographical area under its control. The VLR can be implemented as a separate entity, but usually it is implemented together with the MSC, so that the geographical are

controlled by the MSC corresponds to the one controlled by the VLR, allowing a faster and more convenient access.

- **Equipment Identity Register (EIR)** is one of the NSS registers, used for authentication and security purposes. The EIR is a database which contains a list of all valid ME that belong to the network. As mentioned before, each device is identified by an IMEI, which is marked as invalid in case it has been reported stolen.

- **Authentication Centre (AuC)**, being the other register, is a protected database which stores the secret key also contained in the user's SIM card, used for authentication and encryption over the radio channel, avoiding undesired violations of the system by third parties.

The **OSS** is the implementation of the Operations and Maintenance Center (OMC), which is connected to components of the NSS and the BSC. It is used, by the network operator, to monitor and control the system as well as the traffic load of the BSS.

2.2.4 Radio Transmission Aspects

The GSM system uses a combination of two multiple access techniques, TDMA and Frequency Division Multiple Access (FDMA), to divide the bandwidth among the users. The FDMA element divides the frequency of the (maximum) 25 MHz bandwidth into 124 carrier frequencies of 200 kHz bandwidth, which can be observed in Figure 2.3. One or multiple carrier frequencies are assigned to individual BSs and each carrier is then divided in time into 8 time slots, which form a TDMA frame, through the TDMA scheme, allowing multiplexing of several users on the same carrier frequency, with each user assigned to a different time slot. Each of these time slots is identified by a time slot number, from 0 to 7, being used for both transmission (GSM burst) and reception of data packets. The time slots are separated in time so that a mobile device does not receive/transmit data at the same time, being able to use the same RF channel without interference.

A transmission channel is associated with one time slot position of a TDMA frame and this position defines the type/function of its channel. In a particular carrier frequency, to each frame corresponds a number, with which the MS and the BS are synchronized. Groups of 26 and 51 TDMA frames can form larger frames, existing larger groups as well. There are two types of logical channels: traffic channels and control channels.

Traffic channels provide a bi-directional point-to-point transmission link to a mobile subscriber. The Traffic Channels (TCHs) are divided by several types, depending on the service accessed by the subscribers: voice or data (with several possible data rates).

Control channels are constituted by four different classes: broadcast, dedicated, common and associated, depending on their functions, and they are introduced in the following points:

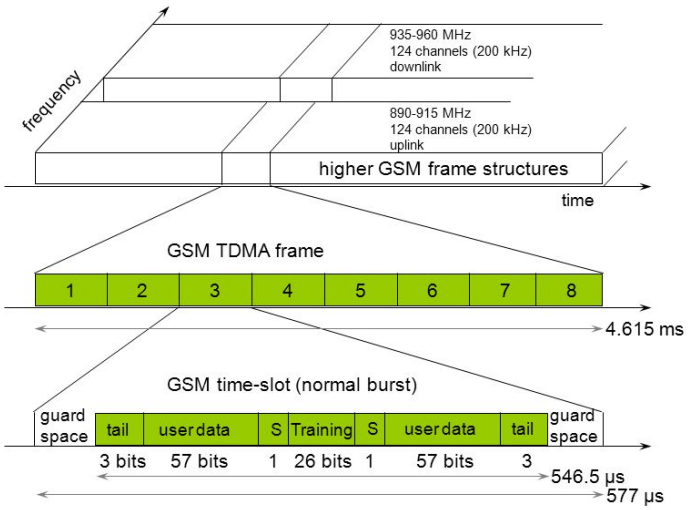


Figure 2.3: TDMA and FDMA in the GSM system - frame structure (adapted from [11]).

- **Broadcast Control Channel (BCCH)** continually broadcasts, on the downlink, information on the cell, including BS identity, frequency allocations and frequency-hopping sequences.
- **Frequency Correction Channel (FCCH)** used by the MS to adjust its local oscillater to the BTS oscillator so that it synchronizes in frequency with the BTS.
- **Synchronization Channel (SCH)** used by the MS in order to occur a time synchronization between the MS and the BTS, and to identify the cell as well.
- **Common Control Channel (CCCH)** utilized to coordinate the control of MS operating in its cell radio coverage area.
- **Stand-alone Dedicated Control Channel (SDCCH)** used to transfer signaling messages to specific mobile devices, such as location updating information, registration, authentication and call setup.

2.2.5 Capacity

Initially, the existing cellular systems (GSM) offered low data rate of 9.6 kbps and lack advanced multimedia services which were subsequently required by the following systems. Therefore, in order to improve data rate, some technologies were developed such as GPRS which achieves higher capacity transmission than GSM through multiple slots in TDMA, providing a data rate of 171 kbps. Finally, EDGE was the successor of GPRS in the evolution towards UMTS, also improving data rate. This improvement was due to the introduction of a different form of modulation and better coding as well, making it possible to achieve, with the appearance of EDGE, data rates of 384 kbps.

For capacity planning purposes, it is particularly important to analyze the traffic volumes at peak times of the day, more specifically ”busy hours” (1 hour of the day in which traffic is maximum), to be able to determine the acceptable service required level. This can be accomplished by the Erlang B traffic model, which is designed to handle the busy hours and to determine this same level, in these specific periods, as well. This model considers that a service request must be immediately serviced or else dropped instantly. The Erlang B formula, which gives the ”blocking probability” (probability that a call will be blocked through congestion), is given by, [12]:

$$P_B = \frac{\frac{A^N}{N!}}{\sum_{X=0}^{N} \frac{A^X}{X!}}, \tag{2.1}$$

where P_B is the blocking probability, A is the offered traffic in Erlangs, N is the number of trunks (servers).

The offered traffic for each user, given by A, can be represented in Erlangs by, [13]:

$$A = \mu \cdot H, \tag{2.2}$$

where μ is the average number of requested calls per hour and H the average call duration in hours.

If the number of channels is not specified, but the amount of users, U, is given, the total traffic offered by the system, A_T, can be given by, [13]:

$$A_T = U \cdot A. \tag{2.3}$$

2.3 UMTS Background

This section presents an overview of UMTS technology, including a chapter about its architecture and the basics of Wideband Code Division Multiple Access (WCDMA). In the end, this section also introduces the spreading process, which leads to a subsection about capacity in UMTS. In order to finalize this introduction to 2G systems, this Section provides also a final High Speed Packet Access (HSPA) overview, connected to the capacity in 2G.

The content of this section is mainly based on the following references: [14, 15] in Subsection 2.3.1; [14, 15] in Subsection 2.3.2; [14–16] in Subsection 2.3.3; [15, 17] in Subsection 2.3.4; [15, 18] in Subsection 2.3.5; [15, 19–21] in Subsection 2.3.6.

2.3.1 Introduction to UMTS

UMTS is a 3G mobile cellular system, developed by the Third Generation Partnership Project (3GPP), which is seen as the successor to the GSM family of standards including GPRS and EDGE. It represents

an evolution in terms of capacity, data rates and new service capabilities from 2G mobile networks.

The UMTS system was firstly initiated by the ETSI in Europe and after this, in 1998, the 3GPP was created by organizations (specialized in standards development) around the world to continue the technical specification work for this system.

3G UMTS offers to its users high speed transmissions allied to global mobility, while provides several services such as voice, messaging, internet access, multimedia transmissions (images/video), although 2G was built with major focus in voice services. The essential upgrades of the UMTS system compared to the 2G technology are the following, [14]:

- Higher data rates;
- Variable data rates allowing bandwidth on demand;
- High spectral efficiency;
- Asymmetric data rates support on downlink and uplink in order to transfer PS traffic;
- QoS provisioning for several services and applications;
- QoS differentiation support.

The radio transmission standard of UMTS is WCDMA, which uses the direct sequence spread spectrum method of asynchronous code division multiple access, in order to provide higher rates and allow simultaneous users, once more compared with GSM radio access technologies: TDMA and FDMA. In this current case, there are two operation modes: Frequency Division Duplex (FDD), where the downlink and uplink are on different frequency bands, and Time Division Duplex (TDD), where the uplink and downlink are split in time with the base stations and then the mobiles transmit alternately on the same frequency.

Through the years new releases emerged, which brought much higher bit rates to the networks for both downlink and uplink, represented in Table 2.1. The need of higher data rates initiated the High Speed Downlink Packet Access (HSDPA) and the High Speed Uplink Packet Access (HSUPA), which together constitute the HSPA, through the Releases 5 and 6, respectively. The last two releases, 7 and 8, introduced the Evolved High Speed Packet Access (HSPA+) and LTE, respectively.

Table 2.1: Peak rates which characterize each 3GPP release [15].

	Release 99	**Release 5**	**Release 6**	**Release 7**	**Release 8**
Downlink peak rate [Mbps]	0.4	14	14	28	160
Uplink peak rate [Mbps]	0.4	0.4	5.7	11	50

2.3.2 Architecture

The UMTS network architecture can be divided into three main elements: User Equipment (UE), UMTS Terrestrial Radio Access Network (UTRAN) and Core Network (CN), which can be observed in Figure 2.4. The 3GPP standards specify the individual functionalities of each part of the network and the logical interfaces between them as well.

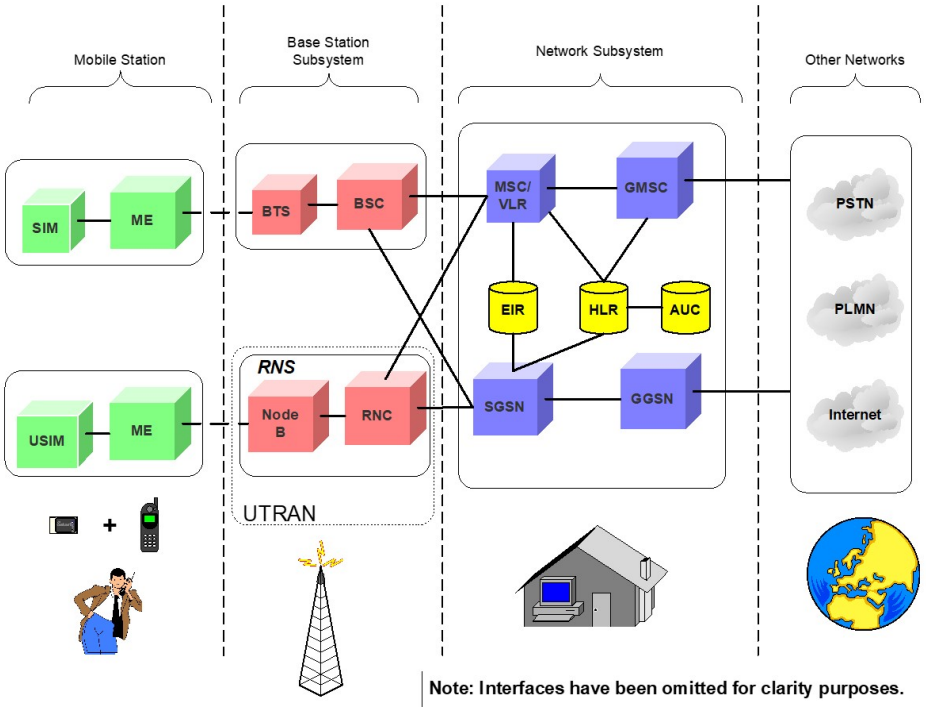


Figure 2.4: General architecture of a UMTS network (adapted from [15]).

The **User Equipment (UE)**, that interfaces with both the user and the radio interface, contains two distinct elements:

- **Mobile Equipment (ME)** is a radio terminal responsible for the radio communications over the Uu interface (radio interface).
- **UMTS Subscriber Identity Module (USIM)**, which communicates with the ME over the internal Cu interface, is a smart card that includes the subscriber identity, authentication algorithms, authentication and encryption keys and also subscription information needed at the terminal.

The **UMTS Terrestrial Radio Access Network (UTRAN)**, which is responsible for all radio funcionalities and connects the UE with the CN, consists of one or more Radio Network Subsystems (RNS), connected to the CN over the Iu interface, each of which includes two parts:

- **Node B**, which converts data flow between both Uu and Iub interfaces, manages the radio resource as well.
- **Radio Network Controller (RNC)**, connected to one or more Node Bs through the Iub interface, is responsible for the control of the radio resources in all attached cells and it is the service access point for all services provided to the CN (connected with through the Iu interface). Two different RNCs connect with each other over the Iur interface.

Finally, the **Core Network (CN)**, responsible for switching and routing calls and data connections to external networks, may be divided, in view of the different ways in which data is carried, into two different domains:

- **Circuit Switched (CS) elements**, which carry data in a CS manner, include the **Mobile Services Switching Centre (MSC)**, the **Visitor Location Register (VLR)** and the **Gateway MSC (GMSC)**. The MSC handles the CS calls (call setup/routing, management of inter-system mobility, etc). The VLR is the database which holds a copy of the subscriber service profile, as well as more precise information on the location of the users who are being served by the MSC. The pair MSC/VLR is connected to the RNC over the Iu-CS interface. The GMSC is basically the interface which links the MSC/VLR with the external CS networks.
- **Packet Switched (PS) elements**, designed to carry packet data, including the **Serving GPRS Support Node (SGSN)** and the **Gateway GPRS Support Node (GGSN)**. The SGSN relays packets from radio networks to the CN and it is also responsible for both mobility and session management. The Iu-PS interface connects the RNC with the SGSN. The GGSN is the entity which links the SGSN with the external PS networks.

The CN of the UMTS system also provides shared (by both CS and PS domains) elements including the **Home Location Register (HLR)**, the **Equipment Identity Register (EIR)** and the **Authentication Centre (AuC)**. The HLR is a central database, located in the subscriber's home system, that contains detailed information of the subscriber, including allowed services, forbidden roaming areas, status of call forwarding and its number. The EIR is also a database which contains the information related to the identity of the ME that can be used to prevent a specific terminal (stolen, unauthorized or defective) from accessing the network.

2.3.3 WCDMA Basics

The air interface of UMTS is based on WCDMA technology that employs the Direct-Sequence Code Division Multiple Access (DS-CDMA) (information bits of the user are spread over a wide bandwidth by multiplying the user data with quasi-random bits, which are derived from Code Division Multiple

Access (CDMA) spreading codes) to provide higher speeds and higher capacity service, allowing more simultaneous users, compared to the TDMA and FDMA utilized by the GSM network. An example of the WCDMA bandwidth allocation is shown in Figure 2.5.

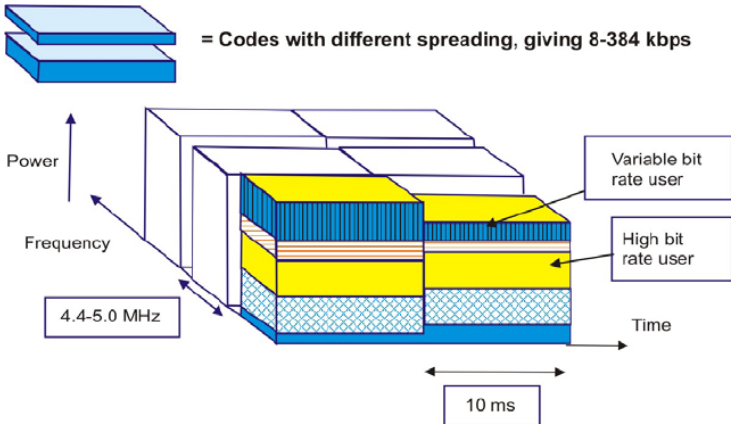


Figure 2.5: WCDMA bandwidth allocation in the time-frequency-code space (adapted from [16]).

For a 5 MHz carrier bandwidth, the used chip rate is 3.84 Mcps. Since this bandwidth is much wider than the ones used for 2G systems (about 200 kHz), it provides higher bit-rates and increased multipath diversity as well. For capacity increasing purposes, network operators are able to deploy several 5 MHz carriers.

WCDMA is able to support variable user data rates. The data rate of the user is constant during each 10 ms frame, but it can change from frame to frame. The network controls the radio capacity allocation in order to achieve optimal throughput for packet data services.

The WCDMA system, developed by the 3GPP, features two modes:

- **Frequency Division Duplex (FDD)** uses separate 5 MHz carrier frequencies the uplink and downlink data transmissions, allowing 100 simultaneous voice calls.
- **Time Division Duplex (TDD)** uses the same frequency band for both directions while adjusts the time domain portion assigned for uplink and downlink dynamically, improving the efficiency.

The operation of asynchronous base station is supported by WCDMA, making global time references (such as GPS) unnecessary. Since there is no need to receive a GPS signal, the deployment of indoor and micro base stations becomes easier.

WCDMA utilizes coherent detection based on the use of pilot symbols or common pilot on both uplink and downlink, while in 2G systems it is only used in downlink, resulting in an increase of coverage and capacity in uplink.

In the WCDMA air interface, advanced CDMA receiver concepts, such as multi-user detection and smart adaptive antennas, can be implemented by the network operator in order to increase capacity and

coverage.

WCDMA is able to modulate each base-band symbol with a binary/quaternary signature with a much higher rate than the one of the original data symbol, thereby artificially increasing the bandwidth of a signal.

2.3.4 Spreading Process

As mentioned before, WCDMA uses a Direct Sequence spreading, where spreading process combines directly the baseband information to high chip rate binary code. The data to be transmitted is encoded using a specific spreading code for a particular user. Thus, only the desired recipient can correlate/decode this specific signal, while the rest of the signals appear as noise, allowing the physical RF channel to be shared by multiple simultaneous users. Each channel is spread with a unique and variable spreading sequence. The degree of this spreading process varies so that the final signal can fill the required bandwidth of the channel.

The chip rate of a code is the number of chips per second at which the code is transmitted/received. The chip rate, for UMTS, is 3.84 Mcps, as refered before. The Spreading Factor (SF) is the ratio of the chip rate (Mcps) to the information (transmission) rate (cps).

In WCDMA, the codes required to spread the signal must be orthogonal in order to provide multiple users/channels operating without mutual interference. Therefore, these codes are Orthogonal Variable Spreading Factor (OVSF), where each signal is spread over a wide spectrum range, before being transmitted, through the use of a user's code. These codes are chosen very carefully to be mutually orthogonal to each other. They derivate from an OVSF code tree, which can be observed in Figure 2.6, where each user has a specific code.

There are two stages of spreading, as represented in Figure 2.7, [15]:

- **Channelization** operation, using OVSF codes, which are used to identify the user services in the uplink and user channels in the downlink.
- **Scrambling** operation, using Pseudo-Noise (PN) codes, which are used to identify the individual Node B in the downlink and a specific UE in the uplink.

2.3.5 Capacity

A Channel Element (CE) are hardware resources of the Node B which describe the software licensed capacity resources required for a dedicated channel. Both in uplink and downlink, channel elements are consumed when a dedicated channel is allocated. The quantity of CEs to be consumed depends on the traffic type, the type of radio bearers which will be used for the dedicated channel and the quantity of simultaneous users for each bearer. This consumption defines the CE cost which is characterized

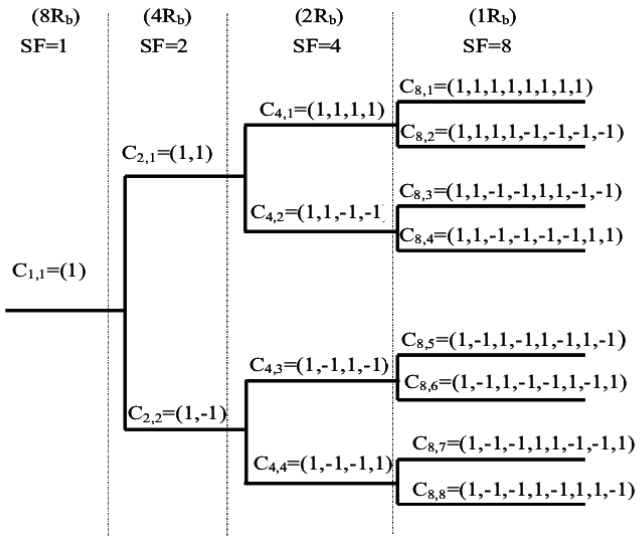


Figure 2.6: OVSF code tree (adapted from [17]).

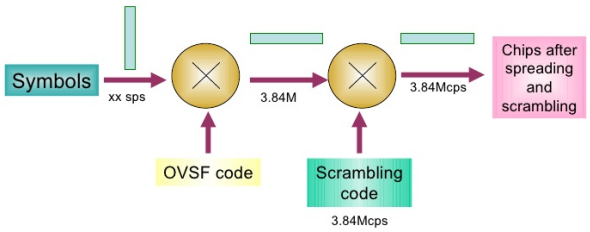


Figure 2.7: Spreading operations (adapted from [15]).

by a pair of channel element factors, one for uplink (Γ_{UL}) and another one for downlink (Γ_{DL}). If the available CEs amount is not enough, a certain user can be switched down to a smaller rate in order to avoid blocking.

The Node B support a certain number of CEs that define the Node B hardware capabilities of channel demodulation. Services at different rates require a distinct quantity of CEs providing proper channel demodulation. Therefore, increasing the CEs supported by a Node B, it will be provided better channel demodulation and capabilities of service processing.

In a RAN, CE resources are managed by the RNC and the Node B as well. The RNC manages the new services considering the number of required CEs and manages the CE resources in a congestion scenario as well. The RNC determines the quantity of CEs required for a certain service based on the SF which matches the service rate. The Node B reports its capacity to the RNC and, taking into account the actual service rate, manages the CE resources adapting the quantity of CEs available to be consumed.

CE may be divided into uplink CEs and downlink CEs: one uplink (downlink) CE needs to be con-

sumed by an uplink (downlink) 12.2 kbps voice service, with SF = 64 (SF = 128), plus 3.4 kbps signalling.

Knowing the number of simultaneous users per site for a certain service i (N_i), the total number of required CEs (n_{CE}) can be calculated through, [18]:

$$\begin{cases} n_{CE,UL} = \sum_i N_i \Gamma_{UL,i} \ , & \text{for uplink} \\ n_{CE,DL} = \sum_i N_i \Gamma_{DL,i} \ , & \text{for downlink} \end{cases} \tag{2.4}$$

where Γ_i is the channel element factor for the service i.

The Release 99 contains several services, (grouped into) including person-to-person services and content-to-person services. Person-to-person services can be CS, where the voice service Adaptive Multi-Rate (AMR) belongs to, and content-to-person services are PS. For these services, the RNC determines the number of CEs and Node B credit resources needed to be consumed based on the SF which matches the maximum bit rate of the respective service, as mentioned before. The Table 2.2 presents an example of channel element factors for these services on different radio bearers.

Table 2.2: Example of channel element factors [18].

Service	Γ_{UL}	Γ_{DL}
AMR12.2	1	1
PS64	4	2
PS128	8	4
PS384	16	8

Usually the voice service is the highest priority service, followed by person-to-person PS services and, finally, content-to-person services.

2.3.6 High Speed Packet Access

As mentioned before, High Speed Packet Access (HSPA) is a fusion of two mobile protocols, HSDPA and HSUPA, which extends and improves the performance of existing 3G mobile networks using the WCDMA protocols.

Release 5 introduced HSDPA, which aims to increase packet data throughput, reduce delay and achieve high peak rate, with adaptive modulation and hybrid Automatic Repeat Request (ARQ). In HSDPA the modulation starts with Quadrature Phase Shift Keying (QPSK), followed by 16-Quadrature Amplitude Modulation (QAM) modulation, being further introduced 64-QAM modulation in Release 7. HSDPA also uses link adaptation, adjusting the modulation and the coding parameters in every Transmission Time Interval (TTI), which is 2 ms long.

In HSDPA, the downlink scheduling was moved from from the RNC to the Node B and, therefore, the scheduling decisions are performed with minimum latency, closer to the radio interface. In addition to

the channels from Release 99 that are also used in Release 5, such as Dedicated Transport Channel (DCH), Downlink Shared Channel (DSCH) and Forward Access Channel (FACH), new physical and transport channels have emerged with this version. High Speed Downlink Shared Channel (HS-DSCH) is the transport channel that carries the data. High Speed Shared Control Channel (HS-SCCH) and High Speed Physical Downlink Shared Channel (HS-PDSCH) are the new downlink physical channels, while High Speed Dedicated Physical Control Channel (HS-DPCCH) is the new uplink physical channel. HS-SCCH is the channel that carries the physical layer information required to decode the data from HS-DSCH, HS-PDSCH is used to carry the HS-DSCH and HS-DPCCH has the information related to the ARQ acknowledgments (Acknowledgement (ACK) and Negative Acknowledgement (NACK)) and downlink Channel Quality Indicator (CQI).

In downlink is carried the data, while in uplink is carried control information which consists in measurements made and reported by the UE to the Node B, i.e, it consists in CQI which has a HS-DSCH sub-frame with the Transport Block Size (TBS), the number of HS-PDSCH codes and the modulation. Both the TBS and the maximum number of HS-PDSCH codes depend also on the mobile category, subsequently introducing the minimum inter-TTI interval as well, which is the time required for the ME to decode, between the beginning of the TTI and the beginning of the following one. On the other hand, the CQI depends on the propagation conditions and the distance between the UE and the Node B.

With the information provided by the CQI, the achievable maximum data rate, $R_{AchievableMax}$, can be calculated by:

$$R_{AchievableMax} = \frac{V_{TBS}}{0.002 \times I_{TTImin}}, \tag{2.5}$$

where V_{TBS} is the TBS size and I_{TTImin} is the minimum inter-TTI interval.

Table 2.3 provides some examples of different mobile categories and the respective maximum number of HS-PDSCH codes, minimum inter-TTI interval and V_{TBS}, representing how the data rate on the HS-DSCH can change with each TTI, according with the reported CQI.

Table 2.3: HSDPA terminal capability categories [15].

Category	Max HS-PDSCH codes	Min inter-TTI interval	V_{TBS} [bits]	$R_{AchievableMax}$ [Mbps]
1	5	3	7298	1.2
3	5	2	7298	1.8
5	5	1	7298	3.6
7	10	1	14411	7.2
10	15	1	27952	14.4
12	5	1	3630	1.8

2.4 LTE Background

This section provides an overview of the LTE system, initially presenting a brief introduction to this technology, followed by an architectural overview, continuing with a subsection about multiple access techniques, where Orthogonal Frequency Division Multiple Access (OFDMA) and Single-Carrier Frequency Division Multiple Access (SC-FDMA) schemes are described, thus finalizing with capacity in LTE and a brief overview to LTE-Advanced.

The content of this section is mainly based on the following references: [22, 23] in Subsection 2.4.1; [24–26] in Subsection 2.4.2; [15, 15, 22, 24, 27] in Subsection 2.4.3; [24, 28–30] in Subsection 2.4.4; [25] in Subsection 2.4.5.

2.4.1 Introduction to LTE

LTE is a 4G wireless communication standard, which quickly became the 4G network technology of choice around the world, and it was developed by the 3GPP, providing increased peak data rates, improved spectral efficiency and reduced user/control plane latency. From the wireless access point of view, LTE results in a packet-switched-only system, which makes wide use of Internet Engineering Task Force (IETF) protocols and practices, having been introduced in the 3GPP Release 8.

For the downlink, LTE uses OFDMA as its multiple access scheme and, for the uplink, it uses SC-FDMA. The first one was chosen due to its high data rate capacity and its high spectral efficiency, while the second one was chosen for its lower peak to average power ratio maximizing the battery life of the terminal, both providing orthogonality between the users, decreasing interference and enhancing the network capacity. In the frequency domain, the resource allocation occurs with a resolution of 180 kHz comprising 12 sub-carriers of 15 kHz each. In the packet scheduling, the frequency dimension is also responsible for the high LTE capacity. In the downlink, resource blocks are used freely from different parts of the spectrum, while, in the uplink, the allocation of a specific user is continuous in order to enable single carrier transmission, allowing efficient terminal power amplifier designed, which is important for the mobile devices battery life, as mentioned before. In the other hand, resource blocks are both frequency and time resources which occupy twelve 15 kHz sub-carriers and one 0.5 ms time slot.

The LTE system allows spectrum flexibility which can range from 1.4 MHz up to 20 MHz, considering the available spectrum. In ideal conditions, the 20 MHz bandwidth can provide up to 172.8 Mbps downlink user data rate with 2x2 Multiple-Input Multiple-Output (MIMO) and 340 Mbps with 4x4 MIMO. On the other hand, the uplink peak data rate is 86.4 Mbps.

2.4.2 Architecture

The LTE technology has been designed to support only PS services. Therefore, its architecture is marked by the removal of the CS domain and by a simplified access network. It provides seamless Internet Protocol (IP) connectivity between the UE and the PDN, without disrupting the applications of the end users during mobility.

The LTE system employs a new network architecture which comprises the User Equipment (UE), the Evolved UMTS Terrestrial Radio Access Network (E-UTRAN) and the Evolved Packet Core (EPC). Figure 2.8 shows the global network architecture including these elements just mentioned as well as standardized interfaces, which connect these same elements, allowing multi-vendor interoperability. LTE includes the evolution of radio access, through the E-UTRAN, along with the evolution of non-radio aspects, nominated by System Architecture Evolution (SAE), including the EPC network. The combination of LTE with SAE constitutes the Evolved Packet System (EPS), which provides the user with IP connectivity to a PDN for internet access purposes, supporting simultaneous services such as Voice over IP (VoIP), web browsing and File Transfer Protocol (FTP).

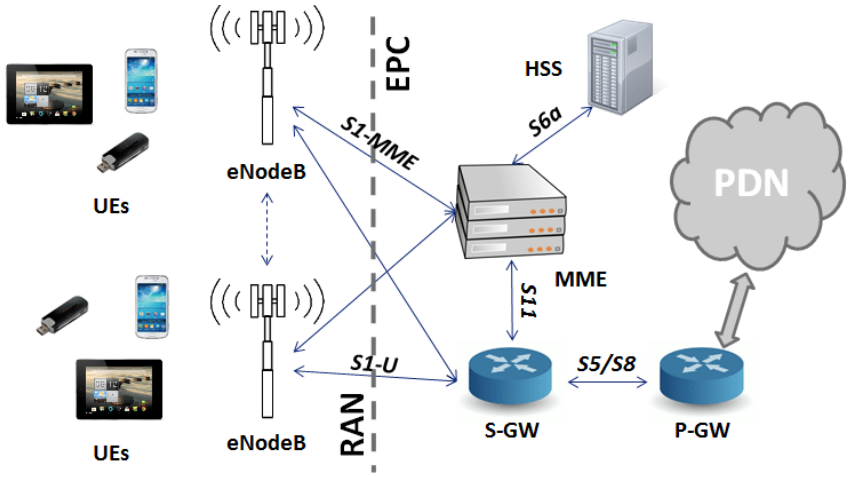


Figure 2.8: General architecture of a LTE network [26].

The **UE** is the interface through which the subscriber communicates with the E-UTRAN. It is composed by the following Network Elements (NEs):

- The **Mobile Equipment (ME)** is the actual communication device and it can be divided into both Mobile Termination (MT), which manages all the communication functions, and Terminal Equipment (TE), that terminates the data streams.
- The **Universal Integrated Circuit Card (UICC)** is a smart card, popularly known as the SIM card, which runs the USIM, which is an application that stores user-specific data such as the phone number and home network identity. Additionally, the UICC executes security procedures using

security keys which are stored by the smart card.

The **E-UTRAN** corresponds to the radio component of the architecture and is responsible for the radio communications between the UE and the EPC. It has only one component, the evolved Node B (eNB), which controls the mobiles in one or more cells, where each mobile communicates with only one base station and one cell at a time. The eNBs are usually inter-connected with each other through the X2 interface and to the EPC over the S1 interface. In particular, they are connected to the Mobility Management Entity (MME) by means of the S1-MME interface and also to the Serving Gateway (S-GW) through the S1-U interface.

The **eNB** contains two main functions:

- In the downlink, it sends radio transmissions to all its UE and, in the uplink, it receives radio transmissions from these devices.
- It controls the low-level operation of all its mobile devices through signalling messages, such as handover commands, which are related to those radio transmissions.

The eNB combines the earlier functions of the Node B and the RNC, from UMTS, in order to reduce the latency that occurs when the mobile exchanges information with the network.

On the other hand, the E-UTRAN is responsible for:

- Radio resource management, controlling the radio bearer, radio admission and radio mobility, and allocating dynamic resources to the UEs in both uplink and downlink.
- IP packet headers compression, ensuring an efficient use of the radio interface, especially for small packets such as VoIP.
- Security, encrypting all the data sent through the radio interface.
- Positioning, providing the required measurements and additional data to the Evolved Serving Mobile Location Centre (E-SMLC) in order to find the UE position.
- Connectivity to the EPC, signalling towards the MME and also the bearer path towards the S-GW.

The **EPC** represents the core network and is responsible for the overall UE control and the establishment of bearers, which are paths used by end users traffic to connect with PDN through the LTE transport network. The EPC is composed by the following logical nodes:

- **Mobility Management Entity (MME)**, responsible for the transition between the E-UTRAN and the EPC, is the main control node that handles the signalling between the UE and the EPC. The main functions supported by the MME are related to: bearer management, including the

establishment, maintenance and release of the bearers; connection management, including the security and the establishment of connection between the UE and the network by cooperating with the Home Subscriber Server (HSS); and inter-working with other networks, including the handover of voice calls to legacy networks.

- **Packet Data Network Gateway (P-GW)** connects the EPS with external PDNs and allocates the IP addresses designated for the UE. However, this allocation can also be performed by an external PDN, where the P-GW tunnels all traffic between the UE and that network. The P-GW is also responsible for the traffic gating and filtering functions that are required to provide the service.
- **Serving Gateway (S-GW)**, which links the eNB with the P-GW, serves as the local mobility anchor for the data bearers when the UE moves between several eNBs and it also retains information about these bearers during the UE's idle state.
- **Evolved Serving Mobile Location Centre (E-SMLC)** is responsible for the management of both overall coordination and scheduling of resources necessary to estimate the UE location. Based on the received estimations, it determines the final location and it also estimates the UE speed and corresponding accuracy.
- **Home Subscriber Server (HSS)** is a central database which contains information about all the network operator's subscribers, such as their QoS profile and any access restrictions for roaming. It holds information about PDNs, to which the user is able to connect, and also dynamic information such as MME identity, to which the user is currently connected to. The HSS may also integrate the AuC, that is responsible to generate the vectors used for both authentication and security keys.
- **Gateway Mobile Location Centre (GMLC)** incorporates the necessary functionalities to support LoCation Services (LCS). After being authorized, it sends positioning requests to the MME and collects the final location estimates.
- **Policy Control and Charging Rules Function (PCRF)** is responsible for policy control decision-making and manages the users' QoS and data charges. It is connected to the P-GW and sends information to it for enforcement.

2.4.3 Multiple Access Techniques Overview

In mobile data communications, the choice of an appropriate modulation and multiple access technique is critical in order to provide a satisfactory system performance. Therefore, with the purpose of fulfilling all the requirements defined for LTE, advances were made in this technology to both the multicarrier and multiple-antenna technology. The first major design choice was to adopt a muticarrier approach. As mentioned before, for the downlink, LTE uses a multiple access scheme based on the OFDMA

and, for uplink, it uses another one based one the SC-FDMA. Both of these schemes, represented in Figure 2.9, present the frequency domain as a new dimension of flexibility which offers an impressive new approach to improve not only the spectral efficiency of the system, but also to reduce both the inter-symbol interference and fading.

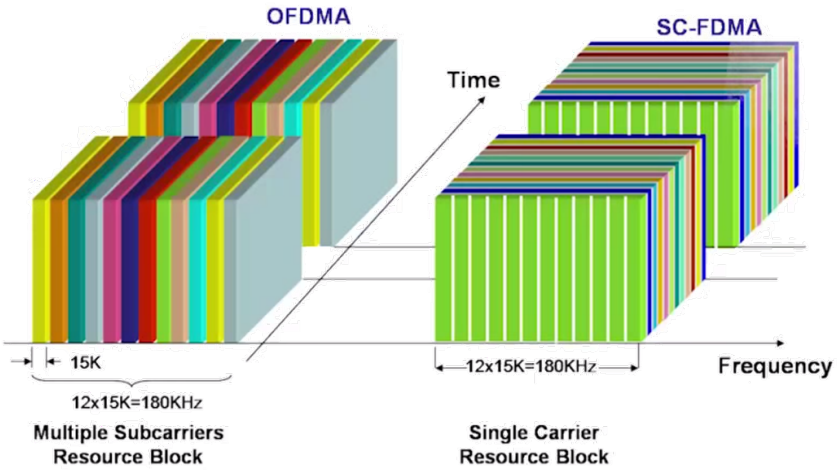


Figure 2.9: OFDMA and SC-FDMA in both frequency and time domains (adapted from [22]).

Here are some basic concepts about both OFDMA and SC-FDMA:

- In LTE transmission with the single carrier approach, this one is modulated in phase and/or amplitude and the corresponding spectrum wave from is a filtered single carrier spectrum which is centered on the carrier frequency.
- In a digital system, the higher the data rate, the higher the symbol rate and therefore the larger the bandwidth that is required for the same modulation. The transmitter is able to change the modulation in order to carry the required number of bits per symbol.
- In a FDMA system, since different carriers and sub-carriers are used, the system can be accessed by different simultaneous users. In this same system, it is important to avoid high interference between carriers without using long guard bands between users.
- In order to explore the spatial domain, multiple antenna technologies were considered to achieve higher spectral efficiency. Thus, in the first LTE release, the MIMO operation was introduced, which includes spatial multiplexing and precoding and transmit diversity as well. Its basic principle is represented in Figure 2.10, where different streams of data are fed to the precoding operation and forwarded to signal mapping and OFDMA signal generation.

In OFDMA systems the original bandwidth is sub-divided into multiple sub-carriers, of 15 kHz each, which are able to be individually modulated. Orthogonality is preserved between all sub-carriers in each

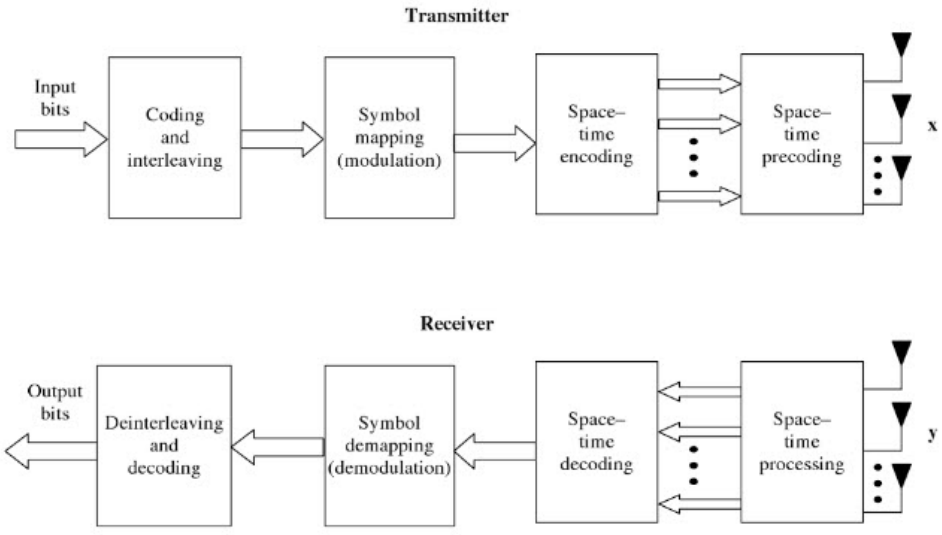


Figure 2.10: Basic block diagram of MIMO system (adapted from [22]).

sampling instant of a certain sub-carrier, as all other sub-carriers have a zero value, which is presented in Figure 2.11.

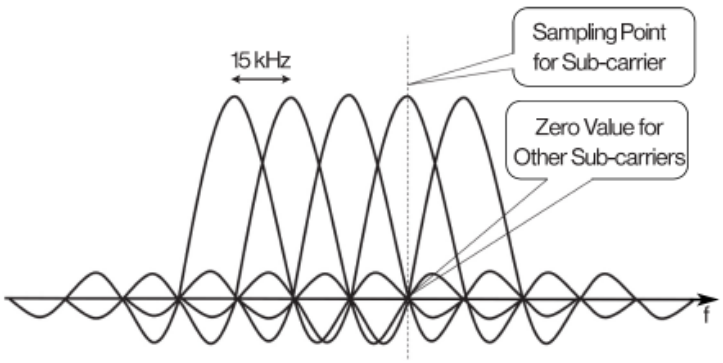


Figure 2.11: Orthogonality between OFDMA sub-carriers [15].

One of the most important advantages of OFDMA is the possible allocation of any of its sub-carriers to users in the frequency domain, providing benefits of frequency diversity to the scheduler. However, the signalling resolution caused by the resulting overhead does not allow the allocation of a single sub-carrier, requiring the use of a Physical Resource Block (PRB), represented in Figure 2.12, which contains 12 sub-carriers with a minimum bandwidth (available for allocation) of 180 kHz. In the time domain, this allocation corresponds to 1 ms, which is known as TTI, even though each PRB only lasts for 0.5 ms. In LTE, each PRB can be modulated through QAM, namely QPSK, 16-QAM, 64-QAM and 256-QAM.

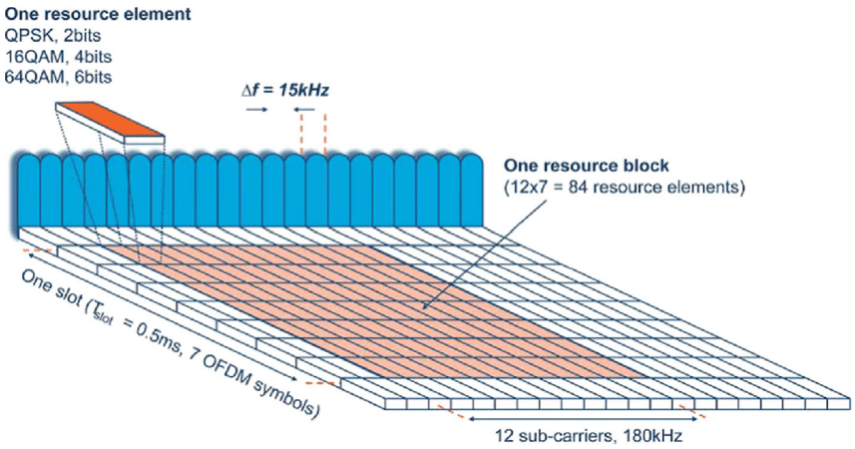


Figure 2.12: PRB in LTE system [27].

2.4.4 Capacity

In the LTE system, radio channels are used by the eNBs for both transmission (downlink) and reception (uplink) of user traffic. The usual radio channel bandwidths, defined by 3GPP, are 5, 10 and 20 MHz. The higher this bandwidth, the higher the volume of user traffic transmitted/received.

In order to maximize the throughput achieved by the subscriber, it is necessary to improve the quality of the radio signal received by the mobile device. This signal quality is affected by several factors: signal path loss, which is the attenuation of the signal as it propagates through the environment in which it is traveling; free space loss that attenuates the signal as the user moves away from the transmitting eNB; and diffraction. The throughput is also affected by the number of simultaneous users, mobility, interference and type of space (indoor/outdoor). The LTE standard defines certain types of modulation schemes: QPSK, 16-QAM and 64-QAM, represented in Figure 2.13. In later releases, the 256-QAM modulation scheme was also added.

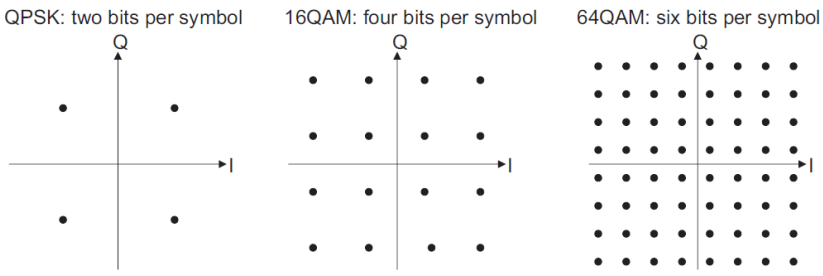


Figure 2.13: Constellations of QPSK, 16-QAM and 64-QAM modulation schemes [24].

If the subscriber approaches the border of the coverage area of the corresponding cell, the modulation will present less robustness, usually QPSK, thereby reducing the data rate received. On the other hand, the use of MIMO antennas, both in the transmitter and in the receiver, improves the system

performance.

The Resource Element (RE) is the smallest modulation structure in LTE technology. A RE is one 15 kHz sub-carrier by one symbol. Several REs aggregate into PRBs, as shown in Figure 2.14.

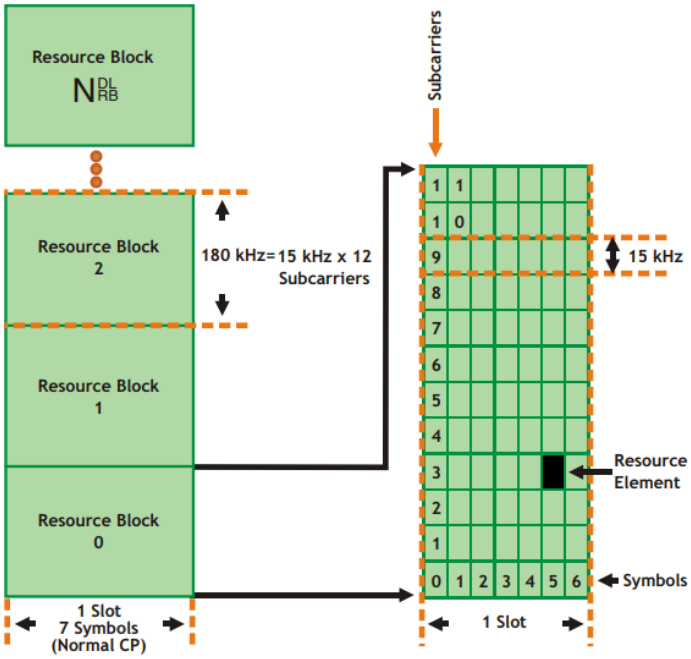


Figure 2.14: PRB structure [29].

The Table 2.4 shows the number of PRBs, the maximum occupied bandwidth, the guard band and the number of RE, depending on the channel bandwidth.

Table 2.4: LTE bandwidth characteristics [22].

Channel bandwidth [MHz]	5	10	20
Number of PRBs	25	50	100
Maximum occupied bandwidth [MHz]	4.5	9.0	18.0
Guard band on each side [kHz]	250	500	1000
Number of REs	4200	8400	16800

Taking into account the scheme presented in Figure 2.14 and assuming a channel bandwidth of 20 MHz, within the ones described in Table 2.4, the total number of REs per subframe is given by, [30]:

$$12 \text{ Subcarriers} \times 7 \text{ LTE symbol} \times 100 \text{ PRBs} \times 2 \text{ slots} = 16800 \text{ REs} \tag{2.6}$$

since 1 transport block contains 2 slots (1 ms). Assuming the utilization of 64-QAM modulation scheme

and no coding, each RE carries 6 bits. Therefore, the throughput is:

$$16800 \text{ REs} \times 6 \text{ bits/Symbol} / 1 \text{ ms} = 100.8 \text{ Mbps}. \tag{2.7}$$

Since there is 25% overhead in LTE, it brings down to:

$$108.8 \text{ Mbps} \times 0.75 = 75.6 \text{ Mbps}. \tag{2.8}$$

In the Release 10 (LTE-Advanced), the MIMO 4x4 is used, so the throughput is finally given by:

$$75.6 \text{ Mbps x } 4 = 302.4 \text{ Mbps}. \tag{2.9}$$

2.4.5 LTE-Advanced

In 2008, the International Telecommunication Union (ITU) publishes a set of requirements for a 4G communication system under the name International Mobile Telecommunications (IMT)-Advanced. According to these requirements, the peak data rate of a compatible system should be at least 600 Mbps in the downlink and 270 Mbps in the uplink, with a 40 MHz bandwidth, which exceed the capabilities of LTE.

3GPP started to study how to enhance the capabilities of LTE, introducing the LTE-Advanced, which was required to deliver a peak data rate of 1000 Mbps in the downlink and 500 Mbps in the uplink. In practice, the system has been designed to deliver peak data rates of 3000 and 1500 Mbps for downlink and uplink, respectively, using a total bandwidth of 100 MHz consisting of 5 separate components of 20 MHz each, however, this is unreachable in any realistic scenario.

This specification also includes targets for the spectrum efficiency, where the comparison with the corresponding figures for WCDMA presupposes a spectral efficiency 4.5 to 7 times higher than that of Release 6 WCDMA in downlink and 3.5 to 6 times higher in uplink. Finally, LTE-Advanced is design backwards compatible with LTE, so that a LTE mobile device can communicate with a eNB that is operating LTE-Advanced and vice-versa.

2.5 5G Background

This section provides an overview of the new technology 5G, presenting a brief introduction to this recent system, followed by a description of the main use cases: enhanced Mobile BroadBand (eMBB), mobile Machine-Type Communications (mMTC) and Ultra-Reliable Low Latency Communications (URLLC), and then an overview of the two different types of 5G New Radio (NR) architectures. Thereafter, there is a subsection about capacity in 5G and, finally, there is a special focus on green networks and their tremendous importance in the upcoming mobile systems.

The content of this section is mainly based on the following references: [31–34] in Subsection 2.5.1; [35–37] in Subsection 2.5.2; [34–36, 38] in Subsection 2.5.3; [39–41] in Subsection 2.5.4; [42–47] in Subsection 2.5.5.

2.5.1 Introduction to 5G

Mobile communication technologies have been evolving over the years with each generation transforming the way people communicate and access information. As the wireless communications market has significantly expanded in the past years, being expected to grow even more in coming years, the network evolution must continue to respond to the user's demand. The predicted market space for the next generation of technology is driven by requirements to improve mobile broadband connectivity, obtain a wide variety of machine-type communications and target services with ultra-reliable and low latency communications.

Unlike previous generations of mobile networks, 5G is expected to essentially transform the role that telecommunications technology plays in the society. Thus, this technology should support significantly faster mobile broadband speeds and heavier data usage than previous generations. It will also enable the full potential of Internet of Things (IoT), with use cases such as smart cities, smart agriculture, autonomous cars, industrial internet, fibre-over-the-air, logistics and public safety agencies.

5G is a system designed to meet the requirements of IMT-2020 set by ITU. It will provide more advanced and improved capabilities compared to 4G LTE (IMT-Advanced), which can be verified by Figure 2.15. More specifically, Table 2.5 summarizes the main key performance parameters of IMT-2020.

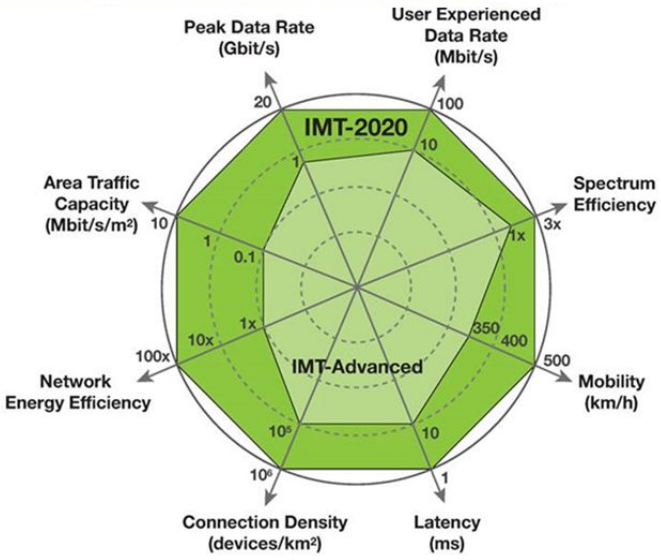


Figure 2.15: Key capabilities enhancement from IMT-Advanced to IMT-2020 [33].

Table 2.5: IMT-2020 key performance requirements [34].

Requirement		Value
Data Rate	Peak	Downlink: 20 Gbps Uplink: 10 Gbps
	User experienced	Downlink: 100 Mbps Uplink: 50 Mbps
Spectral efficiency	Peak	Downlink: 30 bits/s/Hz Uplink: 15 bit/s/Hz
	5^{th} percentile user	Downlink: 0.12∼0.3 bit/s/Hz Uplink: 0.045∼0.21 bit/s/Hz
	Average	Downlink: 3.3∼9 bit/s/Hz Uplink: 1.6∼6.75 bit/s/Hz
Area traffic capacity		10 Mbit/s/m^2
Latency	User plane	1 ms $\sim$ 4 ms
	Control plane	20 ms
Connection density		10^6 devices per km^2
Energy efficiency		Loaded: see average spectral efficiency No data: Sleep ratio
Reliability		1-10^{-5} success probability of transmitting a layer 2 PDU (Protocol Data Unit) of 32 bytes within 1 ms
Mobility		Up to 500 km/h
Mobility interruption time		0 ms
Bandwidth		100 MHz

2.5.2 Use cases

5G will be defined in a set of standardized specifications which will be set most notably by the 3GPP and later by the ITU in 2020. The ITU-R has defined three main use cases classes:

- **eMBB** aims to provide access to multimedia content, services and data, dealing with increased data rates, high user density and very high traffic capacity for hot-spot scenarios as well as consistent coverage and high mobility scenarios with still improved user data rates. This would be accomplished through the installation of millimeter Wave (mmWave) antennas all over the landscape (e.g. on lampposts, sides of buildings, branches of trees, electrical towers and on top of city buses). However, these antennas would cover an area too small, so it would be necessary to implement thousands of them in order to provide good coverage along with high users throughput. The eMBB use case is focused on services characterized by high data rates, such as Augmented Reality (AR), Virtual Reality (VR) and video in all its several forms and formats.
- **mMTC** enables the Machine-to-Machine (M2M) and IoT applications, being able to support at least one million IoT connections per km^2 with extremely long battery life and wide coverage, even inside

buildings. It would seek to restore the service level, once degraded by later generations after 2G in the narrow service bands which were required by M2M devices, by developing a compartmentalized service tier for devices requiring low throughput as 100 kbps, but with low latency (around 10 ms). The concept of mMTC is focused on services characterized by high connection density requirements, such as those applied to smart city and smart agriculture.

- **URLLC** would address safety-critical and mission critical applications, including and end-to-end latency of 1 ms or even less and very high availability, reliability and security, in order to support: real-time applications, avoiding perceptible lags in browsing and videos; drones and robots control; latency-sensitive services, such as autonomous vehicles, remote medical surgery and mobile healthcare. There is actually the possibility, still under discussion, of URLLC making 5G replacing Global Positioning System (GPS) for geolocation.

Figure 2.16 illustrates the usage scenarios of IMT for 2020 and beyond.

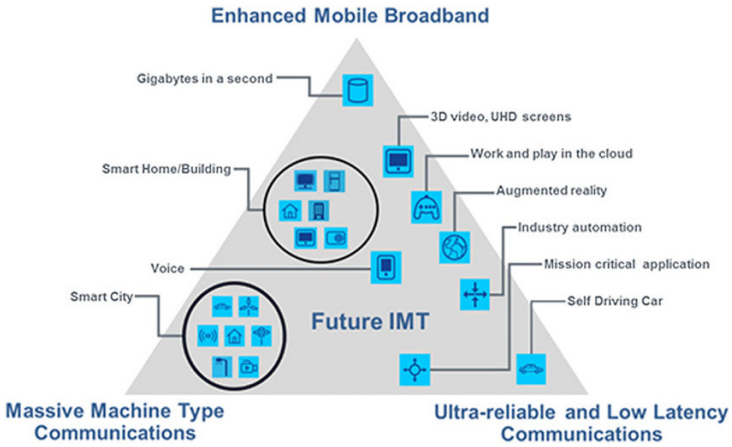


Figure 2.16: Use cases and scenarios of IMT-2020 [37].

2.5.3 Architecture

The current architectures were developed in order to provide personal communication services and content, such as voice, video and web browsing. However, to achieve ultra-low latency the data and control planes can demand substantial enhancements and also new technical solutions, covering aspects not only of radio interface but also of network architecture. Therefore, the architecture evaluation is strictly necessary to be able to provide a wider range of new services and applications with several distinct characteristics.

As in previous generations, 3GPP is defining for 5G both a new core network, called 5G Core (5GC), as well as a new radio access technology, known as 5G NR, which provides substantial throughput and

low latency. In 5G it is possible to integrate elements of distinct generations in distinct configurations, unlike previous generations, which demanded that both access and core networks of the same generation to be deployed.

The two solutions defined by 3GPP for 5G networks are:

- **Non-Standalone (NSA)** combining multiple radio access technologies. In this scenario, the 4G LTE infrastructure will support connecting to a 5G NR base station, referred to as next generation Node B (gNB), enabling dual connectivity. The core network can be either 5GC or 4G EPC, depending on the choice of each mobile operator. Figure 2.17 presents the E-UTRAN New Radio – Dual Connectivity (EN-DC) architecture and shows 5G gNBs, also known as en-gNBs, connected to the LTE EPC. A gNB does not connect to the MME, but it is able to connect to the eNB in order to receive requests to activate/deactivate 5G bearers. Thus, users will be able to support 5G mobile broadband services in a primarily 4G network, considering that their mobile devices support dual connectivity. With this solution, operators are able to provide 5G services to the users with shorter time and lesser cost.

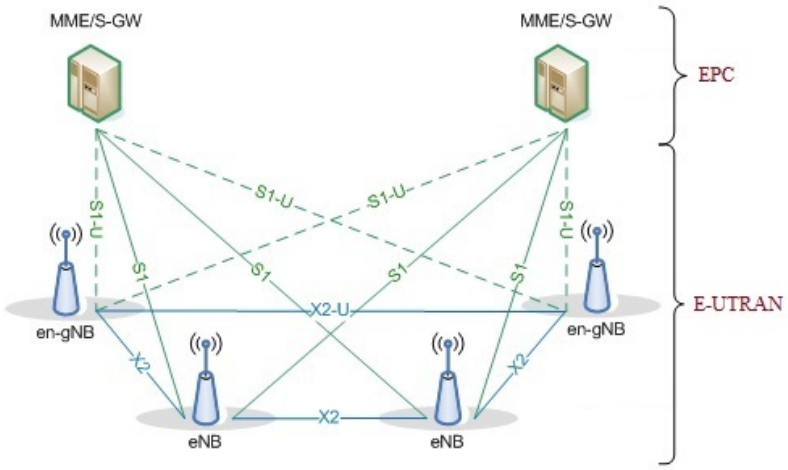


Figure 2.17: 5G EN-DC architecture (adapted from [38]).

- **Standalone (SA)** using only one radio access technology. It is an end-to-end 5G solution, presenting an architecture constituted by 5G NR and 5GC. In this situation, the 5G NR is used for both control plane and user plane. For the network operators, this is a simple solution which may be deployed as an independent network using normal inter-generation handover between 4G and 5G in order to provide service continuity. Figure 2.18 illustrates this present scenario. The Next Generation – Radio Access Network (NG-RAN) will provide higher access performance and support new 5G characteristics, being constituted by several gNBs interconnected with each other by the Xn interface. These gNB are connected by the NG interface to the 5GC as well.

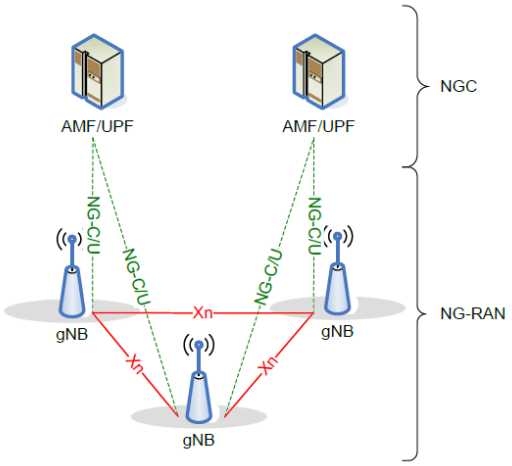


Figure 2.18: 5G NG-RAN architecture [35].

2.5.4 Capacity

In order to support the strong main use cases established for the 5G networks (eMBB, mMTC and URLLC), 5G NR requires a scalable and flexible physical layer design. With this purpose, 3GPP introduced a set of parameters, such as, [39]:

- **Cyclic prefix**, which is required to manage inter-symbol interference due to multiple path signals. Its length is a trade-off between cyclic prefix overhead and inter-symbol interference and its selection is determined by the environment (outdoor/indoor), frequency band and type of service.
- **Subcarrier spacing**, which defines the bandwidth of a single subcarrier, can be represented by Δf = 2μ x 15 kHz, where μ is the NR numerology constant, which can be 0, 1, 2, 3 and 4. It is a trade-off between symbol duration and cyclic prefix overhead.

Since 5G NR supports multiple numerologies, the radio frame structure depends on the type of numerology, μ. In every case the duration of one radio frame is always 10 ms and one subframe is 1 ms, regardless of the numerology. Figure 2.19 shows an example of the radio frame structure. In this configuration, for normal cyclic prefix and numerology equal to 1, a subframe contains 2 slots in it, meaning that a radio frame comprises of 20 slots.

In 5G NR, one PRB contains 14 symbols, in time domain, and 12 subcarriers, in frequency domain. Unlike LTE, where the PRB bandwidth is fixed, in NR it depends on the subcarrier spacing. Therefore, the bandwidth of each PRB is given by, [41]:

$$\text{PRB Bandwidth} = \Delta f \times 12 \text{ Subcarriers.} \tag{2.10}$$

NR is developed in order to increase bandwidth efficiency reaching to 99%, while in LTE reached

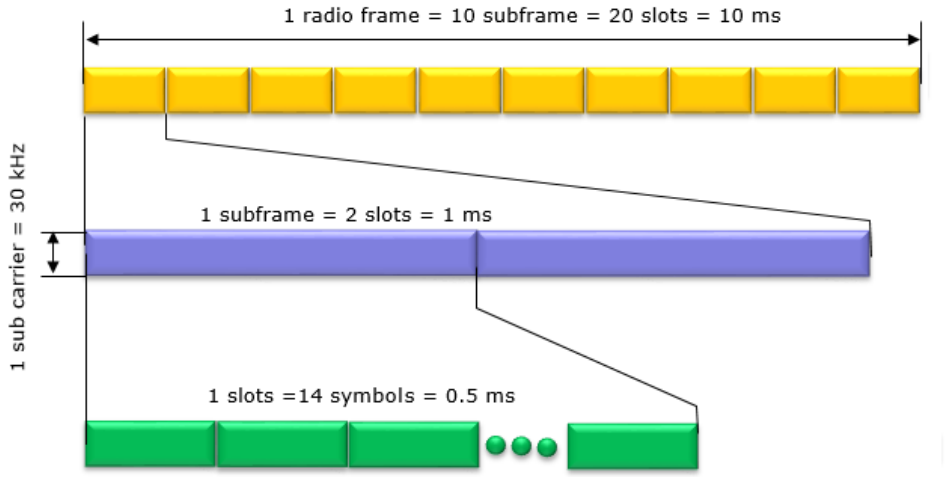


Figure 2.19: Radio frame structure for normal cyclic prefix and μ = 1 (adapted from [40]).

about 90%. Thus, NR will provide 100 MHz channel bandwidth for lower frequency bands (below 6 MHz) and 40 MHz channel bandwidth for higher frequency bands (mmWave ranges). Knowing the bandwidth of one PRB, since each numerology has already set the minimum and maximum number of PRBs, it is possible to determine the minimum and maximum channel bandwidth.

3GPP has already specified a maximum transmission bandwidth configuration for each UE channel and subcarrier spacing. Thus, the maximum number of PRBs can be estimated by:

$$\text{Number of PRB} = (\text{Channel Bandwidth} - 2 \times \text{Guard Bandwidth}) / \text{PRB Bandwidth}. \tag{2.11}$$

5G NR supports multiple modulation schemes, such as QPSK, 16-QAM, 64-QAM and 256-QAM, whose symbol can carry 2 bits, 4 bits, 6 bits and 8 bits, respectively, commonly known as modulation order. Some of these modulation schemes were already represented before in Figure 2.13.

In order to support a wide range of services with distinct requirements and the increase in the number of connected devices, additional spectrum was released. Thus, NR bands are designated for different Frequency Range (FR) and can be defined as FR1, which consists of all band below 6 GHz, and FR2, which consists of all band above 24 GHz, including mmWaves.

The maximum data rate supported by the UE, for both downlink and uplink, is determined by band combinations and baseband processing combinations supported by the UE. In 5G NR, for a given number of aggregated carriers in a band or band combination, the approximate data rate is calculated by, [41]:

$$\text{Data rate (Mbps)} = 10^{-6} \cdot \sum_{j=1}^{J} \left(\nu_{Layers}^{(j)} \cdot Q_m^{(j)} \cdot f^{(j)} \cdot R_{max} \cdot \frac{N_{PRB}^{BW(j),\mu} \cdot 12}{T_s^{\mu}} \cdot \left(1 - OH^{(j)}\right) \right), \tag{2.12}$$

where:

- J is the number of aggregated component carriers in a band or band combination;
- R_{max} = 948/1024;
- For the j-th carrier component:

  $\nu_{Layers}^{(j)}$ is the maximum number of layers;

  $Q_m^{(j)}$ is the maximum modulation order;

  $f^{(j)}$ is the scaling factor and it can take the values 1, 0.8, 0.75 and 0.4;

  μ is the numerology, as mentioned before;

  T_s^μ is the average Orthogonal Frequency Division Multiplexing (OFDM) symbol duration in a subframe for numerology μ, i.e. $T_s^\mu = \frac{10^{-3}}{14 \cdot 2^\mu}$, assuming normal cyclic prefix;

  $BW^{(j)}$ is the UE supported maximum bandwidth in the given band or band combination;

  $N_{PRB}^{BW(j),\mu}$ is the maximum PRB allocation in bandwidth $BW^{(j)}$ with numerology μ;

  $OH^{(j)}$ is the overhead and takes the following values: 0.14 for FR1, 0.18 for FR2 in downlink; 0.08 for FR1 and 0.10 for FR2 in uplink.

The approximate maximum data rate may be computed as the maximum of the approximate data rates computed through Equation (2.12) for each of the supported band or band combinations.

2.5.5 Green Networks

Global wireless data traffic is expected to continue to grow rapidly due to the proliferation of smart devices and applications. Advanced application such as high-resolution video streaming, tactile internet, remote monitoring, road safety and real-time control application, will produce an extensive amount of data traffic. Additionally, services with ultra-reliable and low latency communications such as connected cards and moving robots must be supported in efficient scalable ways. In order to provide all these services, 5G aims to provide, by 2020, a system able to support 1000 times increase in capacity, 1000 times high mobile data volume per area, 100 times more connected devices and 5 times reduced end-to-end delay than the current LTE network. Therefore, it has been predicted that 5G systems would increase network energy consumption 150-170% by 2026 and more than 90% of mobile operators believe 5G will result in higher energy costs.

In order to make the network more energy efficient, new technologies focused on power optimization were developed:

- **Simultaneous Wireless Information and Power Transfer (SWIPT)**, which exploits the same emitted electromagnetic wave field to deliver both energy and information, thereby improving transmission efficiency. Devices that support SWIPT are able to fetch energy when receiving data, thus prolonging their lifetime.
- **mmWaves** that operate in the 30-300 GHz band are a promising technology for the upcoming wireless systems. This technology can significantly increase the RF channel bandwidth, thereby increasing the data capacity.
- **Small cells** are operator-controlled, low-powered and low-cost BSs which operate in licensed spectrum. Reducing the distance between the UE and the BS, they reduce the transmit power required improving energy efficiency for both uplink and downlink communications. It has been predicted that 5G small cell deployments will overtake 4G small cells by 2024, with a 13.1 million installed base of 5G or multimode small cells in 2025.
- **Massive MIMO**, by employing a large number of antennas, the BS is able to communicate with multiple users simultaneously in the same frequency band. Using beamforming, by focusing the radiation pattern, improves signal quality, increases the communication range and also provides lower power consumption. However, the large amount of antennas in massive MIMO consumes high circuit power, resulting in a considerable reduction in energy efficiency.
- **Relays** allow the mobile devices to forward information when they are neither the source nor the destination of this information. As expected, it extends high data rate coverage to the cell edge, increasing the coverage area of transmission and the network throughput, as well as it decreases the transmission power, improving energy efficiency.

Figure 2.20 illustrates a comparison between the current network, with an antenna for 4G BSs, and the new network, with massive MIMO and beamforming for 5G BSs.

Taking into account these aspects and also the fact that the energy consumed by the IoT devices is low, their multiplicity may generate some problems in achieving the desired energy efficiency. Therefore, some algorithms have been developed for BSs of 5G networks in order to improve energy efficiency, reduce carbon footprint and develop a self-sustainable green cellular network. To begin with, the BSs can operate in the following modes:

- **Active mode**, where the BS, as a regular one, fully functions (normal reception and transmission).
- **Sleep mode**, where the BS neither receives nor transmits any user traffic, consuming minimal power since only a wake-up receiver module remains active. This module waits for requests from the other BSs or for power increment requests, helping the current BS to switch again to active mode.

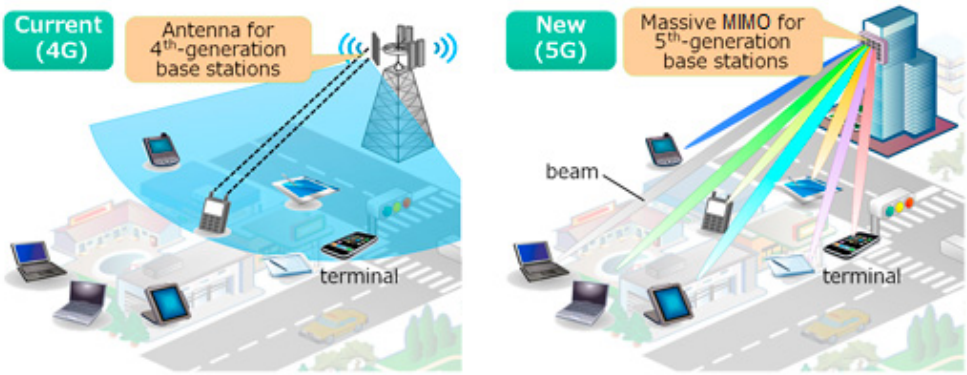


Figure 2.20: Current (4G) and new (5G) BSs (adapted from [47]).

- **Turned-off mode**, where the BS is disconnected from the power supply, becoming a super power saving mode. In this case, the BS is periodically switched on, at regular intervals, with the purpose of checking the neighborhood network environment.

In these algorithms, a BS is able to dynamically decide/switch between active and sleep modes taking into account the network traffic. Each BS, after estimating its own traffic load and one-hop neighborhood traffic load information, decides its mode of operation. However, in the case of BS switching, this new BS checks whether it can satisfy the QoS requirements of its own users and the newly arrived users, before the switching. Particularly, a lightly loaded BS aims to switch to sleep mode by transferring its load to moderately loaded BSs. On the other hand, a heavily loaded BS wants to share part of its traffic load with also moderately loaded neighbor BSs in order to balance the traffic loads and also satisfy the users QoS. These BSs cooperate to optimize the energy saving within the 5G green networks.

2.6 Energy Consumption Monitoring

This section presents various solutions of energy consumption optimization using three distinct approaches: smart metering, focusing in the smart meters and IoT; energy harvesting, using solar power, wind energy, etc.; and, finally, energy saving methods applying different techniques to the BSs.

The content of this section is mainly based on the following references: [48, 49] in Subsection 2.6.1; [50–52] in Subsection 2.6.2; [53–58] in Subsection 2.6.3.

2.6.1 Smart Metering

The current power grid needs to constantly evolve in order to adapt to the increasing technological demands, particularly because of the evolution of renewable energy sources, since it has implied several

changes in the way the energy is both produced and stored. In addition, the need to reduce the total CO_2 emissions requests for a more efficient energy usage. Modern power grids are being adapted providing a more efficient Energy Management (EM) with the support of automated control and modern telecommunication technologies. Therefore, a smart grid implies the handling of an extensive amount of information.

Smart metering transmits energy consumption/generation data towards the utilities and information data towards the smart meters, which play an important role in energy management/saving, allowing the costumers to measure/control the power consumption of their electrical appliances. The smart meter communicates with the router, which sends the data to the energy supplier through Internet. The integration of intelligent measuring devices in a city using IoT is essential to collect all the data needed to become a smart city, keeping the city connected/informed and ensuring each subsystem performs its function. The goal of this system is to provide a balance between demand and consumption. Smart metering enables a two way real-time communication between the meter and the utility central system, providing continuous consumption reading/recording in certain time intervals (at least daily reporting/-monitoring/billing) to the utilities.

From the utilities point of view, this system contains several benefits:

- It provides staff saving (there is no need to physically go to read the meters);
- It permits the detection and subsequent reparation of leaks (which would cause damage and waste energy);
- It is a contribute to the environmental sustainability;
- It improves efficiency in the distribution of electricity, water and gas flows;
- It detects counterfeiting/tampering; enables a quick identification of problems, which may occur, and later a quick problem solving.

On the other hand, from the consumers point of view, the main benefits of smart metering are:

- Efficient consumption and leak detection, which reduces the bill;
- Costumers can observe, in real-time, how much and when they consume;
- They are able to make decisions about their consumption (through the tools provided);
- Detection of electricity, water and gas theft.

Therefore, it is clear that telecommunication companies should invest to associate smart metering to the BSs consumption in order to reduce the energy waste, saving money and improving the environmental sustainability.

2.6.2 Energy Harvesting

Lately, one of the biggest global issues of concern is to obtain a sustainable form of energy to power several autonomous wireless and portable devices and, therefore, numerous alternate sources have been explored.

Energy harvesting collects useful energy from wasted external energy sources (e.g. solar power, thermal energy, wind energy, salinity gradients and kinetic energy) which is captured and stored for small wireless autonomous devices. This system is probably the most promising technology that provides solutions for several today's problems: shortage of a fossil fuel and environmental problems, and improves the energy efficiency in smart grids.

One of the first methods of energy harvesting was the conversion of solar energy to electrical energy, however, this concept has been expanded to other sources: RF energy, piezoelectric energy, thermal energy, etc. Thus, it is possible to use ambient RF energy sources from distinct wireless sources: mobile phones, BSs, televisions, wireless Local Area Networks (LANs), radar and radio broadcasts.

The development of BSs with the support of energy harvesting process reduces operating expenses by means of wind, solar radiation, etc. which provide a part of the required energy for the BS to be able to operate. With the evolution of wireless backhaul, removing the need for wired connections, there is a special interest in autonomous energy harvesting base station (particularly for smart cells or for use as relay stations).

2.6.3 Energy Saving

The evolution of 2G, 3G and 4G mobile communications technologies and the emergence of 5G have increased the variety and quantity of functionalities provided by the network, which has a major impact on energy consumption, along with the gradual increase of wireless technologies subscribers that increments the BSs operation hours and traffic management, in order to guarantee an adequate QoS for end users anywhere and anytime. In order to illustrate this last statement, the Figure 2.21 presents the number of AT&T[1] wireless subscribers each year from 2007 to 2019 and these values are expected to rise further in the coming years.

It is estimated that more than 90% of the wireless networks energy consumption is part of the operating expenses of the network operators, becoming more interested in implementing several energy consumption minimization procedures related to the BSs, such as those already mentioned previously: to produce energy coming from renewable sources, in each BS location, which at the same time reduces the polluting agents emissions in the atmosphere, and to implement intelligent monitoring systems (smart meters) which monitor energy consumption in order to better manage this consumption.

As shown in several studies, these procedures are able to save a considerable amount of energy,

[1] AT&T is an American telecommunications company and the largest in the world [57].

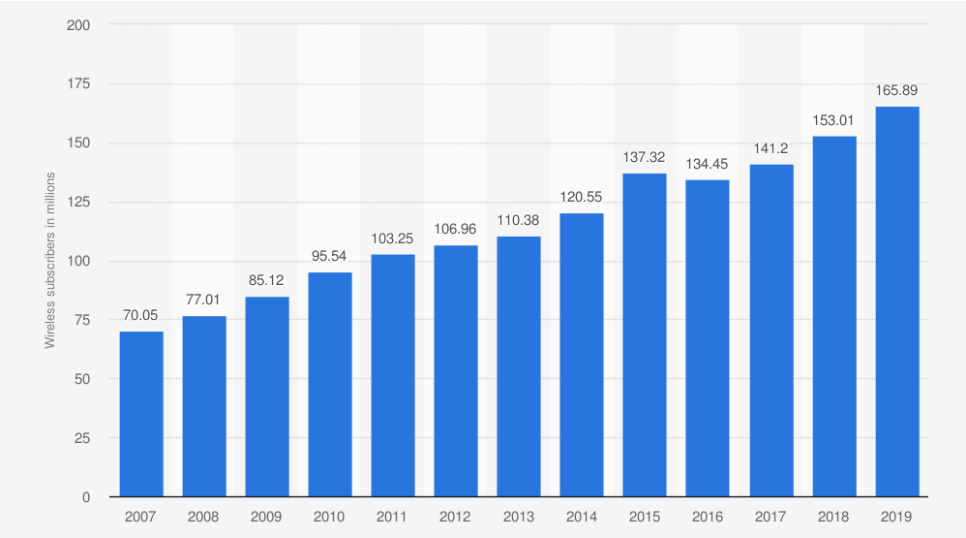


Figure 2.21: Wireless subscribers evolution along the years [57].

since the BSs consume the largest proportion of energy in cellular networks, which can be observed in Figure 2.22.

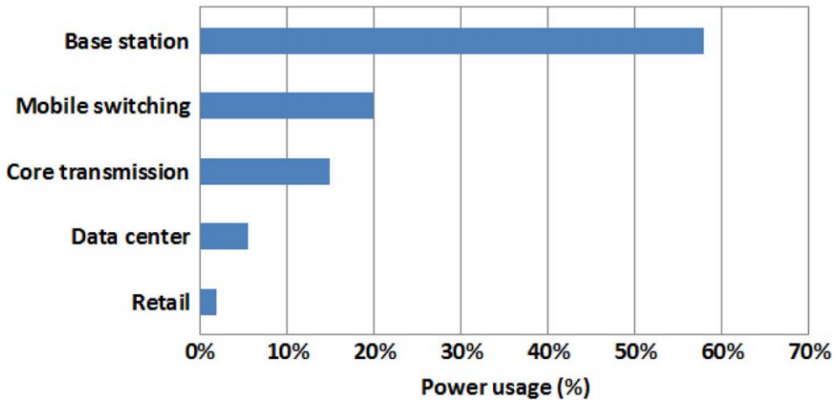


Figure 2.22: Breakdown of energy consumption in cellular networks (adapted from [58]).

This reflects the tremendous important of minimizing BS energy consumption in order to reduce cost and CO_2 emissions. Figure 2.23 shows the energy allocation in the BS per function, which proves that radio equipment is the section with the greatest potential for energy savings.

Analyzing the behavior of energy consumption as a function of the load of mobile traffic of several stations, it is perfectly visible that the first one vary daily along the second one. However, in the stations without active energy management system (traffic load based), during the days of less traffic (weekends, holidays) energy consumption does not change. Therefore, to reduce this same aspect, the transceivers of a BS can be turned off/on according with the traffic needs. ”Sleep mode” techniques intend to turn off BSs during ”off-peak” hours avoiding unnecessary energy consumption.

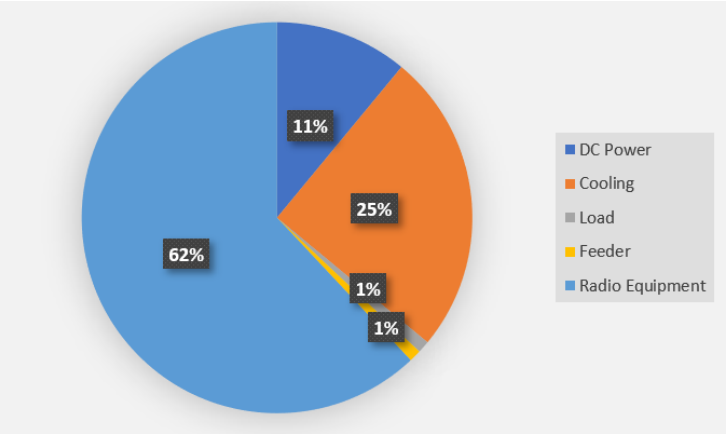


Figure 2.23: Percent BS Energy per function (adapted from [56]).

There are other techniques to minimize the energy consumption through BSs. Some of them only take into account transmission power while others consider the generation (of the radio signal) power consumption in circuitry by applying power models, which represent the amount of power consumed by a transmitter and its dependence on operating parameters.

2.7 Performance Data Collection

It is necessary to develop new managing and monitoring data due to the increased complexity of telecommunications networks. Thus, mobile operators utilize a set of methods which allow to collect data from the networks. These methods are used to plan and optimize the networks and for operators to know if they are providing the demanded quality to their subscribers.

The content of this section is mainly based on the following references: [59, 60] in Subsection 2.7.1; [61] in Subsection 2.7.2; [62] in Subsection 2.7.3.

2.7.1 Performance Management

Performance Management (PM) consists of evaluating and reporting the behaviour and effectiveness of network elements by gathering statistical information, maintaining and analyzing historical logs, determining system performance and adjusting the system modes of operation. This was one of the concepts, along with configuration, security, accounting and fault, that were added to the Telecommunication Management Network (TMN), which is the framework used to manage telecommunications networks and services, defined by the ITU, allowing optimization of the network utilization, which provides an increased QoS for the end users.

PM is the measurement of network and application traffic with the purpose of provide a consistent and predictable service level at a given instance and over a defined time period. It is also a key role

for network operators and vendors since it allows them to detect the deteriorating trend in advance and thereby solve potential threats, which prevents faults.

The architecture of a PM system is composed of four layers:

- **Data Collection and Parsing Layer**, where data is collected from the NEs utilizing network specific protocols, such as FTP and Simple Network Management Protocol (SNMP);
- **Data Storage and Management Layer**, where the parsed data is stored into a database;
- **Application Layer**, which processes the collected and stored data;
- **Presentation Layer**, which provides a web-based user interface where the generated PM results are presented in the form of dashboards and real-time graphs and charts.

One of the main challenges in PM is to perform an efficient administration due to the collection, aggregating, sorting and processing extensive volumes of performance measurement data which is collected over time periods. Performance measurements do not have an unified structure, since each NEs manufacturer contains proprietary protocols and data structures to evaluate the performance in their devices, which is challenging as well.

2.7.2 Key Performance Indicators

A Key Performance Indicator (KPI) is a measurable value that demonstrates how effectively an organization is achieving an intended result. More specifically, in the context of telecommunications, KPIs are measures of the network components performance which the operators use as a measure to evaluate the quality of the services they provide. KPIs are derived from statistical calculations based on NEs installed counters, which register several indicators (e.g. failed handovers, number of voice calls and drop calls). KPIs are fundamental in PM analysis and also in verifying that all elements at various levels of the network are using consistent strategies in order to achieve the shared goals. This allows to accurately identify where an action must be taken to improve network's performance.

There are several KPI categories:

- **Accessibility** indicates the probability of a user to access the requested service from the system (e.g. Call Setup Success Rate);
- **Availability** is the percentage of time that a cell is available (e.g. Radio Network Unavailability Rate);
- **Service Integrity** indicates the E-UTRAN impacts on the service quality provided to the end user (e.g. User Downlink Average Throughput);

- **Mobility** evaluates the successful ratio for several kinds of handover features or service mode changing (e.g. Intra-Frequency Handover Out Success Rate);
- **Retainability** is the ability of a user to retain its requested service for the required duration by itself once connected (e.g. Service Drop Rate);
- **Utilization** evaluates the capability to meet traffic demands in certain conditions (e.g. Resource Block Utilizing Rate);
- **Traffic** is used to measure the traffic volume (e.g. Uplink Traffic Volume).

2.7.3 Configuration Management

Configuration Management (CM) gives the operator the ability to ensure proper and effective operation of the network as it evolves. The purpose of CM actions is to both control and monitor the active configuration of the network resources and elements. These actions can be initiated by the operator or by functions in the Operating System (OS) or NEs and may be carried out as part of implementation programs, optimization programs and to preserve the overall QoS.

CM Service Components

After the first network installation and activation, the operator enhances and adapts the network in order to fulfill short and long requirements and to satisfy costumer needs as well. Thus, CM provides the network operator with a set of capabilities, such as initial system installation, system operation to adapt the system to short term requirements, system update to overcome software bugs or equipment faults and system upgrade to enhance or extend the network by features or equipment respectively. These capabilities are provided by the management system through its service components: system modification, used to adapt the system data to a new requirement due to optimization or new network configurations; and system monitoring, which allows the network operator to receive reports on the configuration of the entire work or parts of it, from managed NEs.

CM Functions

System modification functions were defined due to the CM requirements, such as the creation, deletion and conditioning of network elements and resources. The following requirements are applied for all these previous functions:

- Affected resources should be taken out of service only if necessary, resulting in minimal disruption to the network;
- Independence between physical modifications and related logical modifications should be assured;

- All the actions required to fulfill a certain task should be completed before the resources are brought back to the service;
- Data consistency checks should be performed.

3

Machine Learning Algorithms

Contents

This chapter presents an overview of ML concepts and algorithms thus enabling a better understanding of the work that will be developed along this Thesis. Section 3.1 provides a brief introduction to ML and is followed by a description of the main challenges in the creation of a ML model in Section 3.2. Section 3.3 introduces the MLR approach, which will be used within the development of the present study. Finally, Section 3.6 presents the metrics considered in order to evaluate the hypotheses that will be formulated.

The content of this chapter is mainly based on the following references: [63,64] in Section 3.1; [65,66] in Section 3.2; [64, 67, 68] in Section 3.3; [68–70] in Section 3.4; [71] in Section 3.5; [66, 72] in Section 3.6.

3.1 Machine Learning Overview

In order to solve a problem on a computer, an algorithm is required. An algorithm is a sequence of instructions intended to transform the input to output. However, for some tasks it is not possible to build an algorithm, such as detecting spam from legitimate emails. In this case, it is not known how to transform the input into output (an email document consisting of files with characters in the simplest case) to the output (a yes or no, indicating whether the message is spam or not).

There is a process that explains the observed data, through the analysis of certain patterns in the data. However, the process cannot be completely identified, but a useful approximation. These approximation may not explain everything, but can at least explain some part of the data, over the detection of certain patterns or regularities, which may help in understanding the process or making predictions, assuming the near feature will not be much different from the past (moment when the sample data was collected). These future predictions can also be right.

ML uses the theory of statistics to develop mathematical models, in order to make inference from a sample. It starts with an initial model consisting of several parameters that should be optimized by a learning algorithm which utilizes a set of training data. There are four types of algorithms:

- **Supervised learning** utilizes labeled data, which have both input and output parameters, in order to predict the output from the input data;
- **Unsupervised learning** uses unlabeled data, with only input parameters, to gain knowledge from this data;
- **Semi-supervised learning** uses mainly unlabeled data and a mixture of both supervised and unsupervised techniques;
- **Reinforcement learning** utilizes observations gathered from interactions with the environment in order to take actions which will maximize the reward or minimize the risk.

In the present study, only supervised learning models are addressed, since it is essentially intended to predict a given dependent variable through independent variables in order to find a relationship between them. Supervised learning problems can be further grouped into:

- **Classification problems**, when the output variable is a category (e.g. ”spam” or ”not spam, ”red” or ”blue”);
- **Regression problems**, when the output variable is a continuous value (e.g. price, height, cell throughput).

Thus, the models used and developed are based on the MLR analysis, whose overview will be presented in Section 3.3.

Machine Learning Components

There are a wide variety of available learning algorithms which raises certain doubts in choosing the most suitable one for each specific case. This problem consists of combinations of some components:

- **Representation**: a classifier must be represented in a formal language that is perceptible and manageable by the computer. Choosing a representation for a learner is equivalent to choosing the set of classifiers that it can learn, called the hypothesis space of the learner. Thus, if a classifier is not in this set, it cannot be learned.
- **Evaluation**: an evaluation function is required to score the classifiers and thus differentiate them. The evaluation function that the algorithm uses internally may be distinct from the external function that the classifier will optimize, in order to facilitate it.
- **Optimization**: the key to the learner efficiency is choosing an optimization technique to search for the highest-scoring model among other models, while it helps to determine the classifier produced if the evaluation function has more than one optimum solution.

3.2 Underfitting and Overfitting

When developing a model, it is important to ensure that it can be generalized to different data sets. The overfitting problem is the inability of a model to generalize to different data sets from the one in which it was trained. For example, after continuous adjustments to the parameters, it leads to an accuracy score - ratio of correct predictions from all predictions made - of 100% for the training data and it appears to be a great model. However, after testing the model on test data, it revealed a worse accuracy score than at the beginning. The opposite can also happen when the model can still be further fine tuned to obtain an even better test accuracy score than before. This latter case is called underfitting and occurs when

the model is unable to capture the variability of the data. Figure 3.1 presents three graphs that reflect three different hypothesis. The first one does not fit properly the data which corresponds to a case of underfitting. The second one is a more complex hypothesis and the curve passes through almost all the training samples, but it fits too much the data, therefore represents a case of overfitting. Finally, the last hypothesis seems to properly describe the data, thus being the ideal scenario, since it is more generalized for new data sets, even if it is not the one with the lowest cost when applied to the training set.

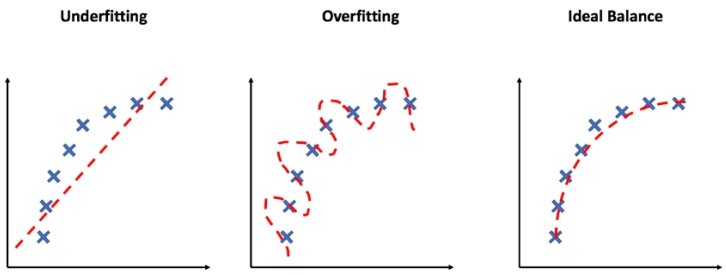


Figure 3.1: Visual demonstration of underfitting, overfitting and ideal balance (adapted from [66]).

The overfitting and underfitting problems can be detected by analyzing the algorithm learning curves, which show the training set and validation set error or correlations evolution as the number of samples of the training set increases, while the number of validation samples is kept constant. With these curves, the evolution of the algorithm is also evaluated. In the presence of an underfitting problem, Figure 3.2(a), the errors in the training and validation sets reach identical values, although higher than the desired one. On the other hand, for overfitting problems, Figure 3.2(b), the training set error converges to lower value than the desired one, while the validation set error converges to a much higher value than the training set, therefore being higher than the desired one.

(a) Underfitting case.

(b) Overfitting case.

Figure 3.2: Learning curves for underfitting and overfitting problems (adapted from [66]).

Usually, the overfitting problem occurs when too many variables are used and thus the hypothesis fits properly the training set, however, it fails to generalize to new data.

3.3 Multiple Linear Regression

In engineering, many applications of regression analysis involve situations where it is necessary to explore the relationship between two or more regressor variables. Thus, a regression model that contains more than one regressor variable is called a multiple regression model.

In first place, a simple linear regression model comprises of only one independent variable or regressor, $\boldsymbol{X}$, and a dependent variable or response, $\boldsymbol{Y}$, whose expected value is a linear function of $\boldsymbol{X}$ and is given by, [68]:

$$\boldsymbol{Y} = \beta_0 + \beta_1 \boldsymbol{X} + \boldsymbol{\epsilon} \tag{3.1}$$

where β_0 is the model intercept, β_1 is the coefficient of the independent variable $\boldsymbol{X}$ and $\boldsymbol{\epsilon}$ is the random error term.

Similarly, in a MLR model, the dependent variable, $\boldsymbol{Y}$, may be related to p independent variables and represented by, [68]:

$$\boldsymbol{Y} = \beta_0 + \sum_{i=1}^{p} \beta_i \boldsymbol{X}_i + \boldsymbol{\epsilon} \tag{3.2}$$

where β_i is the regression coefficient of the variable $\boldsymbol{X}_i$. This parameter represents the expected change in the dependent variable, $\boldsymbol{Y}$, per unit change in $\boldsymbol{X}_i$ when all the remaining independent variables $\boldsymbol{X}_j$ (j $\neq$ i) are held constant.

MLR models are often used as approximating functions, where the true functional relation between $\boldsymbol{Y}$ and $\boldsymbol{X}_i$ is unknown, however, over certain ranges of the independent variables, the linear regression model is an adequate approximation.

Sum of Squared Error

The Sum of Squared Error (SSE) essentially measures the discrepancy between the data and an estimation model. A lower SSE indicates a tight fit of the model to the data. This method allows estimating the model coefficients that minimize the SSE, which can be represented by, [68]:

$$SSE = \sum_{i=1}^{p} (y_i - \hat{y_i})^2 = \sum_{i=1}^{p} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \cdots - \beta_p x_{pi})^2 \ . \tag{3.3}$$

In order to minimize the SSE, the coefficients can by estimated by, [64]:

$$\frac{\partial SSE}{\partial \beta_0} = 0 \wedge \frac{\partial SSE}{\partial \beta_1} = 0 \wedge \cdots \wedge \frac{\partial SSE}{\partial \beta_p} = 0 \ . \tag{3.4}$$

Thus, the estimated coefficients are given by, [64]:

$$\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_p)^T = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{y} \ . \tag{3.5}$$

3.4 Mixed Effects Modeling

A linear regression model with mixed effects provides an alternative approach in order to address the limitations of the common linear regression model with fixed effects, discussed in the previous Section, since it does not verify the cases in which the population is structured by groups, which can lead to incorrect conclusions or wrong estimates. In these cases, observations within the same group may have similar characteristics, indicating that there is no independence between them.

Thus, a mixed-effect model includes both fixed and random effects. The general form of the model, for the i^{th} group, can be given by, [70]:

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i \; ; \; i = 1, ..., M \tag{3.6}$$

where $\boldsymbol{Y_i}$ is the response vector, with dimensions $n_i \times 1$, being n_i the number of observations by group $i = 1, ..., M$; $\boldsymbol{X}_i$ is a $n_i \times p$ design matrix of the p independent variables; $\boldsymbol{\beta}$ is a $p \times 1$ column vector corresponding to the fixed-effect coefficients; $\boldsymbol{Z}_i$ is the design matrix for the q random effects, having size $n_i \times q$; $\boldsymbol{b}_i$ is a $q \times 1$ vector of the random effects. It is considered $\boldsymbol{b}_i \sim N(\boldsymbol{0}, \boldsymbol{G})$, i.e., normal with mean vector $\boldsymbol{0}$ and positive-definite variance-covariance matrix $\boldsymbol{G}$, while the residual component $\boldsymbol{\epsilon}_i$, which is the portion that is not explained by the model, assumes the distribution $\boldsymbol{\epsilon}_i \sim N(\boldsymbol{0}, \boldsymbol{R_i})$, i.e., normal with mean vector $\boldsymbol{0}$ and variance-covariance matrix $\boldsymbol{R_i}$. The components of $\boldsymbol{b}_i$ are independent of each other and of the $\boldsymbol{\epsilon}_i$.

The model in Equation (3.6) has two sources of random variation: $\boldsymbol{b}_i$ and $\boldsymbol{\epsilon}_i$. The former is associated with the variability due to the differences between samples representing the population of interest. These random effects allow to represent a deviation from the fixed-effects general mean $\boldsymbol{\beta}$. Thus, the term $\boldsymbol{\beta}+\boldsymbol{b}_i$ corresponds to the coefficient of the i^{th} group effect, while the variability within the group is represented by $\boldsymbol{\epsilon}_i$.

3.5 Model Assumptions and Diagnostics

During the construction of a regression model, four main assumptions are considered in the model validation. The assumptions verification is performed through the analysis of residual plots, which allow the identification of deviations from the distributional error assumptions. Thus, in order to allow a good performance, the model must verify the following assumptions:

- **Linearity**: the relationship between the independent variables and the dependent variable must be linear, which can be verified using the plot of residuals vs. fitted values. The purpose of the plot is to present a linear trend.

- **Independence**: the residuals must be independent of each other. The presence of autocorrelation

can be verified by the existence of a rectilinear pattern in the residuals plot, such as repeated measures, pairing, cluster sampling or nested factors.

- **Homoscedasticity**: the residuals must have constant variance. To verify this assumption, the residuals vs. fitted values plot is analyzed where the residuals must be randomly distributed around zero without presenting a defined pattern.

- **Normality**: the residuals must follow a normal distribution, which can be verified using the normal Quantile-Quantile (Q-Q) plot, where it is possible to visualize the probability distribution of the observed values vs. expected values, represented by a diagonal. The residuals must approach the diagonal, following that straight line without any significant deviation, in order to confirm the normality assumption.

3.6 Evaluation Metrics

In order to evaluate the hypotheses developed, four metrics are considered: the Root Mean Squared Error (RMSE), the Mean Absolute Percentage Error (MAPE), the Pearson Correlation and the Spearman Correlation.

The RMSE is given by the square root of the average of the square of the differences between the predicted and the original values, being represented by, [66]:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y_i})^2} \tag{3.7}$$

where m measures the number of training samples, y_i and $\hat{y_i}$ are the predicted and observation values of the i^{th} set of parameters, respectively. RMSE gives a bigger importance to the highest errors, presenting greater sensitivity to outliers. This metric has the disadvantage of being an absolute measurement, which makes it dependent on the units of measurement in which the variable is represented.

One of the most commonly used indicators to measure forecast accuracy is the MAPE, which is the sum of individual absolute errors divided by the demand (each period separately), i.e., the average of the percentage errors, and can be given by, [72]:

$$MAPE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{y_i - \hat{y_i}}{\hat{y_i}}\right| 100. \tag{3.8}$$

It measures the normalized error in percentage terms, so its value should be closest to zero. The MAPE error indicator is very useful since, being independent of the measurement scale and easily applicable, it simplifies the interpretation of results. Thus, between RMSE and MAPE it is preferable to use the latter as it is a relative measure.

The Pearson Correlation measures the the strength of a linear association between two variables. This metric is given by, [66]:

$$R = \frac{\sum_{i=1}^{m} \left(\hat{y}_i - \bar{\hat{y}}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{m} \left(\hat{y}_i - \bar{\hat{y}}\right)^2} \sqrt{\sum_{i=1}^{m} \left(y_i - \bar{y}\right)^2}} \tag{3.9}$$

where y_i and $\hat{y}_i$ are the values, corresponding to two distinct quantitative variables, of the i^{th} set of parameters. The $\bar{y_i}$ and $\bar{\hat{y}_i}$ represent their mean values and m measures the number of training samples. The closer to 1 the absolute value of the correlation coefficient is, the greater the association between these two distinct variables. Additionally, the coefficient sign indicates the direction of the association, i.e., if positive, both variables vary in the same direction; if negative, they vary in opposite directions.

The Spearman Correlation is used to measure the degree of association between the original and predicted values. The input parameters of this measurement are first ranked from 1 to N, being N the total of samples of each parameter. After being ranked, Equation (3.9) is applied to those rankings.

On the other hand, in order to summarize the fit quality of a linear regression model, the coefficient of determination, R^2, is analysed, namely for simple linear regression. It reflects the proportion of variation in the forecast variable that is explained by the regression model, thus to quantify the explanatory capacity of the model, ranging from 0 to 1, being expected to be close to 1. However, the fact that $R^2 \simeq 1$ does not mean that the model is adjusted to the data, since adding a variable will always cause the coefficient value to increase regardless of its explanatory capacity.

Therefore, it is preferable to use the coefficient of determination adjusted, R_a^2, for the case of MLR. R_a^2 is an adjusted measure of R^2 that decreases when little explanatory variables are added and is given by, [72]:

$$R_a^2 = 1 - (1 - R^2)\left(\frac{m-1}{m-p-1}\right) \tag{3.10}$$

where p measures the number of independent variables.

4

Power Consumption Models

Contents

As mentioned before, some of the scenarios developed within the Thesis, which will be discussed further, use existing multi-technology energy consumption models, presented in [69] and [73], in order to estimate the RRUs power consumption.

This chapter presents a summary of the procedures and the chosen power consumption models developed for the different RRUs, such as the radios operating on a single technology: 4G, 3G and 2G, presented in this order, and the RRUs shared between 2G and 3G technologies. In addition, a final section regarding error metrics and estimated coefficients of the models is provided, where the parameters related to the RRUs 4G model were discussed in more detail, since for the remaining technologies the analysis is performed in a similar way.

4.1 Introduction

For the energy consumption models development, three different types of data were used: EM data with energy consumption measurements of the BSs equipment, such as RRUs; PM data that reports KPIs; CM data with equipment configurations, such as, hardware model, frequency band and maximum transmit power.

Dataloggers (developed by CELFINET) were used to measure the power consumption. They measure the current flowing through the power cable of each equipment installed in the BS, based on the Hall effect [74]. The power consumption of the RRUs operating on 2G, 3G and 4G technologies was studied over a period of 20 days, being 11 days used for training the model and 9 days for testing it, using 20 distinct BSs located in different cities of Portugal.

Regarding the RRUs 4G power consumption model, 90 units, operating in the frequency bands of 800 MHz, 1800 MHz and 2600 MHz, were considered, where the former has a bandwidth of 10 MHz, while the remaining have a 20 MHz bandwidth; a total of 45 units operating in the 900 MHz and 2100 MHz frequency bands were monitored for the 3G technology, while 15 units were monitored for the 2G technology in the frequency band of 900 MHz, for which the default settings are two transceivers per sector installed inside the cabinet. In addition to the RRUs operating on a single technology, 21 units shared between 2G and 3G technologies in the 900 MHz frequency band (GSM UMTS 900 MHz (GU900)) were also considered. The frequency bands associated with 2G and 3G technologies have a 5 MHz bandwidth.

Both linear fixed effects model and linear mixed effects model are fitted to data from a real network, and thus the final results may be observed subsequently.

4.2 4G RRU Model

In order to obtain more realistic estimates and to adjust the model considering all the equipment diversity, a mixed effects linear regression model was used. The mixed-effect models adapt to the situation in which there is no control over all variables, such as environmental conditions to which the equipment is exposed, energy dissipation losses of the equipment due to the Joule effect, currents induced by the proximity of the cables inside the cabinet, among other factors than can increase the measurement uncertainty.

The independent variables considered to explain the energy consumption in the 4G model are defined by:

- $\boldsymbol{T}_{4G}$: sum of the total traffic volume and signaling bits in 15-minute periods, in [Gb];
- $\boldsymbol{F}_{MicroSleep}$: percentage of activation time of the Micro Sleep feature, which turns the carrier off at symbol time (71.3 μs), in [%];
- $\boldsymbol{F}_{Sleep}$, binary variable (0 = "ON" and 1 = "OFF") that indicates whether a specific carrier is in standby mode ("OFF"), corresponding to the activation of the Sleep Mode energy saving feature.

The proposed 4G model estimates the power consumption (in *Watt*) and is formally described by:

$$\begin{gathered}\boldsymbol{P}_{4G} = \beta_0 + \beta_1 \boldsymbol{T}_{4G} + \beta_2 \boldsymbol{F}_{MicroSleep} + \beta_3 \boldsymbol{F}_{Sleep} + \\ [b_{0i} + b_{0ij} + b_{1i}\boldsymbol{F}_{Sleep} + b_{1ij}\boldsymbol{F}_{Sleep} + b_{2ij}\sin(\boldsymbol{H}_{rad}) + \\ b_{3ij}\cos(\boldsymbol{H}_{rad})] + \epsilon_{ij};\ i = 1, ..., 8; j = 1, ..., 90 \\ b_i \sim N(0, \boldsymbol{G}_1); b_{ij} \sim N(0, \boldsymbol{G}_2); \epsilon_{ij} \sim N(0, \sigma_e^2 \boldsymbol{I}).\end{gathered} \tag{4.1}$$

In this model, β_0 is the intercept coefficient reflecting the baseline power consumption and β_1 is the coefficient associated to the traffic variable, T_{4G} (in Gb). On the other hand, β_2 and β_3 coefficients are associated with energy saving features. The model presents two levels of random variations of *intercepts*: the level 1, given by b_{0i}, representing the equipment configurations, and level 2, with the term b_{0ij}, representing the equipment nested in the configurations, thus $i = 1, ..., 8$ configurations and $j = 1, ..., 90$ equipment. To model the error dependency, *sine* and *cosine* functions were used in each sampled time interval, given by $\boldsymbol{H}_{rad}$, in the random effect term at level 2. In addition, $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ are positive-definite matrices of random effects.

Figure 4.1 presents the evolution of the 4G RRUs power consumption prediction over 9 days of testing.

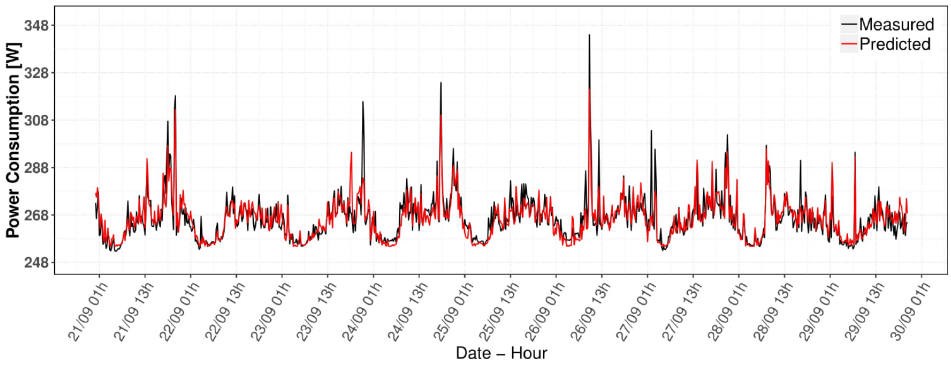


Figure 4.1: 4G power consumption model prediction [73]

4.3 3G RRU Model

Regarding the 3G model, in order to estimate the energy consumption, several independent variables, which are all related to the traffic generated in downlink for a period of 15 minutes, were used:

- $\boldsymbol{T}_{CS_{3G}}$: accumulated voice traffic volume in CS using the sum of each type of service such as 12.2 kbps, 64 kbps and Adaptive Multi-Rate Wideband (AMR-WB) at 16 kbps, in [Mb];
- $\boldsymbol{T}_{PS_{R99}}$: accumulated volume of data traffic in PS with maximum bit rate, in the different services from 16 to 384 kbps, corresponding to the Release 99, in [Mb];
- $\boldsymbol{T}_{HSDPA}$, accumulated traffic volume from the HSDPA transport channel, with a maximum throughput of 14 Mbps, in [Mb].

The 3G model predicts the power consumption (in *Watt*) and is formally described by:

$$\begin{aligned} \boldsymbol{P}_{3G} = \beta_0 + \beta_1 \boldsymbol{T}_{CS_{3G}} + \beta_2 \boldsymbol{T}_{PS_{R99}} + \beta_3 \boldsymbol{T}_{HSDPA} + \\ [b_{0i} + b_{0ij} + b_{1ij} \boldsymbol{T}_{HSDPA} + b_{2ij} \sin(\boldsymbol{H}_{rad}) + \\ b_{3ij} \cos(\boldsymbol{H}_{rad})] + \epsilon_{ij};\ i = 1,..,4; j = 1,..,45 \\ b_i \sim N(0, \boldsymbol{G}_1); b_{ij} \sim N(0, \boldsymbol{G}_2); \epsilon_{ij} \sim N(0, \sigma_e^2 \boldsymbol{I}). \end{aligned} \tag{4.2}$$

The model, given by (4.2), has two levels of random effects variation: level 1 due to different intercepts associated with the configurations and level 2 nested in the level 1, associated with the equipment. The former considered only random intercepts, while in the latter both random intercepts and random slopes were defined, regarding the predictor $\boldsymbol{T}_{HSDPA}$. Both random effect matrices $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ are positive-definite. In addition, similarly to the 4G model, the residuals time periodicity was also detected in the 3G model, thus the transformation with *sine* and *cosine* functions was used in the random effect term at level 2.

Figure 4.2 shows an example of the proposed model application for the 9-day test period.

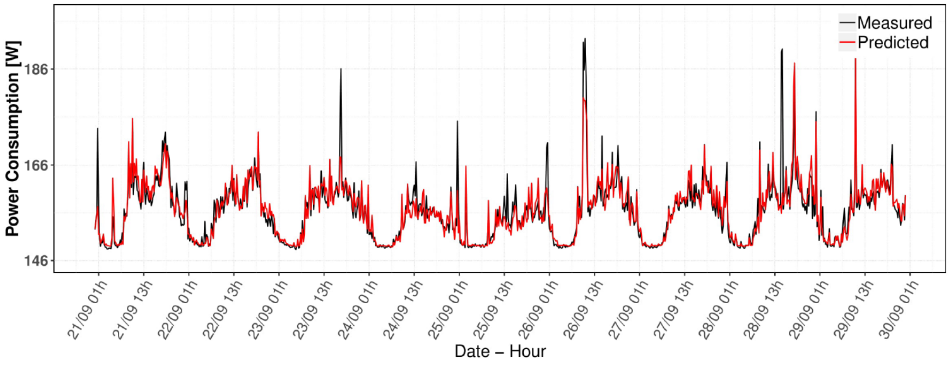


Figure 4.2: 3G power consumption model prediction [73].

4.4 2G RRU Model

In the 2G model, the KPIs related to both voice and data traffic, in the period of 1 hour, are the following:

- $\boldsymbol{T}_{CS_{2G}}$: traffic volume of TCHs allocated to the full-rate and half-rate channels at the cell level, in [Erl];
- $\boldsymbol{T}_{PS_{2G}}$: total EDGE and GPRS data volume carried on downlink, in [Mb].

Similarly to the other technologies, the 2G model estimates the power consumption (in *Watt*) and is formally described by:

$$\begin{gathered}\boldsymbol{P}_{2G} = \beta_0 + \beta_1 \boldsymbol{T}_{CS_{2G}} + \beta_2 \boldsymbol{T}_{PS_{2G}} + [b_{0i} + \\ b_{1i} \boldsymbol{T}_{CS_{2G}}] + \epsilon_i;\ i = 1, .., 15; b_i \sim N(0, \boldsymbol{G}); \epsilon_i \sim N(0, \sigma_e^2 \boldsymbol{I}).\end{gathered} \tag{4.3}$$

Random intercept and slope between equipment were used in the model. The random-effect matrix $\boldsymbol{G}$ is diagonal positive-definite.

Figure 4.3 shows a result of the model application, for the same test period, where it is possible to observe low variability of the power consumption, and yet the prediction approaches the power measurements.

4.5 2G/3G RRU Model

Finally, the power consumption model for RRUs shared between 2G and 3G technologies is given by:

$$\begin{gathered}\boldsymbol{P}_{2G,3G} = \beta_0 + \beta_1 \boldsymbol{T}_{CS_{2G}} + \beta_2 \boldsymbol{T}_{PS_{2G}} + \beta_3 \boldsymbol{T}_{CS_{3G}} + \\ \beta_4 \boldsymbol{T}_{PS_{R99}} + \beta_5 \boldsymbol{T}_{HSDPA} + [b_{0i} + b_{1i} \boldsymbol{T}_{HSDPA} + \\ b_{2i} \sin(\boldsymbol{H}_{rad}) + b_{3i} \cos(\boldsymbol{H}_{rad})] + \epsilon_i \\ i = 1, .., 21; b_i \sim N(0, \boldsymbol{G}); \epsilon_i \sim N(0, \sigma_e^2 \, |\boldsymbol{v_i}|^{2\delta}).\end{gathered} \tag{4.4}$$

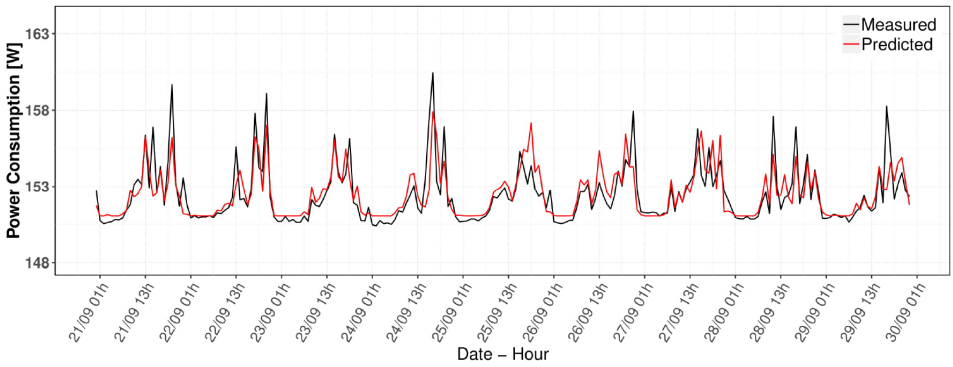


Figure 4.3: 2G power consumption model prediction [73].

The proposed model uses the same independent variables of the previously presented 2G and 3G models, while considering random intercepts and slopes for the predictor $\boldsymbol{T}_{HSDPA}$. The random-effect matrix $\boldsymbol{G}$ is diagonal positive-definite. A non-constant variance was observed given that the residuals increase with the growth of the estimated value. Thus, a power function to model the error variance is presented.

Figure 4.4 presents the evolution of the power consumption prediction of the proposed model, in a 9-day window, which approximates the measured values.

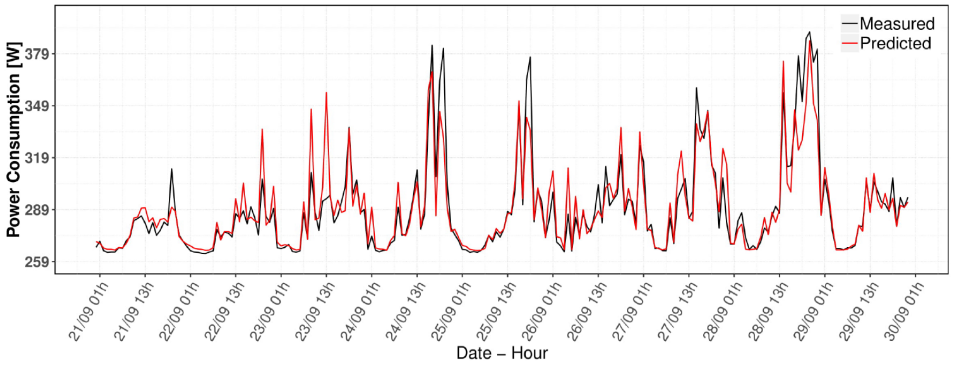


Figure 4.4: 2G/3G power consumption model prediction [73].

4.6 Model Estimates and Metrics

The 4G RRUs power consumption model considers that each equipment has a different baseline power consumption, while equipment of the same configuration has a similar baseline consumption, which allows the random effect to vary by two levels improving the model quality. Random intercepts and a slope coefficient are added and associated with the feature sleep mode. In addiction, *sine* and *cosine* functions were used as a random effect to remove the residuals auto-correlation, thus being possible to

fulfill the assumption of residuals independence. Moreover, a constant variance and zero mean of the residuals are verified at each level of random effects.

The estimates of the model's fixed effect are presented in Table 4.1, for all the different types of RRUs. Among the radios that operate in dedicated mode, 4G RRUs have a higher baseline power consumption of 194.56 W that grows linearly at the rate of 6.11 W/Gb, decreasing with the Micro Sleep feature activation, which, if activated 100% of the time, recorded 25.63 W savings. Additionally, if it is unnecessary to have the equipment turned on, when activating the Sleep Mode feature that turns the equipment off, power consumption is reduced by 96.8 W.

Table 4.1: Power consumption model error metrics and estimated coefficients [73].

Model	R_a^2	MAPE [%]	RMSE [W]	Estimated Coefficients
4G	0.99	1.65	6.14	$\boldsymbol{P}_{4G} \sim 194.564 + 6.110\boldsymbol{T}_{4G} - $ $-25.630\boldsymbol{F}_{MicroSleep} - 96.800\boldsymbol{F}_{Sleep}$
3G	0.99	0.92	3.55	$\boldsymbol{P}_{3G} \sim 168.027 + 0.084\boldsymbol{T}_{CS_{3G}} + $ $+1.177\boldsymbol{T}_{PS_{R99}} + 0.011\boldsymbol{T}_{HSDPA}$
2G	0.88	0.38	1.04	$\boldsymbol{P}_{2G} \sim 152.080 + 1.707\boldsymbol{T}_{CS_{2G}} + $ $+0.031\boldsymbol{T}_{PS_{2G}}$
2G/3G	0.80	1.49	6.84	$\boldsymbol{P}_{2G,3G} \sim 261.123 + 5.063\boldsymbol{T}_{CS_{2G}} + 0.055\boldsymbol{T}_{PS_{2G}} + $ $+0.080\boldsymbol{T}_{CS_{3G}} + 1.114\boldsymbol{T}_{PS_{R99}} + 0.018\boldsymbol{T}_{HSDPA}$

It can be observed in Table 4.1 that both the 2G model and shared radios model present lower R_a^2 values, which may be due to the low variations in the energy consumption of 2G technology equipment and the difficulty of separating the influence of each technology in the shared radios model, since they operate on different amplifiers, despite being configured in the same radio equipment.

5

Capacity Models

Contents

This chapter provides a description of the several steps of the 4G and 5G models development, in which the methods and algorithms used are presented along with the respective results, taking into account distinct examples for each model. Two separate models were developed for 4G technology with different ambitions as well as a model for 5G technology with similar purposes to the second 4G model.

5.1 4G Cell Capacity Model

5.1.1 Introduction

The aim of 4G cell capacity model is to develop cell monitoring methods to detect high load problems and to predict the maximum cell capacity, under different constraints for each traffic environment. The purpose of this model is to increase the QoS and QoE provided by the radio network. The methodology associated with these methods was based on an existing 4G capacity model, developed for a distinct vendor and described in detail in [72]. In addition, a 4G PRB usage prediction model was developed within the scope of the Thesis in order to support or refute the cenarios of downlink traffic migration presented later in Sections 6.1 and 6.2.

Initially, a cell resource monitoring is executed with the purpose of detecting cells with capacity problems by imposing 3 conditions on the data. In the event that cells are detected, an analysis is performed at the level of network performance data in order to indicate improvements. In addition, the maximum cell capacity is estimated for the detected cells, based on throughput prediction, to verify whether the cells have capacity margins for the radio conditions to which they are subjected.

The procedure requires filtering, cleaning and variable selection. After analyzing the network data, different network scenarios were detected. Thus, the study must be carried out separately per cell.

The model was developed using real network performance and configuration data collected from a live LTE network, belonging to a Portuguese mobile operator, over a one-month period with a quarter-hourly granularity. Data were extracted from 114 cells of 22 sites, operating in the frequency bands of 800 MHz, 1800 MHz and 2600 MHz, with system bandwidth of 10 MHz for the first frequency band, and of 20 MHz for the remaining ones.

5.1.2 Cell Resource Monitoring

In this first module, information about how to locate the resource bottleneck and the related handling suggestions are provided. Primarily, the purpose is to detect a possible shortage of resources in each cell that could compromise the level of QoS of the UE, thus the downlink user perception is considered.

The data traffic rise leads to an increase in the PRBs utilization, which reflects the degree of bandwidth usage, over the air interface. On the other hand, the downlink user-perceived rate decreases as the number of users sharing the limited PRBs increases, reflecting the user QoE. Thus, monitoring

these two parameters together may demonstrate the user experience under a certain bandwidth usage over the radio interface.

The used method is based in a strategy, described in [75], where several KPIs are analyzed during the busy hour of each cell to check if given conditions are fulfilled, defined by the vendor, although the values of each threshold are here adapted to the vendor considered in the present study, based on [76], although the current network operator is able to adjust the limits here defined.

Therefore, if the downlink PRB utilization rate reaches or exceeds 80% while the downlink user-perceived rate is below 5 Mbps for three days in a week, then the cell goes through an analysis regarding the following list of thresholds:

- Average CQI lower than 10. CQI is used by the UE to indicate the channel quality to the eNB. The reported CQI value ranges from 0 to 15, indicating the type of modulation and coding that UE can operate.
- Downlink UE latency higher than 9 ms, measuring the impact on the end user. Downlink latency is a parameter that indirectly influences the perceived system data rate and it will increase with more instantaneously active UEs.
- Average 64-QAM scheme usage lower than 10%. This metric represents the percentage of 64-QAM samples, which indicates downlink Signal to Interference plus Noise Ratio (SINR) status along with CQI and Rank Indicator Rank Indicator (RI).
- Average Radio Link Control (RLC) retransmission ratio higher than 1%. The RLC retransmission ratio can be given by the proportion of unsuccessful RLC Protocol Data Unit (PDU) and RLC PDU segment transmissions.

If all these listed conditions are fulfilled, then RF optimization should be performed. Otherwise, it is advised to add carriers or expand the bandwidth of the existing carrier.

In order to increase the accuracy associated with this monitoring method, only samples related to the busy hours are analyzed. Thus, as presented in Table 5.1, 2 cells with capacity problems were detected, i.e., with a bottleneck on the downlink user perception, among the 114 existing ones, characterized by low user throughput values and high utilization rate of cell resources. Thus, it is verified in the 2 detected cells that, for at least 3 days in a week, the PRB utilization rate is above 80% while the user throughput does not exceed 5 Mbps.

Both cells have average CQI values below 10, therefore it is advised to perform a RF optimization (e.g. reconfiguration of antennas tilt and/or azimuth, neighboring cells, etc), thus minimizing the channel interference and, consequently, increase the data rate. It should be noted that cell A presents worse radio conditions than cell B, which may be reflected in the study of cell capacity prediction.

Table 5.1: Cells detected by the resource monitoring module.

Cell	PRB Util Rate [%]	User Thp [Mbps]	Avg CQI	DL UE Latency [ms]	Avg 64-QAM [%]	Avg RLC Retx. [%]
A	99.4	3.47	7.15	4.22	11.4	0.05
B	88.8	4.25	8.01	3.94	15.5	0.06

5.1.3 Downlink Cell Throughput Prediction

The capacity model is based on real data, using MLR algorithms to predict the downlink cell throughput. The MLR model is used to explain a response variable (dependent variable) as a linear function of several input variables (independent variables). The dataset was split in training set (around 70%) for fitting the model, and test set (30%), used to provide an unbiased evaluation of a final model fit on the training dataset. In the present model, the dependent variable is the downlink cell throughput, Thp_{cell}, in Mbps, and is given by:

$$Thp_{cell} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i \tag{5.1}$$

where β_0 is the model intercept, β_i is the coefficient of the variable x_i and n is the number of independent variables.

In the first place, it is advisable to have a wide range of variables, since it may not be obvious which variables are more important. This initial set of variables goes through a variable selection iterative process, where the most relevant describe the cell throughput. This initial set of variables goes through a variable selection iterative process, using stepwise regression, where the most relevant describe the cell throughput. In each step, a variable is considered for addition to or subtraction from the set of explanatory variables based on some pre-specified criterion. In addition, variables that are highly correlated (with Pearson Correlation above 80%) with other variables that prove to be more relevant to the model are eliminated. In order to determine the model accuracy and evaluate which is the model that best fits the real data, several error metrics are used, such as the adjusted determination coefficient R_a^2, Pearson Correlation, MAPE and RMSE.

With zero intercept, the most relevant variables that were chosen to predict the DL cell throughput are the following [77]:

- $\boldsymbol{PRB}_{usage}$: PRB utilization rate, in [%];
- $\boldsymbol{S}_{64-QAM}$: proportion of 64-QAM samples, in [%];
- $\boldsymbol{S}_{16-QAM}$: proportion of 16-QAM samples, in [%];

- $\boldsymbol{U}_{16-QAM}$: percentage of unsuccessful Hybrid Automatic Repeat Request (HARQ) transmissions rate using 16-QAM, in [%];
- $\boldsymbol{Lat}_{agg}$: aggregated downlink latency for a measurement period, in [s];
- $\boldsymbol{RLC}_{tx}$: percentage of successful RLC PDU transmissions, in [%];
- $\boldsymbol{RI}_{rank1}$: proportion of both open and closed loop Spatial Multiplexing (SM) rank 1, regarding the MIMO rank distribution usage, in [%].

As in Table 5.1, cell A and cell B are the detected cells with capacity problems and, therefore, the chosen model was built for them.

Since the purpose of creating the downlink cell throughput prediction model is to estimate the maximum cell capacity, the samples belonging to the highest traffic periods are the most relevant for the model. Thus, in order to increase the model's accuracy, and considering that the number of samples is enough, only samples with PRB utilization rate values above the 70th percentile value were used.

Cell A

The MLR model results are represented in Figure 5.1, where it is possible to visualize both the real cell throughput, given by measurements, and the predicted one, given by the model. In order to improve the visualization of the figure, only 400 samples are displayed, which will be repeated for the results of the predictions subsequently provided. In this way it is possible to notice the evident approximation of the adjusted values with the measured ones.

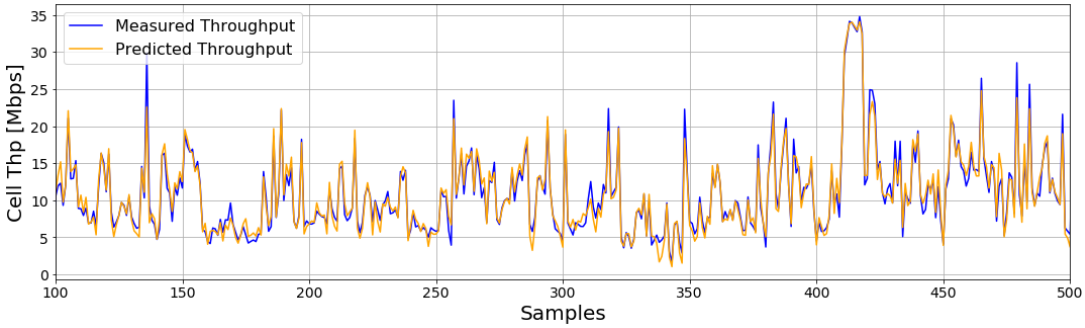


Figure 5.1: Downlink cell throughput model prediction for cell A.

Additionally, diagnostic plots are presented in Figure 5.2, for statistical validation of the present model.

Analyzing the residuals vs fitted values plot, it is observed that the residuals follow a linear pattern, which means that the model has a linear relationship between the independent variables and the de-

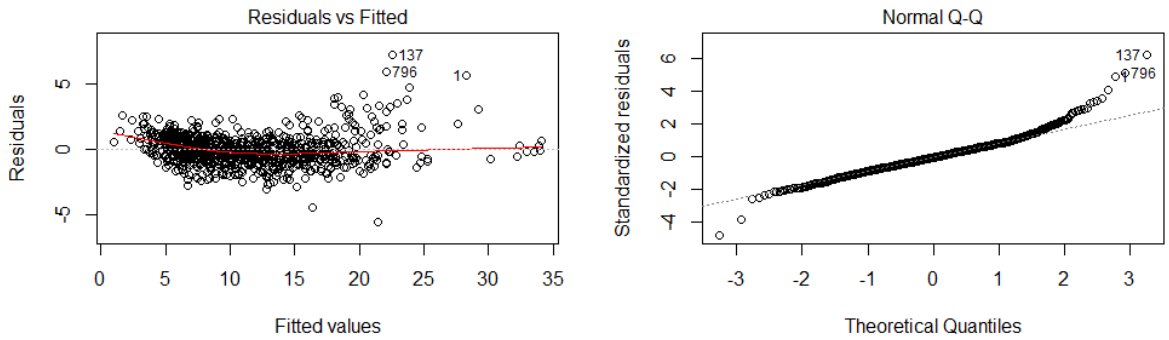


Figure 5.2: MLR model diagnostic plots for cell A.

pendent variable. In addition, it is possible to detect the independence of the residuals. The normal Q-Q plot shows that the residuals follow the straight dashed line, thus verifying the assumption of normality.

Cell B

For the given cell, Figure 5.3 presents the proposed MLR model results, where the measured cell throughput along with the adjusted one are shown, in which the latter describes the former even more efficiently than in cell A.

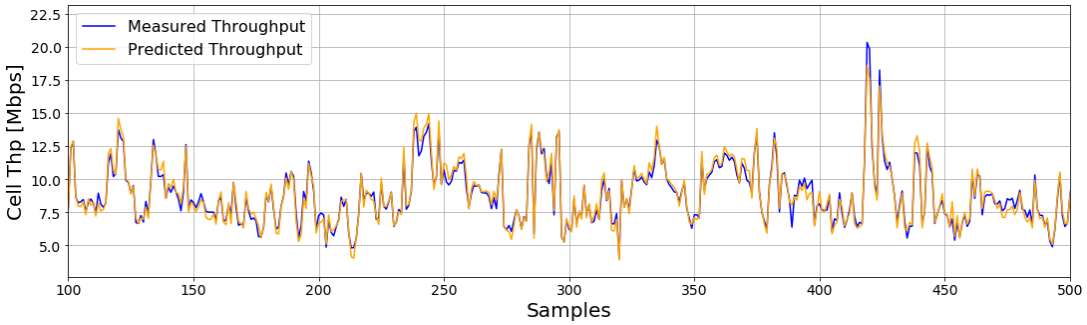


Figure 5.3: Downlink cell throughput model prediction for cell B.

Figure 5.4 presents the residuals vs fitted values and normal Q-Q plots to evaluate the model.

In addition to the independence of the residuals that can once again be observed, the assumption of linearity is more pronounced, compared to the first cell, while homoscedasticity is closer to being verified since, for cell B, the residuals are randomly distributed around the line almost uniformly. The normality of the residuals is demonstrated in the normal Q-Q plot, in Figure 5.4.

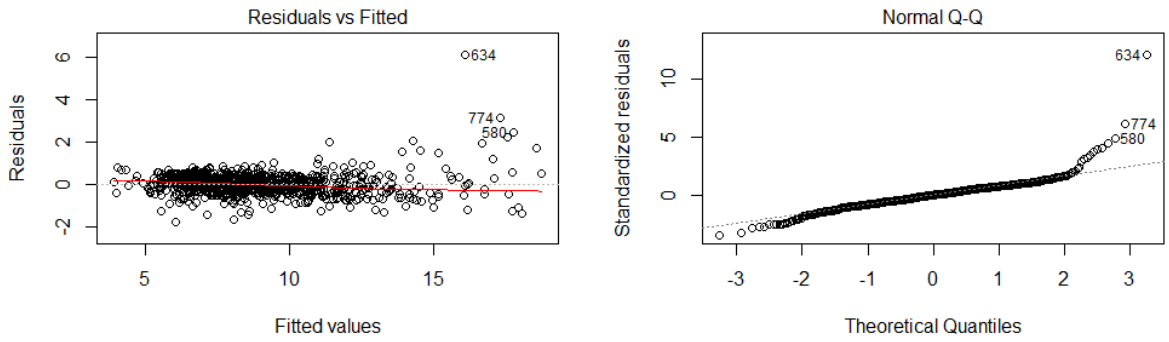


Figure 5.4: MLR model diagnostic plots for cell B.

Error Metrics

Table 5.3 presents the error metrics and the coefficients estimated by the cell throughput prediction model. The error metrics displayed corroborate the high reliability of the model. Cell A has slightly worse results since its downlink cell throughput presents greater variability, which causes some complications in its prediction.

Table 5.2: Cell throughput model error metrics and estimated coefficients.

Cell	R_a^2	Corr [%]	MAPE [%]	RMSE [Mbps]	Estimated Coefficients
A	0.991	97.79	8.91	1.17	$\boldsymbol{Thp}_{cell} \sim 0.016\boldsymbol{PRB}_{usage} + 0.280\boldsymbol{S}_{64-QAM} + 0.071\boldsymbol{S}_{16-QAM} + +0.129\boldsymbol{U}_{16-QAM} - 0.037\boldsymbol{Lat}_{agg} + 0.079\boldsymbol{RLC}_{tx} - 0.082\boldsymbol{RI}_{rank1}$
B	0.997	98.0	4.01	0.51	$\boldsymbol{Thp}_{cell} \sim 0.019\boldsymbol{PRB}_{usage} + 0.199\boldsymbol{S}_{64-QAM} + 0.083\boldsymbol{S}_{16-QAM} + +0.091\boldsymbol{U}_{16-QAM} - 0.019\boldsymbol{Lat}_{agg} + 0.078\boldsymbol{RLC}_{tx} - 0.073\boldsymbol{RI}_{rank1}$

5.1.4 Cell Capacity Estimation

After the detection of cells with capacity saturation and the model development, it is possible to estimate the maximum cell capacity for the cell radio propagation environment using the proposed MLR model to calculate the maximum cell throughput. The independent variable that measures the utilization of cell resources, i.e, the PRB utilization rate, which represents the ratio between the total number of used PRBs and the amount of available PRBs, can act as the cell capacity bottleneck. Forcing the use of all cell resources by setting the PRB utilization rate to 100%, while the remaining variables reflect similar radio conditions, the maximum cell throughput is estimated. Thus, using (5.1), where Thp_{cell} is the downlink cell throughput and x_1 is the PRB utilization rate, the Thp_{cell} is maximum when $x_1 = 100\%$.

The remaining independent variables remain unchanged, reflecting the same radio conditions around the cell.

Once the maximum cell throughput is calculated for all samples, there is a variation of this value over time depending on the radio conditions and interference of each sample. Thus, it is necessary to define a threshold to characterize the maximum capacity of the cell. In order to depict the busy hour, the analysis of the cell capacity will be performed using the values associated with the 95th percentile. The maximum cell throughput is then calculated for the two previous cells (A and B).

Cell A

Figure 5.5 shows, for cell A, the measured cell throughput along with the estimated one assuming that all available PRBs are being used. As mentioned, the cell capacity and measured cell throughput during the busy hour are characterized by their 95th percentile values also displayed in Figure 5.5.

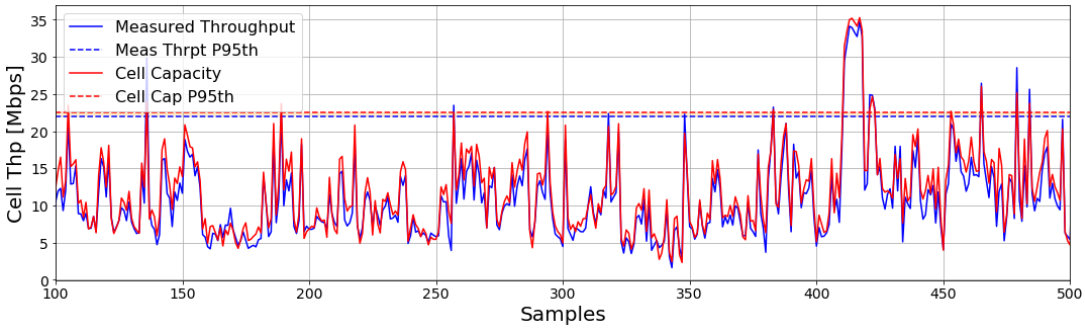


Figure 5.5: 4G cell capacity estimation for cell A.

The calculated 95th percentile values of measured cell throughput and cell capacity are 21.95 Mbps and 22.52 Mbps, respectively, resulting in a cell load of 97.46%, which is consistent with the high PRB utilization rate value indicated in Table 5.1 for cell A. Thus, at the busy hour, there is some capacity margin, although very close to its limit, demonstrating that the cell is approaching an overload situation, as predicted. In the present situation, it is not possible to guarantee an adequate QoE to its users in case of a slight increase in LTE data traffic of the current cell.

Cell B

For cell B, both the measured downlink cell throughput and the predicted maximum one are presented in Figure 5.6 along with the respective 95th percentile values representing the busy hour.

After the maximum cell throughput estimation, the 95th percentile value associated with the real throughput is 13.62 Mbps, while the one related with the cell capacity is 14.47 Mbps, generating a cell

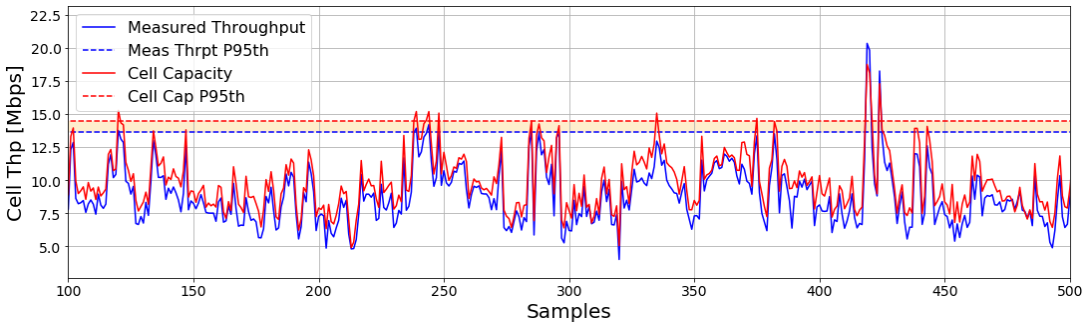


Figure 5.6: 4G cell capacity estimation for cell B.

load of 94.11%. As expected, this cell has a higher capacity margin than cell A, since, according to Table 5.1, the former resources are not being used as abundantly as in the latter, presenting better radio conditions than cell A, as mentioned above. Nevertheless, with its high cell load value, cell B is near its capacity limit as well.

5.1.5 4G Downlink PRB Usage Prediction

In the present study, two traffic volume migration scenarios will be developed for the same vendor but for different sites. Therefore, there was a need to create a scenario validation method in order to find out whether the transfer of traffic volume in question is valid. Predicting the PRB utilization rate, after the traffic volume migration, allows to evaluate the use of physical cell resources, which can be used to validate or not the traffic migration scenarios. Thus, a model that predicts the PRB usage in 4G cells is created similarly to the cell throughput prediction model, described in Section 5.1.3.

In the present MLR model, the dependent variable is the downlink PRB utilization rate, PRB_{usage}, in %, and is given by:

$$PRB_{usage} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i. \tag{5.2}$$

The traffic migration scenarios developed for this vendor involve four 4G cells: cell C in the UMTS 2100 MHz (U2100) migration scenario and cells D1, D2 and D3 in the 3G migration scenario. So that the study does not exclude any cells, the model was developed for all the cells just mentioned, using the same initial set of variables that was used in the development of the downlink cell throughput prediction model, with the aid of residual plots and error metrics analysis.

The independent variables that were detected as relevant in the prediction of PRB usage rate for the four cells, with zero intercept and discarding less important variables with high Pearson Correlation (above 80%), are as follows:

- $\boldsymbol{T}_{4G}$: sum of the total data traffic volume and signaling bits, in [Gb];
- $\boldsymbol{S}_{64-QAM}$: proportion of 64-QAM samples, in [%];

- S_{16-QAM}: proportion of 16-QAM samples, in [%];
- U_{64-QAM}: percentage of unsuccessful HARQ transmissions rate using 64-QAM, in [%].

In the migration scenarios, data traffic will be transferred to these 4G cells, which will be reflected in the data traffic volume independent variable. Thus, the purpose of the current model is to subsequently estimate the PRB usage rate after migration by updating the data traffic volume variable with the new values (the old ones as well as the received volume of traffic also added), assuming the same channel quality.

As expected, the most relevant samples for the current model are those referring to periods of high use of cell resources that are equivalent to periods of high cell throughput. Consequently, analogously to the first 4G capacity model, only samples with downlink cell throughput values above the 70th percentile value are used.

In order to illustrate the model, the results will be presented in detail for one cell on each site. Therefore, for the first scenario, the only hypothesis is cell C, while for the second scenario cell D2 was chosen, as it presents a more reliable estimate.

Cell C

Figure 5.7 presents the proposed MLR model results, where the real PRB usage, given by measurements, along with the predicted one, provided by the model, are displayed, for cell C. An acceptable adjustment is observed, as the latter approximates the former.

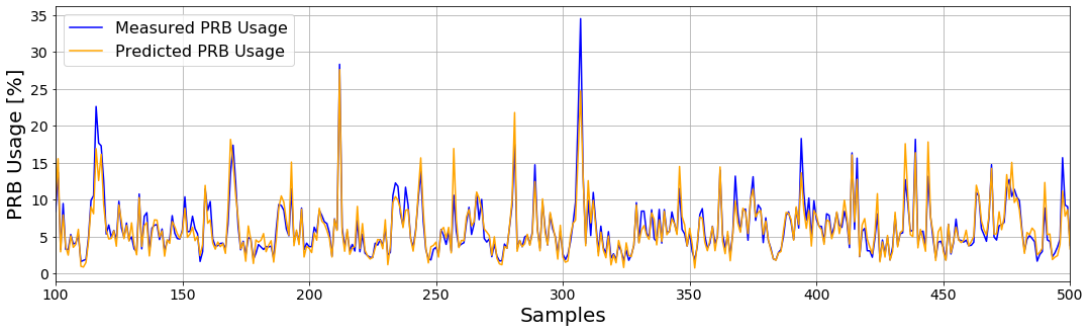


Figure 5.7: 4G PRB usage model prediction for cell C.

A residual plot analysis is performed in Figure 5.8, for statistical evaluation of the current model.

Focusing on the plot of residuals vs fitted values, it is possible to verify a linear trend described by the red line which confirms the existence of linearity between the independent variables and the dependent variable. The absence of rectilinear patterns confirms the independence of the residuals. In addition, the residuals follow a normal distribution, which can be verified by the normal Q-Q plot, in which the

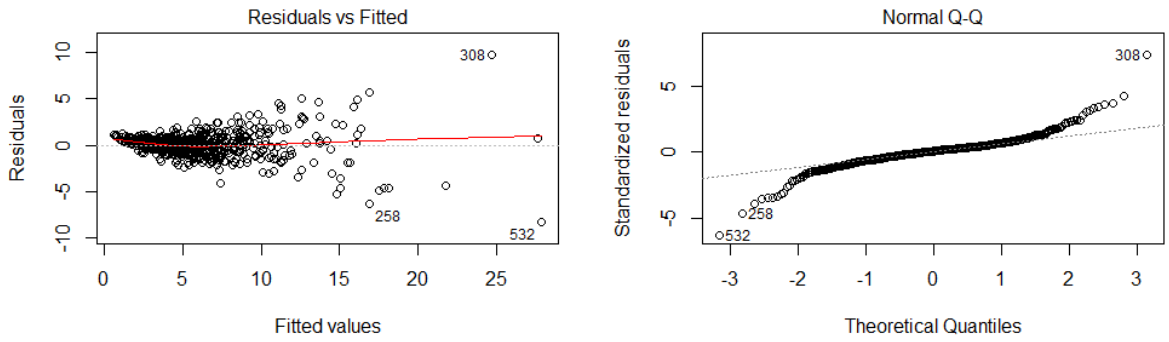


Figure 5.8: MLR model diagnostic plots for cell C.

existence of outliers is also present, although their treatment does not significantly alter the model due to their reduced quantity.

Cell D2

For the given cell, the MLR model results are represented in Figure 5.9, being possible to visualize both the measured PRB utilization rate and the adjusted one, where the latter describes more effectively the former, comparing with the first cell.

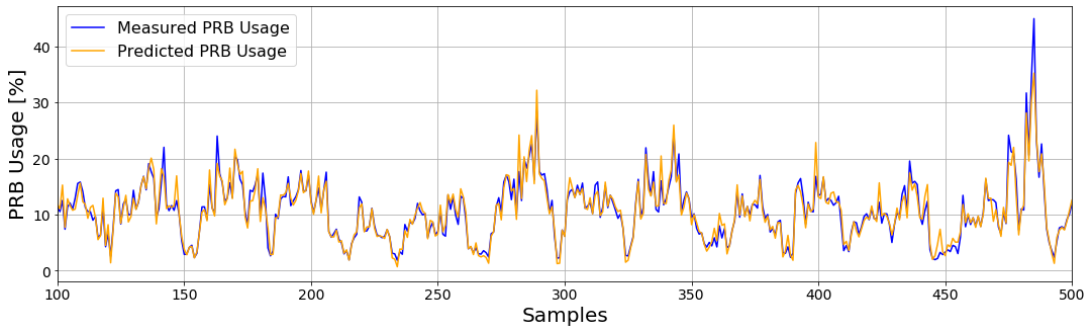


Figure 5.9: 4G PRB usage model prediction for cell D2.

Figure 5.10 displays the residuals vs fitted values and normal Q-Q plots to validate the proposed model.

In the plots generated for the present cell, the linearity is present once again, while the independence of the residuals becomes more evident. As in the first cell, the normality of the residuals is verified in the normal Q-Q plot, in which the residuals follow the diagonal dashed line, this time with less deviations.

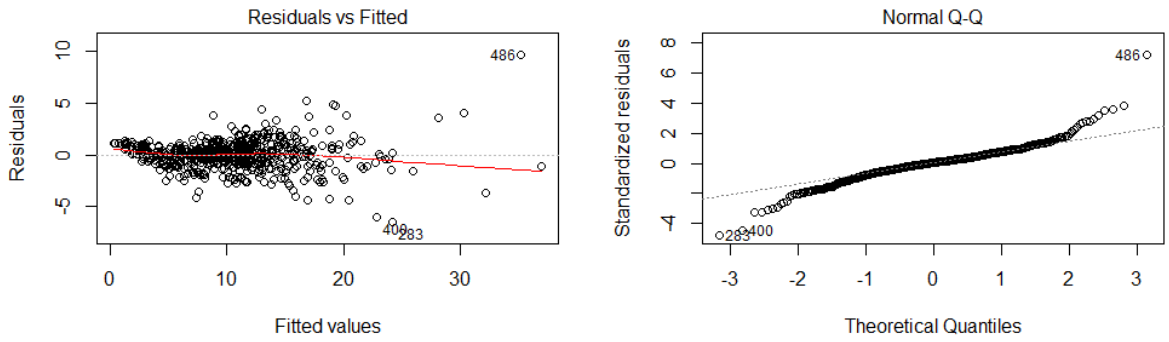


Figure 5.10: MLR model diagnostic plots for cell D2.

Error Metrics

The error metrics associated with the PRB usage prediction model as well as the estimated coefficients, are shown in Table 5.3, confirming its high reliability. It is noticeable that the model presents better metrics for cell D2, which can be justified due to the periodic behavior of its PRB utilization rate, compared to the other cells that present more irregular behaviors and more peak usage times. Cell D3 stands out in the irregularity of its PRB usage, making its prediction more difficult and, therefore, it presents worse results.

Table 5.3: 4G PRB usage model error metrics and estimated coefficients.

Cell	R_a^2	Corr [%]	MAPE [%]	RMSE [%]	Estimated Coefficients
C	0.964	93.32	15.16	1.36	$\boldsymbol{PRB}_{usage} \sim 4.987\boldsymbol{T}_{4G} - 0.019\boldsymbol{S}_{64-QAM} + +0.032\boldsymbol{S}_{16-QAM} + 0.188\boldsymbol{U}_{64-QAM}$
D1	0.942	91.44	19.12	1.67	$\boldsymbol{PRB}_{usage} \sim 3.400\boldsymbol{T}_{4G} + 0.006\boldsymbol{S}_{64-QAM} + +0.072\boldsymbol{S}_{16-QAM} + 0.030\boldsymbol{U}_{64-QAM}$
D2	0.986	96.56	11.05	1.37	$\boldsymbol{PRB}_{usage} \sim 1.911\boldsymbol{T}_{4G} - 0.033\boldsymbol{S}_{64-QAM} + +0.109\boldsymbol{S}_{16-QAM} + 0.097\boldsymbol{U}_{64-QAM}$
D3	0.937	88.86	27.89	0.99	$\boldsymbol{PRB}_{usage} \sim 1.591\boldsymbol{T}_{4G} + 0.014\boldsymbol{S}_{64-QAM} + +0.056\boldsymbol{S}_{16-QAM} + 0.026\boldsymbol{U}_{64-QAM}$

5.2 5G Cell Capacity Model

5.2.1 Introduction

In parallel with the development of MLR models for 4G technology, an MLR model for 5G technology was also created, but for another vendor, taking into account the data available for the present study. The

purpose of this model is to predict the PRB usage in order to ascertain the consequences of transferring downlink data traffic from 4G cells to the respective 5G cell, which is described later in Section 6.3.

The development process of the 5G cell capacity model is similar to the one described for the 4G capacity model, regarding the mentioned methods of downlink PRB usage and downlink cell throughput prediction. For this vendor and technology, KPIs differ slightly from the previous models, although the same analysis is performed.

As mentioned, this model uses a distinct data set, belonging to another vendor, from a real 5G network, collected over a period of two and a half months on an hourly and cell basis. The analyzed area considers 7 cells of 3 sites, operating in the 3.5 GHz frequency band with 80 MHz of system bandwidth.

5.2.2 5G Downlink PRB Usage Prediction

The prediction of the PRB usage rate allows evaluating the use of cell resources, in order to be able to analyze its behavior when receiving 4G data traffic volume.

In order to estimate the downlink PRB utilization rate, in %, the dependent variable is given by:

$$PRB_{usage} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i. \tag{5.3}$$

Although the majority of the 5G traffic volume is residual, since the 5G network is not yet open to public, the cells with the most data traffic volume were chosen (cells E and F), belonging to the same site but different sectors. Thus, the model is developed for these specific cells in order to increase the model's reliability and, subsequently, traffic migration scenarios are performed on them.

With zero intercept, the independent variables that were detected as the most relevant in predicting the PRB usage rate, rejecting variables with less importance and high correlation with more important variables (Pearson Correlation above 80%), are the following:

- $\boldsymbol{T}_{5G}$: total downlink traffic volume in a cell, in [Gb];
- $\boldsymbol{UE}_{NR}$: average number of LTE-5G NR NSA Dual Connectivity (DC) UEs using the current cell as the Primary Secondary Cell (PSCell);
- $\boldsymbol{CCE}_{usage}$: Physical Downlink Control Channel (PDCCH) Control Channel Element (CCE) usage rate, in [%];
- $\boldsymbol{CQI}_{avg}$: average CQI values ranging from 0 to 15.

The purpose of the model is to estimate the cell PRB utilization rate after it receives the LTE traffic volume of the respective cells and, thus, to anticipate the impact of this data transfer on the 5G cell.

Below are presented the results of the proposed MLR model with diagnostic plots and a later analysis of the error metrics.

Cell E

Figure 5.11 presents the model results, for cell E, in which is possible to visualize both the measured PRB usage rate, given by measurements, and the adjusted one, provided by the model. An acceptable adjustment is observed, as the latter approximates the former.

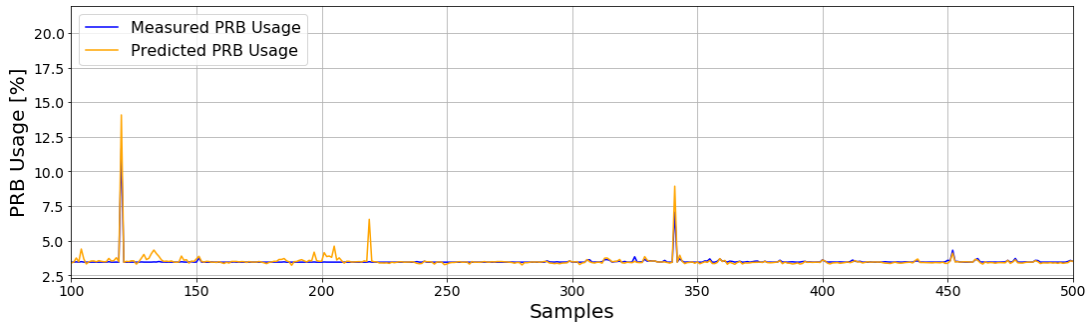


Figure 5.11: 5G PRB usage model prediction for cell E.

Additionally, diagnostic plots are displayed in Figure 5.12, for statistical evaluation of the model.

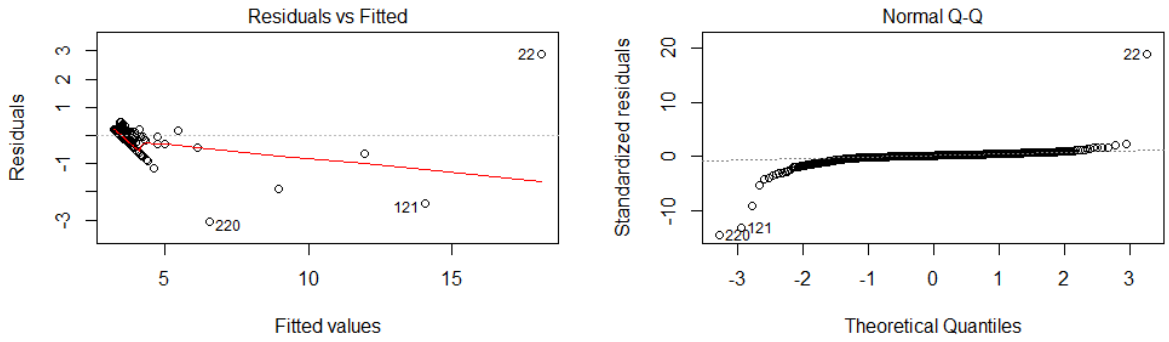


Figure 5.12: MLR model diagnostic plots for cell E.

Analyzing the residuals vs fitted values plot, the red line does not always depict a linear trend, since the existing traffic in 5G cells is mainly residual, presenting rare traffic peak times. For the same reason, the residuals are not uniformly distributed along the line, thus verifying the dependence of the residuals. Despite that, the normal Q-Q plot ensures the normality of the residuals, since they follow the straight dashed line with the presence of some outliers, as expected.

Cell F

For the given cell, the MLR model results are presented in E Figure 5.13, in which the real PRB usage along with the predicted one are displayed. It can be stated that the latter effectively describes the former.

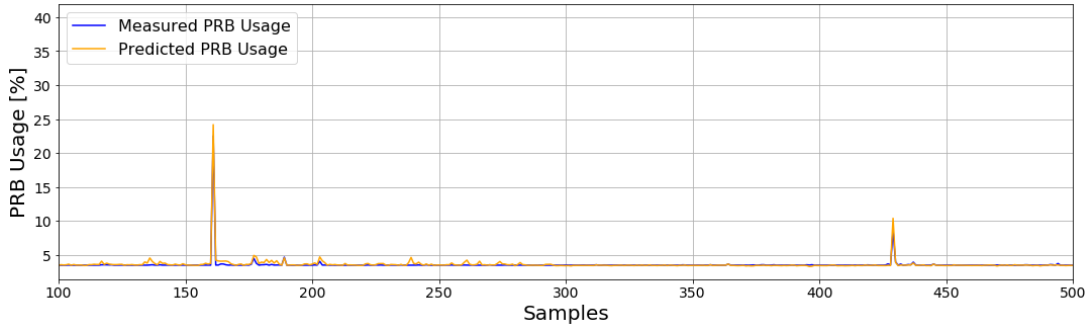


Figure 5.13: 5G PRB usage model prediction for cell F.

Figure 5.14 presents the residuals vs fitted values and normal Q-Q plots to validate the proposed model.

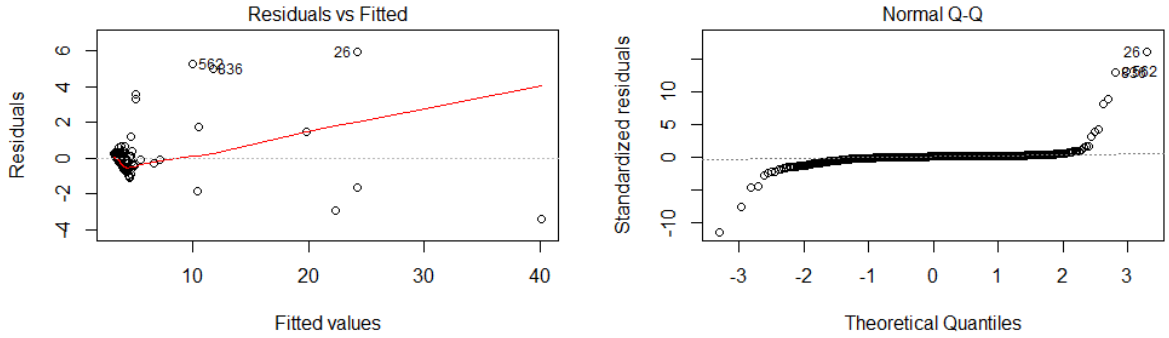


Figure 5.14: MLR model diagnostic plots for cell F.

The traffic volume behavior in this cell is very similar to the one present in cell E and therefore the assumptions of linearity and independence are not yet verified. However, the residuals follow a normal distribution, as can be seen in normal Q-Q plot, in which the residuals follow the diagonal dotted line, except in the presence of the expected outliers.

Error Metrics

Table 5.4 describes the error metrics values as well as the coefficients estimated by the model. Despite the presence of outliers and the unsatisfactory results in the residual plot analysis, the values of the error

metrics are quite positive, presenting the necessary reliability in order to be able to produce sufficiently realistic scenarios thereafter.

Table 5.4: 5G PRB usage model error metrics and estimated coefficients.

Cell	R_a^2	Corr [%]	MAPE [%]	RMSE [%]	Estimated Coefficients
E	0.996	95.41	2.4	0.22	$\boldsymbol{PRB}_{usage} \sim 0.058\boldsymbol{T}_{5G} - 0.144\boldsymbol{UE}_{NR} + +0.653\boldsymbol{CCE}_{usage} + 0.022\boldsymbol{CQI}_{avg}$
F	0.990	97.35	2.48	0.41	$\boldsymbol{PRB}_{usage} \sim 0.057\boldsymbol{T}_{5G} - 0.270\boldsymbol{UE}_{NR} + +0.733\boldsymbol{CCE}_{usage} - 0.011\boldsymbol{CQI}_{avg}$

6

Traffic Migration Scenarios

Contents

This chapter presents three different traffic volume transfer scenarios. Section 6.1 provides a migration scenario through the U2100 frequency band switch-off, whereas Section 6.2 details a more complex transfer scenario with the complete 3G switch-off, based on the development of a link budget for both uplink and downlink. These first two scenarios were built for the same vendor, but distinct sites, and use the power consumption models presented in Chapter 4. In addition, a simple scenario of data traffic volume migration from 4G cells to the respective 5G cell was created in order to predict the impact of 4G traffic on the resources that 5G provides, presented in Section 6.3.

6.1 U2100 Traffic Migration Scenario

6.1.1 Introduction

Cellular networks are changing, and some are sunsetting. With increasing 4G proliferation and 5G already being implemented in some regions, network operators' focus lies in 2G and 3G switch-off to free up the spectrum they occupy and reuse it with newer and more efficient cellular technologies in order to provide a faster and more responsive network to their costumers. However, due to the significant size of the markets for 2G based M2M and IoT applications, many European mobile operators are planning to retire 3G services before 2G [78].

Although the main reason for network shutdowns is that Mobile Network Operators (MNOs) have limited spectrum available for expansion, the legacy networks' switch-off can have a great impact in energy consumption and thus reduce MNOs' costs and ecological footprint, as mentioned. Taking into account that BSs are the main energy consumers of a cellular network and since the energy consumption by the radio equipment represents more than 50% of the total consumption of a BS, energy efficiency solutions applied to RRUs are provided [79].

In addition to the RRUs operating on a single technology, this vendor presents radios shared between 2G and 3G in the 900 MHz band, as already discussed in Chapter 4. Thus, this section provides a traffic volume migration scenario applied to a site with shared radios, based on the switch-off of the U2100 frequency band, whose traffic volume transits to other available frequency bands, whether from GSM, UMTS or LTE, while the traffic volume belonging to the UMTS 900 MHz (U900) band remains intact.

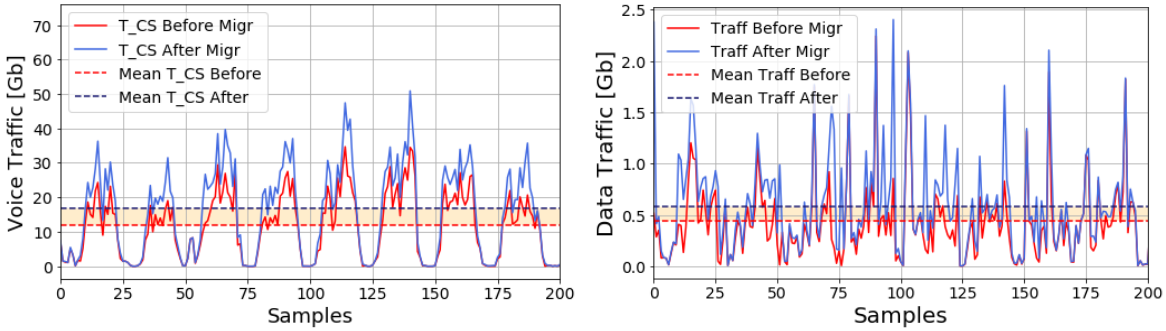
6.1.2 Migration and Capacity Analysis

The site and sector where cell C is located was used in order to present a traffic migration scenario, consisting of a GSM 900 MHz (G900) cell, one U900 cell, two U2100 cells and cell C, which operates in the LTE 800 MHz (L800) frequency band.

Given that the cell coverage of U900 is theoretically larger than U2100, and presupposing that this is still valid when comparing cells of different technologies, i.e., assuming that the cell coverage of L800

is also wider than U2100 and that the distribution of users remains unchanged in order to simplify the scenario, it is possible to transfer the voice traffic volume (CS) from the U2100 cells to the U900 cell and, in turn, their data traffic volume (Release 99 (R99) and HSDPA) to the L800 cell (cell C).

The results of the traffic transfer simulation are presented in Figure 6.1, in which the CS voice traffic volume of U900 cell increases by an average of 39.36% after receiving the voice traffic volume from the U2100 cells, whose data traffic migration leads to an average increase of 30.70% in the data traffic volume of the L800 cell, thus having a major impact on the U900 cell, as expected. For each cell, only 200 samples are illustrated in order to facilitate the visualization and interpretation of the plots.



(a) U900 CS voice traffic.

(b) L800 (cell C) data traffic.

Figure 6.1: Comparison of traffic volume before and after U2100 migration.

With the transfer of traffic volume from the U2100 cells, during the busy hour there is a 37.46% increase in the volume of CS voice traffic from cell U900 as well as an increase of about 22.38% in the volume of data traffic from cell L800. Although the majority of data traffic volume is transmitted in the 4G frequency band at any time of the day, an estimation of the post-migration downlink PRB utilization rate was performed using the PRB usage prediction model developed for cell C, described in Section 5.1.5, and is shown in Figure 6.2. When updating the data traffic volume independent variable with the new values resulting from the transfer, it is possible to obtain a hypothetical estimate of the PRB utilization rate after migration using the model previously proposed, while the remaining independent variables reflect the same radio conditions around the cell.

Figure 6.2 displays the measured PRB utilization rate along with the one estimated after the migration, for the periods of higher traffic, i.e., only for samples whose cell throughput values are above the 70^{th} percentile. During the busy hour of the estimated period, the L800 cell PRB usage rate have increased nearly 4.90%, confirming that the transfer of data traffic from the U2100 cell to the L800 cell does not have a relevant impact on the PRB utilization rate of the latter. In addition, the predicted values never exceed a 100% usage rate, thus allowing its migration.

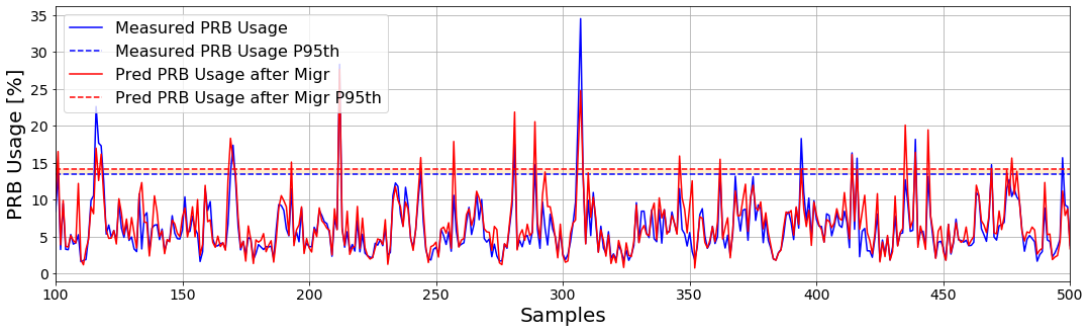
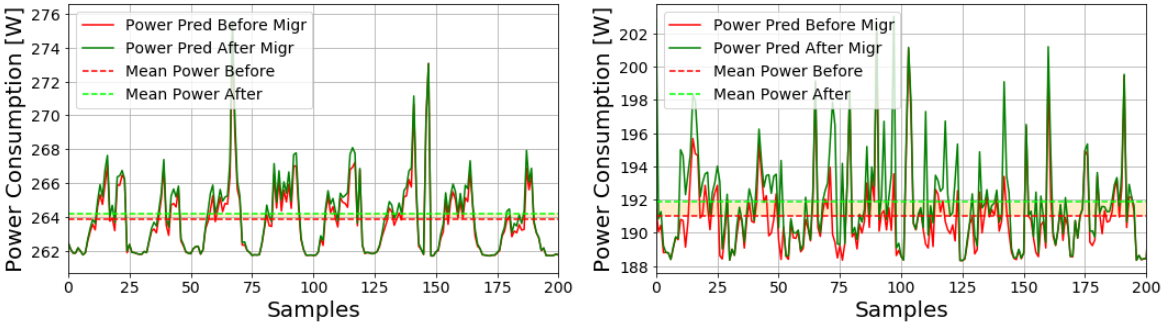


Figure 6.2: L800 cell PRB usage prediction after U2100 traffic migration.

6.1.3 Energy Balance

It is also important to analyze the energetic impact that the traffic volume migration of the U2100 band causes in the receiving frequency bands, which will be determined through the energy consumption models presented in Chapter 4. Figure 6.3 displays the power consumption for the actual traffic data, estimated before the traffic volume migration, as well as the predicted power consumption after the transfer, both estimated by the RRUs power consumption models, for each RRU that receives traffic volume.



(a) For GU900 RRU.

(b) For L800 (cell C) RRU.

Figure 6.3: Comparison of power consumption before and after U2100 migration.

As expected, the traffic volume of U2100 that was migrated to the other bands, does not generate a high increase in the power consumed by the radios present in the current site and sector. Thereby, a study of RRUs power consumption is developed, in order to estimate the energetic impact on the present site and sector due to the U2100 band switch-off. In Table 6.1 it is possible to analyze the difference between the real power consumption (pre-migration) and the expected power consumption (post-migration) in different circumstances. The energy consumption of the remaining BS components is not considered here.

Table 6.1: Balance of RRUs energy consumption within the U2100 scenario.

	$P^{50\%}$ [W]		$P^{95\%}$ [W]		Avg Power [W]	
Band	**Before**	**After**	**Before**	**After**	**Before**	**After**
L800	190.47	191.28	195.63	197.26	191.02	191.85
U2100	212.33	0	220.89	0	213.69	0
GU900	263.56	263.87	267.32	268.04	263.88	264.21
TOTAL	**666.35**	**455.14**	**683.84**	**465.30**	**668.60**	**456.06**

The energy consumption balance, shown in Table 6.1, presents some calculations regarding the power consumed by the RRUs of the current site and sector, such as power values in busy hour (95th percentile), median (50th percentile) and also mean power. Performing a deeper analysis at the level of the mean power, it is estimated an increase of 0.43% in the RRU that operates in the L800 band, corresponding to 0.83 W in energy consumption, and an increase of 0.12% in the average power of the shared GU900 RRU, which corresponds to 0.33 W. On the other hand, there is a reduction of 100% in the overall power of the RRU operating in the U2100 band, since this one has been switched off. Therefore, a 31.79% reduction in the total average power consumed by all the RRUs of that site and sector is estimated, resulting in a decrease of 212.54 W.

6.2 3G Traffic Migration Scenario

6.2.1 Introduction

As already mentioned, in Europe, it is estimated that MNOs will turn-off the 3G networks before 2G, which is primarily due to the extensive roll-out of M2M and IoT types of services based on 2G technology. The cost of migrating a large number of M2M connections remains a cause for concern by the MNOs, being a key factor in maintaining 2G longer than 3G services [78]. Thus, the current scenario suggests a complete 3G shutdown, releasing both U900 and U2100 frequency bands by swapping their traffic to other technologies.

The purpose was to create a more complex scenario than the previous section, since the traffic migration will be based on the coverage areas mapping of each cell belonging to the site and sector, through the development of a link budget that approximately portrays the real radio propagation conditions of the cells in question. After the cell coverage mapping and subsequent traffic volume transfer, the BS energy consumption is analyzed using the models presented in Chapter 4 for RRUs that operate on a single technology. Thus, the study was carried out for a BS composed of radios separated by frequency band, in order to be able to switch-off the U900 band RRU, which was not allowed in the

previous scenario since the respective site owns radios shared between the G900 and U900 frequency bands.

6.2.2 Link Budget

This subsection comprises of a link budget for both uplink and downlink transmission directions held for the site and sector consisting of a G900 cell, one U900 cell, two U2100 cells and, finally, a cell for each 4G frequency band: L800, LTE 1800 MHz (L1800) and LTE 2600 MHz (L2600), already referred as cell D1, D2 and D3, respectively. The presented calculations as well as the typical values adopted during the link budget development are based on approaches detailed in [15] and [80].

In a first phase, the link budget calculations estimate the maximum allowed signal attenuation (path loss) between the mobile and the BS antenna. The maximum path loss allows the maximum cell range to be estimated with a suitable propagation model.

The link budget was built in parallel for each frequency band from the different network technologies, going into detail on 3G services (CS, PS R99 and HSPA). In this subsection, link budget calculations are briefly discussed. Several parameters were obtained by manipulating the actual data from the respective cells, whenever possible, in order to obtain a more reliable estimate.

Data Rate

In order to portray the conditions observed at the cell edge, the data rate was given through the 5^{th} percentile values of the user throughput measured in each cell. For the 3G technology data services, it was calculated the PS interactive High Speed (HS) and DCH/FACH user throughput in both downlink and uplink directions, in a measurement period of 15 minutes, while the AMR 12.2 kbps bit rate was used for both 2G and 3G voice service.

In the case of the LTE network, the 5^{th} percentile of the L800 cell throughput was calculated, while for the remaining bands the user throughput was estimated based on the proportion of bandwidth relative to the L800 cell using its 5^{th} percentile user throughput value. Since the L800 cell bandwidth is 10 MHz and, in turn, the bandwidth of both L1800 and L2600 cells is 20 MHz, the data rates of these last cells will be twice the 5^{th} percentile of the first cell throughput.

Parameter Values

Tables 6.2 and 6.3 introduce the parameters used in downlink and uplink link budgets, respectively. Both provide a description of the parameters along with some parameter assumptions based on [15] and [80]. Other parameters, such as Equivalent Isotropic Radiated Power (EIRP), are obtained based on calculations presented in the tables as well.

Table 6.2: Downlink link budget parameters [15], [80].

Description	Value used
BS maximum transmission power. It is given by the vendor.	-
BS antenna gain depends on the antenna size and the number of sectors. It is given by the manufacturer.	-
Cable loss between the base station antenna connector and the antenna. The cable loss value depends on the cable length, cable thickness and frequency band.	2.0 dB
EIRP [dBm] = BS max tx power [dBm] + BS antenna gain [dBi] − Cable loss [dB]	-
UE RF noise figure. It depends on the frequency band, Duplex separation and on the allocated bandwidth.	7.0 dB
Bandwidth. It is fixed by the MNO.	For G900: 200 kHz
Thermal noise [dBm] = $10\log_{10}(k$(Boltzmann constant) $\times\ T$(290K) $\times$ bandwidth$)$	-
Receiver noise floor [dBm] = UE RF noise figure [dB] + Thermal noise [dBm]	-
Receiver sensitivity [dBm] = Receiver noise floor [dBm] + SINR [dB]	For G900: -104.0 dBm
Interference margin accounts for the increase in the terminal noise level caused by the other cell.	4.0 dB
Control channel overhead.	1.0 dB
UE antenna gain depends on the type of device and on the frequency band.	0 dBi
Body loss is typically included for voice link budget where the terminal is held close to the user's head.	For voice: 3.0 dB

BS Maximum Transmission Power

Although the MNO provides these values, it was necessary to map the 3G power, since the link budget is held for the different services.

The maximum transmit power, P_{Total}, which has to be divided by all the transport channels. With HSDPA, the total power is given by, [81]:

$$P_{Total} = P_{CCH} + P_{DCH} + P_{HSDPA} \tag{6.1}$$

where P_{CCH} is the power allocated to the control channels, P_{DCH} the power of DCH and P_{HSDPA} the power available for the service HSDPA.

Voice and PS R99 users are allocated on dedicated channels (DCH), which carry various services, such as AMR with different bit rates and the various PS R99 (8, 16, 32, 64, 128, 144, 256 and 384 kbps). Hence, a simple MLR model was developed, whose coefficients are used to estimate the respective

Table 6.3: Uplink link budget parameters [15], [80].

Description	Value used
UE maximum transmission power for power class 3. Different power classes would have different power levels.	23.0 dBm
UE antenna gain (see Table 6.2).	0 dBi
Body loss (see Table 6.2).	For voice: 3.0 dB
EIRP [dBm] = UE max tx power [dBm] + UE antenna gain [dBi] − Body loss [dB]	-
BS RF noise figure. It depends on the implementation design.	2.0 dB
Bandwidth. It depends on bit rate, which defines the number of resource blocks. Assuming a two resource block allocation for LTE.	For 3G: 3.8 MHz
Thermal noise (see Table 6.2).	-
Receiver noise floor [dBm] = BS RF noise figure [dB] + Thermal noise [dBm]	-
Receiver sensitivity [dBm] = Receiver noise floor [dBm] + SINR [dB]	For G900: -114.0 dBm
Interference margin accounts for the increase in the terminal noise level caused by the interference from other users. Since LTE uplink is orthogonal, there is no intra-cell interference. However, a margin for the other cell interference is still required.	-
Cable loss between the base station antenna and the low noise amplifier. The cable loss value depends on the cable length, cable type and frequency band.	0 dB
BS antenna gain depends on the antenna size and the number of sectors (see Table 6.2).	-
Fast fading margin is typically used with WCDMA due to fast power control to allow headroom for the power control operation.	For 3G: 1.8 dB
Soft handover.	For 3G: 2.0 dB

powers, and it is described by:

$$P_{Total} = \beta_0 + \beta_1 R_{Voice} + \beta_2 R_{PS_{R99}} + \beta_3 R_{HSDPA}. \tag{6.2}$$

The independent variables are the normalized bit rates corresponding to the services of voice (R_{Voice}), PS R99 ($R_{PS_{R99}}$) and HSDPA (R_{HSDPA}), whose respective powers are given by P_{Voice}, $P_{PS_{R99}}$ and P_{HSDPA}. With the intercept and the coefficients of the independent variables, it is possible to estimate the proportion of power allocated to each transport channel. The intercept will represent the P_{CCH}, while the remaining coefficients will depict the power of the respective services. The approximate values of the

proportions are: 17.5%, 23.1%, 2.4% and 57,0% for P_{CCH}, P_{Voice}, $P_{PS_{R99}}$ and P_{HSDPA}, respectively.

SINR in LTE

The SINR value depends on the modulation and coding schemes, which again depend on the data rate and on the number of resource blocks allocated. The number of PRBs is given by Table 2.4, depending on the bandwidth. As mentioned, the downlink bandwidth is fixed by the MNO, while a 360 kHz uplink LTE bandwidth is assumed, which corresponds to an allocation of two PRBs [15]. Multiplying the data rate on cell edge (given by the 5th percentile of user throughput) by 1 ms, the TBS value is obtained, since it represents the number of bits which can be transmitted per 1 TTI, which is 1 ms long. Giving the TBS value and the number of PRBs, the TBS index can be determined, leading further to the CQI value, which is then used to map SINR. The tables used in these approach are presented in Appendix A.

SINR in UMTS

For UMTS, the procedure was similar to the LTE approach. Once again, the TBS is the number of bits that can be transmitted per 1 TTI, which, in turn, is equal to 2 ms for UMTS. The TBS value can be calculated by multiplying the user throughput by 2 ms. Assuming the UE belong to Category 14 devices, while it uses 64-QAM and does not use MIMO, it is possible to map the CQI with the TBS value, as presented in Appendix A. Finally, the UMTS SINR can be given by, [82]:

$$CQI = SINR + 4.5. \tag{6.3}$$

Propagation Model

A propagation model describes the average signal propagation and converts the maximum allowed propagation loss to the maximum cell range. It depends on conditions such as environment (urban, rural, etc), distance, frequency, indoor/outdoor and atmospheric conditions.

One of the most widely used radio propagation estimate models is the Okumura-Hata model, especially in urban environments, and it is described in dB by, [83]:

$$L_U = 69.55 + 26.16\log_{10}(f) - 13.82\log_{10}(h_{BS}) - C_H + [44.9 - 6.55\log_{10}(h_{BS})]\log_{10}(d) \tag{6.4}$$

where L_U is the pass loss in urban areas, f is the frequency of transmission in MHz, h_{BS} is the effective height of BS antenna in meters and d is the distance between BS and MS in km. In addition, C_H is a correction factor which depends on the environment type. For small and medium-sized cities it is given in dB by, [83]:

$$C_H = 0.8 + (1.1\log_{10}(f) - 0.7)h_{MS} - 1.56\log_{10}(f) \tag{6.5}$$

where h_{MS} is the height of MS antenna. For large cities and frequencies between 150 and 1500 MHz,

the correction factor is presented in dB by, [83]:

$$C_H = \begin{cases} 8.29[\log_{10}(1.54h_{MS})]^2 - 1.1, & f \leq 200 \\ 3.2[\log_{10}(11.75h_{MS})]^2 - 4.97, & f \geq 400 \end{cases}. \tag{6.6}$$

In this scenario, the site is located in Lisbon, thus, considering its size, the correction factor for medium cities was used. The h_{MS} is assumed to be fixed to 1.5 meters [83], while the h_{BS} is provided by the MNO.

6.2.3 Migration and Capacity Analysis

With the development of the link budget for uplink and downlink it is observed that the distances between BS and MS are mostly more restrictive for uplink and, therefore, traffic volume will be transferred according to them. In order to map the traffic proportions to be transferred, the cell coverage areas, A_{Cell} in this direction of transmission are calculated by, [84]:

$$A_{Cell} = \frac{3\sqrt{3}}{2}R^2, \tag{6.7}$$

where R is the distance radius coverage given also by d.

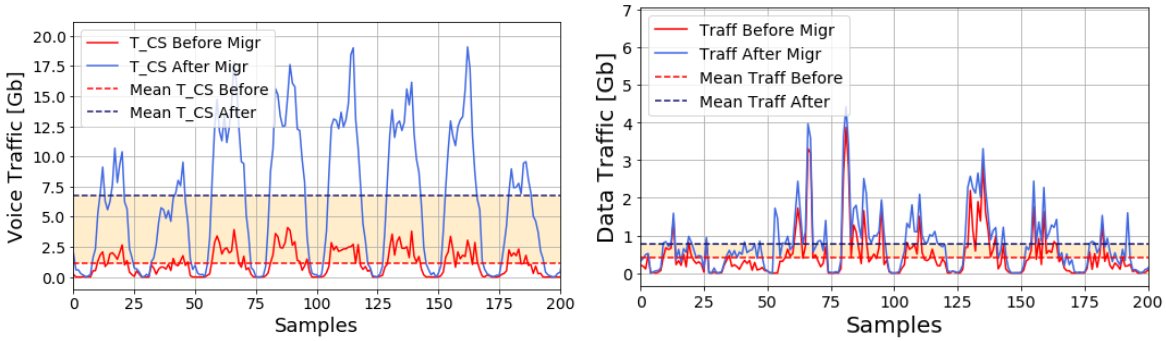
In first place, the coverage area of UMTS voice service is smaller than the coverage area corresponding to the G900 cell, thus, the voice traffic volume of both U900 and U2100 cells is migrated to the G900 cell. In order to determine the UMTS voice traffic in Erlangs, ρ_{Voice}, which is firstly given by Mbps, Equation (6.8) is used, where λ_m is the voice traffic in bps, given by KPIs, and μ_m is the several data rates available on voice services and, consequently, the channel capacity, also in bps, being $m = \{12, 64, \text{AMR-WB}\}$, allowing the voice traffic transfer from UMTS to GSM [85].

$$\rho_{Voice} = \sum \frac{\lambda_m}{\mu_m}. \tag{6.8}$$

In addiction, the coverage area of L2600 involves the areas corresponding to the U2100 data services (comprised of R99 and HSPA), thus their data traffic volume are transferred directly to the L2600 cell. However, the same does not happen with the U900 data services area, which is wider than the L2600 cell coverage area and, thus, traffic migration proportions are calculated for each 4G cell, based on the comparison of each cell area. Through the proportions mapping, 49.88% of the U900 data traffic volume is transferred to the L2600 cell, 33.19% to the L1800 cell and, finally, the remaining 16.93% are migrated to the L800 cell.

The traffic transfer simulation are illustrated in Figure 6.4, in which it is possible to verify a large increase in G900 voice traffic volume, which rises to an average of 483.48%, after receiving the 3G voice traffic. In turn, the L2600 cell is the 4G cell that receives more data traffic from U900 cell while obtaining the entire U2100 data traffic, therefore, its traffic volume increases by an average of 91.79%, whereas the data traffic volume from L800 and L1800 cells increases 5.23% and 1.79%, respectively,

after receiving the remaining U900 data traffic.



(a) G900 CS voice traffic.

(b) L2600 (cell D3) data traffic.

Figure 6.4: Comparison of traffic volume before and after 3G migration.

During the busy hour (95th percentile), there is an increase of about 415.39% in the volume of 2G CS voice traffic as well as a 63.52% increase in the L2600 traffic volume. In turn, for L800 and L1800 cells, traffic increments continue with low significance, being 2.20% and 2.16%, respectively.

An estimate of the post-migration downlink PRB utilization rate was performed using the PRB usage prediction model developed for the present cells, described in Section 5.1.5. When adding the new traffic volume to the previous independent variable value of data traffic volume, the MLR model estimates the PRB utilization rate after migration, whereas the remaining independent variables maintain the same cell radio conditions. Since the L2600 data traffic volume suffers the biggest data traffic increase, compared to other 4G cells, the estimate result is displayed in Figure 6.5 for this cell (cell D3).

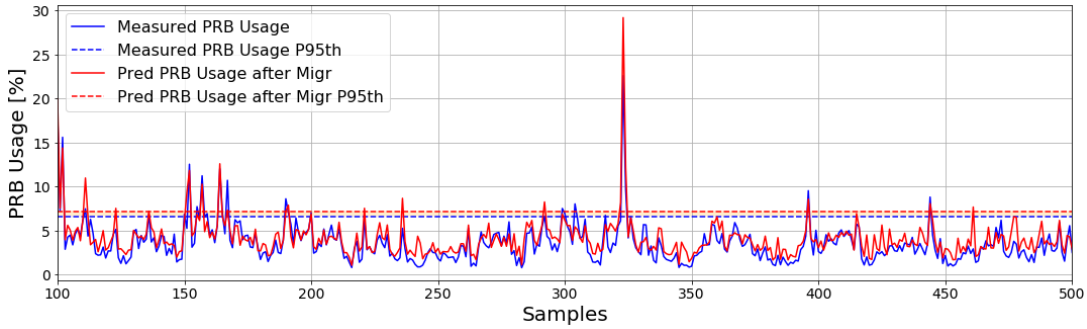


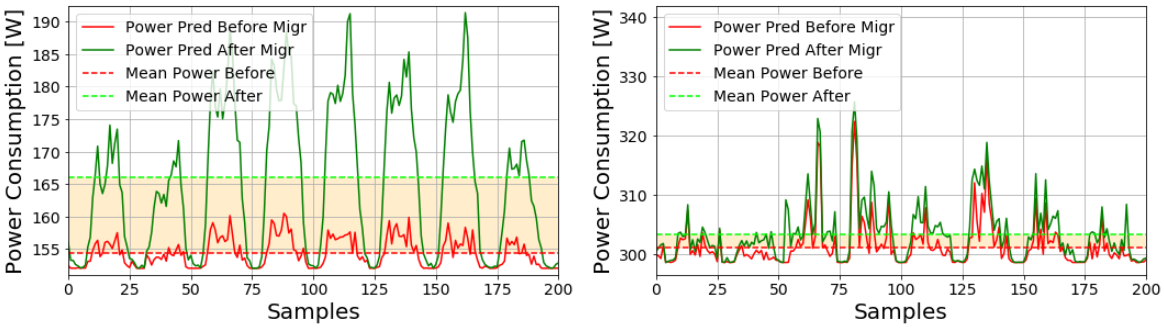
Figure 6.5: L2600 cell PRB usage prediction after 3G traffic migration.

Figure 6.5 shows the real PRB utilization rate as well as the estimated rate after the transfer, for L2600 cell (cell D3), during high traffic periods (for samples with cell throughput above the 70th percentile). Although the cell in question receives a significant amount of data traffic volume, it offers some capacity margin, since the PRB usage rate remains far from reaching 100% (maximum use of resources) with an

increase of nearly 8.76% during the busy hour of the estimated period, confirming that the migration of data traffic does not negatively impact recipient cells.

6.2.4 Energy Balance

Using the energy consumption models presented in Chapter 4, an energy consumption balance is performed, analyzing the impact that the migration detailed above causes on the receiving cells. Figure 6.6 displays the energy consumption for the measured traffic data along with the estimated energy consumption after the migration, both predicted using the RRUs power consumption models, for G900 and L2600 RRUs, once the latter receives more traffic volume than the other 4G cells.



(a) For G900 RRU.

(b) For L2600 (cell D3) RRU.

Figure 6.6: Comparison of power consumption before and after 3G migration.

Since the majority of power consumption is baseline, it is expected that the increase in traffic on the radios that remain on will not have a significant impact on their energy consumption. To prove this statement, an analysis of RRUs power consumption is conducted for the radios present in the current site and sector, thus studying the energy impact generated by the switch-off of 3G RRUs. Table 6.4 details a comparison between the energy consumption before and after the traffic volume transfer. The study does not include the power consumption of the remaining BS components.

The balance of power consumption, presented in Table 6.4, provides several estimates of the power consumed by each RRU in different circumstances such as power values in busy hour (95^{th} percentile), median (50^{th} percentile) and also average power. Regarding the mean values, the L800, L1800 and L2600 RRUs suffer an increase of 0.10%, 0.13% and 0.75% in power consumption, which corresponds to 0.19 W, 0.37 W and 2.27 W, respectively, whereas G900 RRU power consumption increases 7.53%, corresponding to 11.63 W. Nevertheless, the energy consumption of both 3G RRUs is reduced by 100% with their switch-off. Thus, a reduction of 24.45% is finally obtained in the total average power consumed by all the radios belonging to the current site and sector, leading to a global decrease of 307.23 W.

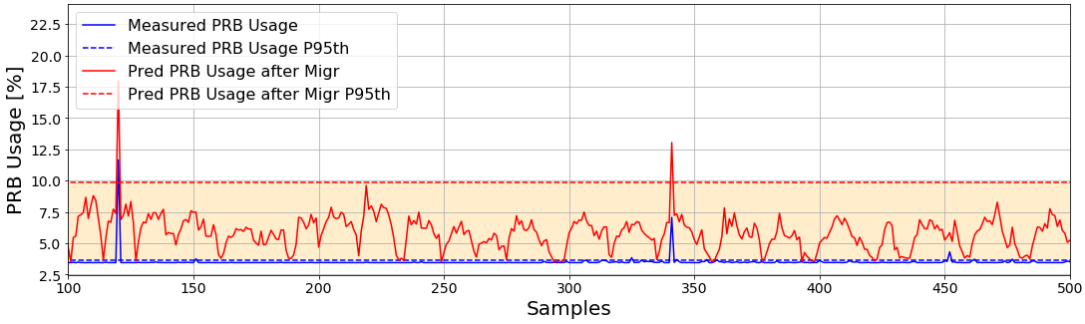
Table 6.4: Balance of RRUs energy consumption within the 3G scenario.

	$P^{50\%}$ [W]		$P^{95\%}$ [W]		Avg Power [W]	
Band	**Before**	**After**	**Before**	**After**	**Before**	**After**
L2600	300.29	302.65	306.45	311.41	301.11	303.39
L1800	289.73	290.05	310.18	311.07	289.51	289.88
L800	188.83	189.07	195.85	196.06	189.63	189.82
U2100	165.68	0	169.87	0	166.03	0
U900	155.48	0	162.89	0	155.65	0
G900	153.94	165.01	158.44	185.23	154.40	166.02
TOTAL	**1253.94**	**946.78**	**1303.68**	**1003.77**	**1256.33**	**949.11**

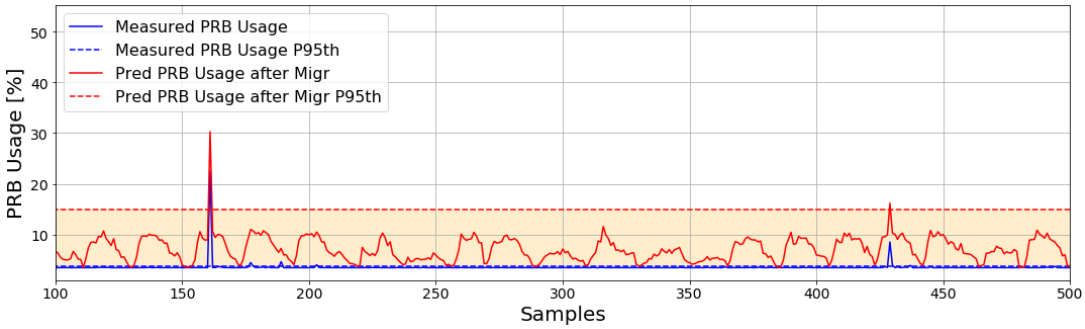
6.3 5G Hypothetical Migration Scenario

For the other vendor, a simple migration scenario was created with the current data, consisting of data traffic volume migration from 4G to 5G, considering the worst possible scenario. Although the coverage area of the 5G cell is narrower than the respective 4G cell areas of the same site, the migration is carried out directly, i.e., without the use of traffic proportions based on the coverage areas, since it is estimated a significant increase in the volume of data traffic with the arrival of 5G. With this approach, it is then assumed that the increase in traffic volume will compensate for the reduction in users of the 5G cell compared to the respective 4G cells.

The data traffic migration results are presented in Figure 6.7 for both cells. The calculated 95th percentile values of measured PRB utilization rate are 3.67% and 3.70% for cells E and F, respectively,representing a mostly residual use of resources, as would be expected since 5G only presents residual traffic volume as mentioned. On the other hand, for the estimated PRB utilization rate after migration, the calculated 95th percentile values are 9.86% and 14.99%, for both cells E and F, respectively; therefore, it can then be concluded that the current 5G cells have sufficient resources to receive the 4G mobile data traffic volume.



(a) For cell E.



(b) For cell F.

Figure 6.7: 5G PRB usage prediction after traffic migration.

7

Conclusions

Contents

This chapter summarizes the work developed within the Thesis scope and enumerates aspects to be developed in future work.

7.1 Summary

The main goal of this Thesis was to develop cell capacity prediction models, in downlink direction, for both 4G and 5G technologies, using MLR algorithms, while building traffic volume migration scenarios using existing energy consumption models.

The LTE capacity model that predicts the downlink cell throughput consists of two modules: one where cells with capacity saturation were detected and another where the maximum cell capacity was estimated. In the first module, the available cells go through an analysis regarding the a list of thresholds, in which two cells with capacity problems were detected, being both advised to perform a RF optimization in order to minimize the channel interference. In the second module, downlink cell throughput was predicted for both detected cells using MLR algorithms, with R_a^2 values of 0.991 and 0.997, which corroborate the high reliability of the model along with the other error metrics provided. It is then estimated the maximum cell capacity for the cell radio propagation environment using the proposed MLR model to calculate the maximum cell throughput, forcing the use of all cell resources by setting the PRB utilization rate to 100%. At the busy hour, both cells present almost no capacity margin, resulting in a cell load of 97.46% and 94.11%, demonstrating that both cells are approaching an overload situation.

The second LTE capacity model predicts the downlink PRB utilization rate, created similarly to the cell throughput prediction model. After each data traffic volume migration, the PRB usage rate of each cell is estimated, based on the new traffic volume that was received, allowing to evaluate the hypothetical use of physical cell resources, which can be used to validate or not the traffic migration scenarios performed for this vendor. Once more, the model error metrics confirm its high reliability, having R_a^2 values of 0.964 and 0.986, for example.

Two traffic migration scenarios were developed for these cells: the first is based on the U2100 frequency band switch-off, while the second one is more complex since it performs the complete 3G switch-off including the development of a link budget for both uplink and downlink. Within the U2100 traffic migration scenario, the voice traffic volume (CS) from the U2100 cells was transferred to the U900 cell, leading to an average increase of 39.36% in its traffic volume, and, in turn, their data traffic volume (Release 99 (R99) and HSDPA) migrated to the L800 cell, whose traffic volume increased by an average of 30.70%. These traffic transfers generated a 31.79% reduction in the total average power consumed by all the RRUs of that site and sector, resulting in a decrease of 212.54 W.

Regarding the 3G traffic migration scenario, the transfer was based on the coverage areas mapping of each cell, through the development of a link budget that approximately portrays the real radio propagation conditions of the cells in question. The link budget calculations provided the cell coverage

areas, allowing to map the traffic proportions to be transferred. With the migration, there was a significant increase in G900 voice traffic volume, as expected. The PRB utilization rate was estimated for the 4G cells, taking into account the addition of data traffic volume, which remains far from reaching 100% (maximum use of resources) with an increase of 8.76% in the L2600 cell, which suffers the biggest data traffic increase, during the busy hour of the estimated period, confirming that the migration of data traffic does not negatively impact recipient cells. Once more, the energy consumption balance was performed resulting in a reduction of 24.45% in the total average power consumed by all the radios belonging to the current site and sector, leading to a global decrease of 307.23 W.

Finally, a 5G cell resource prediction model has been proposed and developed for two cells which present R_a^2 values of 0.996 and 0.990, despite the predominance of residual traffic in the measured data that were used in the model development. The development process of this model is similar to the one described for the 4G PRB utilization rate model, evaluating the use of cell resources, in order to be able to analyze its behavior when receiving 4G data traffic volume. Additionally, a simple scenario of data traffic volume migration from 4G cells to the respective 5G cell was created in order to predict the impact of 4G traffic on the resources that 5G provides, concluding that 5G technology has enough cell resources.

7.2 Future Work

Further work can be made to evolve this research. With the emergence of 5G technology, to enhance the energy efficiency of 5G and provide a green future cellular networks, Artificial Intelligence (AI) usage and ML techniques could be applied in the context of wireless networks to improve the overall performance and efficiency.

The development of a module for monitoring and evaluating energy efficiency in 5G networks, accompanied with further solutions to reduce energy consumption and carbon footprint in the different network segments, would be advantageous.

Finally, investigating the importance of KPIs which may impact the QoS and QoE, traffic prediction models could be developed in order to constitute an AI-based network optimization which would optimize network sites power consumption through several parameters.

Bibliography

[1] Ericsson, "Ericsson Mobility Report," Tech. Rep., June 2020. [Online]. Available: https://www.ericsson.com/49da93/assets/local/mobility-report/documents/2020/june2020-ericsson-mobility-report.pdf

[2] J. A. Fernández-Segovia, S. Luna-Ramírez, M. Toril, and J. J. Sánchez-Sánchez, "Estimating Cell Capacity From Network Measurements in a Multi-Service LTE System," *IEEE Communications Letters*, vol. 19, no. 3, pp. 431–434, March 2015.

[3] K. Chang and R. P. Wicaksono, "Estimation of network load and downlink throughput using RF scanner data for LTE networks," *2017 International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, pp. 1–8, July 2017.

[4] M. H. Alsharif, R. Nordin, and M. Ismail, "Survey of Green Radio Communications Networks: Techniques and Recent Advances," *Journal of Computer Networks and Communications*, 2013.

[5] ZTE. (2012) "BeamHop: A New Member of LTE Base Station Family". Accessed 26-March-2019. [Online]. Available: https://www.zte.com.cn/global/about/magazine/zte-technologies/2012/6/en_561/370650

[6] Forsk. (2013) "RAN Planning". Accessed 28-March-2019. [Online]. Available: https://www.forsk.com/ran-planning

[7] E. Seurre, P. Savelli, and P.-J. Pietri, *GPRS for Mobile Internet*. Artech House Publishers, 2003.

[8] J. Scourias, "Overview of the GSM Cellular System, Extended Abstract," in *1991 GSM Protocol Architecture: Radio Sub-system Signalling, IEEE 41st Vehicular Technology Conference*, August 1997.

[9] L. Harte, B. Bramley, and M. Davis, *Introduction to GSM*, 2nd ed. Althos, 2011.

[10] A. S. Mondal, *Mobile IP: Present State and Future*, 1st ed., ser. Series in Computer Science. Springer US, 2003.

[11] M. A. Shamma, “GSM/GPRS Erlang capacity analyses and simulations under air traffic loading conditions,” in *2005 IEEE Aerospace Conference*, March 2005, pp. 1931–1943.

[12] E. Katsaros and K. Vlachodimitropoulos, “Capacity and Dimensioning Problems of a GSM Radio Access Network after Seismic Activity,” in *2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, September 2007, pp. 1–5.

[13] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice Hall PTR, 2002.

[14] X. Li, *Radio Access Network Dimensioning for 3G UMTS*. Vieweg+Teubner, January 2011.

[15] H. Holma and A. Toskala, *WCDMA for UMTS: HSPA evolution and LTE*, 4th ed. Wiley, 2007.

[16] A. Kang, “TMS320C6713DSP Implementation of pulse shaping filter,” *European Scientific journal*, January 2013.

[17] R. J. Susan and S. S. Pillai, *Frequency Domain Equalisation for OFDMA System in Multipath Fading Channels*, January 2010, vol. 90, pp. 369–377.

[18] Ericsson, “Channel Element Dimensioning Guideline, Ericsson WCDMA Radio Access Network,” March 2007.

[19] 3GPP. ”HSPA”. Accessed 29-April-2019. [Online]. Available: https://www.3gpp.org/technologies/keywords-acronyms/99-hspa

[20] H. Holma and A. Toskala, *HSDPA/HSUPA for UMTS: high speed radio access for mobile communications*. John Wiley, 2006.

[21] A. Masmoudi, D. Zeghlache, and S. Tabbane, “Resource and scheduling optimization in HSDPA based UMTS networks,” June 2019.

[22] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, 2nd ed. Wiley, 2011.

[23] Tektronix Communications, “LTE Networks: Evolution and Technology Overview,” Tektronix Communications, Whitepaper, September 2010.

[24] S. Sesia, I. Toufik, and M. Baker, *LTE – The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed. Wiley, 2011.

[25] C. Cox, *An Introduction to LTE: LTE, LTE-Advanced, SAE, VoLTE and 4G Mobile Communications*, 2nd ed. Wiley, 2014.

[26] J. Jermyn, R. Piqueras Jover, I. Murynets, M. Istomin, and S. Stolfo, “Scalability of Machine to Machine systems and the Internet of Things on LTE mobile networks,” July 2015.

[27] M. Abdulhasan, M. Salman, C. K. Ng, N. Noordin, S. Jahari Hashim, and F. Hashim, “An Adaptive Threshold Feedback Compression Scheme Based on Channel Quality Indicator (CQI) in Long Term Evolution (LTE) System,” *Wireless Personal Communications*, vol. 82, June 2015.

[28] Motorola, “Realistic LTE Performance – From Peak to Subscriber Experience,” Motorola, Whitepaper, 2009.

[29] Anritsu, “LTE Resource Guide,” Anritsu, Tech. Rep., 2009.

[30] DCU School of Electronic Engineering. ”Architecture of LTE frame (FDD)”. Accessed 20-May-2019. [Online]. Available: http://www.eeng.dcu.ie/~longhao/material/LTE-FDD.pdf

[31] MediaTek and du, “5G design concepts towards the next generation networks,” MediaTek and du, Whitepaper, 2016.

[32] GSMA, “5G Spectrum - GSMA Public Policy Position,” GSMA, Tech. Rep., November 2018.

[33] ITU-R, “Recommendation M.2083-0, IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond,” 2015.

[34] GSMA, “Road to 5G: Introduction and Migration,” GSMA, Tech. Rep., April 2018.

[35] GTI, “5G Network Architecture White Paper,” GTI, Whitepaper, February 2018.

[36] ITU Radiocommunication Sector, “Recommendation ITU-R M.2083-0,” ITU Radiocommunication Sector, Tech. Rep., September 2015.

[37] K. Liolis, A. Geurtz, R. Sperber, D. Schulz, S. Watts, G. Poziopoulou, B. Evans, N. Wang, O. Vidal, B. Tiomela Jou, M. Fitch, S. Diaz Sendra, P. Sayyad Khodashenas, and N. Chuberre, “Use cases and scenarios of 5G integrated satellite-terrestrial networks for enhanced mobile broadband: The SaT5G approach,” *International Journal of Satellite Communications and Networking*, November 2017.

[38] Cisco, “5G Non Standalone Solution Overview,” 2018.

[39] A. A. Zaidi, R. Baldemair, H. Tullberg, H. Bjorkegren, L. Sundstrom, J. Medbo, C. Kilinc, and I. Da Silva, “Waveform and Numerology to Support 5G Services and Requirements,” *IEEE Communications Magazine*, vol. 54, no. 11, November 2016.

[40] X. Lin, J. Li, R. Baldemair, J. T. Cheng, S. Parkvall, D. C. Larsson, H. Koorapaty, M. Frenne, S. Falahati, A. Grovlen, and K. Werner, “5G New Radio: Unveiling the Essentials of the Next Generation Wireless Access Technology,” *IEEE Communications Standards Magazine*, vol. 3, no. 3, pp. 30–37, 2019.

[41] 3rd Generation Partnership Project. (ETSI, 2018) ETSI TS 138 306 V15.3.0. Accessed 1-June-2019. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/138300_138399/138306/15.03.00_60/ts_138306v150300p.pdf

[42] M. M. Mowla, I. Ahmad, D. Habibi, and Q. V. Phung, "A Green Communication Model for 5G Systems," *IEEE Transactions on Green Communications and Networking*, vol. 1, no. 3, pp. 264–280, September 2017.

[43] A. Abrol and R. K. Jha, "Power Optimization in 5G Networks: A Step Towards GrEEn Communication," *IEEE Access*, vol. 4, pp. 1355–1374, 2016.

[44] J. Huang, C. Xing, and C. Wang, "Simultaneous Wireless Information and Power Transfer: Technologies, Applications, and Research Challenges," *IEEE Communications Magazine*, vol. 55, no. 11, pp. 26–32, November 2017.

[45] R. T. Prabu, M. Benisha, V. T. Bai, and V. Yokesh, "Millimeter wave for 5G mobile communication application," in *2016 2nd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, February 2016, pp. 236–240.

[46] U. K. Dutta, M. A. Razzaque, M. Abdullah Al-Wadud, M. S. Islam, M. S. Hossain, and B. B. Gupta, "Self-Adaptive Scheduling of Base Transceiver Stations in Green 5G Networks," *IEEE Access*, vol. 6, pp. 7958–7969, 2018.

[47] IEEE Communications Society. (2018, August) Massive MIMO. Accessed 3-June-2019. [Online]. Available: https://techblog.comsoc.org/tag/massive-mimo/

[48] N. Andreadou, M. Olariaga, and G. Fulli, "Telecommunication Technologies for Smart Grid Projects with Focus on Smart Metering Applications," Energies, Tech. Rep., 2016.

[49] J. Lloret, J. Tomás, A. Canovas, and L. Parra, "An Integrated IoT Architecture for Smart Metering," *IEEE Communications Magazine*, vol. 54, pp. 50–57, December 2016.

[50] M.-G. Kang, W.-S. Jung, C.-Y. Kang, and S.-J. Yoon, "Recent Progress on PZT Based Piezoelectric Energy Harvesting Technologies," *Actuators*, vol. 5, p. 5, February 2016.

[51] S. Sojan and R. Kulkarni, "A Comprehensive Review of Energy Harvesting Techniques and its Potential Applications," *International Journal of Computer Applications*, vol. 139, pp. 14–19, April 2016.

[52] P. D. Diamantoulakis, K. N. Pappi, G. K. Karagiannidis, and H. V. Poor, "Autonomous Energy Harvesting Base Stations With Minimum Storage Requirements," *IEEE Wireless Communications Letters*, vol. 4, no. 3, pp. 265–268, June 2015.

[53] A. Spagnuolo, A. Petraglia, C. Vetromile, R. Formosi, and C. Lubritto, “Monitoring and optimization of energy consumption of base transceiver stations,” *Energy*, January 2015.

[54] J. Wu, Y. Zhang, M. Zukerman, and E. K. Yung, “Energy-Efficient Base-Stations Sleep-Mode Techniques in Green Cellular Networks: A Survey,” *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 803–826, Secondquarter 2015.

[55] H. Holtkamp, G. Auer, V. Giannini, and H. Haas, “A Parameterized Base Station Power Model,” *IEEE Communications Letters*, vol. 17, no. 11, pp. 2033–2035, November 2013.

[56] C. Lubritto, A. Petraglia, C. Vetromile, F. Caterina, A. D’Onofrio, M. Logorelli, G. Marsico, and S. Curcuruto, “Telecommunication power systems: Energy saving, renewable sources and environmental monitoring,” in *INTELEC 2008 - 2008 IEEE 30th International Telecommunications Energy Conference*, September 2008, pp. 1–4.

[57] AT&T Inc, “2019 Annual Report Form 10-K,” AT&T Inc, Tech. Rep., 2020.

[58] C. Han, T. Harrould, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. Thompson, I. Ku, C.-X. Wang, T. Le, M. Nakhai, J. Zhang, and L. Hanzo, “Green radio: Radio techniques to enable energy-efficient wireless networks,” *IEEE Communications Magazine*, vol. 49, pp. 46–54, June 2011.

[59] N. Shankar and S. Nayak, “Performance Management in Network Management System,” *International Journal of Science and Research (IJSR)*, vol. 4, pp. 2505–2507, May 2015.

[60] Cisco, “Performance Management Best Practices and Broadband Service Providers,” Cisco, Tech. Rep., June 2008.

[61] Huawei, “eNodeB V100R005C00 KPI Reference,” March 2012.

[62] 3GPP. (2010) Technical Specification Group Services and System Aspects; Telecommunication management; Configuration Management (CM); Concept and high-level requirements. Technical Specification (TS) 32.600, 3rd Generation Partnership Project (3GPP). Accessed 25-September-2019. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2440

[63] P. Domingos, “A Few Useful Things to Know About Machine Learning,” *Commun. ACM*, vol. 55, no. 10, pp. 78–87, October 2012.

[64] E. Alpaydin, *Introduction to Machine Learning*, 2nd ed. The MIT Press, 2010.

[65] P. Domingos, “A Unified Bias-Variance Decomposition and its Applications,” *Proc. 17th International Conf. on Machine Learning*, pp. 231–238, 2000.

[66] V. Pedras, “Identifying Quality of Experience (QoE) in 3G/4G Radio Networks based on Quality of Service (QoS) Metrics,” Master’s thesis, Instituto Superior Técnico, November 2017.

[67] D. Montgomery and G. Runger, *Applied Statistics and Probability for Engineers*, 3rd ed. John Wiley and Sons, 2003.

[68] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*. Springer, 2014.

[69] T. Saraiva, “Desenvolvimento de um Modelo de Consumo Energético Multi-Tecnologia e Multi-Vendor para Nós de Acesso Rádio em Redes Móveis,” Master’s thesis, Instituto Superior de Engenharia de Lisboa, October 2019.

[70] K. van Montfort, J. Oud, and A. Satorra, *Longitudinal Research with Latent Variables*, 1st ed. Springer, 2010.

[71] P. K. Dunn and G. K. Smyth, *Generalized Linear Models With Examples in R*. Springer, 2018.

[72] D. Parracho, “Desenvolvimento de Modelos de Capacidade para Redes Móveis 3G e 4G Usando Dados de Desempenho Reais,” Master’s thesis, Instituto Superior de Engenharia de Lisboa, December 2018.

[73] T. Saraiva, D. Duarte, I. Pinto, and P. Vieira, “An Improved BBU/RRU Energy Consumption Predictor for 4G and Legacy Mobile Networks using Mixed Statistical Models,” *2020 International Conference on Computing, Networking and Communications (ICNC)*, pp. 320–325, 2020.

[74] CELFINET, “RAN Energy Optimizer, Product Description,” 2018.

[75] Huawei, *eRAN Capacity Monitoring Guide*, Huawei Technologies Co., Ltd., 2016.

[76] Ericsson Internal, *LTE Post Launch Optimization Problem Cause, Counter Pegging, Solution & Case Study*, Ericsson AB, 2016.

[77] P. Vieira, P. Queluz, and A. Rodrigues, “Improving MIMO spectral efficiency in 4G macro-cellular networks,” *2nd Symposium of the Portuguese Committee of URSI*, pp. 320–325, November 2008.

[78] GSMA, “Legacy Mobile Network Rationalisation,” April 2020.

[79] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer, and L. van der Perre, “Challenges and Enabling Technologies for Energy Aware Mobile Radio Network,” *IEEE Communications Magazine*, vol. 48, no. 11, pp. 66–72, November 2010.

[80] H. Holma and A. Toskala, *LTE for UMTS - OFDMA and SC-FDMA Based Radio Access*, 1st ed. Wiley, 2009.

[81] P. Zanier and D. Soldani, “A simple approach to HSDPA dimensioning,” in *2005 IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2, September 2005, pp. 883–887.

[82] C. Kurnaz, B. K. Engiz, and M. O. Esenalp, “Challenges and Enabling Technologies for Energy Aware Mobile Radio Network,” *Turkish Journal of Electrical Engineering & Computer Sciences*, November 2017.

[83] J. T. J. Penttinen, *The Telecommunications Handbook: Engineering Guidelines for Fixed, Mobile and Satellite Systems*, 1st ed. John Wiley and Sons, 2015.

[84] F. Richter, A. J. Fehske, and G. P. Fettweis, “Energy Efficiency Aspects of Base Station Deployment Strategies for Cellular Networks,” in *2009 IEEE 70th Vehicular Technology Conference Fall*, 2009, pp. 1–5.

[85] P. A. García, A. A. González, A. A. Alonso, B. C. Martínez, J. M. A. Pérez, and A. S. Esguevillas, “Automatic UMTS system resource dimensioning based on service traffic analysis,” *EURASIP Journal on Wireless Communications and Networking*, October 2012.

[86] 3rd Generation Partnership Project. (ETSI, 2016) 3GPP TS 36.213 V13.0.0. Accessed 5-March-2020. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/136200_136299/136213/13.00.00_60/ts_136213v130000p.pdf

[87] A. Ghosh and R. Ratasuk, *Essentials of LTE and LTE-A (The Cambridge Wireless Essentials Series)*, 1st ed., ser. The Cambridge Wireless Essentials Series. Cambridge University Press, 2011.

[88] 3rd Generation Partnership Project. (ETSI, 2009) 3GPP TS 25.214 V8.6.0. Accessed 7-March-2020. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/125200_125299/125214/08.06.00_60/ts_125214v080600p.pdf

A

SINR Mapping Tables

These presented tables allow to map the TBS index, the CQI and SINR values.

I_{TBS}	N_{PRB}									
	1	2	3	4	5	6	7	8	9	10
0	16	32	56	88	120	152	176	200	232	248
1	24	48	88	120	160	200	232	272	304	344
2	32	72	120	160	200	248	296	336	376	424
3	40	104	152	208	272	320	392	440	504	568
4	48	120	200	264	320	408	488	552	632	696
5	72	152	232	320	424	504	600	680	776	872
6	320	176	288	392	504	600	712	808	936	1032
7	104	232	320	472	584	712	840	968	1096	1224
8	120	248	392	536	680	808	968	1096	1256	1384
9	136	296	456	616	776	936	1096	1256	1416	1544
10	152	320	504	680	872	1032	1224	1384	1544	1736
11	176	376	584	776	1000	1192	1384	1608	1800	2024
12	208	440	680	904	1128	1352	1608	1800	2024	2280
13	232	488	744	1000	1256	1544	1800	2024	2280	2536
14	264	552	840	1128	1416	1736	1992	2280	2600	2856
15	280	600	904	1224	1544	1800	2152	2472	2728	3112
16	320	632	968	1288	1608	1928	2280	2600	2984	3240
17	336	696	1064	1416	1800	2152	2536	2856	3240	3624
18	376	776	1160	1544	1992	2344	2792	3112	3624	4008
19	408	840	1288	1736	2152	2600	2984	3496	3880	4264
20	440	904	1384	1864	2344	2792	3240	3752	4136	4584
21	488	1000	1480	1992	2472	2984	3496	4008	4584	4968
22	520	1064	1608	2152	2664	3240	3752	4264	4776	5352
23	552	1128	1736	2280	2856	3496	4008	4584	5160	5736
24	584	1192	1800	2408	2984	3624	4264	4968	5544	5992
25	616	1256	1864	2536	3112	3752	4392	5160	5736	6200
26	648	1320	1992	2664	3368	4008	4584	5352	5992	6712

Figure A.1: LTE TBS index mapping table (for 2 PRBs adapted from [86]).

I_{TBS}	N_{PRB}									
	41	42	43	44	45	46	47	48	49	50
0	1128	1160	1192	1224	1256	1256	1288	1320	1352	1384
1	1480	1544	1544	1608	1608	1672	1736	1736	1800	1800
2	1800	1864	1928	1992	2024	2088	2088	2152	2216	2216
3	2408	2472	2536	2536	2600	2664	2728	2792	2856	2856
4	2984	2984	3112	3112	3240	3240	3368	3496	3496	3624
5	3624	3752	3752	3880	4008	4008	4136	4264	4392	4392
6	4264	4392	4584	4584	4776	4776	4968	4968	5160	5160
7	4968	5160	5352	5352	5544	5736	5736	5992	5992	6200
8	5736	5992	5992	6200	6200	6456	6456	6712	6968	6968
9	6456	6712	6712	6968	6968	7224	7480	7480	7736	7992
10	7224	7480	7480	7736	7992	7992	8248	8504	8504	8760
11	8248	8504	8760	8760	9144	9144	9528	9528	9912	9912
12	9528	9528	9912	9912	10296	10680	10680	11064	11064	11448
13	10680	10680	11064	11448	11448	11832	12216	12216	12576	12960
14	11832	12216	12216	12576	12960	12960	13536	13536	14112	14112
15	12576	12960	12960	13536	13536	14112	14688	14688	15264	15264
16	13536	13536	14112	14112	14688	14688	15264	15840	15840	16416
17	14688	15264	15264	15840	16416	16416	16992	17568	17568	18336
18	16416	16416	16992	17568	17568	18336	18336	19080	19080	19848
19	17568	18336	18336	19080	19080	19848	20616	20616	21384	21384
20	19080	19848	19848	20616	20616	21384	22152	22152	22920	22920
21	20616	21384	21384	22152	22920	22920	23688	24496	24496	25456
22	22152	22920	22920	23688	24496	24496	25456	25456	26416	27376
23	23688	24496	24496	25456	25456	26416	27376	27376	28336	28336
24	25456	25456	26416	26416	27376	28336	28336	29296	29296	30576
25	26416	26416	27376	28336	28336	29296	29296	30576	31704	31704
26	27376	27376	28336	29296	29296	30576	31704	31704	32856	32856

Figure A.2: LTE TBS index mapping table (for 50 PRBs adapted from [86]).

I_{TBS}	N_{PRB}									
	91	92	93	94	95	96	97	98	99	100
0	2536	2536	2600	2600	2664	2664	2728	2728	2728	2792
1	3368	3368	3368	3496	3496	3496	3496	3624	3624	3624
2	4136	4136	4136	4264	4264	4264	4392	4392	4392	4584
3	5352	5352	5352	5544	5544	5544	5736	5736	5736	5736
4	6456	6456	6712	6712	6712	6968	6968	6968	6968	7224
5	7992	7992	8248	8248	8248	8504	8504	8760	8760	8760
6	9528	9528	9528	9912	9912	9912	10296	10296	10296	10296
7	11064	11448	11448	11448	11448	11832	11832	11832	12216	12216
8	12576	12960	12960	12960	13536	13536	13536	13536	14112	14112
9	14112	14688	14688	14688	15264	15264	15264	15264	15840	15840
10	15840	16416	16416	16416	16992	16992	16992	16992	17568	17568
11	18336	18336	19080	19080	19080	19080	19848	19848	19848	19848
12	20616	21384	21384	21384	21384	22152	22152	22152	22920	22920
13	23688	23688	23688	24496	24496	24496	25456	25456	25456	25456
14	26416	26416	26416	27376	27376	27376	28336	28336	28336	28336
15	28336	28336	28336	29296	29296	29296	29296	30576	30576	30576
16	29296	30576	30576	30576	30576	31704	31704	31704	31704	32856
17	32856	32856	34008	34008	34008	35160	35160	35160	35160	36696
18	36696	36696	36696	37888	37888	37888	37888	39232	39232	39232
19	39232	39232	40576	40576	40576	40576	42368	42368	42368	43816
20	42368	42368	43816	43816	43816	45352	45352	45352	46888	46888
21	45352	46888	46888	46888	46888	48936	48936	48936	48936	51024
22	48936	48936	51024	51024	51024	51024	52752	52752	52752	55056
23	52752	52752	52752	55056	55056	55056	55056	57336	57336	57336
24	55056	57336	57336	57336	57336	59256	59256	59256	61664	61664
25	57336	59256	59256	59256	61664	61664	61664	61664	63776	63776
26	59256	61664	61664	61664	63776	63776	63776	63776	66592	75376

Figure A.3: LTE TBS index mapping table (for 100 PRBs adapted from [86]).

Modulation and TBS index table for PDSCH					Modulation and TBS index table 2 for PDSCH				
I_{MCS}	Modulação	Ordem Mod. (Q_m)	I_{TBS}	CQI	I_{MCS}	Modulação	Ordem Mod. (Q_m)	I_{TBS}	CQI
0	QPSK	2	0		0	QPSK	2	0	
1		2	1	1	1		2	2	1
2		2	2		2		2	4	2
3		2	3	2	3		2	6	3
4		2	4		4		2	8	4
5		2	5	3	5	16QAM	4	10	
6		2	6		6		4	11	
7		2	7	4	7		4	12	5
8		2	8		8		4	13	
9		2	9	5	9		4	14	6
10	16QAM	4	9		10		4	15	
11		4	10	6	11	64QAM	6	16	7
12		4	11		12		6	17	
13		4	12	7	13		6	18	8
14		4	13		14		6	19	
15		4	14	8	15		6	20	9
16		4	15		16		6	21	
17	64QAM	6	15	9	17		6	22	10
18		6	16		18		6	23	
19		6	17	10	19		6	24	11
20		6	18		20	256QAM	8	25	
21		6	19	11	21		8	27	12
22		6	20		22		8	28	
23		6	21	12	23		8	29	13
24		6	22		24		8	30	
25		6	23	13	25		8	31	14
26		6	24		26		8	32	
27		6	25	14	27		8	33/33A	15
28		6	26/26A	15	28	QPSK	2	Reserv.	Reserv.
29	QPSK	2	Reserv.	Reserv.	29	16QAM	4		
30	16QAM	4			30	64QAM	6		
31	64QAM	6			31	256QAM	8		

Figure A.4: CQI mapping table for LTE (adapted from [86]).

CQI index	Modulation	Spectral efficiency (bps/Hz)	Reference SINR (dB)
0	Out of range		
1	QPSK	0.15	−6.7
2	QPSK	0.23	−4.7
3	QPSK	0.38	−2.3
4	QPSK	0.60	0.2
5	QPSK	0.88	2.4
6	QPSK	1.18	4.3
7	16-QAM	1.48	5.9
8	16-QAM	1.91	8.1
9	16-QAM	2.41	10.3
10	64-QAM	2.73	11.7
11	64-QAM	3.32	14.1
12	64-QAM	3.90	16.3
13	64-QAM	4.52	18.7
14	64-QAM	5.12	21.0
15	64-QAM	5.55	22.7

Figure A.5: SINR mapping table for LTE [87].

Category	Used CQI mapping table					
	MIMO not configured		MIMO configured			
	64QAM not configured	64QAM configured	64QAM not configured		64QAM configured	
			In case of type B or single transport block type A CQI reports	In case of dual transport block type A CQI reports	In case of type B or single transport block type A CQI reports	In case of dual transport block type A CQI reports
1-6	A	N/A				
7 and 8	B	N/A				
9	C	N/A				
10	D	N/A				
11 and 12	E	N/A				
13	C	F	N/A			
14	D	G	N/A			
15	C	N/A	C	H	N/A	
16	D	N/A	D	I	N/A	
17	C	F	C	H	N/A	
18	D	G	D	I	N/A	
19	C	F	C	H	F	J
20	D	G	D	I	G	K
21	C	N/A				
22	D	N/A				
23	C	F	N/A			
24	D	G	N/A			

Figure A.6: Applicability of CQI mapping tables [88].

CQI value	Transport Block Size	Number of HS-PDSCH	Modulation	Reference power adjustment Δ	N_{IR}	X_{RV}
0	N/A	Out of range				
1	136	1	QPSK	0	43200	0
2	176	1	QPSK	0		
3	232	1	QPSK	0		
4	320	1	QPSK	0		
5	376	1	QPSK	0		
6	464	1	QPSK	0		
7	648	2	QPSK	0		
8	792	2	QPSK	0		
9	928	2	QPSK	0		
10	1264	3	QPSK	0		
11	1488	3	QPSK	0		
12	1744	3	QPSK	0		
13	2288	4	QPSK	0		
14	2592	4	QPSK	0		
15	3328	5	QPSK	0		
16	3576	5	16-QAM	0		
17	4200	5	16-QAM	0		
18	4672	5	16-QAM	0		
19	5296	5	16-QAM	0		
20	5896	5	16-QAM	0		
21	6568	5	16-QAM	0		
22	7184	5	16-QAM	0		
23	9736	7	16-QAM	0		
24	11432	8	16-QAM	0		
25	14424	10	16-QAM	0		
26	15776	10	64-QAM	0		
27	21768	12	64-QAM	0		
28	26504	13	64-QAM	0		
29	32264	14	64-QAM	0		
30	38576	15	64-QAM	0		

Figure A.7: CQI mapping table G [88].