



Analysis of the Implementation of Private Networks in 5G Using Network Slicing

Marco Antonio Sá Failache Filho

Thesis to obtain the Master of Science Degree in

Electrical and Computer Engineering

Supervisor: Prof. Luís Manuel de Jesus Sousa Correia

Examination Committee

Chairperson: Prof. José Eduardo Charters Ribeiro da Cunha Sanguino

Supervisor: Prof. Luís Manuel de Jesus Sousa Correia

Members of Committee: Prof. António José Castelo Branco Rodrigues

Eng. Bruno Miguel Gaspar

July 2022

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

To my family and friends.

Acknowledgements

Firstly, I would like to express my great appreciation and gratitude to Professor Luís M. Correia. For all of the meetings through the whole period of developing the thesis. For the opportunity of working with NOS and getting in touch with the working environment of telecommunications. For understanding and giving all the possible support during the difficult times. For all of the knowledge shared and taught, not only on the telecommunications field, but also for life.

To Eng. Bruno Gaspar, from NOS, for all the support, advices and information. Also, for the opportunity of working closely with the industry and understanding their needs and priorities.

To all GROW members, especially to my partners Beatriz Ferreira, Sofia Patrício, Afonso Carvalho and Maria Ferreira, for always being there for the meetings and contributing for the good environment.

To my family, my mother, Diana Failache, father, Marco Failache, sister, Hanna Failache, brother, Renzo Failache, brother-in-law, Daniel Barclay for all the love and emotional support. For also being with me since the beginning and always motivating me to achieve my goals.

To my friend Emanuel Pascoal, for the partnerships during the courses and for all the friendship that helped me through the whole engineering course.

To my friends, António Souto, Daniel Ascen, David Caldeira, David Estrela, Duarte Baltazar, Lucas Philipovsky and Mário Nobre, for all the good times and friendship that helped to ease the process, and for always believing in me.

To my girlfriend, Emily Amaral, for the love and support during the final period and during difficult times.

Abstract

The goal of this thesis was to provide an analysis of 5G systems' performance in Private Networks, using Network Slicing. A model was developed and adapted to simulate 5G data rates and latency, consisting of two stages, one for data rate calculation and allocation and another for latency calculation and analysis. Several Input Parameters are considered for both stages, most of the variations having been done in the data rate stage, as it is the focus of this work. MIMO Layers and the Service Mix were varied to analyse their influence. The model main outputs are the Service's Data Rates, the VRRM Capacity Share, User's Satisfaction and the Network Delay. Different Private Network scenarios were defined and deployed, each of them with services specific for their sector. Mission Critical, Healthcare and Smart Factory sectors were considered. The Private Network analysis look into 5G sub-6GHz licensed bands and the unlicensed 5.9 GHz band. In addition to that, three different Private Network Architectures were considered, i.e., Network Slicing Deployment, Shared RAN Deployment and Standalone Deployment. The results show that the SLAs with highest priority are always the last to be delayed and, when served, have at least the minimum data rate defined. It is also possible to analyse the Shared RAN performance for the Hospital and Smart Factory scenarios, results showing that it is possible to serve all users in both scenarios in a satisfactory manner with only a percentage of the total available capacity.

Keywords

5G, Private Networks, Data Rate, Network Slicing, Latency, SLA.

Resumo

O objetivo desta tese foi de fornecer uma análise do desempenho de sistemas 5G em Redes Privadas, utilizando *Network Slicing*, tendo-se desenvolvido e adaptado um modelo para simular as taxas de dados e latência em 5G. O modelo consiste em duas etapas, uma para cálculo e alocação da taxa de dados e outra para cálculo e análise de latência. Vários Parâmetros de Entrada são considerados para ambas as etapas, a maioria das variações tendo sido feitas na etapa de taxa de dados, pois é o foco deste trabalho. As Camadas MIMO e o Mistura de Serviços foram variadas para analisar sua influência. As principais saídas do modelo são as Taxas de Dados do Serviço, o Compartilhamento de Capacidade do VRRM, a Satisfação do Utilizador e o Atraso da Rede. Diferentes cenários de Rede Privada foram definidos e implantados, cada um deles com serviços específicos para o seu setor. Foram considerados os setores de Missão Crítica, Saúde e Fábrica Inteligente. A análise da rede privada de 5G focou-se nas bandas licenciadas sub-6GHz e não licenciada de 5,9 GHz. Além disso, três diferentes arquiteturas de rede privada foram consideradas: *Network Slicing Deployment*, *Shared RAN Deployment* e *Standalone Deployment*. Os resultados mostram que os SLAs com maior prioridade são sempre os últimos a serem atrasados e, quando atendidos, possuem pelo menos a taxa de dados mínima definida. Também foi possível analisar o desempenho da RAN Compartilhada para os cenários Hospitalar e Fábrica Inteligente, mostrando os resultados que é possível atender todos os utilizadores em ambos os cenários de forma satisfatória com apenas uma certa percentagem da capacidade total disponível.

Palavras-chave

5G, Redes Privadas, Ritmo de Transmissão, *Network Slicing*, Latência, SLA.

Table of Contents

Abstract.....	ix
Resumo.....	x
Table of Contents.....	xi
List of Figures.....	xiii
List of Tables.....	xvi
List of Acronyms.....	xviii
List of Symbols.....	xxi
Chapter 1 Introduction	1
1.1. Overview	2
1.2. Motivation and Contents.....	3
Chapter 2 Fundamental Concepts	5
2.1 Network Architecture	6
2.2 Radio Interface.....	8
2.3 Network Aspects	11
2.3.1 Virtualisation and Slicing Architecture	11
2.3.2 Cloud and Edge Networking.....	13
2.4 Services and Applications	16
2.5 Private Networks Architectures	19
2.6 State of the Art	20
Chapter 3 Model and Simulation.....	23
3.1. Model Overview	24
3.2. Cell Capacity.....	27
3.3. Admission Control and Delay Process	29
3.4. VRRM Optimisation.....	30
3.5. Aggregation Process.....	32
3.5.1. Latency	32
3.5.2. Node Processing Power	35
3.5.3. Aggregation Process Workflow	37
3.6. Bandwidth	38
3.8. User Experience and Network Performance.....	40
3.9. Model Assessment.....	41
Chapter 4 Results Analysis.....	47
4.1. Reference Scenarios.....	48

4.2. Hospital Scenario Analysis	50
4.2.1. Hospital Network Slicing Results	50
4.2.2. Hospital Shared Ran Results	54
4.2.3. Hospital Standalone Results	56
4.2.4. Mix and MIMO Layers Variation Results.....	57
4.2.5. Hospital Latency Results	60
4.3. Smart Factory Scenario Analysis.....	61
4.3.1. Smart Factory Network Slicing Results.....	61
4.3.2. Smart Factory Shared Ran Results	63
4.3.3. Smart Factory Standalone Results	64
4.3.4. Smart Factory Latency Results	65
4.4. Mission Critical Scenario Analysis	66
4.4.1. Mission Critical 700 MHz Band Analysis.....	66
4.4.2. Mission Critical 3600 MHz Band Analysis.....	69
4.4.3. Mission Critical Latency Results	71
Chapter 5 Conclusions	75
Annex A 5G QoS Characteristics	81
Annex B Numerology Selection Algorithm	84
Annex C Military Applications Key Performance Indicators.....	86
Annex D Scenarios Development	88
Annex E Additional Results	92
References.....	101

List of Figures

Figure 1.1 - 5G use cases and their requirements (extracted from [FaPE17]).	2
Figure 1.2 - Network Slices under the same physical infrastructure (extracted from [OIFa17]).	3
Figure 2.1 - Basic EPS architecture (extracted from [SaKn18]).	6
Figure 2.2 - 5G System Architecture (extracted from [HUAW17]).	8
Figure 2.3 - 3GPP UE Connectivity Options (extracted from [5GAm18]).	8
Figure 2.4 - Resource blocks for two subcarrier spacings (extracted from [DaPS18]).	9
Figure 2.5 - NR-U Deployment Scenarios (extracted from [GMBC19]).	10
Figure 2.6 - ONF SDN Architecture (extracted from [YBSS17]).	12
Figure 2.7 - Network Slicing Layers Framework (extracted from [FPEM17]).	12
Figure 2.8 - Cloud and Edge Computing models (extracted from [HaYW19]).	14
Figure 2.9 - C-RAN Architecture (extracted from [HNHS19]).	15
Figure 2.10 - MEC and 5G integrated framework (extracted from [PFHP20]).	15
Figure 2.11 - 5G Private Networks Functional Architecture (extracted from [Aija20]).	20
Figure 3.1 - General Model Workflow.	26
Figure 3.2 - Allowed SCSs for each 5G scenario (extracted from [VaMA21]).	27
Figure 3.3 - Admission Control and Delay Process workflow.	30
Figure 3.4. Delay contributions on the network (extracted from [Domi19]).	33
Figure 3.5 - Aggregation Process Workflow.	38
Figure 3.6 - Cell Data Rate with possible MIMO Layers.	43
Figure 3.7 - VRRM Capacity Share with the different MIMO Layers possibilities.	43
Figure 3.8 - Numerology Selected according to the Scenario Topology.	44
Figure 3.9 - Cell Data Rate response to Bandwidth variation.	44
Figure 3.10 – VRRM Capacity Share with Remote Surgery of users' variation.	45
Figure 4.1 - Service's Tuning Weight.	49
Figure. 4.2. Services Data rate with different number of users for the Hospital Network Slicing Private Deployment.	51

Figure. 4.3. VRRM Capacity Share with different number of users for the Hospital Network Slicing Private Deployment.	52
Figure. 4.4. Cell's used capacity versus the total number of users for the Hospital Network Slicing Private Deployment.	52
Figure. 4.5. Each service's served users versus the total number of users for the Hospital Network Slicing Private Deployment.....	53
Figure 4.6. Service's Data Rates versus Shared RAN Percentage.	55
Figure 4.7. Service's User Satisfaction versus Shared Ran Percentage.	55
Figure 4.8. VRRM Capacity Share versus Shared RAN percentage for the Hospital Shared RAN scenario.	56
Figure 4.9. Service's Data Rate for both 3.5 GHz and 5.9 GHz frequency bands.	57
Figure 4.10. Service's Data Rates versus MIMO Layers variation for the Hospital Network Slicing Private Deployment.	58
Figure 4.11. Service's Data Rates versus Mix Variation for the Hospital Network Slicing Private Deployment.	59
Figure 4.12. RU sites provided by NOS on the Hospital Scenario.....	60
Figure 4.13. Data rate for Smart Factory Network Slicing Private Deployment Scenario.....	62
Figure 4.14. VRRM Capacity Share versus the evolution of Users.	62
Figure 4.15. Service's data rate evolution with Shared RAN percentage increments for the Smart Factory Shared Ran scenario.....	64
Figure 4.16. Service's data rate on both 3.5 GHz and 5.9 GHz frequency bands for the Smart Factory Standalone deployment.....	65
Figure 4.17. RU sites provided by NOS for the Smart Factory Scenario.....	66
Figure 4.18. Data Rate for Mission Critical 700 MHz scenario.	68
Figure 4.19. VRRM Capacity Share for Mission Critical Scenario at the 700 MHz band.	68
Figure 4.20. User's Satisfaction for Mission Critical Scenario at 710 MHz.	69
Figure 4.21. Data rate for Mission Critical Scenario at 3600 MHz.	70
Figure 4.22. VRRM Capacity Share versus the Number of Users for the 3.6 GHz Mission Critical Scenario.....	71
Figure 4.23. RU sites provided by NOS.for the 700 MHz Mission Critical Scenario at Pedrogão Grande.....	72
Figure E.1.Each service's number of users versus the total number of users for the Hospital Network Slicing Private Deployment.....	93

Figure E.2 GBR VNO User's Satisfaction versus the number of users for the Hospital Network Slicing Private Deployment.	93
Figure E.3. Non-GBR VNO User's Satisfaction versus the Number of users for the Hospital Network Slicing Private Deployment.....	94
Figure E.4. GBR VNO User's Satisfaction for Smart Factory Network Slicing Private Deployment Scenario.....	94
Figure E.5. Non-GBR VNO User's Satisfaction for Smart Factory Network Slicing Private Deployment Scenario.....	95
Figure E.6. DelayC-GBR User's Satisfaction for Smart Factory Network Slicing Private Deployment Scenario.....	95
Figure E.7. User's Satisfaction versus the Number of Users for the 3.6GHz Mission Critical Scenario.	96
Figure E.8. User's Satisfaction for both 3.5 GHz and 5.9 GHz frequency bands.	96
Figure E.9. User's Satisfaction for both 3.5 GHz and 5.9 GHz frequency bands	97
Figure E.10. User's Satisfaction versus Shared RAN percentage for the Smart Factory Shared RAN scenario.	97
Figure E.12. VRRM Capacity Share versus Mix Variation for the Hospital Network Slicing Private Deployment.	98
Figure E.13. RU sites provided by NOS for the 3600 MHz Mission Critical Scenario at city of Porto..	99

List of Tables

Table 2.1 - Key performance requirements for healthcare applications (adapted from [SFRD17]).	17
Table 2.2 - Key performance requirements for energy related services (adapted from [GSMA18]).	18
Table 2.3 - Railway networks key performance requirements (adapted from [GMVA17]).	18
Table 2.4 - Industrial applications key performance requirements (adapted from [Aija20] and [3GPP19]).	19
Table 3.1 - Input and Output Parameters, along with model stages.	24
Table 3.2 – Modulation schemes and maximum code rate.	28
Table 3.3 – RB per Numerology per Bandwidth for Frequency Range 1 [3GPP17].	29
Table 3.4 – Available frequency bands.	39
Table 3.5 – Scaling Factor for Delay Spread calculation for the frequencies and topologies considered (adapted from [VaMA21]).	39
Table 3.6 – Voice codecs and respective MOS [Dini21].	41
Table 3.7 – List of model assessment tests.	42
Table 3.8 – Cell Input Parameters.	42
Table 3.9 – Service Input Parameters.	42
Table 3.10 – Latency Results.	46
Table 4.1 – Services with Mix, SLA and VRRM parameters.	48
Table 4.2. Service's Data Rates.	49
Table 4.3. Network Input Parameters.	50
Table 4.4. Cell Input Parameters for the Hospital Network Slicing scenario.	51
Table 4.5. Service Mix for the Shared Ran Hospital Scenario.	54
Table 4.6. Different Service Mixes for Mix Variation Analysis.	58
Table 4.7. Hospital Scenario Delay Results.	60
Table 4.8 Smart Factory Shared RAN Scenario Mix.	63
Table 4.9. Smart Factory Scenario Delay Results.	66
Table 4.10. Cell Input parameters for Mission Critical Scenario.	67

Table 4.11. Service Mix for the 700 MHz band Mission Critical Scenario.	67
Table 4.12. Service Mix for the 3600 MHz band Mission Critical Scenario.	70
Table 4.13. Mission Critical Rural Scenario Delay Results.	72
Table 4.14. Mission Critical Urban Scenario Delay Results.	72

List of Acronyms

3GPP	3 rd Generation Partnership Project
4G	4 th Generation
5G	5 th Generation
5GC	5G Core
AF	Application Function
AMF	Access and Mobility Management Function
AR	Augmented Reality
AUSF	Authentication Server Function
BBU	Baseband Unit
BS	Base Station
CAC	Connection Admission Control
CAPEX	Capital Expenditures
CAGR	Compound Annual Growth Rate
CCA	Clear Channel Assessment
CN	Core Network
CNF	Core Network Functions
CP	Cyclic Prefix
CPU	Central Processing Unit
C-RAN	Cloud Radio Access Network
CU	Centralized Unit
DelayC	Delay Critical
DDoS	Distributed Denial-of-Service
DL	Downlink
DU	Distributed Unit
eMBB	Enhanced Mobile Broadband
eNodeB	Evolved Node B
EPC	Evolved Packet Core Network
EPS	Evolved Packet System
ETSI	European Telecommunications Standard Institute
FDD	Frequency Division Duplex
FS	File Sharing
GBR	Guaranteed Bit Rate

gNodeB	Generalised NodeB
GSM-R	Global System for Mobile - Railways
HSS	Home Subscription Server
ICI	Inter Carrier Interference
IMS	IP Multimedia Subsystem
IMT-2020	International Mobile Telecommunications 2020
IoT	Internet of Things
IO	Indoor Office
IP	Internet Protocol
ISM	Industry, Scientific and Medical
ISI	Inter Symbol Interference
ITU-R	International Telecommunication Union Radiocommunication Sector
LAA	Licensed Assisted Access
LBT	Listen-Before-Talk
LTE	Long Term Evolution
MCPTT	Mission Critical Push to Talk
MEC	Multi-Access Edge Computing
MECO	MEC Orchestrator
MIMO	Multiple Input Multiple Output
MILP	Mixed Linear Integer Programming
MME	Mobility Management Entity
mMTC	massive Machine-Type Communications
MNO	Mobile Network Operators
NEF	Network Exposure Function
NIC	NATO Communications and Information
NFV	Network Function Virtualisation
NR	New Radio
NRF	Network function Repository Function
NR-U	New Radio Unlicensed
NSA	Non-Standalone
NSSF	Network Slice Selection Function
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
ONF	Open Networking Foundation
OPEX	Operating Expenditures
O2I	Outdoor to Indoor
PCF	Policy Control Function
PDN-GW	Packet Data Network Gateway
QAM	Quadrature Amplitude Modulation
QoE	Quality of Experience

QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RAN	Radio Access Network
RB	Resource blocks
RE	Resource Element
RNF	Radio Network Functions
RRH	Remote Radio Head
RS	Remote Surgery
RU	Radio Unit
SA	Standalone
SBA	Service-Based Architecture
SC-FDMA	Single Carrier Frequency Division Multiple Access
SDN	Software-Defined Networks
SFI	Slot Format Indicator
S-GW	Serving Gateway
SLA	Service Level Agreement
SMF	Session Management Function
SN	Social Networking
SNR	Signal to Noise Ratio
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TETRA	Terrestrial Trunked Radio
TPMA	Third Party Monitoring Application
UE	User Equipment
UL	Uplink
UDM	Unified Data Management
UPF	User Plane Function
URLLC	Ultra-Reliable and Low Latency Communications
Uma	Urban Macro
Umi	Urban Micro
VM	Virtual Machine
VNF	Virtual Network Function
VNO	Virtual Network Operator
VR	Virtual Reality
VRRM	Virtual Radio Resources Management
WB	Web Browsing

List of Symbols

δ_{App}	Maximum latency depending on what application is chosen
$\delta_{\text{BH,C}}$	Backhaul to core transmission Latency
$\delta_{\text{BH,MEC}}$	Backhaul to MEC transmission Latency
δ_{Cor}	Core processing delay
$\delta_{\text{C-RAN}}$	C-RAN associated Latency
$\delta_{\text{CU,DL}}$	CU DL processing delay
$\delta_{\text{CU,UL}}$	CU UL processing delay
$\delta_{\text{DU,DL}}$	DU DL processing delay
$\delta_{\text{DU,UL}}$	DU UL processing delay
δ_{E2E}	End to End Latency
δ_{EN}	External Data centre contribution delay
δ_{FH}	Transmission delay between the RU to the DU
δ_{HARQ}	HARQ protocol requirement latency
$\delta_{\text{MEC,DL}}$	MEC DL processing delay
$\delta_{\text{MEC,UL}}$	MEC UL processing delay
δ_{MH}	Transmission delay between the DU to the CU
$\delta_{\text{Node,que}}$	Queuing delay on the RU
δ_{Node}	Processing delay on the Node
$\delta_{\text{RU,DL}}$	RU DL processing delay
$\delta_{\text{RU,UL}}$	RU UL processing delay
$\delta_{\text{Node,proc}}$	BS function processing delay on the RU
δ_{Tran}	Transport transmission delay from the core to the external data centres
γ_v	Priority defined by InP and assigned to VNO v
δ_s	Serving weight, assigned to service s
λ_{v_s}	Tuning weight associated with service s , provided by VNO v , to prioritise data rate assignment
μ_L	Lagrange multiplier corresponding to the inequality constraint

d_{BH}	Backhaul distance
d_{E2E}	Maximum E2E distance
d_f	Fibre link distance
d_{FH}	Fronthaul maximum distance
d_{MH}	Middlehaul maximum distance
d_{Tran}	Distance between the core and external data centre
E	Spectral efficiency
E_{ref}	Reference spectral efficiency
$F_{DC,ref}$	Reference system load in the frequency-domain
F_{DC}	System load in the frequency-domain
$f^{(j)}$	Scaling factor
J	Number of aggregated component carriers in a band or band combination
M_{UMIMO}	Multi-user MIMO factor
$N_{v_s}^{usr}$	Number of users performing service s , from VNO v
$N_{A,ref}$	Reference number of antennas in the BS
N_A	Number of antennas in the BS
$N_{PRB}^{BW(j),\mu}$	Maximum number of RB allocation in bandwidth $BW^{(j)}$ with numerology μ
N^{srv}	Number of services
$N_{streams,ref}$	Reference number of transmission streams
$N_{streams}$	Number of transmission streams
N_{SY}	Number of symbols
$N_{U,max}$	Maximum number of users in the network
$O_h^{(j)}$	Overhead
$Q_m^{(j)}$	Maximum supported modulation order
$p_{v_s[\%]}^{usr_{net}}$	Percentage of served users
$p_{VRRM[\%]}^{tot}$	Percentage of total assigned data rate
P_{BBm}	Baseband modulation/demodulation processing component
P_{BH}	Processing power required for the backhaul interface

P_{CN}	Processing power used by the CN
P_{Code}	FEC function processing component
P_{CU}	Processing power used by the CU
P_{DU}	Processing power used by the DU
P_{MAC}	Processing power required for the MAC layer
P_{Map}	Mapping and demapping functions processing component
P_{MEC}	Processing power used by the MEC
P_{MIMO}	MIMO encoding/decoding processing component
P_{Node}	Processing power on the aggregator node
$P_{Node,Cap}$	Processing capacity assign to the aggregator node
$P_{Node,Cap,ref}$	Reference processing capacity on the node
$P_{Node,fix}$	Fixed processing power required for scheduling and signalling
P_{OFDM}	Frequency domains function for OFDM modulation processing component
$R_{v_s,i}^{usr}$ [Mbps]	Data rate of each user
$R_{v_s}^{srv_{tot}}$ [Mbps]	Total data rate of each service
$R_{VRRM}^{VNO_v}$ [%]	VRRM capacity share
R_{cell} [Mbps]	Total available data rate of the cell
R_{max}	Maximum coding rate
$R_{vs}^{srv_{max}}$ [Mbps]	Maximum assignable data rate to the user of service s , from VNO v
$R_{vs}^{srv_{min}}$ [Mbps]	Minimum assignable data rate to the user of service s , from VNO v
R_{vs}^{srv} [Mbps]	Total served data rate of service v_s
\mathbf{R}^{srv}	Vector of serving data rates
$S_{v_s,i}^{usr}$	Users' satisfaction
T_s^μ	Average OFDM symbol duration in a subframe for numerology μ
$v_{Layers}^{(j)}$	Maximum number of supported MIMO layers
$w_{v_s,i}^{usr}$	Assigned weight to user i , performing service s , from VNO v
\mathbf{w}^{usr}	Vector of users' weights, to obtain the long-term average data rate of users

Chapter 1

Introduction

This chapter presents a brief overview of the Thesis. It starts with 5G and Private Network Aspects. Then, the motivation for the work is presented, ending with the structure and contents of the thesis.

1.1. Overview

Mobile devices have become essential to today's society way of living, in a way that mobile networks have become critical and are in constant evolution. New services are being deployed and developed, such as Industrial Internet of Things, Virtual and Augmented Reality, Mobile Online Gaming, Remote Surgery etc. Aiming to attend the new demands of performance imposed by these services, such as high data rates, low latency, and massive capacity the fifth generation (5G) of mobile networks is bringing new concepts. It has a service-oriented view of development, that has a goal of delivering several services with different Service Level Agreements (SLA) and top-notch Quality of Service (QoS) under the same physical infrastructure.

In addition to that, one can define three broad use case families: enhanced mobile broadband, massive machine-type communications and critical communications. They all have very different requirements but will be deployed under the same physical network; it shows how important network flexibility is. Figure 1.1 illustrates these use cases and their requirements, the further the distance of a requirement from the centre, the more important it is to the corresponding use case.

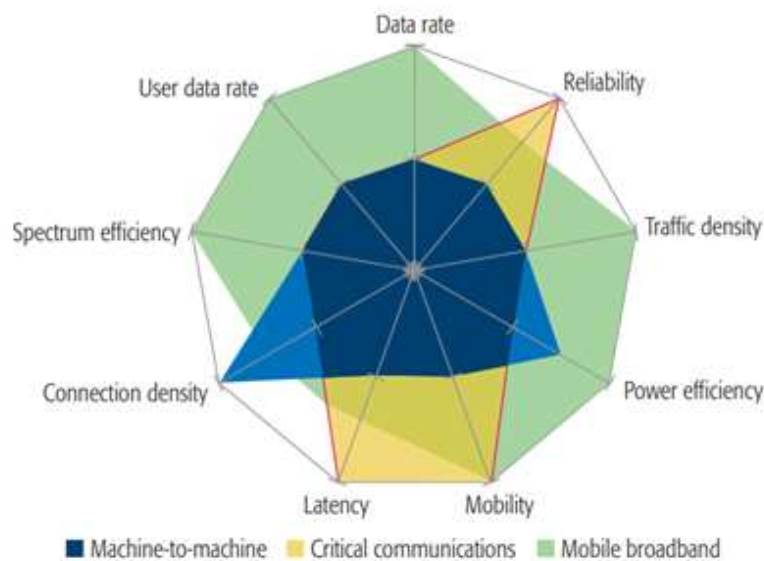


Figure 1.1 - 5G use cases and their requirements (extracted from [FaPE17]).

To attend all these different use cases concurrently over the same physical network, 5G is bringing a new feature called Network Slicing, that will be deployed by means of Network Function Virtualisation (NFV), Software Defined Networking (SDN), cloud computing and edge computing [FaPE17]. Figure 1.2 illustrates network slices running over a single physical infrastructure.

The shared physical infrastructure includes Radio Access Network (RAN), edge and cloud computing servers, satellites, Wi-Fi access points and other radio access technologies. NFV allows the use of generic hardware for cost-effective implementation of different network functions, and SDN offers separation of control plane from the data plane, thus enabling easier network management [KYTH20].

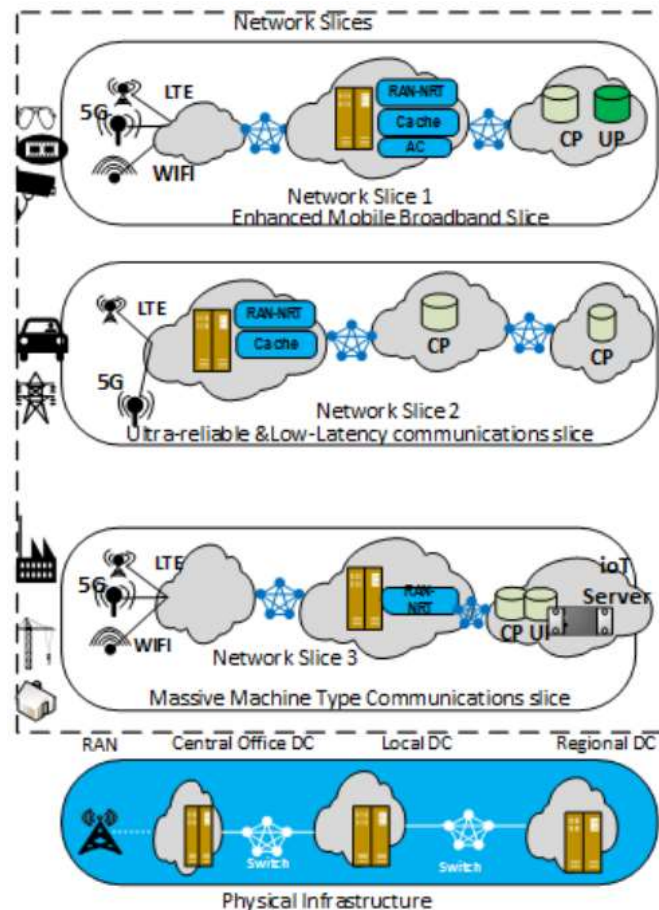


Figure 1.2 - Network Slices under the same physical infrastructure (extracted from [OIFa17]).

Private mobile networks are radio systems for exclusive use of their owners, like the military, emergency services, transport companies, etc. [Corr20]. In Europe, one has the Terrestrial Trunked Radio (TETRA) standard that sets the SIRESP network owned by the Portuguese Government and responsible for the national emergency and security communications. Nowadays, with the growing presence of industry applications of mobile networks, such as Industrial IoT, the implementation of private 5G networks will soon become a reality.

Various reports suggest the growing trend of private 5G market, one of them forecasts a growth at approximately 30% CAGR between 2018 and 2021 to more than \$5 billion by the end of 2021. SIRESP is an example of a Standalone Private Network with its own structure and separated from the operations of the public network. 5G Private Networks have the possibility of being deployed as public network integrated private network, through the network slicing feature, being a novel and profitable product [PoOM20].

1.2. Motivation and Contents

5G will enable the appearance of private networks running over the infrastructure of licenced

operators, through the network slicing feature. These networks may have heterogeneous requirements, depending on the specific service they will provide. Therefore, there is a challenge to ensure that the SLAs are met and that the networks are properly isolated from the public ones. In addition to that, there is well management of the radio component so that priority and capacity are well handled and services are delivered appropriately

The current thesis is motivated by the necessity of analysing the conditions in which private networks can be implemented, how requirements are handled and how strict they can be, and how they can coexist with public ones. The objective of this thesis is to develop a model for the analysis of the implementation of private networks in 5G using the approach of network slicing.

This thesis is composed of 5 chapters.

Chapter 1 introduces the thesis by providing a brief overview of the subject and the work's motivation and contents.

Chapter 2 revises the theoretical aspects of the reports subject. It describes the main aspects of the 5G networks. First, the network architecture of 5G Non-Standalone and Standalone is provided. Considering LTE importance for Non-Standalone 5G deployment, some LTE aspects are also addressed. Then, the Radio interface for both LTE and 5G is described. Following that, novel 5G network aspects are described. Software Defined Networking, Network Function Virtualisation and Network Slicing are addressed. Then the Cloud applied to the Radio Access Network and the Edge architecture for Multi-Access Edge Computing are described. After that, 5G use cases are split into three broad categories and described along with their needs. Private Networks are then addressed along with their use cases and performance requirements. The chapter ends with the state of the art of the report subject.

Chapter 3 gives the methodology that the thesis work will follow until its end, followed by the description of the expected results.

Chapter 4 presents the thesis results and their analysis.

Chapter 5 concludes the thesis with the conclusions made and proposal for future work.

Chapter 2

Fundamental Concepts

This chapter provides an overview of the 5G systems, mainly focussing on radio and network aspects.

2.1 Network Architecture

As defined by 3GPP in [3GPP18], 5G initial deployment will be by using the existing 4G infrastructure, this solution being called Non-Standalone (NSA) 5G and allowing a quicker and less costly implementation. At the next phase, one will have the Standalone 5G, with different core network and radio access.

NSA 5G architecture relies on the existing LTE 4G network, which comprises an IP based Evolved Packet System (EPS), with a core network called Evolved Packet Core (EPC) and an access network made of Base Stations (BSs), called evolved NodeB (eNB). Such a flat architecture allows the efficient handling of data traffic. Figure 2.1 shows the LTE Network Architecture with the representation of interfaces between eNBs and the EPC, these connections between eNBs distributing the intelligent control of the network, so that the connection set-up is speed up and handover takes less time.

The EPC encompasses four network elements [3GPP08]:

- Serving Gateway (Serving GW): serves the User Equipment (UE) by routing the incoming and outgoing IP Packets
- Packet Data Network Gateway (PDN GW): interconnects the EPC and external IP networks, performs IP address/IP prefix allocation, policy control and charging.
- Home Subscriber Server (HSS): database for user and subscriber related information.
- Mobility Management Entity (MME): responsible for the control plane, handles the signalling related to mobility and security for the radio access.

Figure 2.1 shows a basic architecture of the network with the representation of eNBs interfaces and the EPC's network elements and its connection to external networks.

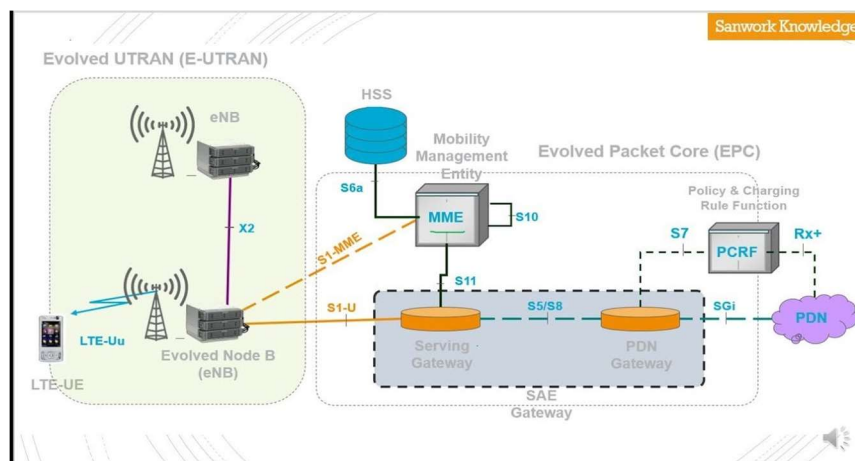


Figure 2.1 - Basic EPS architecture (extracted from [SaKn18])

The Standalone 5G adopts a new concept of networking, empowered by virtualisation technologies, so that a service-adapted network is deployed. The main difference to the previous generation is the new 5G Core Network, based on a Service-Based Architecture (SBA), such that network functions are modularised and can interact between themselves directly, via a common framework of interfaces.

And the base stations are now called Generalised NodeB (gNB).

The 5G core network emphasises the separation between Control and User planes, both with independent capacity scaling. The User Plane Function (UPF) is equivalent to GWs in 4G and is a gateway between the Radio Access Network and external networks. The Control-Plane consists of several network functions that are listed below [HUAW17]:

- Access and Mobility Management Function (AMF): Controls signalling between the core network and the user equipment (UE).
- Session Management Function (SMF): establishes and administers sessions by managing IP address allocation for the UE among other general session-management functions.
- Policy Control Function (PCF): Provides policy rules, incorporating network slicing, roaming and mobility management
- Unified Data Management (UDM): stores access authorisation and authentication credentials, works like HSS in 4G.
- Network Exposure Function (NEF): a gateway that allows external users to monitor, provision and enforce application policy for users inside the network
- NF Repository Function (NRF): Provides information so that NFs can discover each other and communicate between themselves
- Authentication Server Function (AUSF): Provides authentication to the UE regarding the requested network function.
- Application Function (AF): Interacts with the core network to provide services for applications.

The interfaces N1, N2 and N4 connect the AMF and the SMF to the user-plane to manage subscribers and mobility. The N2 and N3 interfaces depend on how the 5G radio presents itself to the core network, as on it has multiple possible radio access technologies, including WiFi.

Figure 2.2 shows the 5G system architecture with 5G Core inside the dashed lines, RAN stands for Radio Access Network and DN is the Data Network. Therefore, the connection to the User Equipment can be done by a few LTE and New Radio (NR) combination options as follows [5GAm18], Figure 2.3:

- Option 1: EPC and LTE and is listed for reference
- Option 2: standalone NR connected to the 5G Core Network
- Option 3: Non-standalone NR based on LTE-anchored dual connectivity with NR only for User Plane.
- Option 4: Non-standalone NR based on NR-anchored dual connectivity with LTE only for User Plane.
- Option 5: LTE connected to the 5G Core Network.
- Option 6: NR connected to EPC
- Option 7: Non-standalone NR based on LTE-anchored dual connectivity with NR only for User Plane. Connected to 5G Core Network
- Option 8: Non-standalone NR based on 11NR-anchored dual connectivity with LTE only for User Plane. Connected to EPC.

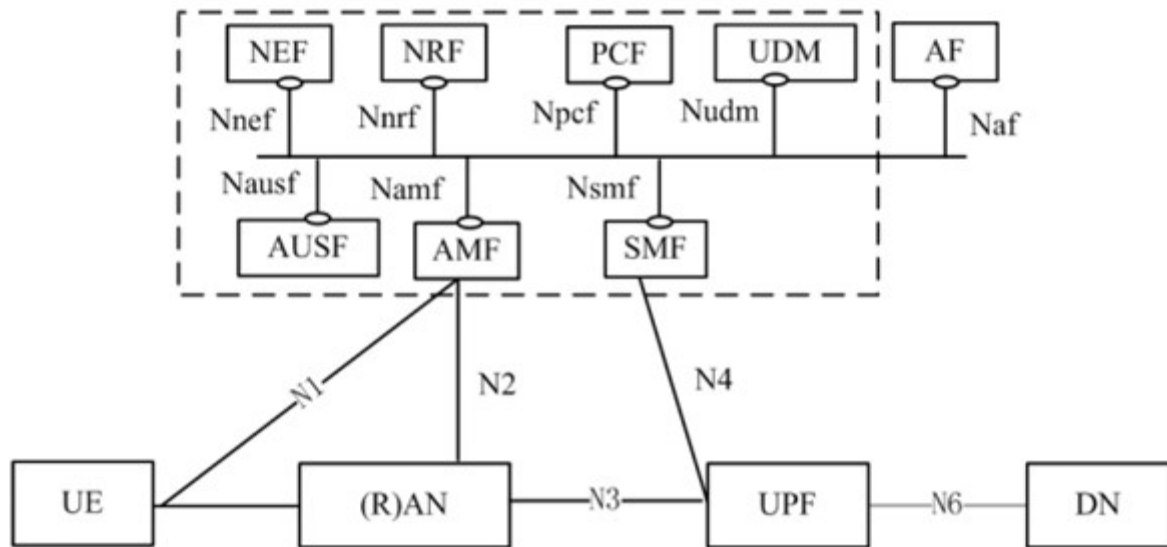


Figure 2.2 - 5G System Architecture (extracted from [HUAW17]).

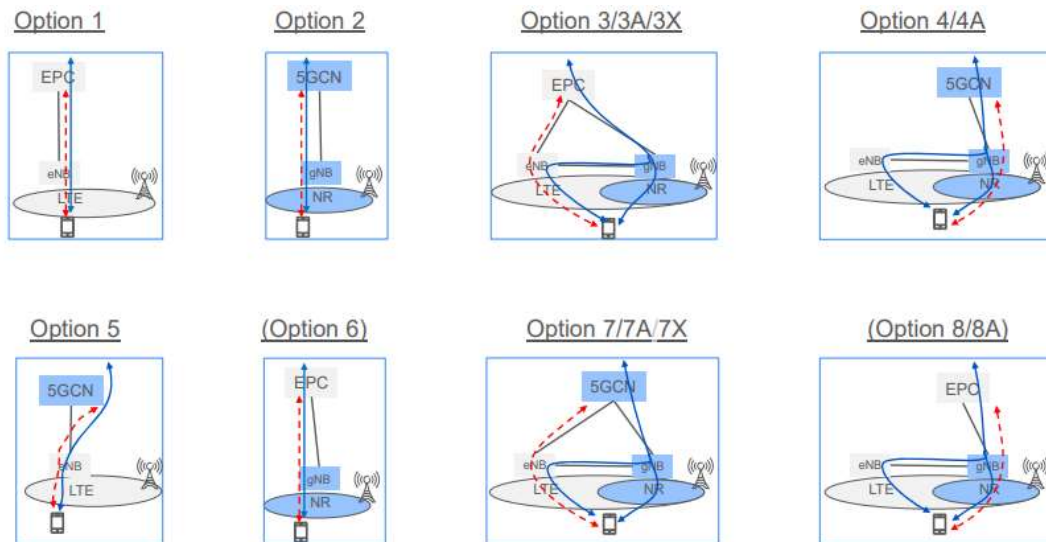


Figure 2.3 - 3GPP UE Connectivity Options (extracted from [5GAm18]).

2.2 Radio Interface

This subsection is based on [DaPS18], [BsSJ20], [GMBC19], [5GAm20]. 5G NR radio technology has been defined in 3GPP Release 15 and is the same for both Standalone and Non-Standalone. As already discussed, NR shares some of the 4G LTE infrastructure and technology, and it also can be deployed in the same spectrum as LTE, such that the overall spectrum is dynamically shared, which is known as Spectrum Coexistence. Besides, LTE is in constant development, parallel with NR. So, it is relevant to discuss LTE characteristics.

5G NR transmission scheme consists of Orthogonal Frequency-Division Multiplexing (OFDM) for the

downlink and Single-Carrier Frequency Division Multiplexing (SC-FDMA) for the uplink. NR uses the same transmission approach as LTE, but with a flexible OFDM numerology that spaces subcarriers according to the deployment scenario, such as large cells with low band frequencies or mm waves with very wide spectrum. The subcarriers bandwidth can range from 15 kHz to 240 kHz, according to $2^n \times 15$ kHz, opposed to the sole 15 kHz of LTE. As a consequence of this flexible approach, there is a proportional change in the cyclic prefix and the symbol time, such that the different delay spreads of the different scenarios can be well handled.

The time domain structure consists of frames, subframes and time slots. A frame has duration of 10 ms and a subframe of 1 ms. A time slot consists of 1 ms for LTE and can vary from 1 ms to 0.0625 ms for NR, depending on the subcarrier spacing. NR also supports the possibility of mini-slot transmission, in which only part of a time-slot is used.

Therefore, Physical Channels comprise resource blocks in both of technologies, which are sets of subcarriers and time slots that are allocated to each user depending on their service. Resource Blocks are composed of 12 subcarriers and one time slot, composed of 14 OFDM symbols in LTE. In NR, resource blocks are defined only in the frequency domain with 12 subcarriers, like LTE. NR is flexible in the time domain, as transmission can occupy less than one slot.

Figure 2.4. shows the corresponding resource blocks of two subcarrier spacings, one being the double of the other.

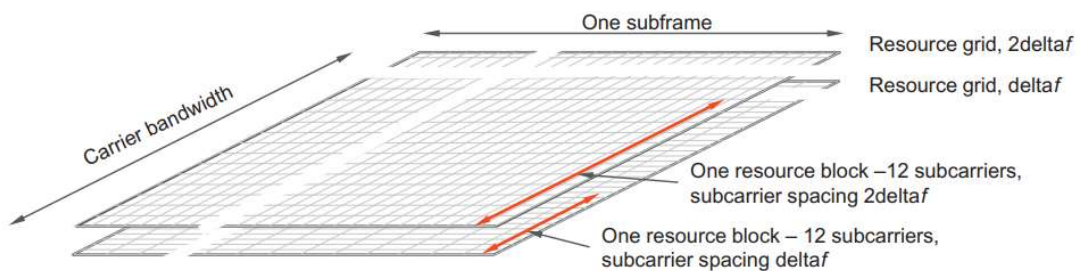


Figure 2.4 - Resource blocks for two subcarrier spacings (extracted from [DaPS18]).

Spectrum flexibility is essential, and in addition to the flexible numerology of the transmission scheme, NR duplex scheme consists of both FDD and TDD, being operated as half or full duplex with the same single frame structure. Besides, time slots allocation in NR changes over time, which is called dynamic TDD. LTE also supports both duplex schemes, but NR has three different frame structures: type 1 for FDD, type 2 for TDD and later on type 3 for operation in unlicensed spectra; and does not support dynamic TDD.

Both LTE and NR support adaptive modulation and coding, so that the modulation and coding rate change according to the channel conditions. LTE supports QPSK, 16 QAM and 64 QAM modulation schemes for uplink and downlink, NR supports 256 QAM in addition to those of LTE. In the case of NR, there is also support for $\pi/2$ -BPSK for the uplink when SC-FDMA is used. LTE uses Turbo coding and NR uses LPDC. NR also supports non-3GPP radio access.

Unlicensed spectrum also represents an important role in NR deployment, especially in the Private

Networks environment. Therefore, 3GPP Release 16 introduces NR Unlicensed (NR-U) in the frequency bands of 5 GHz and 6 GHz.

NR-U uses a License Assisted Access (LAA) based solution. This solution was also used in LTE unlicensed operation, in which the unlicensed spectrum access is only possible united with the licensed one. LAA can be divided into two modes: Carrier Aggregation and Dual Connectivity. In Carrier Aggregation mode the uplink and control plane signals are sent by the licensed spectrum, whereas the unlicensed spectrum manages the data plane downlink, augmenting its capacity. Dual Connectivity mode supports data plane traffic through unlicensed spectrum in both uplink and downlink, while the control plane uses the licensed spectrum. NR-U supports both modes, with the novelty of a Standalone NR-U that works in the unlicensed spectrum without the need of a licensed carrier. Therefore, one can define four main deployment scenarios, in which one of them leverages on the NR and LTE coexistence.

Figure 2.5 exemplifies those possible scenarios.

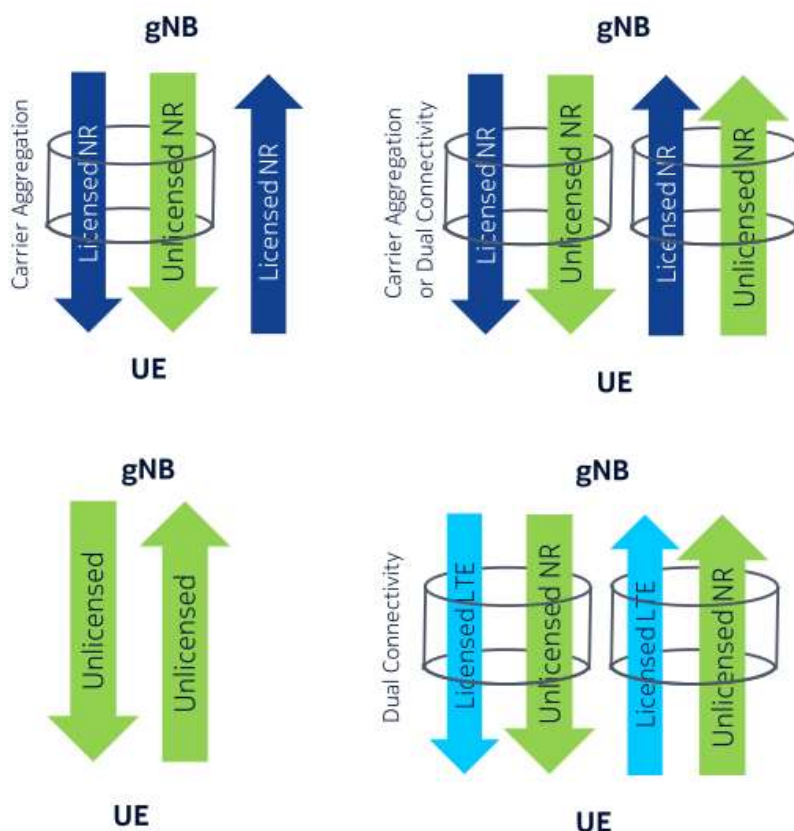


Figure 2.5 - NR-U Deployment Scenarios (extracted from [GMBC19]).

The main challenge of Unlicensed Spectrum is to manage the interference between all the technologies that may use the spectrum (e.g., Wi-Fi), so that these different networks can coexist. As of that, regulatory requirements are usually issued by the government of a given region to ensure network fairness and coexistence. For the 5 GHz and 60 GHz bands, Europe and Japan require the use of the Listen-Before-Talk (LBT) access protocol, which is a spectrum sharing mechanism that analyses the channel using a Clear Channel Assessment (CCA) check before acceding to it. CCA

measures the interfering energy in all directions to make the channel assessment. In the case of NR, in which beam forming is supported, the directionality of beams can be used to determine a direction-oriented energy measurement. In certain regions that LBT is not required, there is a regulatory constraint for the transmitted power level, in that way only small cells can be deployed.

2.3 Network Aspects

2.3.1 Virtualisation and Slicing Architecture

This subsection is based on [YBSS17], [FPEM17], [BoDF19], [OIFa17]. 5G vision aims to deliver top notch QoS and QoE to heterogeneous service requirements. In that sense, Network Slicing is of great importance, as it allows the concurrent deployment of multiple logical networks in a shared physical infrastructure, orchestrated in different ways according to the SLAs of the users. These network slices are owned by tenants or verticals, as they work over multiple layers of the network, integrating it vertically [YBSS17], and as enablers of this feature, the key technologies are Software Defined Networking (SDN) and Network Function Virtualisation (NFV).

SDN is a network paradigm created to enable easier network management with the support of multi-tenancy, by decoupling control and data plane. The control plane is represented by the virtualised SDN Controller, which is responsible for the handling of the network traffic. It runs separated from the network equipment and devices, which represent the data plane.

SDN implements a logically centralised intelligence (SDN Controller) that can create new network services by dynamically chaining network functions according to the necessity of the tenant. This dynamic chaining is called Service Chaining and is a network functionality that allows a service-based networking through ordered connection of network functions. In that way, network operators can create, scale, modify, remove, or add network functions based on dynamic demands of the network.

The Open Networking Foundation (ONF) proposes a SDN architecture that consists of a centralised SDN Controller that is responsible for provisioning, managing and controlling services and related resources. The controller is connected to Applications and Resources through northbound and southbound interfaces, respectively. These interfaces allow the direct interaction of users and applications with the network, therefore adjusting it according to the situation. Figure 2.6 shows the ONF SDN Architecture. NFV virtualises classical network functions such as routers, firewalls, intrusion detection and evolved packet cores. They run as multiple software applications in virtual machines on a cloud infrastructure as Virtual Network Functions (VNFs). This network deployment offers several benefits, such as: eliminating the dependency on specific and dedicated hardware, rapid implementation and deployment of new services, support for multi-tenancy scenarios and reduction in Capital Expenditure (CAPEX) and Operational Expenditure (OPEX) with efficient energy usage and automation of operational processes.

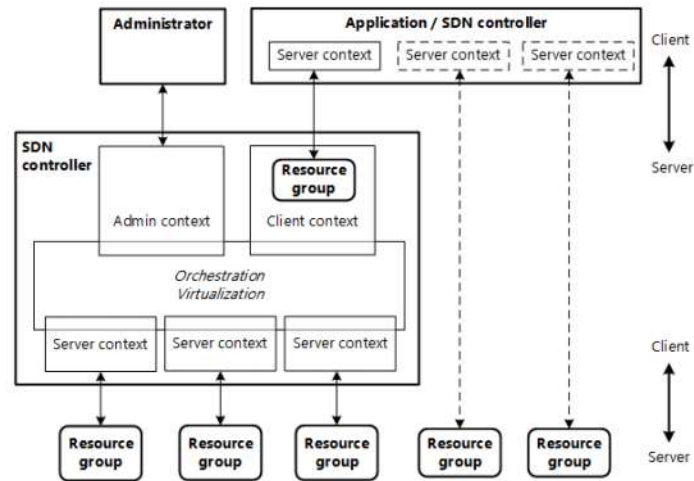


Figure 2.6 - ONF SDN Architecture (extracted from [YBSS17]).

Therefore, 5G networks can connect and manage network resources in order to create network slices for specific and diverse use cases. A generic Network Slicing framework based on 3GPP's and NGMN's proposals is presented by [FPEM17], being composed of three layers:

- **Infrastructure Layer:** Broadly refers to the physical network infrastructure (RAN and CN), It includes deployment, control and management of infrastructure, allocation of resources to slices and how these resources are presented to higher layers.
- **Network Function Layer:** Encompasses the network functions, their configuration and life cycle management. Enabled by SDN and NFV, it handles the optimal placement and chaining of network functions, to deliver an end to end service that meets certain constraints and requirements.
- **Service or Business layer:** Handles the high-level description of the service's business model and its mapping to the appropriate infrastructural element and network function. The service level description can be a set of traffic characteristics, SLA requirements and additional services.

Figure 2.7 shows the generic framework of the Network Slicing layers considering various architecture proposals, in addition to the layers previously addressed.

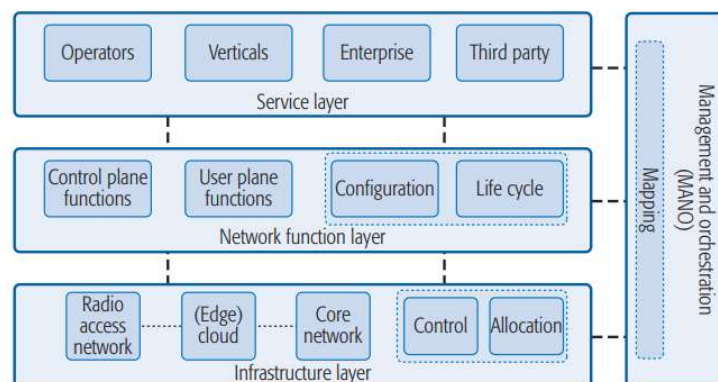


Figure 2.7 - Network Slicing Layers Framework (extracted from [FPEM17]).

There is a slice Management and Orchestration (MANO) entity that maps and stitches together the components available on the various layers of the architecture to create the end-to-end slice. This entity can gather information from the SLA of the service and then link them to the network functions and infrastructure element, or analyse the vendors' implementation and then deliver the best possible slice configuration.

2.3.2 Cloud and Edge Networking

This subsection is based on [HaYW19], [HNHS19]), [PFHP20]). Another important technology in 5G is Cloud Computing. It offers storage and computing resources on demand, leveraging on servers and network virtualisation. It allows higher flexibility and efficiency. In order to assure ultra-low latency, high data rate and massive communications, the concept of Edge Computing is also adopted, which is a decentralised model that enables edge servers in mini clouds at the edge of the network.

There are four key requirements for the deployment on edge computing on 5G Networks: real-time interaction, local processing, high data rate and high availability. These Edge servers can be embedded in the Base Stations, so that the access to edge clouds is easy and faster, therefore good for real-time interaction. Besides, as data and user requests can be processed by these servers, there is a decrease on the traffic in the Core Network and cloud services become available at the Edge. In addition to that, a centralised cloud model leaves the cloud too far away from users, which also increases energy consumption. Edge computing in 5G environment is known as Multi-Access Edge Computing (MEC).

The decrease in latency is of great importance for Mission Critical Communications, which represents one of the main use cases for Private Networks deployment. Military, Security forces, Fire-fighters and Healthcare professionals can benefit from this. They work with different kinds of information about an incident and have to make fast and critical decisions that must be shared in real-time with an emergency response team.

Figure 2.8 shows simplified models for cloud and edge computing.

Given the amount of user-data and heterogeneous requirements of the services, there is a need to centralise and virtualise RAN. This is done by means of Cloud Radio Access Network (C-RAN) architecture. In 4G, the radio and baseband processing functions are separated into different nodes, called Remote Radio Head (RRH) and the Baseband Unit (BBU). RRH is responsible for transmitting, receiving and digitising radio signals. This separation is called Distributed RAN. C-RAN moves the BBU from distributed BSs to a centralised BBU pool. Then every RRH is connected to a BBU pool through a fronthaul link and the BBU pool is connected to the Core Network through a backhaul link. In that way, cloud computing is embedded into the RAN architecture by centralising and virtualising the baseband processing, this also allows resource pooling to perform the network slicing according to the service requirements. This architecture has the challenge of high bandwidth requirements between RRH and BBU, so the fronthaul would be necessarily fibre optics. As a solution to this, a partially centralised C-RAN architecture is proposed in which functions such as sampling, modulation, resource block mapping, quantisation are integrated into the RRH, with that the bandwidth requirements are

lower but also lowers the flexibility in network upgrades. Figure 2.9. shows the general C-RAN architecture.

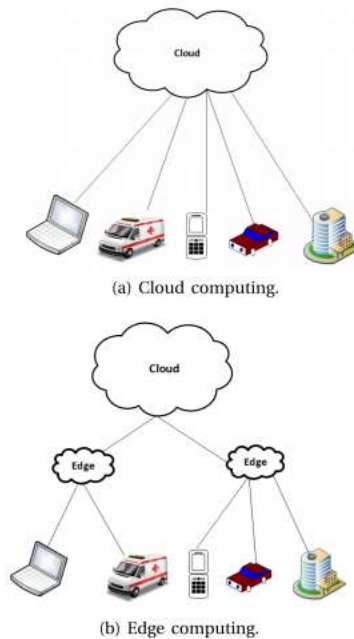


Figure 2.8 - Cloud and Edge Computing models (extracted from [HaYW19]).

5G NR proposed solution for high throughput requirements on the fronthaul is a C-RAN implementation in which the network functions are split among three units: Radio Unit (RU), Distributed Unit (DU) and Centralized Unit (CU). This solution provides the possibility for more functions to be processed in the DU closer to the user before being transmitted to the CU, where the processing power is higher. The connection between the DU and CU is called Middlehaul (MH). The main splitting options considered in this work are Splitting Option 7.2, where the resource element mapping is included in the RU and the data is transported on subcarrier symbols. The data symbols are only exchanged when data is available, so the transport capacity demand is reduced and scaled with the cell load.

MEC integration into the 5G system is in constant development and standardisation by the European Telecommunications Standards Institute (ETSI), with focus on reference architecture, NFV applied to MEC, and C-RAN integration. 3GPP recently included MEC in the 5G standardisation with the technical specification 3GPP TS 23.502. The reference framework is composed of the 5G Core Network service-based architecture (SBA), which has already been discussed in Section 2.1, and reference MEC architecture.

The MEC reference architecture comprises MEC system level and host level. The MEC system level main component is the MEC Orchestrator (MECO), which holds information on deployed MEC servers (hosts) and services, available resources and system topology. The MECO is also responsible for the selection of the MEC servers, assigning them for specific application tasks. The MEC host level is composed of the MEC Platform Manager and the virtualisation infrastructure manager (VIM). The MEC platform manager manages the life cycle of applications, provisions element management

functions, controls application rules and requirements, and also processes fault reports and performance measurements. In its turn, the VIM allocates virtualised resources, provisions MEC applications and monitors application faults and performance, sending them to the MEC platform manager. At last, the MEC server or host consists of an MEC platform and a virtualisation infrastructure. The MEC platform holds functionalities necessary to run MEC applications on a given virtualisation infrastructure. And the virtualisation infrastructure executes data plane functionalities, such as execution of traffic rules sent by the MEC platform and directing the traffic among applications and networks.

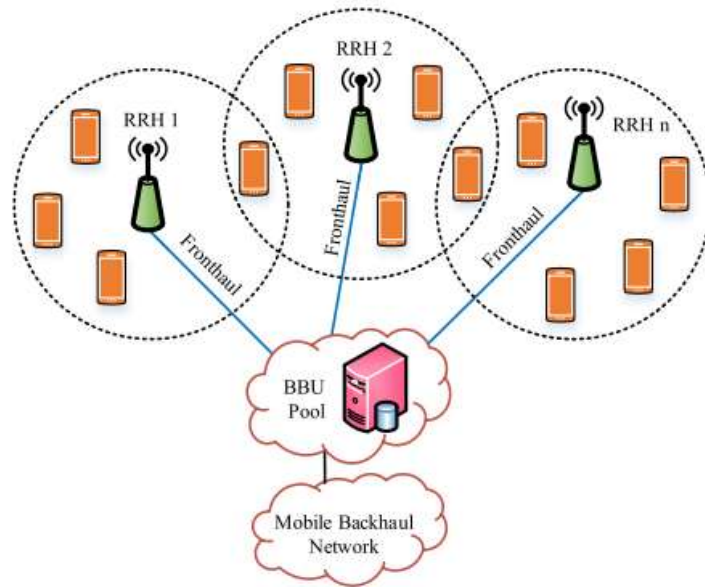


Figure 2.9 - C-RAN Architecture (extracted from [HNHS19]).

Figure 2.10 shows the reference framework of MEC and 5G integration, with the 5GCN at the left and the MEC system at the right.

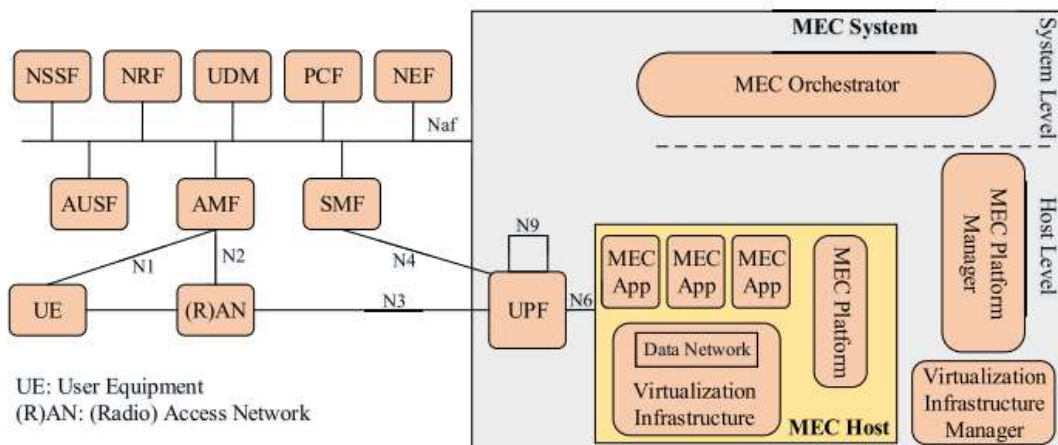


Figure 2.10 - MEC and 5G integrated framework (extracted from [PFHP20]).

2.4 Services and Applications

This subsection is based on [3GPP18], [3GPP19], [GSMA18], [5GAm17]. 5G's main goal is to provide several services with heterogeneous requirements, which are defined by the application to be supported and are highly variable. Different Key Performance parameters can be necessary, such as mobility, reliability, security, latency or data rate. ITU defines three broad classes of use cases:

- Enhanced Mobile Broadband (eMBB): it corresponds to the evolution of the current mobile broadband use cases growing scenario of a fully connected society, with improved capacity, coverage and enhanced user experience, and includes Ultra High Definition video streaming, 3D applications are part of this class.
- Massive Machine Type Communications (mMTC): it is composed of high user density services, usually low-power, low-cost devices connect in an energy efficient way, and includes Industrial Internet of Things, smart cities, environmental monitoring and traffic control.
- Ultra Reliable, Low Latency Communications (URLLC) or Critical Communications: it comprises services that require extreme availability (reliability) and a strong demand in real-time interaction, and includes factory automation, military applications and self-driving cars.

This classification aims to simplify the variety of 5G use cases, therefore according to [GSMA18], one can define two new service types that would be between those previously defined:

- eMBBLLC: it corresponds to use cases that require higher data rates and lower latency at the same time, availability not being as important as URLLC, which includes Cloud Gaming and AR/VR gaming.
- mMTC/URLLC: it comprises use cases in which several devices send time sensitive and critical data, therefore requires high usage density, low latency and high reliability, and includes motion control of machine parts in the Industry 4.0 context.

In addition to that, 3GPP Release 15 maps services according to their QoS characteristics, which are:

- Resource Type: it determines if dedicated network resources related to Bit Rate are permanently allocated which can be Guaranteed Bit Rate (GBR), non-GBR and Delay Critical GBR.
- Priority Level: it indicates the priority of resource allocation among QoS management.
- Packet Delay Budget: it defines an upper bound for the delay time between the UE and the UPF.
- Packet Error Rate: it defines an upper bound for packets that are not successfully delivered between protocol layers.

The table in Annex A maps services according to their QoS characteristics, GBR means Guaranteed Bit Rate.

There are several possible use cases for Private Networks with different classes and corresponding requirements, one of them being Healthcare services. Given the global pandemic and the need of keeping health workers safe from infections, these services come with great importance. The applications are very diverse and include remote surgery, wireless service robots for health monitoring

and smart rehab homes, etc. Remote Surgery must provide the surgeon with tactile feedback, so that he/she can sense hard tissue or nodules, therefore one must have low delay, extreme reliability and broadband for the rendering of images. Wireless Service Robots would join labour roles, being used for health monitoring, cleaning and logistic. The robots reasoning system would be located at the edge cloud, therefore extreme low delay, reliability and high data rates would also be required. Table 1.1 gives the approximate value of the key performance requirements of the Healthcare applications described [SFRD17].

Table 2.1 - Key performance requirements for healthcare applications (adapted from [SFRD17]).

Application	Latency	Failure Rate	Bit rate
Remote Surgery	<100ms (audio/video feedback) <25ms (haptic feedback) <5ms (interactive live holographic feedback)	10^{-7} (3.17 s of outage per year)	1 Gbps
Service Robots	<5ms	10^{-7} (3.17 s of outage per year)	450 Mbps

Military applications are in constant and rapid development. Information based warfare brings the need for high performance requirements. According to [LiOu20], there are four types of typical applications scenarios: battle command, training exercise, logistics support and special equipment support. These scenarios have different but very stringent requirements. In [BaCK20], the NATO Communications and Information (NIC) Agency also makes a technical assessment of the potential 5G technologies on the military environment. They developed four 5G military application domains, which are: Deployable Communications and Information Systems (CIS) for Expeditionary Operations; Land Tactical Operations, Maritime Operations and Static Communications. The NIC Agency specifies and studies these application domains. They can be included on the four application scenarios that [LiOu20] defined.

Therefore, one can enumerate the key performance indicators for military applications in a generalised way. Annex C does this and presents estimate values for these requirements.

Power generation and distribution solutions must be considered. Some of these solutions are Smart Grid, Smart Meters and Electricity traffic scheduling. Smart Grid consist of an intelligent and efficient power distribution control, in a way that loads can be better balanced and faults can be handled faster. With multiple sensors, Smart Meters supply constant measurements to the energy provider and allow

real-time optimisation of the network, such as electricity traffic scheduling, which permits building a map of power consumption and improving the power schedule, in addition to faster fault location and isolation. Table 2.2 gives some reference requirements for energy related services [GSMA18].

Table 2.2 - Key performance requirements for energy related services (adapted from [GSMA18]).

Bandwidth	1 kbps per residential user
Latency	<5 ms
Packet-Loss	<10 ⁻⁹
Reliability	>99.999% (5 minutes downtime per year)
Failure Convergence Time	Seamless failover required
Handling of crisis situations	(Surviving long power downtimes on a large scale, assuring black start capability): mandatory.
Connection Density	>1 000 devices/km ²

Railway networks also present a possible use case for 5G Private Networks. GSM-R is a standard developed for railway communications that is used in Europe and most of the world. More recently LTE-R has been developed so that higher data rates could be delivered. For the future smart railways, services like real time ultrahigh definition video transmission, train-to-track secured closed circuit television (CCTV) and massive sensor coverage of the train and track will be deployed. These services require data rates to the order of the Gbps and high connection density, therefore are not met by LTE-R [AMRZ20]. 5G comes as a strong substitute for GSM-R and LTE-R. [GMVA17] splits the possible railway communications into three: Train-to-Ground communications, Train Communication Network and Passenger Communications. The first two are mission critical communications and are related to security and reliability of the railways transportation systems. Passenger communications are not critical, but require high capacity. Table 2.3 summarises the three scenarios key performance requirements.

Table 2.3 - Railway networks key performance requirements (adapted from [GMVA17]).

Communication Network	Data Rate	Availability	Mobility	Delay	Connection Density
Train-to-Ground	<10 Mbps	>99.99%	Up to 500 km/h	<50 ms	-
Train Communication Network	1 - 10Gbps	>99.999%	-	<10 ms	-
Passenger comms	7.5 - 15 Gbps per train	>95%	Up to 400km/h	<500 ms	1000/train

The emergency of Industry 4.0 brings several new applications with diverse and stringent requirements. These applications aim to monitor, control and connect Industry processes. The applications include Monitoring, Safety Control, Closed-Loop Control, Motion Control, Process automation. Most of these can be classified as mMTC/URLLC, as they comprise several sensors that must have low delay and high reliability. Table 2.4 summarises some key requirements for Industrial Applications [Aija20] [3GPP] [GSMA18].

Table 2.4 - Industrial applications key performance requirements (adapted from [Aija20] and [3GPP19]).

	Reliability	Latency	User Experienced Data Rate	Connection Density
Monitoring	>99.9%	50 ms – 100 ms	0.1 Mbps – 0.5 Mbps	10 000/km ²
Safety Control	>99.999%	5 ms – 10 ms	0.5 Mbps – 1 Mbps	1 000/km ²
Closed-Loop Control	>99.999%	2 ms – 10 ms	1 Mbps – 5 Mbps	1 000/km ²
Motion Control	>99.9999%	0.5 ms – 2 ms	1 Mbps – 5 Mbps	1 000/km ²

Unlicensed Operation also faces the problem of Interference. There are methods for interference mitigation, such as the LBT protocol and the Dynamic Frequency Selection (DFS), but they also lower spectral efficiency, which brings a trade-off to the table.

2.5 Private Networks Architectures

Private Network Slicing and NR-U present a big opportunity for Private Networks provisioning in the mobile communications systems. By means of SDN and NFV, Network Slicing allows the deployment of isolated logical networks under the same physical infrastructure. It lowers costs and energy usage and gives Mobile Network Operators (MNO) the possibility of new business model. NR-U comprises a standalone operation that allows the deployment of Private 5G networks by non-MNOs, as there is no anchoring to the licensed spectrum. In the License Anchored operation the unlicensed spectrum is dependent on the licensed one. Such operation can be used to augment the capacity of private networks deployed by MNOs.

One can also classify 5G Private Networks according to their functional architectures, [Aija20]:

- **Standalone Deployment:** corresponds to a Private Network independent of the Public Network. All data flows and network functions take place inside the premises of the service area and can be used by the Unlicensed Spectrum. One can connect the Private Network to the Public Network through a firewall, in case public network services are needed.
- **Public-Private Shared Ran Deployment:** corresponds to RAN sharing of private and public networks. Yet all network functions are separated and all data flows are confined to the premises of the service area.
- **Shared RAN and Control-Plane Deployment:** in this case, RAN is shared between public and

private network and all control-plane network functions are handled on the public network, therefore leveraging on the public network infrastructure and on network slicing

Figure 2.11 illustrates the three architectures.

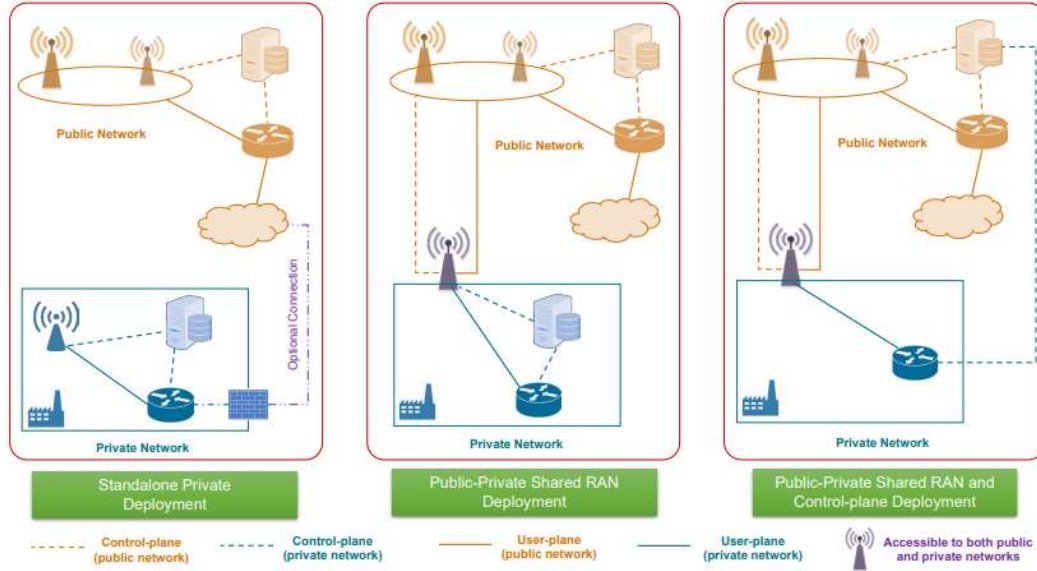


Figure 2.11 - 5G Private Networks Functional Architecture (extracted from [Aija20]).

2.6 State of the Art

In the context of 5G Private Networks, parameters like isolation, security and priority are of great importance, as Private Networks cannot be influenced by the public network traffic and usually deal with classified or sensitive data. Resource allocation referring to the RAN and Core slicing is also an essential issue that impacts the logical network performance in a private network slice. Furthermore, Unlicensed Operation brings forward the coexistence challenge. This subsection addresses research related to these issues and specific use cases scenarios.

Isolation between slices or Inter-Slice Isolation has to do with the status maintenance of the network given that one of the slices is under heavy load, i.e., changes affecting users in one slice should not affect users of other slices. [HMML19] proposes isolation between slices by using the Flex Ethernet (FlexE) technology, which uses a TDMA-like resource subdivision between different slices. An algorithmic framework based on a Column Generation routine configures the FlexE interfaces in order to achieve hard isolation and QoS guarantees. A SDN Controller can be used to implement the algorithm.

[YLWW19] proposes a Connection Admission Control (CAC) mechanism to isolate slices and allocates RBs to users according to a game theory algorithm that aims to minimise the Signal-to-Interference-plus-Noise-Ratio (SINR) in all the RBs allocated to a user. They also define implicit-

isolation as the use of radio resource management in order to ensure that changes in one slice do not deteriorate the QoS of users in other slices. With the CAC mechanism, new connections are only admitted when established ones are not affected and present QoS guarantees. This work considers interference between slices in the RAN. In [YLWW20], the same authors now propose a Real-Encoded Genetic Algorithm to achieve optimal inter-slice resource allocation and minimise the inter-slice interference. They now define a new parameter called Isolation Factor, which is calculated by the sum of the total differences of allocated resources between slices in all neighbouring cells. The goal is to minimise this parameter.

[SaMa19b] proposes a model that optimises slice allocation and satisfies end-to-end delay constraints by using a Mixed-Linear Integer Programming (MILP) as optimisation model. In their work, they guarantee end-to-end delay, provide intra-slice isolation and find a minimum delay path between the slice elements. Intra-Slice isolation stands for the physical isolation between VNFs and might be required for higher reliability, as it allows the slice operator to recover from partial compromise or unavailability of a network slice. Their network model considers that each slice request is associated with a computing demand, a bandwidth requirement, end-to-end delay and the intra-slice isolation. The model allocates the slice request to the least used server. Their evaluation results show that when there is little or none intra-slice isolation, the bandwidth requirement reduces and CPU utilisation increases. On the other hand, if intra-slice isolation is high, more bandwidth is required and less CPU is used.

[SaMa19a] addresses security concerning Distributed Denial-of-Service (DDoS) attacks by using Inter and Intra-Slice Isolation as security constraints for the mitigation of these attacks. Inter-slice isolation mitigates DDoS attacks because hardware resources are not shared between slices, therefore an attack on one slice does not affect the others. This work uses a MILP approach like [SaMa19b] to address optimal resource allocation, but now with security-related constraints related to the slices isolation. The results show that complete intra-slice isolation provides the best mitigation to DDoS attacks. Inter-slice isolation provides reliable resource isolation, but can reduce the efficiency of their utilisation. Therefore, intra-slice isolation presents better control over the trade-off security, availability and resource utilisation. Both [SaMa19a] and [SaMa19b] works focus on Core Slicing.

[MZLZ20] proposes a core network for network slicing and secure communication; it handles the problem of unencrypted user data with strong connectivity, security and scalability. The solution consists of adding two network elements to the existing core network, which are: Bootstrapping Function (BSF) and Network Application Function (NAF). BSF guides the service and gives it a 5-tuple of Authentication Vector that corresponds to the private network user, and performs a two-way authentication with the UE, forming a shared key. NAF is used to encrypt, decrypt and transmit data in the uplink and downlink to the UE. The authors say that this solution is scalable enough to work on 5G networks, but has only been tested on LTE Networks.

[PMKL19] proposes a secure keying scheme for network slicing, when the slices access is made by third party applications. The authors consider several attack scenarios, like DDoS, Data Tampering, attempt to take administrative control, data transportation attacks and key-compromise impersonate

attacks. This work considers that services are monitored by a Third-Party Monitoring Application (TPMA). The retrieval of data and key management are done by using multi-party computation mechanisms. The TPMA interfaces with several network functions such as AMF, AUSF and UPF. The cryptographic keys are provided by a Key Distribution Centre. The results show satisfactory performance, with flexibility for different use cases scenarios and integration with the 5G architecture.

[VeNN20] envisions a Network Slicing architecture able to provide reliability for smart healthcare. The authors introduce a Framework for Fingerprint smart-health Apps Traffic and Providing Network Resource Slicing, called FLIPER. This framework fingerprints health applications according to their network traffic behaviour and provides network resources customised to the application requirements. The FLIPER architecture consists of four modules: Pre-Processing, Feature Extraction, Fingerprinting and Network Slicing Configuration. The results show that this framework can fingerprint network traffic with 90% accuracy.

In [HaWW19], a network slicing scheme is proposed to attend private power communication networks. The proposed scheme introduces Access Gateway and Application Edge Gateway to the usual network slicing scheme. The Application Edge Gateway comprises five application related instances (i.e., power system load control unit, on-site inspection terminal, distribution terminal unit, energy meter and power related sensor) that are constructed on Linux Virtual Machines, separating application related software from hardware. The Access Gateway has the same architecture as the Application Edge Gateway but comprises three application instances with individual protocols, each for smart meter service, distribution terminal service and sensor terminals. These gateways augment communication efficiency and flexibility.

Chapter 3

Model and Simulation

In this chapter, the methodology for the thesis development is broken down into main tasks and each one of them is described.

3.1. Model Overview

In order to analyse private networks performance, one must consider many aspects. Annex A summarises the important parameters for QoS assurance in these networks, as this work focus on the analysis of latency and data rate considering services requirements. It also aims to optimise resource allocation in the network slicing environment and aggregate the closest nodes in the network. Table 3.1 shows the input and output parameters and the model stages.

Table 3.1 - Input and Output Parameters, along with model stages.

Input	Model	Output
<ul style="list-style-type: none"> • Cell <ul style="list-style-type: none"> ○ MIMO layers ○ Bandwidth ○ Frequency band ○ Scenario Topology • Network <ul style="list-style-type: none"> ○ VNO ○ QCI Value • Services <ul style="list-style-type: none"> ○ Service Data rates ○ Service Latency • User <ul style="list-style-type: none"> ○ Number of Users ○ Service Mix 	<ul style="list-style-type: none"> • Numerology Selection • Maximum Achievable cell data rate calculation • Admission control and delay process. • Maximise the usage of the available capacity, using VRRM optimisation. • Aggregation Process 	<ul style="list-style-type: none"> • E2E Latency • User's Data Rate • VRRM Capacity Share • Total Data Rate of Each Service • User's Satisfaction • Percentage of served users

The model inputs consist of four classes: cell, network, services and user. Each class has different parameters that characterise the scenario and the relevant variables.

The first class parameters are used to calculate the cell's maximum available data rate, being: MIMO Layers, Numerology, Bandwidth, Frequency Band. These parameters are used in an equation provided by 3GPP and may change according to the service and frequency band. Details and full description of these parameters are in Subsection 3.2.1.

The second class, Network, is composed of parameters that characterise the slices and the network architecture. These parameters are VNO, QCI Value, Nodes Locations, Nodes Processing Capacity and Nodes Specifications. VNO is the identification tag of the VNO. QCI value identifies the type of service and its priority. The Nodes Locations were provided by NOS. Nodes Processing Capacity provide the maximum processing power of each node, important for the aggregation process. Nodes Specifications define the C-RAN Architecture and the percentage of nodes for each kind of unit.

The third and fourth classes are Services and User, respectively. Services parameters are Data Rate and Latency, both with an acceptable data range for each service. User's parameters are Mobility, Number of Users and Service Mix. Mobility is given as the UE velocity in km/h, Number of Users is the total number of users in a given scenario, and Service Mix is the percentage of users assigned to each service.

The model starts by the aggregation process, in which nodes are interconnected considering a maximum distance, defined by the services' maximum latency, and their processing capacity. Then, the maximum achievable cell data rate is calculated with the Cell parameters. After that, admission control and delay process is done, as it is essential for the solution of the VRRM problem and for prioritising delay critical services. When the network is congested, as for the processing capacity of the nodes or the VRRM capacity available, users with no minimum latency or data rate must be delayed. Then, users with the lowest priority are delayed one by one until the capacity is enough. This process uses the cell capacity calculated, network, services and users' parameters, specifically it analyses the QCI values of two services, comparing them and then admitting the higher priority. The next stage is the VRRM Optimisation, whose goal is to maximise the usage of available capacity considering services' priority, while distributing capacity in a fair manner, subject to some constraints, including maximum achievable capacity, predefined SLA thresholds [Rouz19]. This stage uses the calculation of the maximum achievable cell capacity and the network and services input parameters. Finally, the total E2E Latency is calculated considering the delay process, processing delays and link latencies.

The model output parameters are:

- Percentage of total assigned data rate: total network throughput in terms of data rate, meaning that if it is close to 100% it reflects an optimal VRRM performance.
- VRRM capacity share: total capacity assigned to each slice out of the total capacity available to VRRM.
- Total data rate of each service.
- Percentage of served users: is the percentage of users that are allocated to a given service out of the total number of users of that service
- Data rate of each user: has a direct impact on the satisfaction level of the served users.
- Users' satisfaction: classification method that intends to determine if the data rate allocated to a user will have a good or a bad experience.
- E2E Latency: total end to end latency of the network.

The overall model flowchart is given in Figure 3.1. The used algorithms are presented and further explained on the following subsections. The model is implemented in MATLAB, and CVX for the VRRM Optimisation.

The first step is reading the input parameters. The input parameters are loaded from an excel data sheet. Then, the Numerology Selection algorithm determines the best numerology so that the Maximum Achievable Cell data rate can be calculated. This is done using (3.1) and following all the necessary steps detailed in the Section 3.2.1. Then the programme proceeds to check if there is

enough capacity to serve all users and if not, the delay process algorithm starts. After delaying users from lower priority services, a matrix containing the services delayed is generated. Once it is verified that all remaining users can be served with at least the minimum demanded capacity for their service, the VRRM optimisation algorithm is initiated. After running the VRRM optimisation, the programme computes the several output parameters that reflect the overall network performance and user satisfaction, regarding the data rate: percentage of total assigned data rate, VRRM capacity share, total data rate of each service, percentage of served users, data rate of each user, the users' satisfaction.

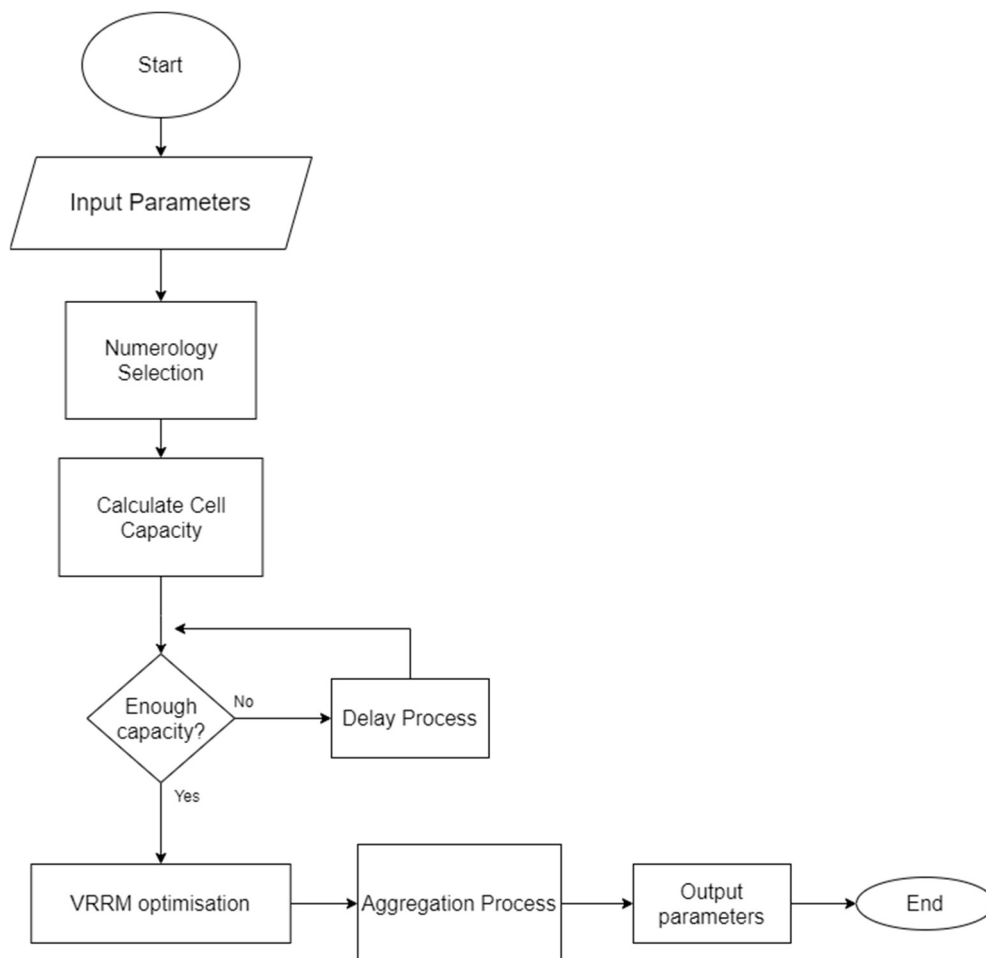


Figure 3.1 - General Model Workflow.

After that, the programme starts the aggregation process, based on the chosen architecture and network information. RU, DU, CU, CN or MEC are all connected in sequence, respecting distance and node capacity requirements. At the end, total latency is calculated, considering contributions from the network architecture and delay process.

3.2. Cell Capacity

In order to achieve maximum performance, one needs to appropriately choose the used numerology, as it defines the TTI and Cyclic Prefix. Consequently, it influences Inter Symbol (ISI) and Inter Carrier Interferences (ICI), and one must consider the scenario topology, services requirements and frequency band in order to make the best choice. This model does not take into consideration Cell Range details.

[VaMA21] quantifies interference in different scenarios and related it with the available numerologies, Figure 3.2 showing the allowed subcarrier spacings in different frequencies. This work focuses on sub-6 GHz frequency bands, so the algorithm only considers these bands. It first analyses the frequency band, and then the scenario topology. After that, if there is the possibility of using more than one numerology, the algorithm checks if there is a service with high data rate and/or low latency requirements. If there is such a service, the algorithm selects the highest possible numerology. The workflow of the Numerology Selection algorithm is presented in Annex B.

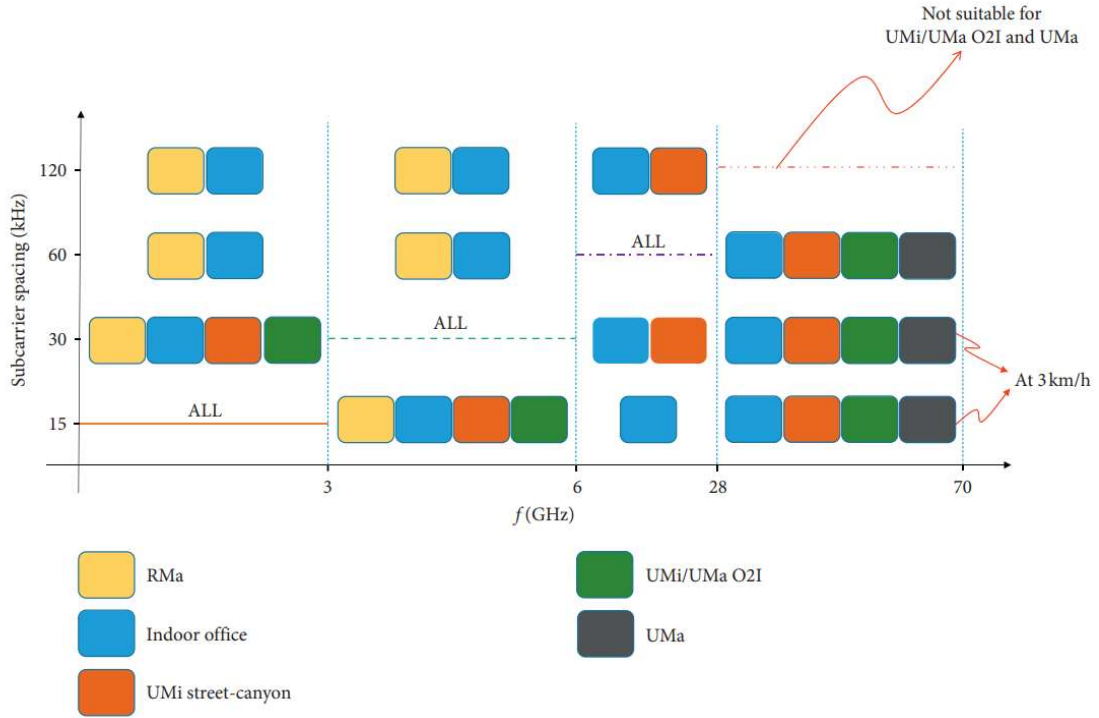


Figure 3.2 - Allowed SCSs for each 5G scenario (extracted from [VaMA21]).

In order to analyse the data rate distribution among slices and services, one must compute the maximum achievable data rate of the cell. The following expression defines the peak data rate for both UL and DL transmission as with FDD and TDD techniques [3GPP20].

$$R = \sum_{j=1}^J (\alpha^{(j)} \cdot v_{Layers}^{(j)} \cdot Q_m^{(j)} \cdot f^{(j)} \cdot R_{max} \cdot \frac{N_{PRB}^{BW(j), \mu} \cdot 12}{T_S^{\mu}} \cdot (1 - O^{(j)})) \quad (3.1)$$

where:

- R : Peak cell data rate.
- J : total number of aggregated carriers in a frequency band.
- R_{max} : Maximum Coding Rate.
- For the j -th carrier:
 - $\alpha^{(j)}$: Normalised scaling factor related to FDD and TDD.
 - $v_{Layers}^{(j)}$: Maximum number of MIMO layers.
 - $Q_m^{(j)}$: Maximum supported modulation order.
 - $f^{(j)}$: scaling factor
 - $N_{PRB}^{BW(j),\mu}$: Maximum RB allocation in bandwidth $BW(j)$ with numerology μ .
 - T_s^μ : Average OFDM symbol duration in a subframe for numerology μ .
 - $O^{(j)}$: Overhead.

In this work, carrier aggregation can be considered when working under unlicensed spectrum. MIMO layers have a maximum value of 8 in DL and 4 in UL. The modulation order, $Q_m^{(j)}$, is determined by the number of different transmittable symbols, consequently QPSK, 16 QAM, 64 QAM and 256 QAM have modulation orders respectively, 2, 4, 6 and 8. The scaling factor, $f^{(j)}$, reflects the capability mismatch between baseband and RF capability for both SA UE and NSA UE, and it can take the values of 0.4, 0.75, 0.8 and 1. The normalised scaling factor relates to the proportion of resources used in the DL/UL ratio for the j -th carrier; for FDD, it takes value of 1 for UL and DL, but for TDD one must calculate it based on the frame structure and the Slot Format Indicator (SFI); for this work one considers the scaling factor as 0.857.

The values for the maximum coding rate, R_{max} , are the best coding rate values for the modulation to be used. Table 3.2 shows the Modulation Scheme and the respective maximum code rate.

Table 3.2 – Modulation schemes and maximum code rate.

Modulation Scheme	Maximum Code Rate
QPSK	449/1024
16QAM	616/1024
64QAM	873/1024
256QAM	948/1024

The overhead, $O^{(j)}$, is calculated as the average ratio of the number of Resource Elements occupied by L1/L2 control, synchronisation signals, PBCH, reference signals and guard bands with respect to the total number of REs for the effective bandwidth in a 5G NR frame time product. It varies according to the frequency range and link and assumes the following values:

- 0.14, for frequency range FR1 for DL
- 0.08, for frequency range FR1 for UL

In order to support different services and frequencies, NR uses a scalable OFDM in which the subcarrier spacing depends on a given numerology μ and is given by $15 \times 2^\mu$ kHz. As such, the available numerologies for FR1 are 0, 1, and 2 that correspond to a subcarrier spacing (SCS) of 15 kHz, 30 kHz and 60 kHz respectively. With μ one can calculate T_s^μ as shown in

$$T_s^\mu = \frac{10^{-3}}{14 \times 2^\mu} \quad (3.2)$$

To obtain the number of RBs it is necessary to know both the BW as well as the numerology being used. Table 3.3 presents, according to 3GPP, the number of RB available for a given bandwidth and numerology.

Table 3.3 – RB per Numerology per Bandwidth for Frequency Range 1 [3GPP17].

Bandwidth [MHz]	SCS [kHz]		
	15 ($\mu = 0$)	30 ($\mu = 1$)	60 ($\mu = 2$)
5	25	11	n/a
10	52	24	11
15	79	38	18
20	106	51	24
25	133	65	31
30	160	78	38
40	216	106	51
50	270	133	65
60	n/a	162	79
70	n/a	189	93
80	n/a	217	107
90	n/a	245	121
100	n/a	273	135

3.3. Admission Control and Delay Process

This process aims to delay low priority users in order to achieve a state where the high priority users can be served. The first users to be delayed are the Best Effort (BE) ones, as they do not have any minimum contracted level of data rate. Next, both VNO slices and services are ordered from lowest to highest priority. Starting from the lowest priority slice and service, 1 user is delayed. After every delayed user it is checked if it is possible to serve the remaining users with minimum demands. In case it is not possible the programme continues to delay users until all Non-GBR users are delayed. Next starting again from lowest to highest priority, the programme delays all GBR users, and then the same process for Delay Critical Non-GBR users until there is enough capacity. This process only makes sense when the capacity available is not enough to serve all users.

Figure 3.3 shows the workflow of the admission control and delay process.

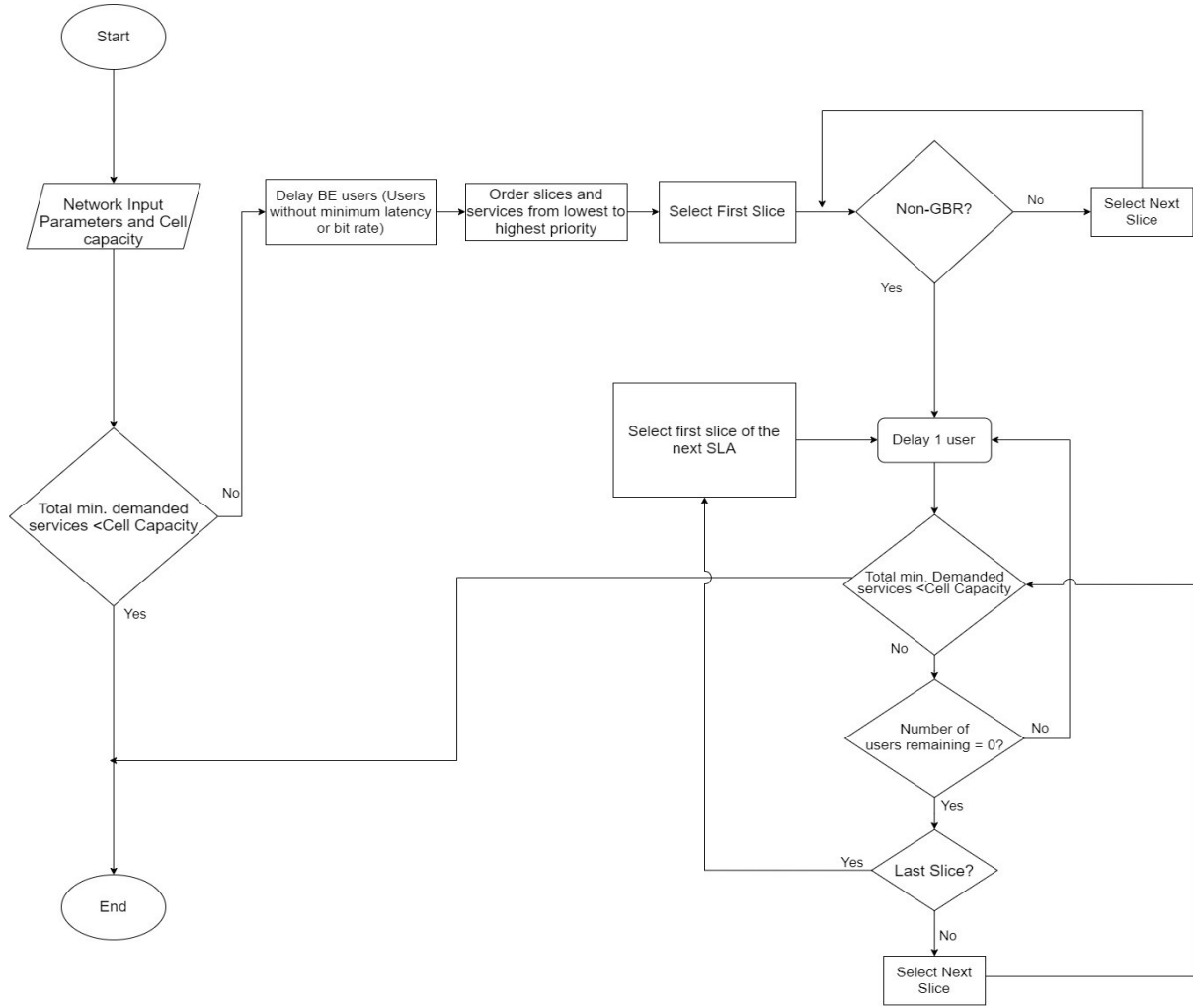


Figure 3.3 - Admission Control and Delay Process workflow.

3.4. VRRM Optimisation

In order to maximise the usage of aggregated capacity, Virtual Radio Resource Management (VRRM) model is used. It also allocates the available radio resources to the diverse services and their corresponding VNOs. The model used in this work was developed by [Rouz19].

The problem is formulated as a constrained concave optimisation problem. An objective function, f_{VRRM} , is defined in order to balance between efficiency and fairness when allocating resources in a network with heterogeneous services. The utility function used in this work is the Logarithmic one, as it provides satisfactory handling of proportional fairness considering the priority of the requested services. The goal is to balance the aggregated capacity, providing maximum capacity to a high priority user and at the same time providing a minimum level of data rate to all users. The objective function is formulated as the logarithm of the normalised weighted sum of the different services total data rate:

$$f_{VRRM}(R^{srv}) = \sum_{v_s=1}^{N^{srv}} \lambda_{v_s} \log \frac{R_{v_s}^{srv} [\text{Mbps}]}{R_{cell} [\text{Mbps}]} \quad (3.3)$$

where:

- R^{srv} : Vector of serving data rates, which can be written as $R^{srv} = [R_1^{srv}, \dots, R_{N^{srv}}^{srv}]^T$.
- N^{srv} : Number of services.
- λ_{v_s} : Tuning weight associated with service s , provided by VNO v , to prioritise data rate assignment.
- $R_{v_s}^{srv} [\text{Mbps}]$: Total served data rate of service.
- $R_{cell} [\text{Mbps}]$: Total available capacity at the cell.

In order to solve this optimisation problem, one needs to find the maximum of the objective function subject to inequality constraints. A standard technique based on Lagrange multipliers is used to do so. One also needs to put an upper bound to the total assigned data rate, which is the maximum cell data rate. Since both objective and constraint have continuous first partial derivatives, a new variable, the Lagrange multiplier, is introduced to form the $L(R^{srv}, \mu_L)$ Lagrangian:

$$L(R^{srv}, \mu_L) = \sum_{v_s=1}^{N^{srv}} \lambda_{v_s} \log \frac{R_{v_s}^{srv} [\text{Mbps}]}{R_{cell} [\text{Mbps}]} + \mu_L \left(1 - \sum_{v_s=1}^{N^{srv}} \frac{R_{v_s}^{srv} [\text{Mbps}]}{R_{cell} [\text{Mbps}]} \right) \quad (3.4)$$

where:

- μ : Lagrange multiplier corresponding to the inequality constraint.

by taking the derivatives with respect to the variables, one obtains:

$$\begin{aligned} \left\{ \frac{\partial L(R^{srv}, \mu_L)}{\partial R_{v_s}^{srv}} = 0 \quad \frac{\partial L(R^{srv}, \mu_L)}{\partial \mu_L} = 0 \right\} &\rightarrow \left\{ \frac{\lambda_{v_s}}{R_{v_s}^{srv} [\text{Mbps}]} = \mu_L \quad v_s \right. \\ &\in \{1, \dots, N^{srv}\} \quad R_{cell}^{cell} [\text{Mbps}] - \sum_{v_s=1}^{N^{srv}} R_{v_s}^{srv} [\text{Mbps}] = 0 \end{aligned} \quad (3.5)$$

Then, by solving the two equations one obtains the allocated data rate of each service proportional to its serving weight:

$$R_{v_s}^{srv} [\text{Mbps}] = \frac{\lambda_{v_s}}{\sum_{v_s=1}^{N^{srv}} \lambda_{v_s}} R_{cell} [\text{Mbps}] \quad (3.6)$$

Therefore, (3.3) is rewritten as (3.7) in order to further differentiate users' weights in each slice. Then the goal becomes to find the vector w^{usr} that maximises f_{VRRM} :

$$\text{Max } f_{VRRM}(w^{usr}) = \text{Max } \sum_{v_s=1}^{N^{srv}} \lambda_{v_s} \log \left(\sum_{v_s=1}^{N^{usr}} w_{v_s,i}^{usr} \frac{R_{v_s}^{srv_{max}} [\text{Mbps}]}{R_{cell} [\text{Mbps}]} \right) \quad (3.7)$$

where:

- w^{usr} : Vector of users' weights, to obtain the long-term average data rate of users, which can

be written as $\mathbf{w}^{usr} = [w_{1,1}^{usr}, \dots, w_{N_1^{usr}}^{usr}, \dots, w_{N^{srv},1}^{usr}, \dots, w_{N^{srv},N^{usr}}^{usr}]$

- $N_{v_s}^{usr}$: Number of users performing service s , from VNO v .
- $R_{v_s}^{srvmax}$ [Mbps]: Maximum assignable data rate to the user of service s , from VNO v .
- $w_{v_s,i}^{usr}$: Assigned weight to user i , performing service s , from VNO v , ranging in $[0,1]$.

Regarding the tuning weight, this serves the purpose of prioritising data rate assignment to each service. This parameter also isolates InP policies from VNOs decisions for capacity sharing. It is the combination of two independent positive integer numbers: γ is defined by InP and assigned to VNO v according to the type of its SLA agreements to VNOs' priorities in capacity sharing; δ is a serving weight, assigned to service s , performed by VNO v , to project the internal policy of each VNO in distributing capacity among the services provided by that VNO [Rouz19]. This way, the higher the serving weight number, the higher its priority is regarding capacity allocation and the lowest value provided by each VNO is always 1. The services priority can be related to their QCI values.

$$\lambda_{v_s} = \gamma_v \delta_s \quad (3.8)$$

There are two constraints associated with the problem of VRRM and the objective function has to be solved respecting these constraints. The first constraint considers that the average long-term data rate assigned to each user has to fall within this acceptable data rate interval due to VNO policies:

$$R_{v_s}^{srvmin} \leq w_{v_s,i}^{usr} R_{v_s}^{srvmax} \leq R_{v_s}^{srvmax} \quad (3.9)$$

where:

- $R_{v_s}^{srvmin}$ [Mbps]: Minimum assignable data rate to the user of service s , from VNO v .

The second constraint is a logical constraint, which indicates that the whole bandwidth allocated to all users cannot exceed the total aggregated cell capacity. Therefore, the entire VRRM bandwidth assigned to all users are subject to an upper bound defined by the InP:

$$\sum_{v_s=1}^{N^{srv}} \sum_{i=1}^{N_{v_s}^{usr}} w_{v_s,i}^{usr} R_{v_s}^{srvmax} \leq R_{cell} \text{ [Mbps]} \quad (3.10)$$

3.5. Aggregation Process

3.5.1. Latency

The E2E latency is based on the delay of packet transmission through the network. Two extreme scenarios are considered: one without the implementation of MEC that takes C-RAN with independent RU-DU-CU, Core backhaul, core network, and external data centre delays into account, whose delay contribution to the network is presented in (3.11). In this second scenario, information does not go to

the CN and takes just the C-RAN, MEC backhaul, and the MEC processing delays into account, whose delay contribution in the network is presented in (3.12). In this scenario, C-RAN may be collocated, depending on the available nodes distance and latency requirements. Figure 3.4 illustrates the delay contributions from the different nodes and links of the network. This model was developed by [Domi19].

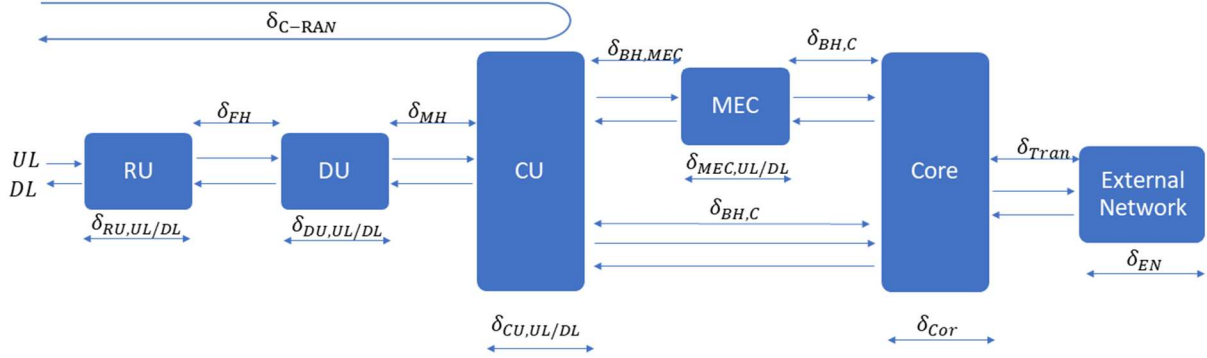


Figure 3.4. Delay contributions on the network (extracted from [Domi19]).

$$\delta_{E2E}[ms] = \delta_{C-RAN}[ms] + 2\delta_{BH,C}[ms] + \delta_{Cor}[ms] + \delta_{Tran}[ms] + \delta_{EN}[ms] \quad (3.11)$$

$$\delta_{E2E}[ms] = \delta_{C-RAN}[ms] + 2\delta_{BH,MEC}[ms] + \delta_{MEC,UL/DL}[ms] \quad (3.12)$$

where:

- δ_{E2E} - End to End Latency.
- δ_{C-RAN} - C-RAN associated Latency.
- $\delta_{BH,C}$ - Backhaul to core transmission Latency.
- $\delta_{BH,MEC}$ - Backhaul to MEC transmission Latency.
- δ_{Cor} - Core processing delay.
- δ_{Tran} - Transport transmission delay from the core to the Internet data centres.
- δ_{EN} - External Data centre contribution delay.
- $\delta_{MEC,UL/DL}$ - MEC processing delay.

The C-RAN delay represents the latency contribution from the network edge, delay contributions coming from the RU, DU, and CU processing delays and the transmissions ones from FH and MH.

$$\begin{aligned} \delta_{C-RAN}[ms] = & \delta_{RU,UL}[ms] + \delta_{RU,DL}[ms] + 2\delta_{FH}[ms] + 2\delta_{MH}[ms] + \delta_{DU,UL}[ms] + \delta_{CU,UL}[ms] + \\ & \delta_{DU,DL}[ms] + \delta_{CU,DL}[ms] \end{aligned} \quad (3.13)$$

where:

- $\delta_{RU,UL/DL}$ - RU processing delay on UL and DL.
- δ_{FH} - Transmission delay between the RU to the DU.

- $\delta_{DU,UL/DL}$ - DU processing delay on UL and DL.
- δ_{MH} - Transmission delay between the DU to the CU.
- $\delta_{CU,UL/DL}$ - CU processing delay on UL and DL.

The processing delay, (3.14), in the nodes depends on two factors: first, the delay from the process of the BS function, which is directly related to the number of functions that are addressed in the node; second, the delay from the admission control and delay process. It is also important to note that in Unlicensed Operation it is necessary to account for the processing delay regarding channel assessment and/or aggregated carriers in LAA. The implemented model does not consider Delay from Admission Control as it would be relevant for a dynamic and not static delay process.

$$\delta_{Node,UL/DL[ms]} = \delta_{Node,proc[ms]} + \delta_{N,admis[ms]} \quad (3.14)$$

where:

- δ_{Node} - Processing delay in the node.
- $\delta_{Node,proc}$ - BS function processing delay in the node.
- $\delta_{Node,admission}$ – Delay from Admission Control and Delay Process.

The C-RAN latency is limited by two factors: application latency and HARQ protocol requirements. In this work, HARQ is implemented in the DU so the retransmission process restriction needs to be taken into account, considering a retransmission maximum latency of 3 ms,

$$\begin{aligned} \{\delta_{C-RAN[ms]} < \delta_{HARQ[ms]}, \quad & \text{if } \delta_{HARQ[ms]} < \delta_{App[ms]} \\ \delta_{C-RAN[ms]} < \delta_{App[ms]}, \quad & \text{if } \delta_{HARQ[ms]} > \delta_{App[ms]} \} \end{aligned} \quad (3.15)$$

where:

- δ_{HARQ} - HARQ protocol requirement latency.
- δ_{App} - Maximum latency depending on what application is chosen.

The latency of the network is essential to determine the length of the links in the network. The distance of an E2E communication is determined by the time between application delay requirements and network delay, expressed by

$$d_{E2E[km]} = (\delta_{App[ms]} - \delta_{E2E[ms]}) \frac{v[m/s]}{2} \quad (3.16)$$

where:

- d_{E2E} - Maximum E2E distance.
- v - Propagation speed in the link is $2 * 10^8$ m/s, as only fibre links are considered.

The total distance in the network is divided into four parts:

$$d_{[km]} = d_{FH[km]} + d_{MH[km]} + d_{BH[km]} + d_{Tran[km]} \quad (3.17)$$

where:

- d_{FH} - Fronthaul maximum distance.
- d_{MH} - Middlehaul maximum distance.
- d_{BH} - Backhaul distance.
- d_{Tran} - Distance between the core and external data centre.

3.5.2. Node Processing Power

In order to achieve the maximum network performance, it is important to balance the processing capacity among RU, DU and CU specific for each use case requirements. The processing power in the node is one of the two parameters that define the node processing capacity, and processing power requirements are directly correlated with the splitting option of the BS function, so it is important to analyse the processing required for each BS function, measured in GOPS, being based on [DDLo15] and [Domi19].

The model presented in [DDLo15] estimates the processing power used in each node instance (i.e. RU, DU, CU, MEC, and CN) for DL and UL, considering the multiple physical layer functions processing power, the processing power associated with the data flow management and system control of the MAC and RLC layers, the processing of the PDCL, and the processing power used for the transmission to the core network:

$$P_t [\text{GOPS}] = P_{RF} [\text{GOPS}] + P_{PHY} [\text{GOPS}] + P_{MAC} [\text{GOPS}] + P_{RLC} [\text{GOPS}] + P_{PDCL} [\text{GOPS}] + P_{BH} [\text{GOPS}] \quad (3.18)$$

where:

- P_t - Total processing power required for each node.
- P_{RF} - Processing power required for the RF front-end.
- P_{PHY} - Processing power required for the physical layer functions.
- P_{MAC} - Processing power required for the MAC layer.
- P_{RLC} - Processing power required for the RLC layer.
- P_{PDCL} - Processing power required for the PDCL layer.
- P_{BH} - Processing power required for the backhaul interface depending on the data rate.

The processing power of the physical layer depends on the complexity of the multiple digital processing components:

$$P_{PHY} [\text{GOPS}] = P_{OFDM} [\text{GOPS}] + P_{MAP} [\text{GOPS}] + P_{MIMO} [\text{GOPS}] + P_{BBm} [\text{GOPS}] + P_{Code} [\text{GOPS}] \quad (3.19)$$

where:

- P_{OFDM} - Frequency domains function for OFDM modulation processing component including FFT and IFFT.
- P_{Map} - Mapping and demapping functions processing component.
- P_{MIMO} - MIMO encoding/decoding processing component.
- P_{BBm} - Baseband modulation/demodulation processing component.
- P_{Code} - FEC function processing component.

The processing power associated with each component can be calculated by:

$$P = P_{ref} \left(\frac{B_{[MHz]}}{B_{ref[MHz]}} \right)^{e1} \left(\frac{E_{[bps/Hz]}}{E_{ref[bps/Hz]}} \right)^{e2} \left(\frac{N_A}{N_{A,ref}} \right)^{e3} \left(\frac{F_{DC[\%]}}{F_{DC,ref[\%]}} \right)^{e4} \left(\frac{N_{streams}}{N_{streams,ref}} \right)^{e5} \left(\frac{N_Q[bits]}{N_{Q,ref[bits]}} \right)^{e6} \quad (3.20)$$

where:

- P_{ref} - Complexity associated with each function, measured in GOPS.
- B - Bandwidth used in the BS.
- B_{ref} - Reference bandwidth used in the BS.
- E - Spectral efficiency dependent on the modulation and coding rate used.
- E_{ref} - Reference spectral efficiency dependent on the modulation and coding rate used.
- N_A - Number of antennas in the BS.
- $N_{A,ref}$ - Reference number of antennas in the BS.
- F_{DC} - System load in the frequency-domain.
- $F_{DC,ref}$ - Reference system load in the frequency-domain.
- N_{strea} - Number of transmission streams, up to the number of antennas.
- $N_{streams,ref}$ - Reference number of transmission streams, up to the number of antennas.
- N_Q - Number of bits used in quantisation.
- $N_{Q,ref}$ - Reference number of bits used in quantisation.

The processing capacity in each node can be computed from:

$$P_{RU} [GOPS] = \sum_1^N P_{i[GOPS]} \quad (3.21)$$

$$P_{DU} [GOPS] = \sum_1^{N_{RU} \text{ Connected}} \sum_1^N P_{i[GOPS]} \quad (3.22)$$

$$P_{CU} [GOPS] = \sum_1^{N_{DU} \text{ connected}} \sum_1^{N_{RU} \text{ Connected}} \sum_1^N P_{i[GOPS]} \quad (3.23)$$

$$P_{MEC/CN} [GOPS] = \sum_1^{N_{CU} \text{ connected}} \sum_1^{N_{DU} \text{ connected}} \sum_1^{N_{RU} \text{ Connected}} \sum_1^N P_{i[GOPS]} \quad (3.24)$$

where:

- P_{RU} - Processing power used by the RU.
- P_{DU} - Processing power used by the DU.
- P_{CU} - Processing power used by the CU.
- $P_{MEC/CN}$ - Processing power used by the MEC or CN.
- P_i - Function i assign to the node.

One considers that the total processing power is always divided among nodes, without existing any additional process required:

$$P_t [GOPS] = P_{RU} [GOPS] + P_{DU} [GOPS] + P_{CU} [GOPS] + P_{MEC/CN} [GOPS] \quad (3.25)$$

The calculation of the load of the aggregation node is based on the functions processing power

assigned by each connected node and a fixed component independent of the number of connected nodes required for scheduling and signalling:

$$\mu_{Node} = \frac{P_{fix,Node[GOPS]} + P_{Node[GOPS]}}{P_{Node,Cap[GOPS]}} \quad (3.26)$$

where:

- μ_{Node} - Node load.
- $P_{Node,fix}$ - Fixed processing power required for scheduling and signalling, independent of the number of connected nodes.
- P_{Node} - Processing power on the aggregation node assigned by the connected nodes.
- $P_{Node,Cap}$ - Processing capacity assigned to the aggregation node.

In order to analyse the impact of the load of the node on network performance, one considers a multiplier factor of the processing capacity assign to the nodes:

$$P_{Node,Cap[GOPS]} = P_{Node,Cap,ref[GOPS]} M_P \quad (3.27)$$

where:

- $P_{Node,Cap,ref}$ – Reference processing capacity on the node.
- M_P – Processing capacity multiplier.

3.5.3. Aggregation Process Workflow

Private Networks may be deployed as three different functional architectures, as described in Chapter 2. Therefore, the model must also consider these architectures when analysing Private Network's performance and feasibility. In that sense, the three architectures are:

- **Standalone Deployment:** All network nodes are located inside the premises of the private network operator, which stands for a collocated RU-DU-CU architecture. This architecture makes it possible to use the Unlicensed Spectrum, considering interference from Non-3GPP Radio Access Technologies. In order to analyse this private network deployment, one considers the usual VRRM optimisation but with only one VNO.
- **Public-Private Shared RAN Deployment:** In this case the RU belongs to the Operator and uses Licensed Spectrum. Less capacity is available, therefore not all Private Networks may be possible to deploy. The RU node is located outside the Private Network Operator premises, but DU or CU can be located inside it. Therefore, one may have Collocated CU-DU Independent RU or Independent RU-CU-DU. In order to analyse this architecture, only a given configuration presents higher available capacity with the possibility of delivering ~1 ms of end-to-end latency, and only a given percentage of the maximum cell data rate will be available.
- **Shared RAN and Control-plane Deployment:** This is the case in which Network Slicing is used to separate Private and Public Networks. RAN is shared and Control-Plane network

functions are all handled by the licensed operator. Therefore, capacity is shared with the public network and later allocated through VRRM. The architecture depends exclusively on the licensed operator. To analyse this architecture, the usual VRRM optimisation process is done.

After choosing the Private Network architecture, the aggregation process happens accordingly, choosing the C-RAN architecture and then aggregating the nodes. Based on [Domi19], a minimise delay algorithm is used, along with the balance number of connections algorithm. The first one aims to aggregate the closest nodes with the required processing capacity, the model analyses all the possible connection distances and choose the smallest one. The second algorithm balances the number of aggregations for each node, checking the nodes capacity in order to aggregate new ones. In addition to that, considering that this work focuses on 5G and that the splitting options are not under study, only fibre links are made, and the default splitting option is 7.1. Figure 3.5 shows the workflow of this stage.

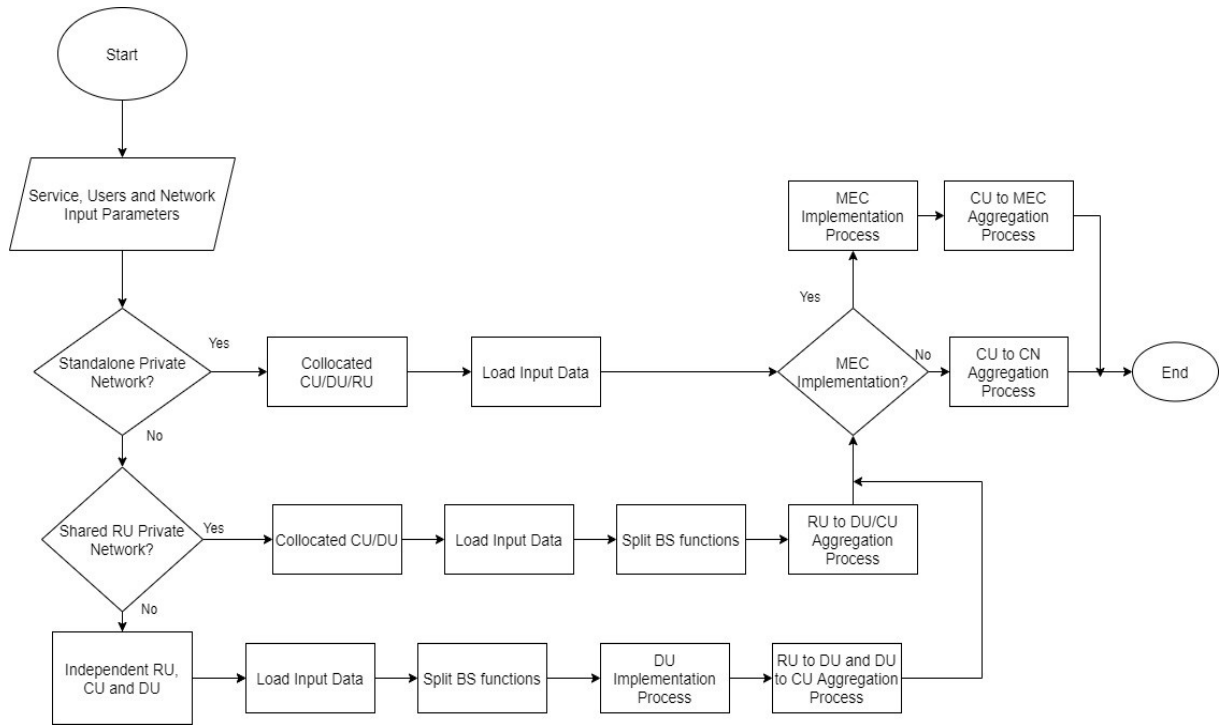


Figure 3.5 - Aggregation Process Workflow.

3.6. Bandwidth

As described in Chapter 2, Private Networks may be deployed under the Unlicensed Spectrum, through a Standalone or LAA operation. In addition to that, some services may have specific frequency bands, such as the Industry, Scientific and Medical (ISM) band. Therefore, one needs to choose appropriate numerology values, according to the band and service. The available frequency bands both from Operators and Unlicensed Spectrum are listed below:

Table 3.4 – Available frequency bands.

Designation	Band [MHz]	Available Numerologies	Harmonisation	Bandwidth [MHz]	Duplexing
ISM	2400-2500,	0, 1, 2	CCA	100	TDD
	5725-5830,	0, 1, 2		150	TDD
	5850-5925	0, 1, 2			TDD
Military	862-880,	0, 1	No CCA	9	FDD
	915-925	0, 1		5	FDD
ITS	5850-6700	0, 1, 2	CCA	30	TDD
Railways	876-880	0, 1	No CCA	4	FDD
	921-925,	0, 1		4	FDD
Operators	700-710	0, 1	No CCA	10	FDD
	1850-1910	0, 1, 2		20	FDD
	2010-2025	0, 1, 2		15	TDD
	2620-2690	0, 1, 2		20	FDD
	3500-3600	0, 1, 2		100	TDD

The goal is to provide low latency and high data rates. Regarding latency, one wants to lower the Transmission Time Interval (TTI), which corresponds to the transmission of one slot. Therefore, in order to achieve low TTI, highest available numerologies may be used.

But then there is the issue of ISI and ICI. Numerology selection is of great importance in this sense, as the subcarrier spacing defines the Cyclic Prefix (CP) length and insufficient CP causes ISI and ICI. [VaMA21] quantifies interference and analyses which numerologies can be used in different scenario topologies. The scenarios considered there and in this work are: Indoor Office (IO), Urban Micro (UMi), Urban Macro (UMa), Rural Macro (RMa) and UMi/UMa Outdoor to Indoor (O2I). A scaling factor, that is multiplied by a normalized Delay Spread is defined for each scenario topology and shown on Table 3.5. These values were obtained from several measurements and estimations.

Table 3.5 – Scaling Factor for Delay Spread calculation for the frequencies and topologies considered (adapted from [VaMA21]).

Scenarios' topologies		Frequency [GHz]	
		2	6
Indoor office	Median Delay Spread	39	30
	90 th percentile RMS delay spread.	59	53
Umi Street-canyon	Median Delay Spread	129	93
	90 th percentile RMS delay spread.	634	316
Uma	Median Delay Spread	363	363
	90 th percentile RMS delay spread.	1148	1148
RMa	Median Delay Spread	37	37
	90 th percentile RMS delay spread.	153	153
UMi/Uma O2I	Median Delay Spread	240	240
	90 th percentile RMS delay spread.	616	616

3.8. User Experience and Network Performance

It is important to know if the overall user experience and network performance are working as intended. Several evaluation metrics are defined with the goal of measuring these performance requirements. The metrics used for network performance evaluation are as follows and were taken from [Mari20]:

- Percentage of total assigned data rate – shows the total network throughput in terms of data rate allocation. It is important to note that once this value goes below 100%, the network saturates and starts to delay users.

$$p_{VRRM}^{tot}[\%] = 100 \frac{\sum_{v_s=1}^{N^{srv}} \sum_{i=1}^{N_{v_s}^{usr}} w_{v_s,i}^{usr} R_{v_s}^{srvmax} [\text{Mbps}]}{R_{cell} [\text{Mbps}]} \quad (3.28)$$

- VRRM capacity share – the percentage of capacity allocated to each VNO, out of the total available VRRM one:

$$R_{VRRM}^{VNO_v}[\%] = 100 \frac{\sum_{v_s=1}^{N_{VNO_v}^{srv}} \sum_{i=1}^{N_{v_s}^{usr}} w_{v_s,i}^{usr} R_{v_s}^{srvmax} [\text{Mbps}]}{R_{cell} [\text{Mbps}]} \quad (3.29)$$

- Total data rate of each service – it shows the total data rate assigned to each service slice of a VNO.

$$R_{v_s}^{srvtot} [\text{Mbps}] = \sum_{i=1}^{N_k^{usr}} w_{v_s,i}^{usr} R_{v_s}^{srvmax} [\text{Mbps}] \quad (3.30)$$

- Percentage of served users – the percentage of served users performing a specific service out of the total number of users from that service.

$$p_{v_s}^{usrnet}[\%] = 100 \frac{N_{v_s}^{usr}}{N^{usrtot}} \quad (3.31)$$

The metrics that are important from a users' viewpoint are as follows:

- Data rate of each user – the data rate allocated to a user is an important QoS metric from both users' and VNOs' viewpoints, having a direct impact on the satisfaction level of the served users:

$$R_{v_s,i}^{usr} [\text{Mbps}] = w_{v_s,i}^{usr} R_{v_s}^{srvmax} [\text{Mbps}] \quad (3.32)$$

- Users' satisfaction, $S_{v_s,i}^{usr}$ – it is important from a provider perspective because it reflects the user satisfaction. This metric is measured differently according to the service type. For voice, one uses AMR-WB codecs with the respective voice quality mean opinion score (MOS)

values provided by NOS, presented in Table 3.6. For the remaining services, a similar way of classification is used. According to the service data rate, it is defined five levels of user satisfaction, where just like MOS one represents bad quality and five represents excellent quality.

Table 3.6 – Voice codecs and respective MOS [Dini21].

Codec Name	Nominal MOS
AMR WB Mode 0 (6.6k)	3.39
AMR WB Mode 1 (8.85)	3.81
AMR WB Mode 2 (12.65)	4.04
AMR WB Mode 3 (14.25)	4.09
AMR WB Mode 4 (15.85)	4.11
AMR WB Mode 5 (18.25)	4.14
AMR WB Mode 6 (19.85)	4.18
AMR WB Mode 7 (23.05)	4.18
AMR WB Mode 8 (23.85)	4.18

3.9. Model Assessment

This subsection aims to validate the developed model by using a set of tests whose results are known and calculated. Table 3.7 describes the tests done.

As this model was adapted from [Mari20] and [Domi19], part of the model is already validated. It was necessary to verify the Numerology Selection and how it affects the model as a whole.

In order to confirm the Numerology Selection algorithm, a test scenario is defined. For that, cell input parameters are defined as follows on Table 3.8. In addition to that, one will only consider two services, as done by [Mari20]. Here these services are taken as Web Browsing and Remote Surgery, and will be used to verify the output parameters and their variations. These services were chosen in order to represent extreme cases, as can be seen from their SLA, Non-GBR and DelayC-GBR. The Service Input Parameters are presented in Table 3.9.

In order to assess and validate the model, the input parameters will suffer variations and the output parameters will be shown in order to verify the results. The first round of tests consists of changing the cell input parameters in order to verify the Cell Data Rate and how it affects the services and users. MIMO Layers can reach values of 2, 4 and 8, and these values cause the following response on the cell data rate, as shown in Figure 3.6.

Table 3.7 – List of model assessment tests.

Number	Description.
1	Validation of the input file read, by verifying if the type of variable is correct.
2	Validation of the input variables, by verifying if the parameters are correctly stored in memory.
3	Validation of the computation of the cell capacity: <ul style="list-style-type: none"> • Check the Numerology Selection. • Check the computation of T_s^μ. • Check if the overhead, the number of resource blocks, and the modulation order are correct. • Check if the computation R_{cell} is correct.
4	Validation of the admission control and delay process: <ul style="list-style-type: none"> • Check if the available capacity is not enough to serve all users with minimum demanded data rate. • Check if all BE users are delayed. • Check if the rest of delayed users were delayed based on their slice and service priorities. Check if the new minimum demanded capacity is equal or approximate to the previously calculated cell capacity.
5	Validation of the VRRM optimisation: <ul style="list-style-type: none"> • Check if the size and values of the created array with the minimum demanded capacity for all users are correct. • Check if the CVX status is solved. • Check if the computed values of the users' weights give the optimal solution to the problem. • Check if the imposed restrictions are being complied with.
6	Validation of Latency Contributions: <ul style="list-style-type: none"> • Check if the number of unserved users is correct.
7	Validation of Private Networks deployment: <ul style="list-style-type: none"> • Check if the Private Network deployment corresponds to the input file • Verify that the C-RAN architecture of the Private Networks deployment corresponds to the scenario.
8	Validation of the output files, by checking if they are correctly printed and plotting the output results.

Table 3.8 – Cell Input Parameters.

Link	MIMO Layers	Bandwidth [MHz]	Scenario Topology	Frequency Band [GHz]
Downlink	4	100	Indoor Office	3.6

Table 3.9 – Service Input Parameters

Parameters	Web Browsing	Remote Surgery
SLA	Non-GBR	DelayC-GBR
QCI Value	6	85
$R_{\min} [Mbps]$	0.5	500
$R_{\max} [Mbps]$	$R_{\text{cell}} [Mbps]$	$R_{\text{cell}} [Mbps]$
λ	1	1
Number of Users	1	1

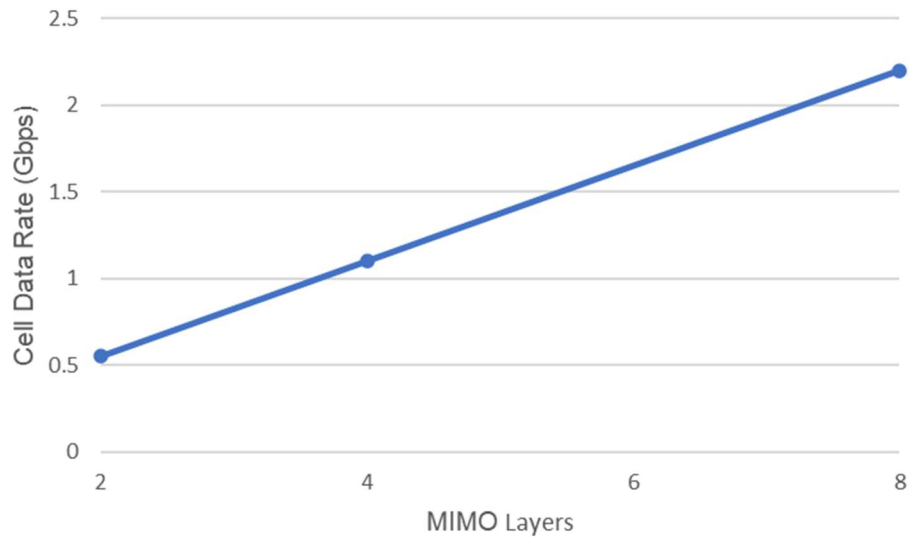


Figure 3.6 - Cell Data Rate with possible MIMO Layers.

The Cell Data Rate is one of the most important parameters in order to validate the model, as it dictates the available capacity for all users and services. This can be seen in Figure 3.7, where the VRRM Capacity Share is given according to the MIMO Layers variation.

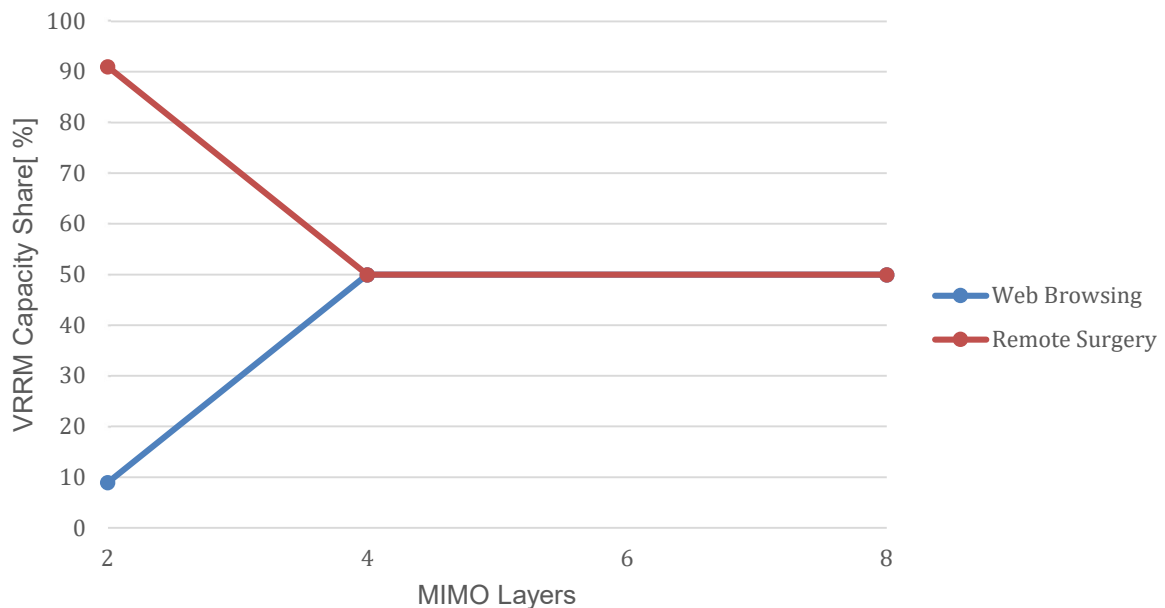


Figure 3.7 - VRRM Capacity Share with the different MIMO Layers possibilities.

It can be stated that the VRRM model works accordingly to what was expected. For 2 MIMO Layers the Cell Data Rate was enough to attend Remote Surgery at a minimum data rate, and the Web Browsing at a rate of 49 Mbps. From 4 MIMO layers, the Cell Data Rate reaches 1 Gbps, and with that the model verifies that it is possible to serve both services with the minimum data rate, therefore it starts to split the capacity equally. Following that, one varies the other cell input parameters. It is also important to verify the Numerology Selection for each Scenario Topology. Figure 3.8 shows the

Numerology selected for the different Scenario Topologies. It validates the Numerology Selection model as all Numerologies are selected according to what was expected from subsection 3.2.

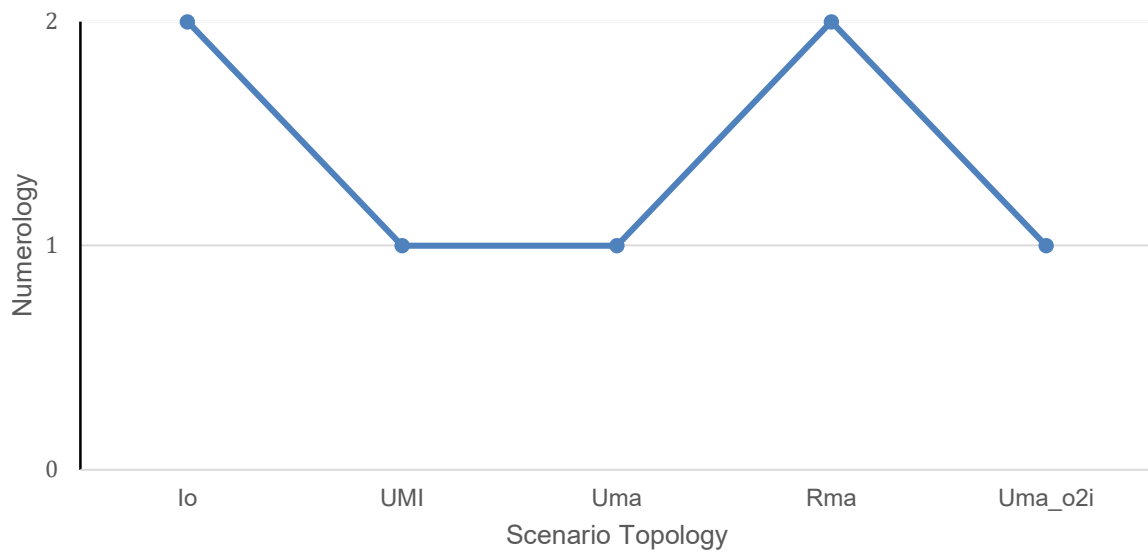


Figure 3.8 - Numerology Selected according to the Scenario Topology.

Finally, one varies the Frequency Band and Bandwidth parameters to validate the Cell Model as a whole. Figure 3.9 shows the Cell Data Rate variation according to the Bandwidth and consequently the Frequency Band.

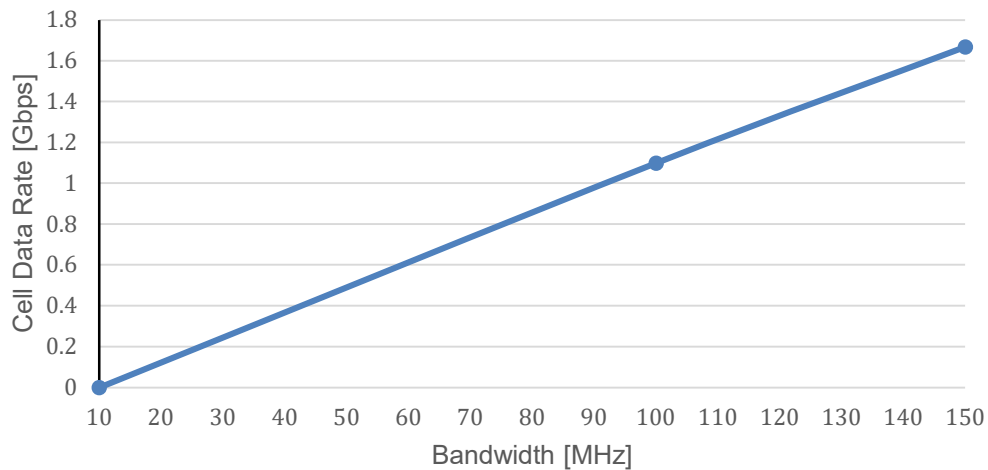


Figure 3.9 - Cell Data Rate response to Bandwidth variation.

Figure 3.10 shows us how the VRRM model manages the different Cell Data Rate received, therefore the other Cell Input Parameters variation shows only the Cell Data Rate, as the behaviour of the VRRM model will be the same, given the variation of capacity.

The next step is to change the number of users of the Remote Surgery in order to overload the network and verify how the model handles capacity distribution. Figure 3.10. shows the model behaviour when increasing the number of Remote Surgery users one by one. As the Remote Surgery service demands a minimum of 500 Mbps, when one user of that service is added the Capacity Share of the Web Browsing VNO decreases around 40%. When that number of users goes to 3, the Web Browsing user and one remote surgery user are delayed. This translates into increasing the latency for these 2 users at least 0.25 ms, which is due to the Sub-Carrier Spacing of 60 kHz that comprises a TTI of 0.25 ms. With that, both the VRRM model and the delay process were verified. In addition to that, it is important to note the saturation of the network with 3 Remote Surgery users. With the delay of the Web Browsing and Remote Surgery users, the algorithm delivers 500 Mbps of minimum data rate for the remaining users. With that 108 Mbps are unused, and therefore only 91.02% of the available capacity is used.

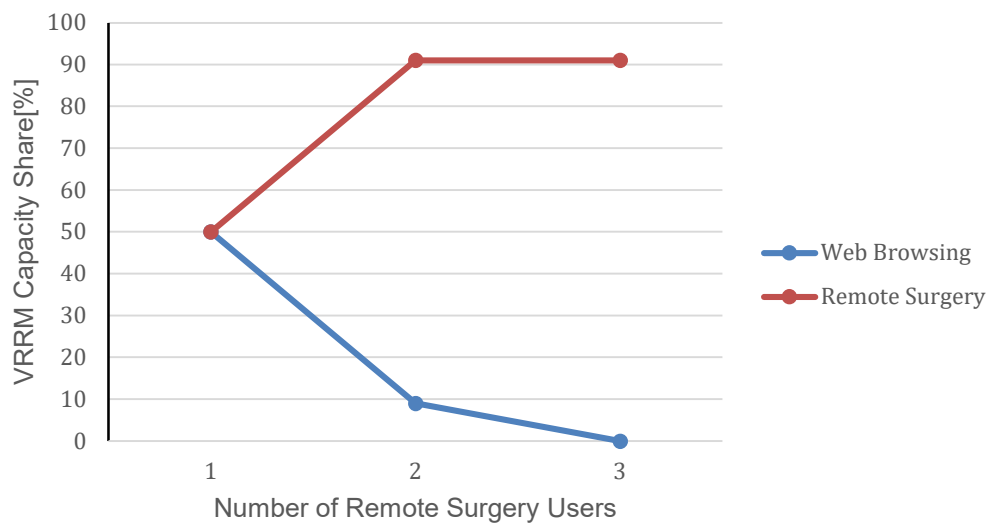


Figure 3.10 – VRRM Capacity Share with Remote Surgery of users' variation.

After that, two different network architectures were tested in order to confirm the lower latency in each C-RAN Architecture so that they can be linked to each private network architecture. The aggregation process connects all the available nodes that attend the latency requirements, and in this case the Portugal network provided by NOS is considered. The latency algorithm analyses one service at a time; therefore, the chosen service was Factory Automation, in order to restrict the latency requirements. With that three C-RAN architectures were tested, being them: Collocated RU-DU-CU, Independent RU with Collocated DU-CU and Independent RU+DU+CU. The results are presented in Table 3.11.

Even though the median total link delay is lower in the Independent RU, Collocated DU-CU architecture than Collocated RUDUCU, the number of links and the total distance is larger on the

second. And if one divides the total link delay for the total distance, one will get lower Delay/Distance ratio at the Collocated RUDUCU scenario, compared to all of the others.

Table 3.10 – Latency Results.

	Number of Links	Distance [km]	Total Network Delay [ms]
Collocated RUDUCU	5510	186.51	0.95
Independent RU, collocated DU-CU	2047	158.41	0.84
Independent RU+DU+CU	2577	171.71	0.96

Chapter 4

Results Analysis

This chapter provides the description of the reference scenario and all the variations made to it. Then it provides the results obtained by the developed model and a study of all these variations.

4.1. Reference Scenarios

The reference scenarios, chosen in collaboration with NOS, consider the main private network use cases: Hospital, Smart Factory and Mission Critical. All of these use cases are analysed in different scenarios, considering relevant frequency bands and number of users. The Mix parameter shows the users' distribution among services. The SLAs are defined as recommended by 3GPP and presented in Chapter 3.

Table 4.1 – Services with Mix, SLA and VRRM parameters.

Sector	Service	γ_s	δ_s	SLA	Mix Hospital [%]	Mix Smart Factory [%]	Mix Mission Critical [%]
Public	Voice	10	10	GBR	30	5	10
	Video		8		30	5	15
	Music		9		9	5	5
	Social networking		5	Non-GBR	10	5	5
	Web browsing		6		8	5	5
	File Sharing		4		5	5	5
Emergency	Mission Critical Video	40	9	GBR	-	-	15
	MCPTT		10		-	-	30
	Drone Supervision		8		-	-	5
	VR/AR	40	10		5	5	5
Hospital	Service Robots		11	DelayC-GBR	2	-	-
	Remote surgery		12		1	-	-
Industry	Closed Loop Control	40	11		-	15	-
	Monitoring		8		-	25	-
	Motion control		10		-	15	-
	Safety Control		9		-	10	-

The values of γ_s and δ_s are based on [Mari20]. The data rates of the private network services are taken from the table at Annex D and the public ones from [Mari20]. Table 4.2 represents the services and their respective data rates.

Table 4.2. Service's Data Rates

Service	Class	Data rate [Mbps]
Voice/MCPTT	Conversational	[0.0066, 0.024]
Music	Streaming	[0.015, 0.32]
Web browsing (Web)	Interactive	[0.5, R_{cell}]
File Sharing (FS)	Interactive	[1, R_{cell}]
Social networking (SN)	Interactive	[2, R_{cell}]
Video/ Mission Critical	Streaming	[2, 13]
VR/AR	Streaming	[50, R_{cell}]
Drone Supervision	Streaming	[20, R_{cell}]
Closed Loop Control	Background	[1, 5]
Monitoring	Background	[0.1, 0.5]
Safety Control	Background	[0.5, 1]
Motion Control	Background	[1, 5]
Remote surgery (RS)	Interactive	[150, 500]
Service Robots (SR)	Streaming	[60, 450]

Regarding the data rate calculation, as described in [Mari20] the maximum achievable cell data rate is the result of averaging the values of the maximum achievable cell data rate of all four considered modulations (QPSK, 16-, 64-, 256QAM). This means that $Q_m^{(j)}$ assumes the values of 2, 4, 6 and 8, and R_{max} will assume the respective values of 449/1024, 616/1024, 873/1024 and 948/1024. This is done so that the value is as close to reality as possible.

In addition to that, each service has a VRRM priority given by their Tuning Weight, as depicted in equation (3.8). Figure 4.1. shows all of the service's tuning weight, which translates to their priority in terms of radio resources allocation. The services are grouped based on their SLA.

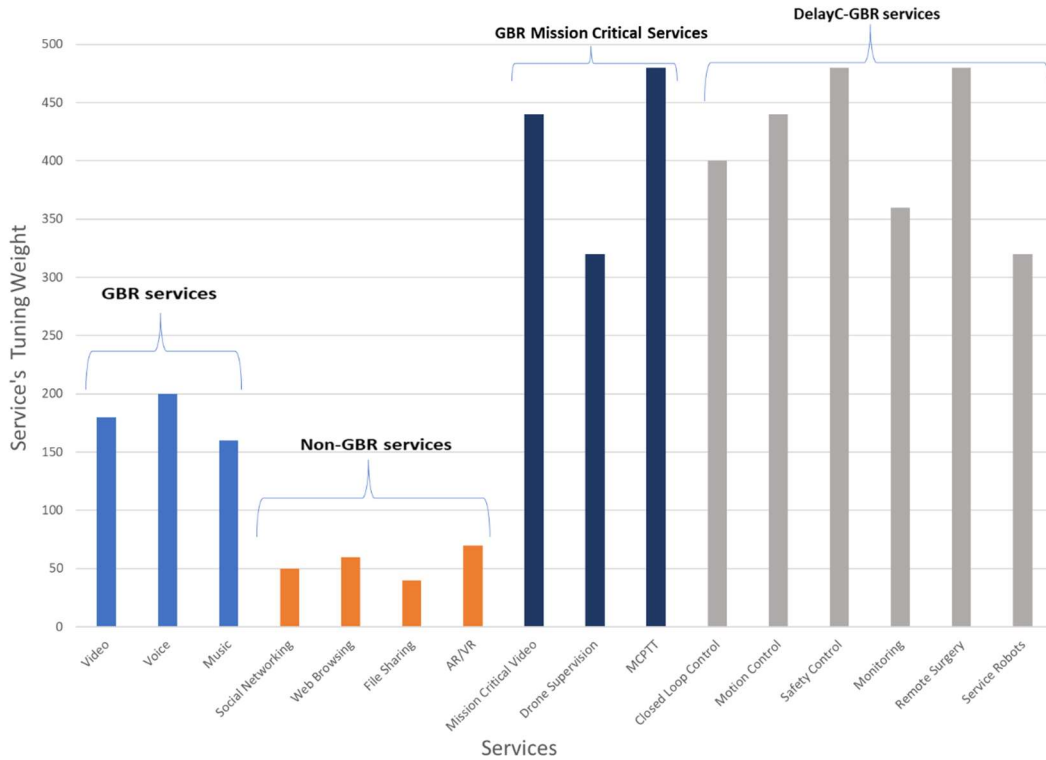


Figure 4.1 - Service's Tuning Weight.

The hospital scenario is divided into two sub-scenarios: Low and High demand (LD and HD). The most demanding application in the Hospital Scenario is the 3D camera flow, present in the Service Robots and Remote Surgery services. It can get Data Rate from 137 Mbps to 1.6 Gbps [ZhLZ18], which depends on the range of vision. Considering surgeries that do not need the whole field of view and low complexity service robots, the low demand scenario has a minimum bit rate of 150 Mbps for the Remote Surgery and 60 Mbps for the Service Robots. The High demand scenario considers a minimum bit rate of 500 Mbps for the Remote Surgery and 150 Mbps as minimum bit rate for the Service Robots.

Regarding the UE's position, the Hospital and Smart Factory scenarios consider users inside the buildings, with the cell located at the city outside for the Network Slicing private deployment, and small cells of 250 meters inside the building for the Standalone Private Deployment. For the Mission Critical scenarios, the rural one considers users at a random rural location, with now buildings nearby and low population density. The urban mission critical scenario considers users in the streets of city of Porto surrounded by buildings.

The latency calculation is done once per scenario with the respective scenario's service mix. Two C-RAN Architectures were simulated and analysed, depending on the scenario. The latency model parameters are fixed as they were not approached in Chapter 3 and their analysis is not under the scope of this thesis, being shown in Table 4.3. The maximum allowed latency is defined by the DelayC-GBR services, as they are the strictest ones. The simulation aggregates nodes from Portugal and therefore gives, minimum, maximum, and mean latency.

Table 4.3. Network Input Parameters

Architecture	{RU-DU+CU; RU+DU+CU}
Splitting Option	{7.2}
RU nodes converted to DU nodes [%]	20
CU nodes converted to MEC nodes [%]	15
Node Processing Capacity multiplier	[1; 10 ²]
Usage and Penetration ratio [%]	{10;30}

4.2. Hospital Scenario Analysis

4.2.1. Hospital Network Slicing Results

This first Hospital Scenario considers a low demand Network Slicing cell. The 3.5 GHz frequency band is used as the network is public and deployed on 5G. The topology for this scenario is Uma_O2I,

which represents a hospital located in a city centre, receiving the signal from an antenna located outside. This 3.5 GHz band is analysed as a Network Slicing Private Network Deployment and is considered Low Demand. The cell input parameters are as follows:

Table 4.4. Cell Input Parameters for the Hospital Network Slicing scenario.

MIMO layers, $v_{Layers}^{(j)}$	{2,4,8}
Frequency Band [MHZ]	3500
Bandwidth [MHz]	100
Scenario Topology	Uma_O2I

This scenario considers one remote surgery and two service robots. The service mix used is the reference one and is adapted as the number of users is increased so that there is always the same number of remote surgeries and service robots. The first simulation uses the reference scenario mix and is considered as a Network Slicing Private Network Deployment. Because of that, one has AR/VR in the service mix, which greatly uses the network capacity. As expected, the VRRM optimisation distributes the data rate for each service accordingly to their tuning weight parameter. Starting with 100 users and incrementing their number by 50 until services stop being served. At 300 users, all Web Browsing, Social Networking and File Sharing users are delayed. 13,3% of the AR/VR users are delayed. Video reaches its minimum data rate at 250 users, and Voice and Music reach their minimum at 300 users, along with the Non-GBR services delay. All of the DelayC-GBR services are served with fixed data rate. AR/VR that are not delayed, are served with fixed data rate of 50 Mbps. Figure 4.2 shows the Service's Data rates against the number of users being served.

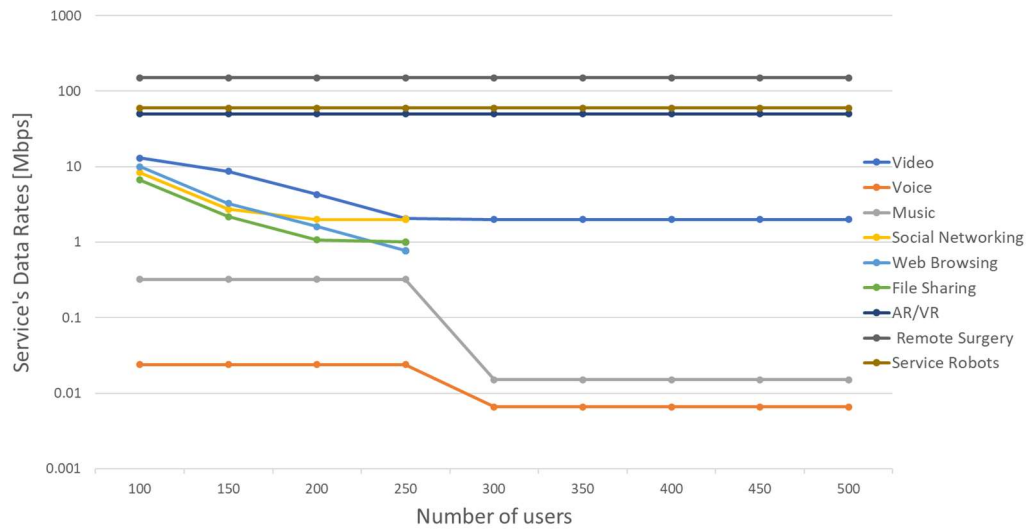


Figure. 4.2. Services Data rate with different number of users for the Hospital Network Slicing Private Deployment.

After that, one verifies the VRRM Capacity Share. As the DelayC-GBR services have fixed data rates, their capacity share percentage is also constant. With the increment of users, due to the Service Mix, from 150 until 250 users there is an increase on the Non-GBR Services capacity share and consequent decrease on the GBR services capacity share. At 300 users, with the delay of the Non-GBR services, the GBR capacity share increases in detriment of the Non-GBR one. Figure 4.3 shows the VRRM capacity share versus the Number of Users.

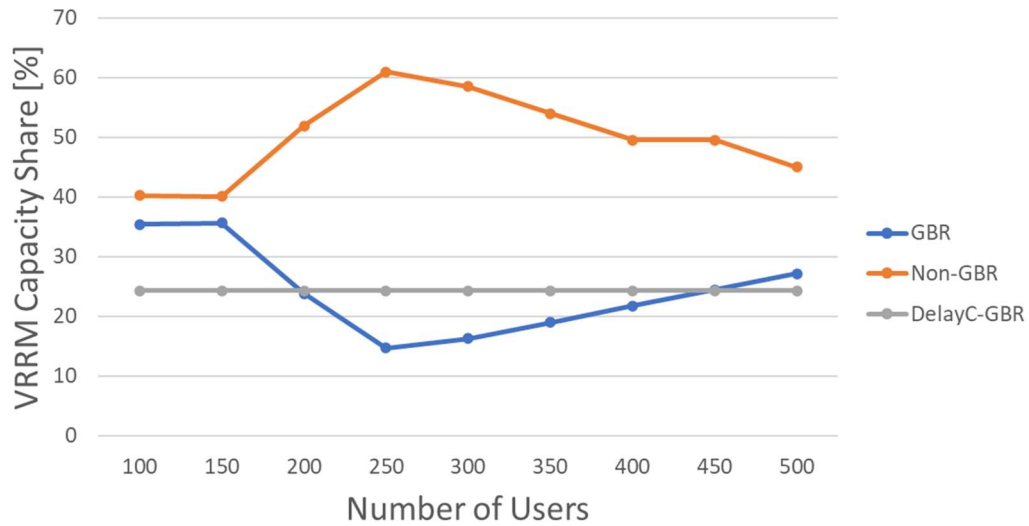


Figure. 4.3. VRRM Capacity Share with different number of users for the Hospital Network Slicing Private Deployment.

Furthermore, it is important to note that once the services are delayed, the network saturates and its network capacity is not fully used. With every increment of the number of users, the used capacity decreases until the 400 users mark. From 400 to 450 users, used capacity increases and then goes back down. Figure 4.4. shows the Used Capacity versus the Number of users on the cell.

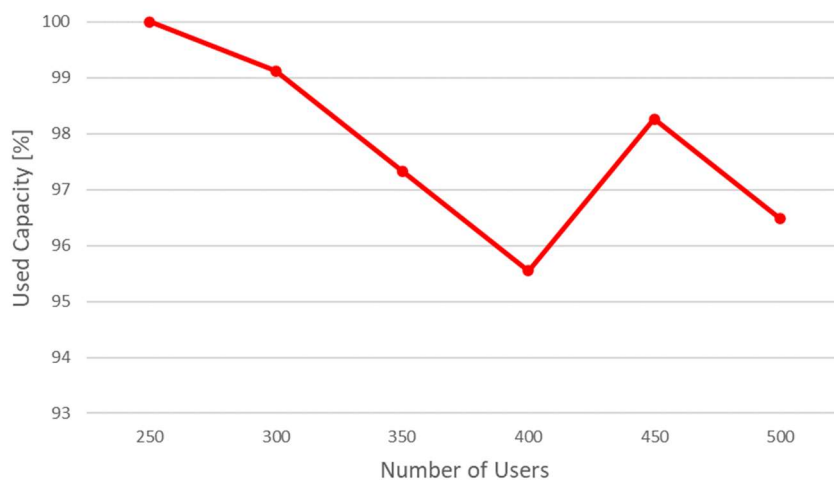


Figure. 4.4. Cell's used capacity versus the total number of users for the Hospital Network Slicing Private Deployment.

This phenomenon can be well explained by verifying the served users percentage. From 300 users the AR/VR users start to be delayed at a rate of almost 15% at each increment of 50 users. The rate of decrease changes from 400 to 450 users, at that point the AR/VR served users go from 55% to 50%. That translates into the same amount of 11 AR/VR users being served. Therefore, with the increase of total number of users, the GBR services occupy a larger share, increasing the used capacity at that number of users. Figure 4.4 shows the service's served users percentage versus the network's total number of users.

In addition to that, Figure E.1. shows each service's number of users versus the total number of users being served. It is important to note that Remote Surgery and the Service Robots are fixed with 1 and 2 users, respectively. In order to keep the mix pattern, the spare users that would go to the DelayC-GBR services are transferred to Voice, as it is a low demand service, therefore the network performance does not change.

Following the data rate and VRRM Capacity Share analysis, it is also important to verify the service's user's satisfaction. Figure E.2 starts with the GBR VNO User's Satisfaction. As Video has a higher range of data rate, its user's satisfaction decreases from 5 to 1 from 100 to 250 users. At 300 users Music and Voice reach their lowest User's Satisfaction, which is where the network saturates and Non-GBR services are delayed. Next, Figure E.3 shows the Non-GBR VNO user's satisfaction. AR/VR user's satisfaction is fixed at 5 due to it being always served with minimum data rate of 50 Mbps. Web browsing, File Sharing and Social Networking all have their user's satisfaction decreased until 2 before being delayed at 300 users and going to 0.

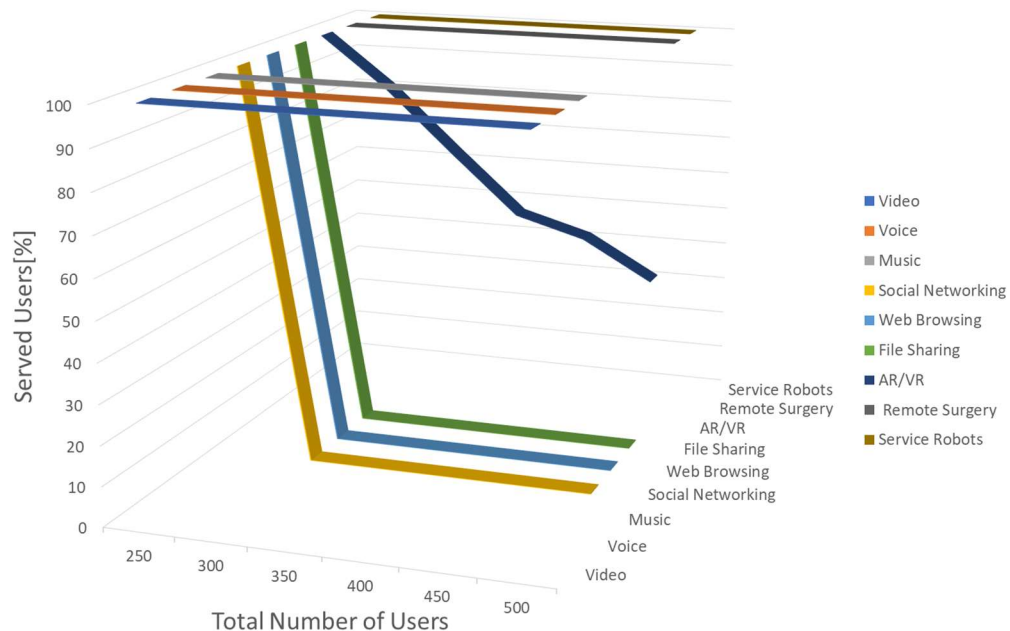


Figure. 4.5. Each service's served users versus the total number of users for the Hospital Network Slicing Private Deployment.

Regarding the DelayC-GBR services, as Remote Surgery is a critical and life dependant service, its minimum data rate is defined as the best one possible, therefore whenever it is served, user satisfaction is at its maximum. The same happens for the Service Robots, for that reason no chart of DelayC-GBR VNO User's Satisfaction is shown at this scenario.

4.2.2. Hospital Shared Ran Results

The next step is analysing the results of the Shared Ran scenario. This scenario considers a Shared Ran Private Deployment, where the Private Network uses a public RU and shares the capacity with the public. In order to analyse this scenario, a given percentage of the Cell Capacity was considered on the simulations, so that a Shared Ran scenario is represented. It is important to note that a Shared Ran private deployment serves only the hospital authorised users, being then employees or customers. Therefore, the service mix is adapted so that no AR/VR is present in the mix, as a hospital private network would have no need of this service. In this case, Network Slicing is used inside the Private Network so that each slice attends the specificities of each service. Table 4.6 shows the updated Service Mix.

Table 4.5. Service Mix for the Shared Ran Hospital Scenario.

Service	Mix [%]
Video	28
Voice	30
Music	10
Social Networking	10
Web Browsing	10
File Sharing	9
Remote Surgery	2
Service Robots	1

The analysis for this scenario consists of verifying the percentage of the Cell Capacity that could attend the proposed services. This simulation verifies each Cell Capacity percentage at a time and the number of users is fixed at 100. DelayC-GBR services start to be served at 20% of shared capacity, with all other services delayed. At 30%, GBR Services and Web Browsing are also served, but at their lowest data rate. From 40% of shared capacity Voice and Music are served with maximum data rate and all Non-GBR services are served. Until 80% Video and the Non-GBR services have their data rates increased. Voice and Music data rates do not increase as they do not require higher data rates for best user satisfaction. Figure 4.6 shows the Service's Data Rate versus the Shared RAN Percentage.

By verifying the User's satisfaction results, one can see that once the DelayC-GBR services are served, they have maximum user's satisfaction. Voice and Music reach their maximum user's satisfaction at 40% of Shared RAN capacity. Video gradually increases its user's satisfaction until

reaching its maximum at 70% of shared capacity. Non-GBR services increase their User's Satisfaction until reaching the maximum at 80% of Shared Ran. With this, one verifies that depending on the number of users, at least 40% of the capacity of a 3.5 GHz Cell should be used for these applications. Figure 4.7. shows the Service's User's Satisfaction evolution related to the Shared RAN Percentage.

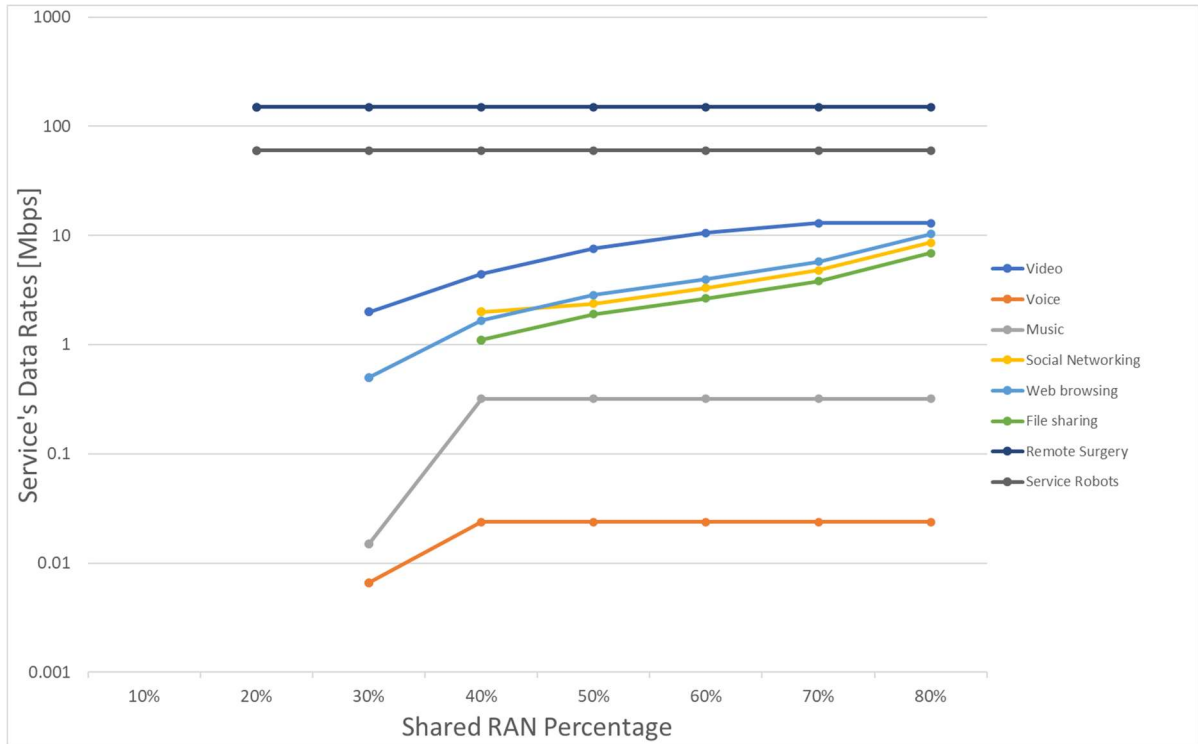


Figure 4.6. Service's Data Rates versus Shared RAN Percentage.

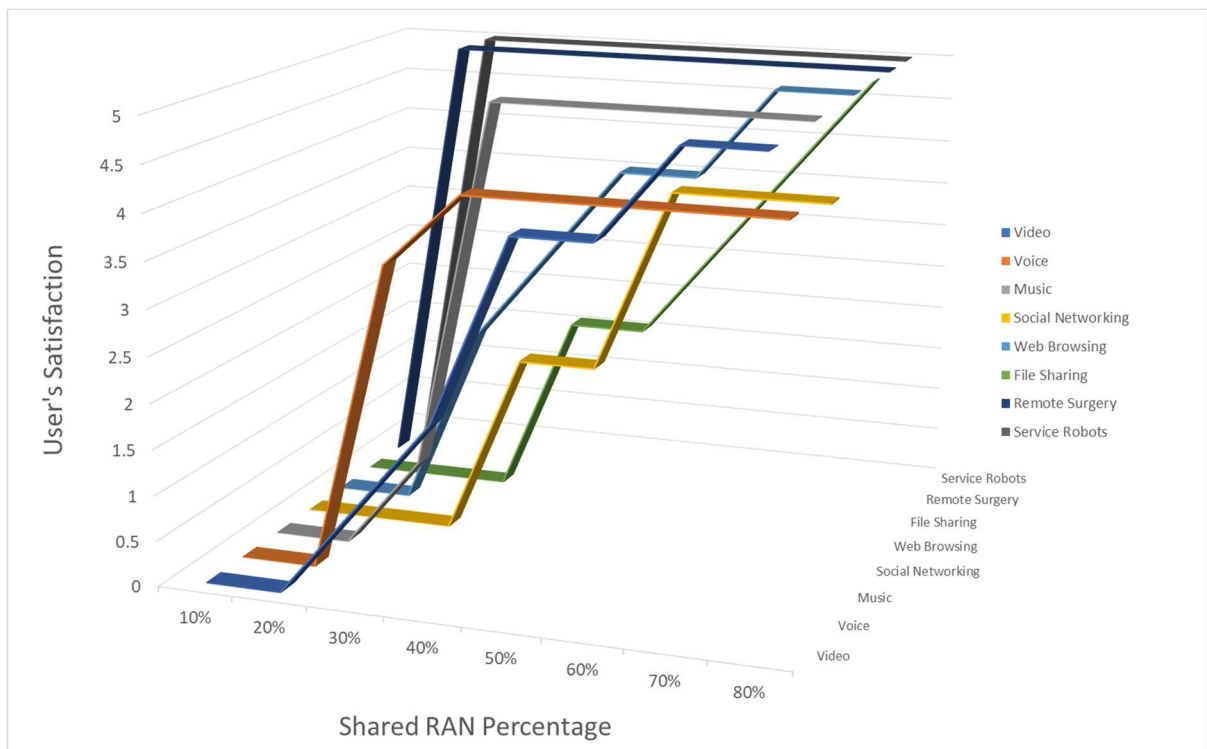


Figure 4.7. Service's User Satisfaction versus Shared Ran Percentage.

Finally, one analyses the VRRM Capacity Share. At 20%, only DelayC-GBR services are served and the network is saturated, as all of the other services and users were delayed, 94.523% of the available capacity is used. From 20% to 70% of Shared capacity, DelayC-GBR VNO share decreases and both GBR and Non-GBR shares increase, as there is an increase in the available capacity. From 70% to 80%, Non-GBR share continues to increase, as the Non-GBR services upper limit of data rate is the Cell Capacity. Consequently, the GBR share decreases, as Video reaches its maximum data rate and Voice and Music do not need more data in order to deliver the best experience. Figure 4.8 shows the VRRM Capacity Share evolution versus the Shared RAN Percentage.

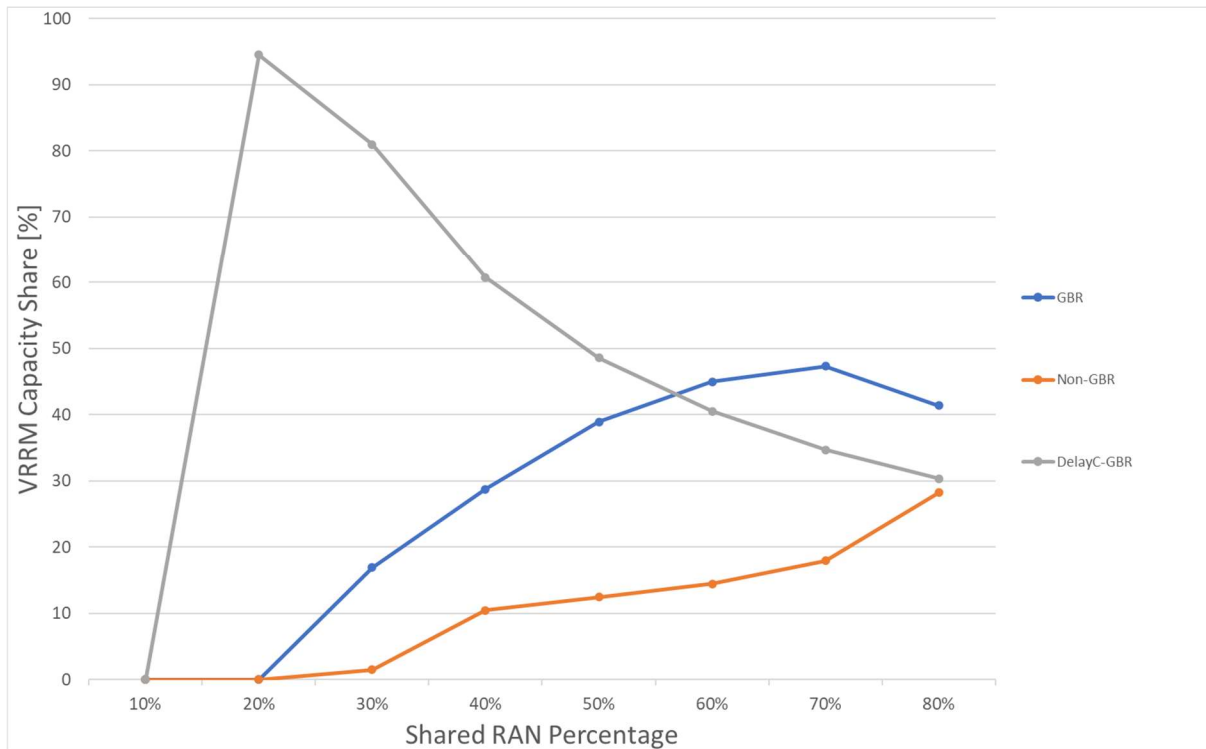


Figure 4.8. VRRM Capacity Share versus Shared RAN percentage for the Hospital Shared RAN scenario.

4.2.3. Hospital Standalone Results

For the 5.9 GHz, the topology is an Indoor Office, and would represent a Standalone Private Network deployment. The 5.9 GHz is unlicensed, in order to make use of it, our scenario considers the cell with a power limitation, so that interference is minimised. Therefore, a small cell of 250 m is considered. In addition to that, the scenario is taken as a high demand one, therefore Remote Surgery has an increase of the minimum data rate to 500 Mbps and Service Robots to 250 Mbps. The Service Mix used for this scenario is the same one used for the Shared RAN scenario. As the standalone scenario comprises of a larger capacity, the number of users was increased and fixed to 500 users, in order to better compare results between the 3.5 GHz and the 5.9 GHz frequency bands.

Simulation results show that the 3.5 GHz is saturated at 500 users, therefore Social Networking users are delayed, with the remaining services being served with minimum data rate. For the 5.9 GHz band, no user is delayed, and all users are served with higher data rates. Figure 4.9. shows the Service's Data Rate for both 3.5 GHz and 5.9 GHz frequency bands.

Regarding User's Satisfaction, one verifies on the 5.9 GHz band Non-GBR services served with higher user satisfaction, namely Web Browsing and Social Networking, with the last one not being served on the 3.5 GHz band. As for the GBR Services at the 5.9 GHz band, Music and Voice are served with their best possible user's satisfaction on this case, and Video goes from 1 to 2. At 3.5 GHz, all GBR services users are served with minimum satisfaction. As DelayC-GBR services are extremely critical, they are served in both cases with maximum user's satisfaction, allowing the best performance. Figure E.8. shows User's Satisfaction for both 3.5 GHz and 5.9 GHz frequency bands.

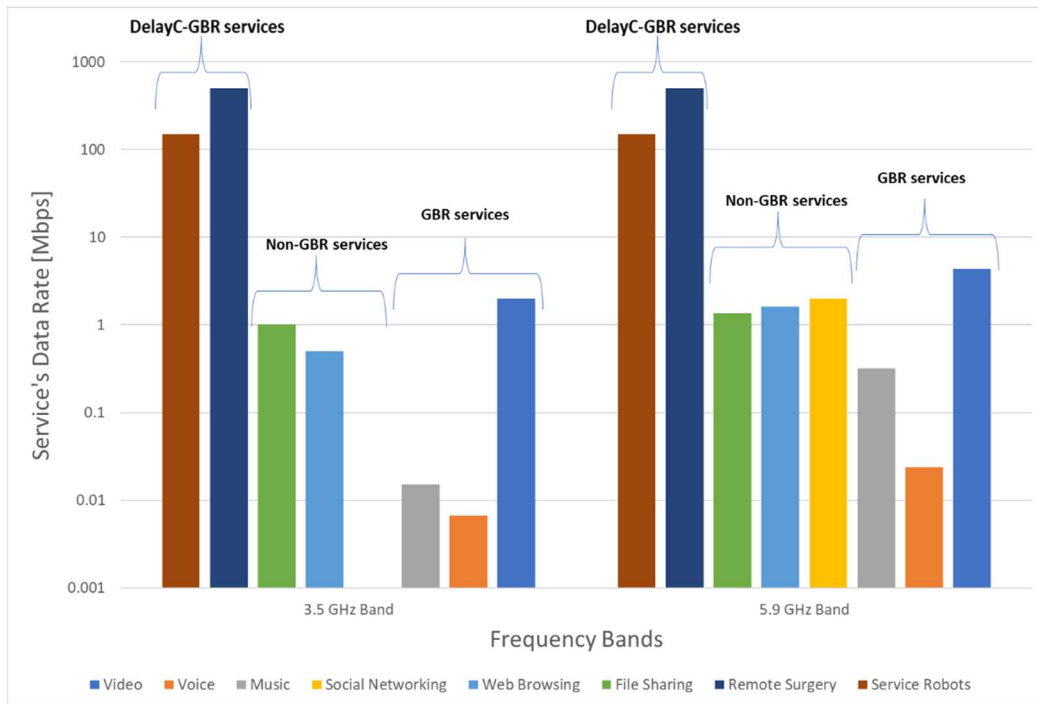


Figure 4.9. Service's Data Rate for both 3.5 GHz and 5.9 GHz frequency bands.

4.2.4. Mix and MIMO Layers Variation Results

In order to continue the analysis, one varies the MIMO Layers input parameter, using the reference mix and fixing the number of users at 100. Figure 4.10 shows the results for the MIMO Layers variation. With 2 MIMO Layers, the network is saturated and all low priority Non-GBR services are delayed. 4 MIMO Layers was chosen for the previous simulations and serves well all users. By finally increasing to 8 MIMO Layers, all GBR services are served at their maximum data rate. The DelayC-GBR and Non-GBR services have their rate increased proportionally to their priority and the available spare capacity.

Following that, one considers the possibility of Mix Variation on the hospital environment. Three other mixes are considered, apart from the reference one. First variation considers a Mix focused on the normal user, with more users allocated to Social Networking, Web Browsing and AR/VR. Second variation considers Professional AR/VR use, with 3 times the amount of users of the reference mix. Also, users for Voice and Video are decreased and Music increased. The last variation considers different hospital resources, with a significant decrease on the video users, slight increase on AR/VR considering a small hospital use and Remote Surgeries increased to 2 at a time and 5 Service Robots. Table 4.6 Shows the different mixes considered.

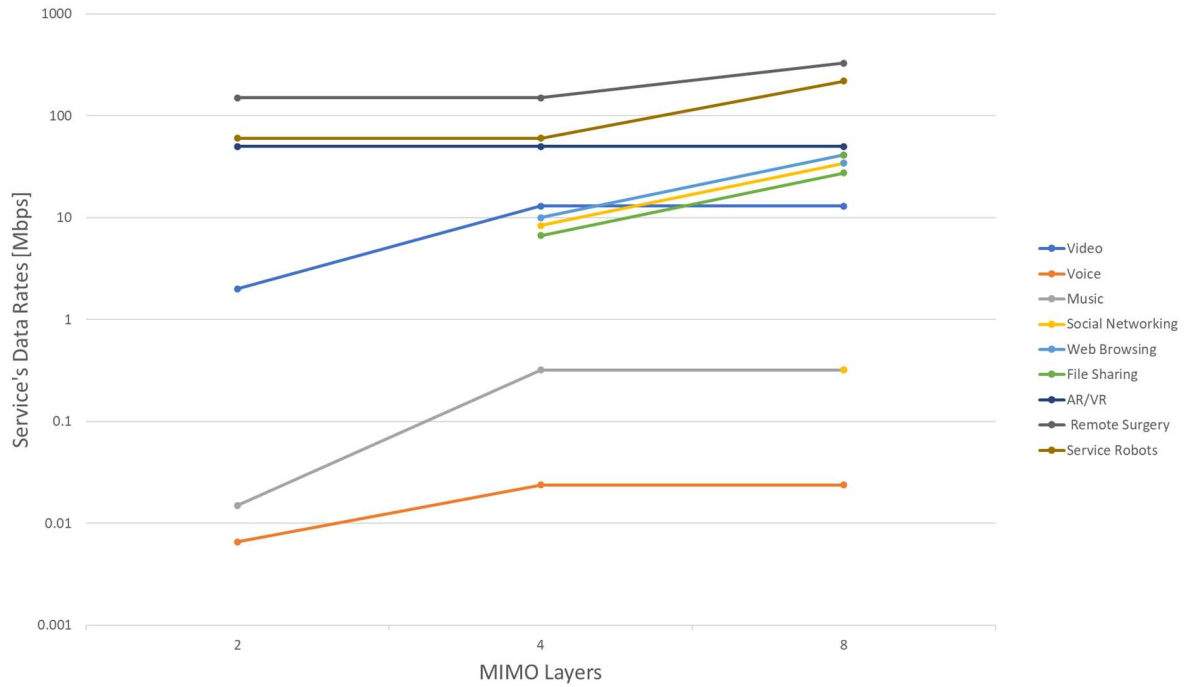


Figure 4.10. Service's Data Rates versus MIMO Layers variation for the Hospital Network Slicing Private Deployment.

Table 4.6. Different Service Mixes for Mix Variation Analysis.

Service	Reference Mix [%]	Focused on normal user Mix [%]	Professional AR/VR [%]	Different Hospital Resources [%]
Video	30	15	10	15
Voice	30	15	10	31
Music	9	4	10	4
Social Networking	10	25	20	15
Web Browsing	8	25	20	15
File Sharing	5	5	12	5
AR/VR	5	8	15	8
Remote Surgery	1	1	1	2
Service Robots	2	2	2	5

In order to perform Mix Variation simulation, total number of users was again fixed at 100 users. First result considers the Service's Data Rate. The first mix variation focused on the normal user shows a decrease on the data rate of the low priority Non-GBR services, Web Browsing, File Sharing and Social Networking. This decrease was expected and happens because of the increase in the number of users for those services. The other services maintain their data rate. For the second variation in the mix, considering Professional user of AR/VR, one verifies that AR/VR occupies a lot of capacity, and therefore the Non-GBR low priority services' data rates decrease even more, and the Video data rate goes to its minimum, while the other services maintain theirs. For the final mix variation, there is a small increase on the data rate of the low priority Non-GBR services and video, if compared to the Professional AR/VR scenario, this shows that a network loaded with AR/VR must have more attention to network capacity. Figure 4.11 shows the Service's Data Rates for the different service mixes considered.

It is also relevant to verify the VRRM Capacity Share of this Mix Variation. As expected, the Focused on normal user mix has an almost 20% increase of the Non-GBR VNO share, with the equivalent decrease on the GBR VNO Share. The Professional AR/VR mix pulls even more the share to Non-GBR VNO due to the increase of the AR/VR mix. In both cases presented the DelayC-GBR VNO Share is the same, as the number of Remote Surgeries and Service Robots was not changed. The different hospital resources mix shows a great increase in the DelayC-GBR VNO share, which was expected as the amount of users of that VNO increased, the Non-GBR VNO share is the same as the reference mix, and the GBR VNO share decreased. Figure E.12 shows the VNO's VRRM Capacity Share versus Mix Variation for the Hospital Network Slicing Private Deployment.

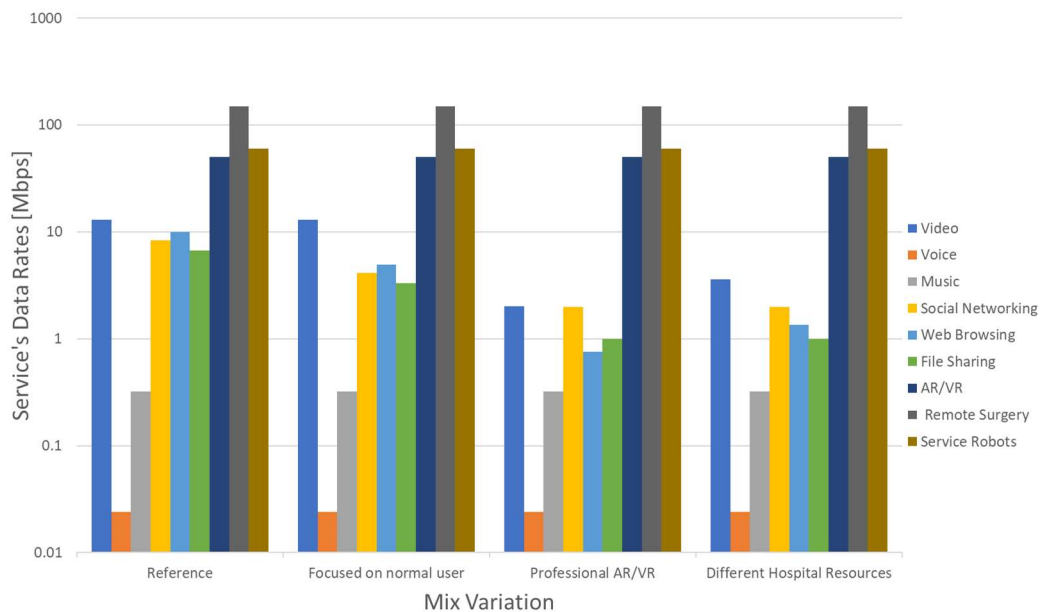


Figure 4.11. Service's Data Rates versus Mix Variation for the Hospital Network Slicing Private Deployment.

4.2.5. Hospital Latency Results

Latency simulations considered only Radio Units close to the Santa Maria Hospital located in Lisbon. Figure 4.12. shows the RU site locations. Collocated RU-DU-CU with MEC architecture was defined for this simulation. Only one C-RAN architecture was considered. The traffic profile used was from 4G and adapted to 5G with multipliers. This shows that the latency results are more of an approximation. Next, by analysing the delay results, one remembers that the minimum Max latency allowed in the Hospital scenario is 5ms from Remote Surgery Haptic Feedback. Therefore, even with maximum latency, the Hospital could be served in a satisfactory manner. The maximum value relates to the RUs farther from the hospital and the minimum relates to the RUs closer to the hospital. The MEC implementation also decreases the delay and allows serving all users Table 4.7 shows the simulation Delay results.

Table 4.7. Hospital Scenario Delay Results.

Total Network Delay	Delay [ms]
Mean	0.4991
Minimum	0.3802
Maximum	0.7998
Standard Deviation	0.16



Figure 4.12. RU sites provided by NOS on the Hospital Scenario.

4.3. Smart Factory Scenario Analysis

4.3.1. Smart Factory Network Slicing Results

The first Smart Factory scenario considers the Network Slicing Private Network Deployment. This scenario works under the 3.5 GHz band, with the topology being UMa_O2I, considering an Industry close to the city centre, more specifically Auto Europa industry in Setubal. The Service Mix considered takes into account a Smart Factory scenario. AR/VR is used for remote expertise support and training of employees. Sensors for Motion, Closed Loop and Safety control are considered for the factory automation and monitoring sensors are represented in higher quantity with less strict latency requirements. The cell input parameters are the same as the ones for the Hospital Scenario, listed in Table 4.2.

The reference scenario Mix shown in Table 4.1 is used for the Network Slicing Private Network Deployment, considering users in the vicinity of the factory. Simulations were done starting in 200 users, with increments of 50 users. Low Priority Non-GBR services stop being served at the 400 users mark, which is also when the network saturates and the delays cause the Cell Capacity to not being fully used. SN, WB and FS are all delayed while Voice, Music and Video are served with the minimum data rate, which implies in lowest user's satisfaction. The DelayC-GBR services are all served. Motion Control and Closed Loop Control have data rates of 5 Mbps at 200 users, which is their maximum data rate. Their data rate decreases until it reaches its minimum of 1 Mbps at 350 users. Monitoring sensors are served with the maximum data rate of 0.5 Mbps until 350 users, when it starts to decrease to 0.44 Mbps, then goes to its minimum at 400 users. Safety Control sensors are served with the maximum data rate of 1 Mbps until 350 users, at 400 users it goes to the minimum of 0.5 Mbps. AR/VR is always served with 50 Mbps. Figure 4.13 shows the Service's Data Rates as users are incremented.

The VRRM Capacity share shows Non-GBR VNO with the highest capacity share, that is due to AR/VR high data requirements. GBR VNO share decreases from 200 to 250 users, due to Video data rate lowering from 13 to 2 Mbps. From 250 users until 500 users, GBR VNO share fluctuates between 3% and 4%. DelayC-GBR VNO share increases from 200 to 250 users, as their services are served with their maximum data rate. Beyond 250 users, DelayC-GBR share decreases until 400 users. At 400 users the DelayC-GBR services reach their minimum data rate and the Non-GBR services are all delayed. With that, Non-GBR share starts to decrease and DelayC-GBR increases. Figure 4.14 shows the VRRM capacity share of all VNOs with the increment of the number of users.

After that, Figure E.4 shows the GBR Services User's Satisfaction. As seen previously with the Data Rate, Video User's Satisfaction goes from 5 to 1 between 200 and 250 users and remains there until 450 users. Voice has a decrease in its User's Satisfaction at 400 users, when the data rate goes to the minimum. Music is served with maximum user's satisfaction until 350 users, when it decreases to 4, then at 400 users it decreases to its minimum of 1.

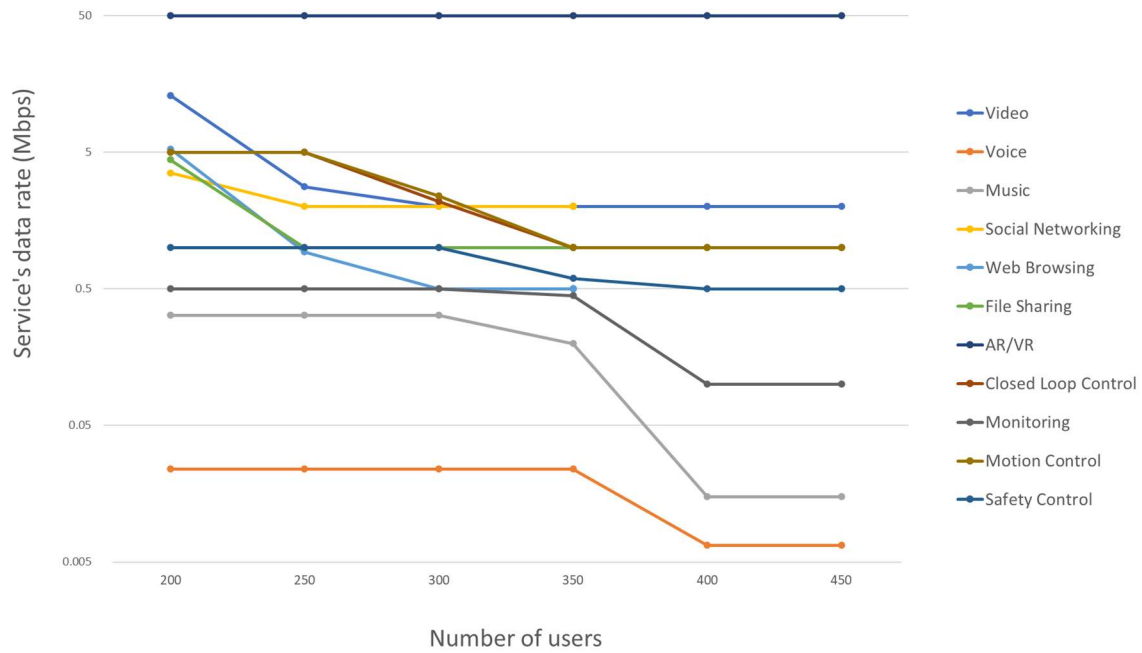


Figure 4.13. Data rate for Smart Factory Network Slicing Private Deployment Scenario

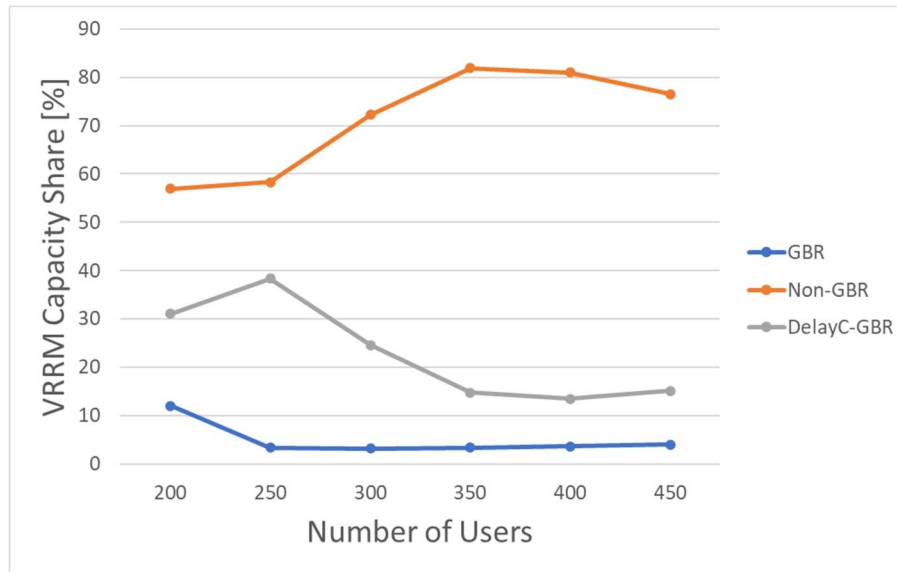


Figure 4.14. VRRM Capacity Share versus the evolution of Users.

Non-GBR VNO User's Satisfaction versus increment of users is shown at Figure A.5 As can be seen, AR/VR is always served with maximum user's satisfaction, as its minimum data rate was set to a satisfactory level. The Low priority Non-GBR services (WB, FS and SN) are served with minimum user's satisfaction from the 250 until 400 users, when the network saturates and all users of those services are delayed. AR/VR users start to be delayed at 400 users, at rate close to 10% per increment of 50 users.

As for the DelayC-GBR VNO, all of their services' users are served with maximum satisfaction until 300 users, when Closed Loop and Motion Control start to decrease and go from 5 to 2. Monitoring User's Satisfaction decreases from 5 to 4 only at the 350 users mark. This happens due to Monitoring

having a lower data rate requirement, even though its priority is also lower. Safety Control decreases its User's Satisfaction from 5 directly to 1 at 350 users. One should also note that Safety Control has the highest Tuning Weight. Figure E.6 shows the DelayC-GBR VNO User's Satisfaction versus the evolution of number of users.

4.3.2. Smart Factory Shared Ran Results

The Shared Ran scenario characteristics are reflected at the service mix and the available data rate. As in the hospital scenario, the service mix is adapted so that only employees are using the network. In that sense, GBR and Non-GBR services have their Mixes lowered by 4%, except for the AR/VR service that has a decrease of 1%, since the AR/VR service is present in the Smart Factory mix. The percentage taken out of those services were redirected to the sensors mix, in order to represent a higher quantity of them. Users are fixed at 500. Table 4.8 shows the Mix for the Shared RAN scenario.

Table 4.8 Smart Factory Shared RAN Scenario Mix.

Service	Mix [%]
Video	1
Voice	1
Music	1
Social Networking	1
Web Browsing	1
File Sharing	1
AR/VR	4
Closed Loop Control	20
Monitoring	30
Motion Control	20
Safety Control	20

The results show that all users are only served at 50% of Shared Ran. At 10%, all GBR and DelayC-GBR users are served with minimum data rate, as the network is saturated. At 20%, AR/VR users start to be served. With each 10% increment of the Shared RAN percentage, AR/VR served users increase 25%, until all users are served at the 50% mark. DelayC-GBR services reach their maximum data rate at 80% of Shared RAN, as do Voice and Music. Figure 4.15. shows the Service's Data Rate versus the Shared RAN percentage.

Furthermore, in order to better understand the behaviour of the network, one verifies the User's Satisfaction results. As on previous occasions, once AR/VR users are served they have maximum

User's Satisfaction, even though the network is saturated. Besides, GBR services are all served with minimum User's Satisfaction, as Non-GBR users are being delayed. At 50% of Shared RAN usage, low priority Non-GBR users are served with a low user's satisfaction, while Voice and Music get to their best quality. Regarding DelayC-GBR services, Monitoring and Safety Control reach high user's satisfaction at 50% shared RAN usage, while Motion and Closed Loop Control reach it only at 80% of shared RAN usage. That happens because Safety Control has the highest Tuning Weight and Monitoring the lowest data rate requirements of the DelayC-GBR services. Due to video high data rate requirements it is always served with low user's satisfaction. Figure E.10 shows the Service's User's Satisfaction with relation to the Shared RAN percentage.

Finally, VRRM Capacity Share shows DelayC-GBR VNO with most of the share at 10%. At 20%, AR/VR users are served and the Non-GBR VNO share increases greatly. At 50% of shared RAN usage, all users are served and the Non-GBR share starts to decrease as the DelayC-GBR one increases, as was expected with the DelayC-GBR VNO being served with the best User's Satisfaction at that stage. GBR VNO share starts with 3% then fluctuates between 0,6% and 1%, this is due to the lower mix of the GBR services. Figure E.11 shows the VNO's VRRM Capacity Share versus Shared RAN percentage.

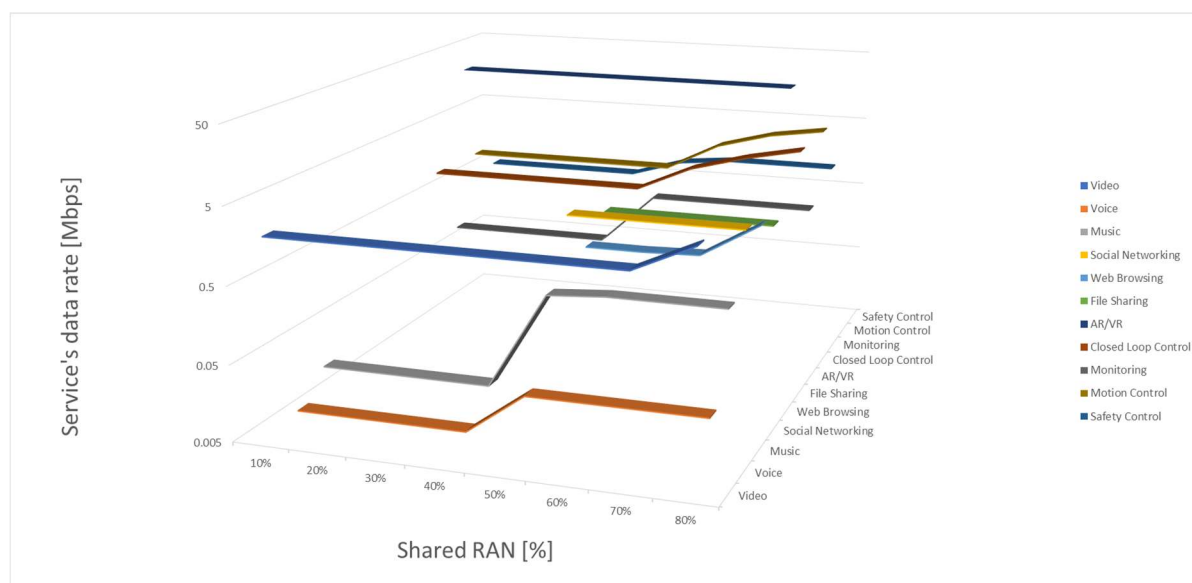


Figure 4.15. Service's data rate evolution with Shared RAN percentage increments for the Smart Factory Shared RAN scenario.

4.3.3. Smart Factory Standalone Results

The Standalone deployment works under the 5.9 GHz ISM band, in an Indoor Office topology. Therefore, the standalone scenario consists of the analysis of one small cell located in the factory premises. This cell would have an approximated range of 250 m, therefore a number of these cells would be allocated on the factory floor to serve all sensors and employees. At this scenario analysis,

the number of users is fixed to 600.

Starting with the Service's Data Rate, one verifies that the 3.5 GHz band is saturated and does not serve all users. SN, WB and FS are all delayed and the other services work at their minimum data rate. Opposed to that, 5.9 GHz band serves all users. Voice and Music are served at their maximum data rate. And all DelayC-GBR services have an increase in their data rates. Figure 4.16 shows the Service's Data Rate at the 3.5 GHz and 5.9 GHz frequency bands.

Regarding User's Satisfaction, one verifies Safety Control and Monitoring being served with maximum User's Satisfaction at the 5.9 GHz band, opposed to them being served with their minimum at the 3.5 GHz band. Motion and Closed Loop control also increase their values to a satisfactory level when the two bands are compared. All low priority Non-GBR services are served with medium to low user's satisfaction, but the main services all have high satisfaction scores. Video does not change its User's Satisfaction from one band to the other, this happens due to the relative low priority considering the other services. Figure E.9 shows the Service's User's Satisfaction at both 3.5 GHz and 5.9 GHz bands.

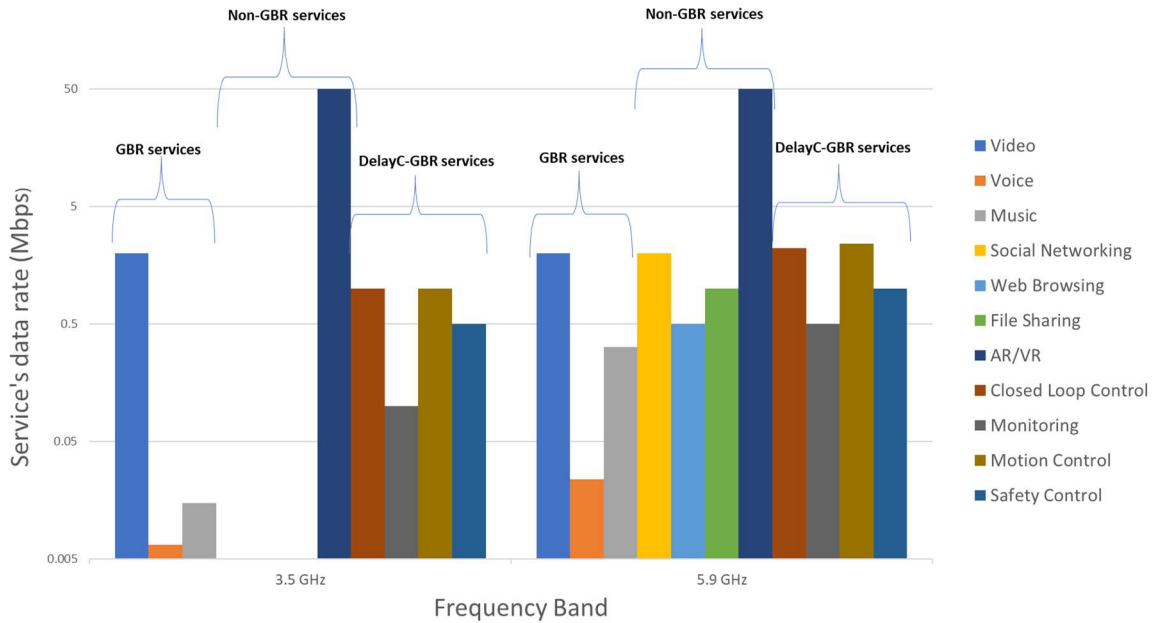


Figure 4.16. Service's data rate on both 3.5 GHz and 5.9 GHz frequency bands for the Smart Factory Standalone deployment.

4.3.4. Smart Factory Latency Results

Latency simulations considered only Radio Units close to the Auto Europa Factory located in Lisbon. Figure 4.27. shows the RU site locations. Collocated RU-DU-CU with MEC architecture and 7.2 Splitting Option were defined for this simulation. Only one C-RAN architecture was considered. The traffic profile used was the same from the Hospital Scenario and adapted to the Smart Factory case

with increase of the Factory Automation mix percentage. Next, by analysing the delay results, one remembers that the minimum Max latency allowed in the Smart Factory scenario is between 0.5 and 2ms from the Motion Control service. The results show a minimum of 0.3591 ms and maximum of 0.8141 ms. This means that for the closest RUs the most strict service can be deployed with maximum user's satisfaction latency-wise. As for the farthest ones, Motion Control would still be served but with a lower user's satisfaction, given the higher delay. The MEC implementation also decreases the delay and allows serving all users. The cell sites considered are shown on Figure 4.17 and are located close to the Auto Europa factory. Table 4.9 shows the simulation Delay results.

Table 4.9. Smart Factory Scenario Delay Results.

Total Network Delay	Delay [ms]
Mean	0.5090
Minimum	0.3591
Maximum	0.8141
Standard Deviation	0.15

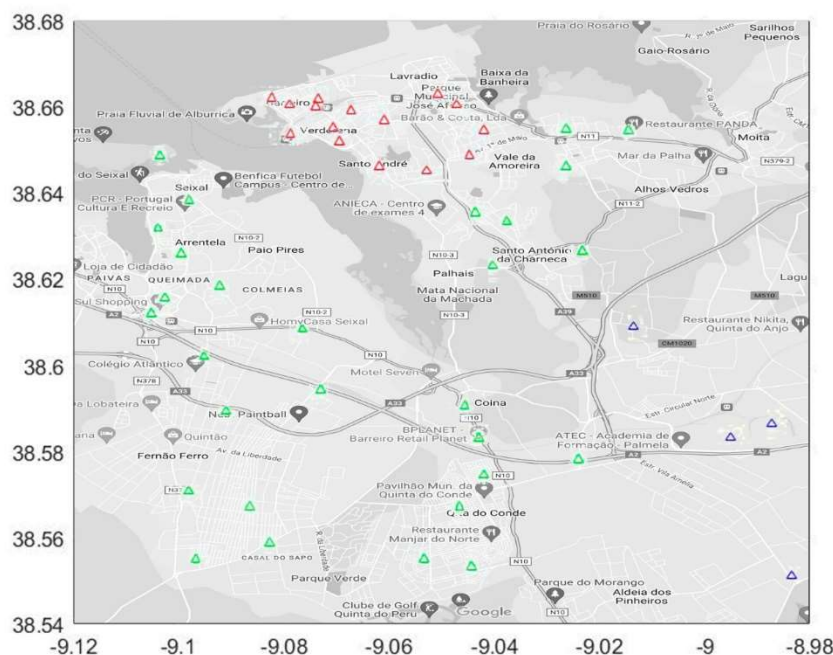


Figure 4.17. RU sites provided by NOS for the Smart Factory Scenario.

4.4. Mission Critical Scenario Analysis

4.4.1. Mission Critical 700 MHz Band Analysis

For the Mission Critical Scenario, Mission Critical Video, Mission Critical Push to Talk (MCPTT) and Drone Supervision are added as services, as depicted in Section 4.1. AR/VR is considered for training

and remote support. A Network Slicing approach is considered, as most emergency services may occur in diverse sites. The frequencies to be considered are the 700 MHz and 3.6 GHz frequency bands. Table 4.10 shows the Cell Input Parameters for this scenario variation.

Table 4.10. Cell Input parameters for Mission Critical Scenario

MIMO layers, $v_{Layers}^{(j)}$	2
Frequency Band [MHZ]	{710, 3600}
Bandwidth [MHz]	{10, 100}
Scenario Topology	{RMa, UMa}

Lower Bandwidths have less available capacity, and therefore eMBB services are harder to attend, given the users connected to the BS. It is important to note that there is a maximum of three drones per cell, according to NOS. Therefore, as the users are incremented, the Mix is adapted, by adding to mix to Voice, so that a maximum of 3 Drone Supervision users are present. The first scenario considered is the 700 MHz frequency band. This Scenario considers the location of Pedrogão Grande in Portugal. This is an area known for fires, and therefore of great relevance for Mission Critical services. Only 10 MHz of Bandwidth is available for the 700 MHz. Therefore, services with high demand of data rates cannot be well served. With that in mind, an adapted service mix is defined. MCPTT is prioritised with 40% of the Mix, Video Mix is lowered to 5% and AR/VR is removed from the Mix, as it has high data rate requirements. Table 4.11 shows the Service Mix for the 700 MHz band mission critical scenario.

Table 4.11. Service Mix for the 700 MHz band Mission Critical Scenario.

Service	Mix [%]
Video	5
Voice	15
Music	5
Social Networking	5
Web Browsing	5
File Sharing	5
Mission Critical Video	15
MCPTT	40
Drone Supervision	5

As the 700 MHz band has a low bandwidth, the number of users increment was decreased to 10 users in order to better analyse the network behaviour. One verifies that the network saturates at 100 users with File Sharing users being delayed. MCPTT and Voice are served with max data rate until the 80 users mark, then go to their minimum when the network saturates. Mission Critical Video starts with 5.1 Mbps of data rate and Video with 2.1 Mbps at 50 users, by 100 users both move to their minimum data rate. Mission Critical Video has higher data rate when there is free capacity due to its higher priority than Video. Drone Supervision is always served and reaches its minimum data rate at 80

users. Figure 4.18 shows the Service's Data rate versus the Number of Users for the 700 MHz Mission Critical Scenario.

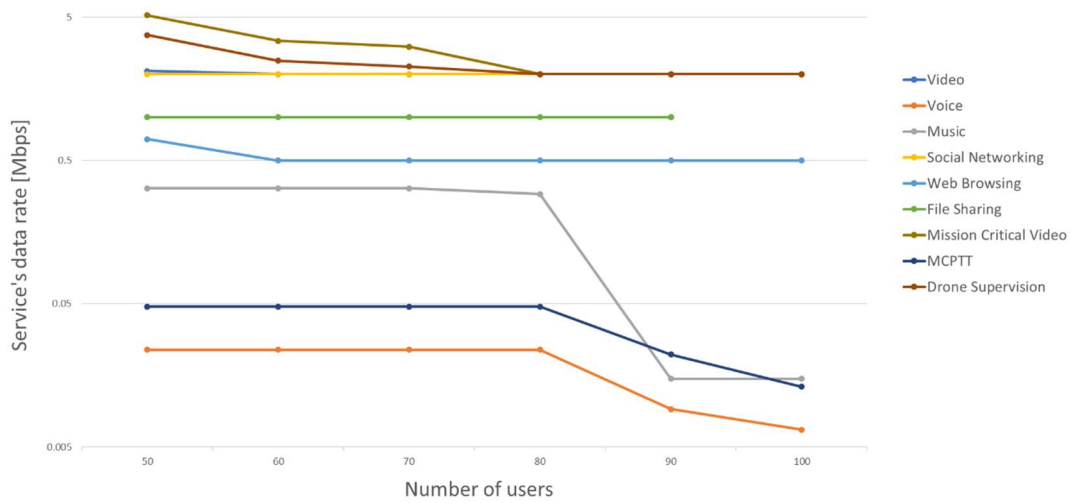


Figure 4.18. Data Rate for Mission Critical 700 MHz scenario.

Following the data rate analysis, the VRRM Capacity Share is analysed. As would be expected, most of the share belongs to Mission Critical GBR services, followed by Non-GBR and then GBR. Until 90 users, GBR and Non-GBR VNO share increased around 10%. When the network saturates, at 100 users, Non-GBR VNO share decreases due to their users delayed, while GBR and Mission Critical GBR VNO's Share increases. Figure 4.19 shows the VRRM Capacity Share versus the Number of Users for the 700 MHz band Mission Critical Scenario.

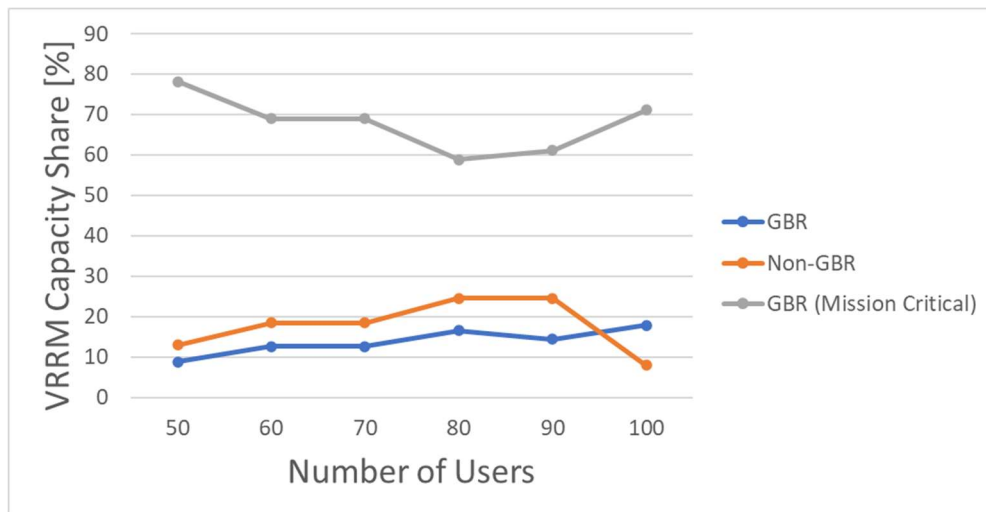


Figure 4.19. VRRM Capacity Share for Mission Critical Scenario at the 700 MHz band.

Finally, in order to end this scenario analysis, one verifies the Service's User's Satisfaction. Video is always served with minimum satisfaction, opposed to Mission Critical Video that starts with a value of 3, then goes to 1 at 80 users. Non-GBR user's are always served with satisfaction of value 2, which is low but enough for its use. MCPTT is served at the maximum possible User's Satisfaction until at 100 users, when the network saturates, after that it still has a Nominal MOS of 4.04, which is higher than

the Voice minimum of 3.39. Music is always served with good user's satisfaction until the network saturation, when it goes to the minimum data rate and consequently user's satisfaction. Drone Supervision starts with the best possible User's Satisfaction but decreases to its minimum at 70 users. It is important to note that Mission Critical GBR service have higher priority, and consequently have their user's satisfaction decreased in a lower rate, considering the increment of users. Figure 4.20 shows the Service's User's Satisfaction versus the Number of Users for the 700 MHz Mission Critical Scenario.

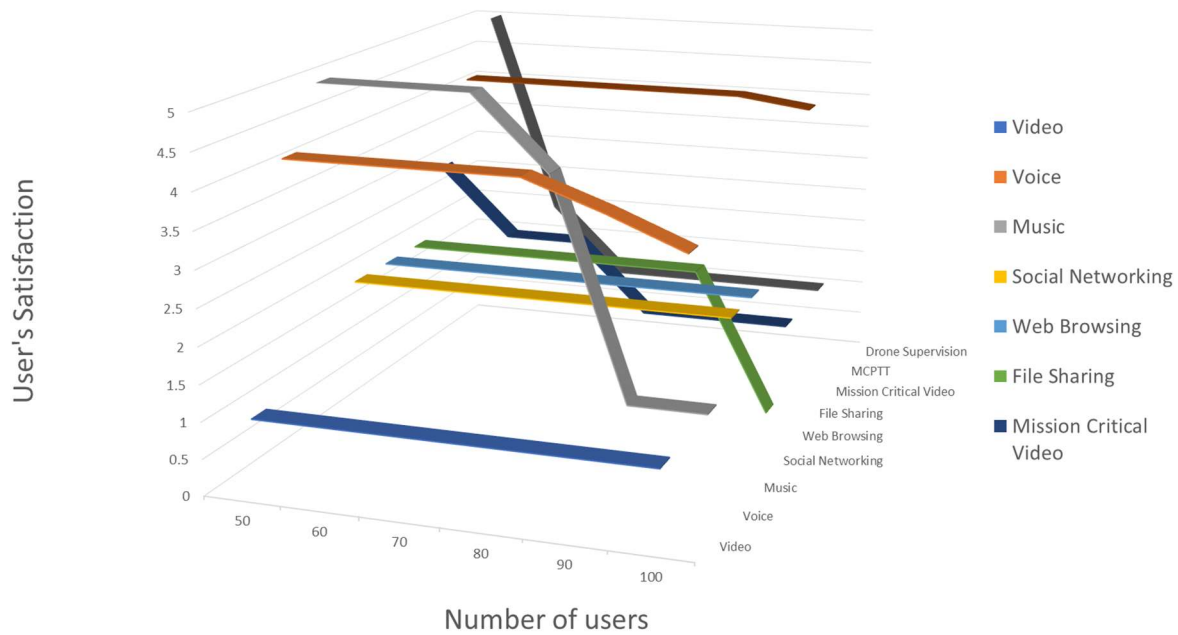


Figure 4.20. User's Satisfaction for Mission Critical Scenario at 710 MHz.

4.4.2. Mission Critical 3600 MHz Band Analysis

As for this scenario, it is an Urban Macro and would be considering cells inside the City of Porto. As the 3.6GHz frequency band has 10 times the bandwidth of the 700MHz frequency band, even though only two MIMO layers are used, it is relevant to once again alter the Service Mix in order to attend eMBB services. Therefore, AR/VR was introduced to the mix, and Video was increased by 5%, while MCPTT mix was lowered by 10%. Table 4.12 shows the updated Service Mix for the 3.6 GHz band Mission Critical Scenario.

In order to analyse the network performance, one verifies the service's data rates with an increment of 50 users per simulation. Given that two MIMO layers are used, the available capacity is lower than the one from the Hospital and Factory scenarios, therefore the number of possible users is also reduced. The network saturates at 200 users, and at that stage all low priority Non-GBR users are delayed. AR/VR is always served with maximum data rate of 50 Mbps. At 100 users, Mission Critical Video is served with 13 Mbps while video with 5.5 Mbps, both decay to their minimum of 2 Mbps once the

network saturates. Music, Voice and MCPTT are served with maximum data rate until 150 users, at 200 users all of them go to their minimum data rate. Drone Supervision starts with 4 Mbps then decays to its minimum of 2 Mbps once the network saturates. At 200 users, 20% of AR/VR users are delayed, at 250 users 41.6% of its users are delayed and at 300 users, 53.3% of AR/VR users are delayed. Figure 4.21. shows the Service's Data Rate versus the Number of Users evolution.

Table 4.12. Service Mix for the 3600 MHz band Mission Critical Scenario.

Services	3600 MHz Mix [%]
Video	10
Voice	15
Music	5
Social Networking	5
Web Browsing	5
File Sharing	5
AR/VR	5
Mission Critical Video	15
MCPTT	30
Drone Supervision	5

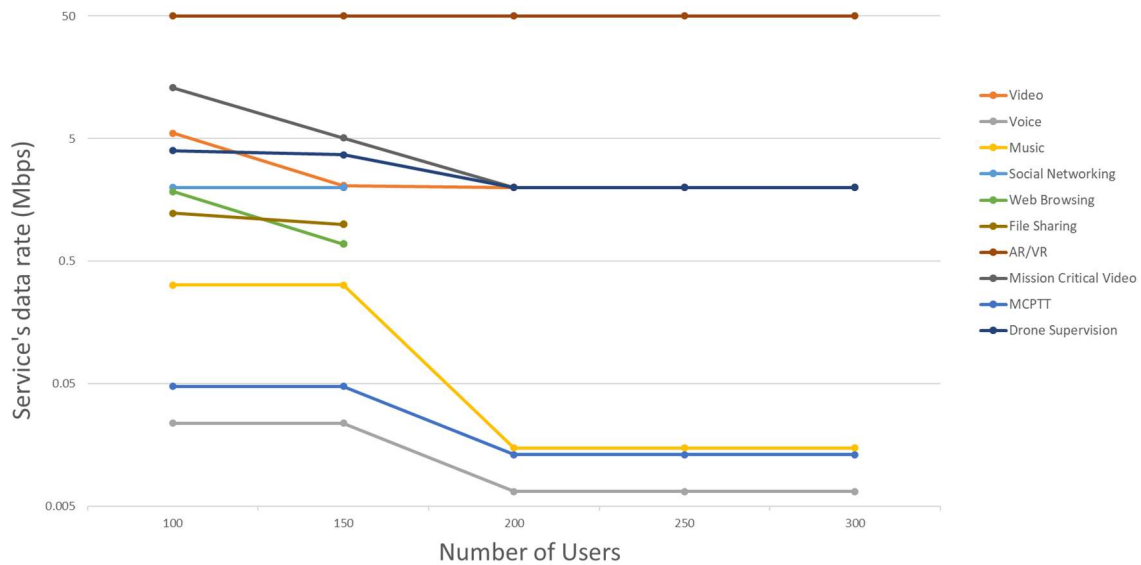


Figure 4.21. Data rate for Mission Critical Scenario at 3600 MHz.

Next, one verifies the VRRM Capacity share. Non-GBR VNO share is the highest in this scenario due to AR/VR having the highest data rate demands. Non-GBR share increases while GBR Mission Critical decreases until 200 users, when the network saturates and the Non-GBR users are delayed. With that, both GBR VNO shares increase while the Non-GBR share decreases. From 250 to 300 users, the same phenomenon that happened at the Hospital scenario happens here with the Non-GBR share maintaining its value and not decreasing, even though the percentage of served AR/VR

users is decreasing. Figure 4.22 shows the VRRM Capacity Share versus the number of users evolution.

To finalise the data rate analysis of the Mission Critical Scenario, User's Satisfaction results are verified. Mission Critical Video, AR/VR and Drone Supervision start with maximum user's satisfaction. AR/VR maintains its high user's satisfaction while users are served. Mission Critical and Drone Supervision reach their minimum satisfaction when the network saturates. MCPTT is served with satisfactory data rate even when the network saturates due to its higher data rate requirements. Music is served with maximum satisfaction until the network saturation when it goes to the minimum. The same happens with Voice. Low Priority Non-GBR users are served with medium to low satisfaction (Values between 2 and 3) until they are all delayed at the 200 users mark. Figure A.7. shows the User's Satisfaction evolution versus the number of users for the 3.6 GHz Mission Critical Scenario.

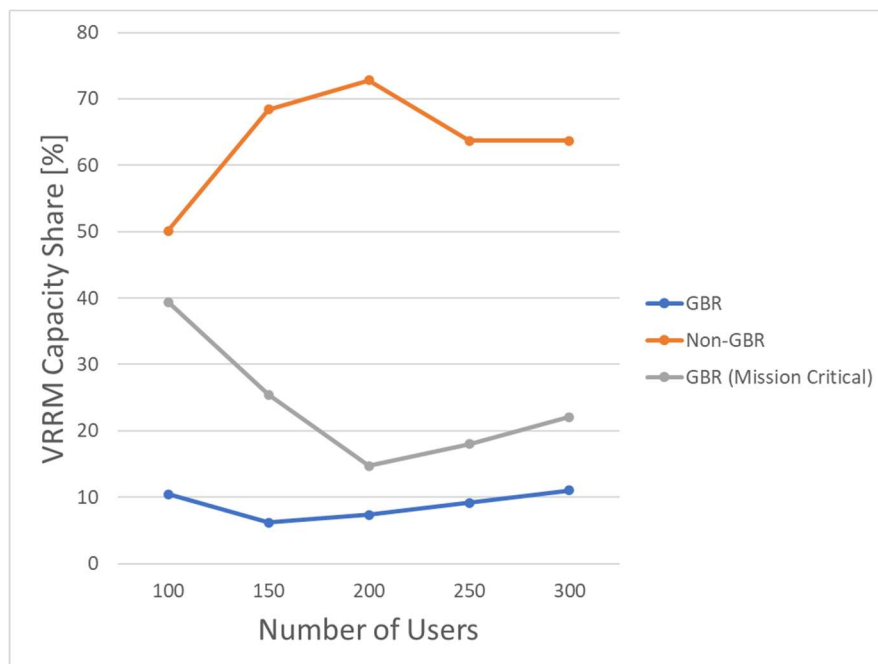


Figure 4.22. VRRM Capacity Share versus the Number of Users for the 3.6 GHz Mission Critical Scenario.

4.4.3. Mission Critical Latency Results

Latency simulations considered two cases, one Rural and one urban. The Rural scenario considers the 700 MHz band and is located at Pedrogão Grande. The Urban scenario considers the 3.6 GHz band and is located at City of Porto. RU-DU+CU architecture and 7.2 Splitting Option were defined for this simulation. Only one C-RAN architecture was considered. The traffic profile used was the same from the previous Scenarios and adapted to the Mission Critical case with increase of the Voice and Video mix percentage. Figure 4.32. shows the RU site locations for the Rural Scenario

Table 4.13. Mission Critical Rural Scenario Delay Results.

Total Network Delay	Delay [ms]
Mean	2.9592
Minimum	2.4799
Maximum	3.3624
Standard Deviation	0.25

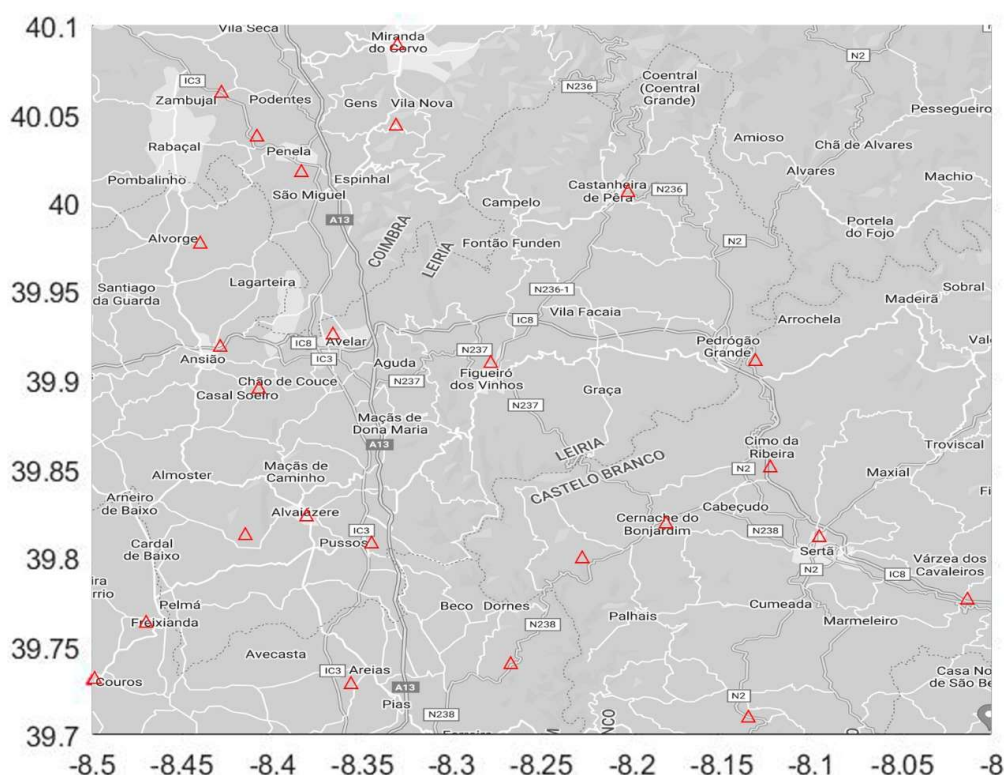


Figure 4.23. RU sites provided by NOS.for the 700 MHz Mission Critical Scenario at Pedrogão Grande.

For the Urban Scenario, the minimum delay is 0.1919 ms and the maximum 2.6729 ms. One verifies that the delay results for the Urban scenario are considerably lower than the Rural one. This difference happens due to City of Porto having a higher density of RUs and as there is no MEC, the distance to the Core Network also poses big influence on the delay results, as there is a Core node at City of Porto, the delay there is considerably lower. Table 4.14 shows the Mission Critical Urban delay results. Figure E.13 shows the RU sites at the City of Porto for the Mission Critical Urban Scenario results.

Table 4.14. Mission Critical Urban Scenario Delay Results.

Total Network Delay	Delay [ms]
Mean	1.107
Minimum	0.192
Maximum	2.673
Standard Deviation	0.600

Next, by analysing the delay results, one verifies that both cases attend well the network with the services chosen for the Mission Critical Scenario. The results show a minimum of 2.480 ms and maximum of 3.362 ms for the Rural scenario. Table 4.14 shows the simulation Delay results for the Mission Critical Rural Scenario Delay Results.

Chapter 5

Conclusions

This chapter summarises the purpose of this thesis and its content.

The main goal of this work is to analyse private networks implementation in 5G using Network Slicing. Different scenarios were analysed regarding data rate and latency in order to analyse the Radio Network performance. The model was adapted and developed in order to receive cell and network input parameters, then select the best numerology based also on the topology considered.

Chapter 1 starts by summarising the main 5G use cases and introducing NFV and SDN technology. It also introduces private networks deployment and their relevance. Then it gives motivation to their study and structures the thesis along with its contents.

After that, Chapter 2, addresses the theoretical subject relevant to Network Slicing and Private Networks in 5G, essential to understand the thesis. Network architecture is described, emphasising deployment options and the core network. Radio interface addresses LTE, 5G and unlicensed operation, which is of great importance for Private Networks deployment. Then, Network Aspects are deepened by tackling SDN, NFV and Network Slicing architecture. C-RAN and MEC are also addressed as they represent important technologies for Network Slicing deployment. Use cases and QoS Characteristics are addressed at the Services and Applications subsection. Private Networks architectures and performance parameters are referred. Finally, the state of the art finishes Chapter 2.

Chapter 3 describes the model developed for this thesis. First, an overview is given with the inputs and outputs of the model and brief description of its stages along with some considerations made and a general workflow of model. Then the next section describes the Numerology Selection algorithm along with its theoretical reference, considering interference. After that, it is explained how the cell capacity is calculated with an expression given by 3GPP. The expressions for the parameters are explained and their possible values are shown. After that, the Admission Control and Delay process is explained, in which users are delayed when there is not enough capacity to serve all users. Users are delayed according to their SLA, which include GBR, Non-GBR and Delay Critical GBR. After that the VRRM Optimisation is explained. It consists of a constrained concave optimisation problem. The objective function is formulated as the logarithm of the normalised weighted sum of the different services total data rate: Following that, the aggregation process is detailed. It starts by receiving nodes input as network info, and that the model makes their connections based on their capacity. This capacity is calculated with a Node Processing Power model explained in this Chapter. With the nodes connected, latency is calculated based on queuing, processing and transmission delay on the C-RAN and core network. Then a workflow of the whole aggregation process is presented and the Private Network considerations are defined. On the following section, bandwidth concerns are addressed. The available frequencies are presented with their duplexing technique and available numerologies. Also, a study of the Mean Delay Spread in different scenario topologies in order to compute the ISI and ICI for numerology selection concerns. Then on the next section User experience and network performance parameters are defined and explained. Finally, on the final subsection, model assessment is made.

Chapter 4 contains all simulation results and their analysis, using the developed model. The first section starts by presenting the reference service Mix for all scenarios. Then all service's data rates, and their respective tuning weight are presented, and finally the input parameters for the aggregation

process are presented, along with some considerations regarding the latency results, as the latency model is not the focus of this work. Each scenario has a specific cell input parameter; therefore the next section comprises of the Hospital Scenario and starts by present its Cell Input parameter. First simulations consider a Network Slicing Private Network deployment and the reference Hospital Mix is used. The 3.5 GHz band is analysed for this scenario, and Low Demand Remote Surgery and Service Robots are considered. The influence of user is analysed with an increment of 50 users, starting at 100 users. For this scenario, the network saturates at 250 users, by delaying the low priority Non-GBR users. AR/VR users are delayed in a higher percentage with each increment. But from 400 to 450 users, even though the percentage of served users decreases, the absolute number of served user is the same, therefore the total capacity used increases in this increment, when a decrease would be expected. When the network saturates, the remaining users are served with minimum (bad) user's satisfaction. That is different for the DelayC-GBR users, which are always served with best user's satisfaction, even when the network is saturated-

The following subsection analyses a Shared Ran Private Network deployment. The simulation consists of an increment in the Shared RAN capacity percentage used by the Private Network. The Mix is slightly changed, as it now considers only Hospital employees or patients, therefore AR/VR is removed and the Non-GBR services mix increased. Users are fixed at 100 users and the 3.5 GHz band is considered. The Shared Ran percentage starts at 10% and increases until 80%. DelayC-GBR services start to be served at 20% with maximum data rate, at the network has enough capacity to serve all users at 40% of Shared RAN capacity. At 40%, Non-GBR user's are served with medium user's satisfaction, while Voice and Music with good satisfaction. All services have good user's satisfaction from 60% of Shared Ran. Next subsection considers a Hospital Standalone Private Network deployment with unlicensed operation. The 5.9 GHz band is considered with a small cell deployment indoor, so that the is no need of a Clear Channel Assessment mechanism. The service mix is also adapted, introducing high demand remote surgery and service robots. Users are fixed to 500 and the results consist of a comparison for the high demand mix between the 3.5 GHz and the 5.9 GHz bands. And the analysis shows that the network is saturated at the 3.5 GHz band, while the 5.9 GHz serves all users with medium to good user's satisfaction.

After that, returning to the 3.5 GHz band to the reference one and a variation of the MIMO Layers and the Service Mix is made. MIMO Layers assume values of 2, 4 and 8. Users were fixed to 100 for both variations. Results show the network saturated with 2 MIMO Layers, showing that this Mix would not work for this cell parameters. 4 MIMO Layers is the parameter used in the first simulations. For 8 MIMO Layers, one verifies enough capacity to double the Remote Surgery data rate and increase by almost 4 times the service robots data rate. Therefore 8 MIMO Layers would allow for a service mix with higher data rate requirements. As for the service mix variation, three new mixes are defined apart from the reference one: Focused on Normal user, Professional AR/VR and Different Hospital Resources. The main result can be taken as the VRRM Capacity Share one, as one verifies the Non-GBR Share increase from 40% to 57% in the Normal User mix and to 73% on the Professional AR/VR one, as would be expected from the increase in the Non-GBR services' mix. Different Hospital Resources mix sees an increase on the numbers of Remote Surgeries and Service Robots, therefore

the DelayC-GBR share goes from 24% on the previous mixes, to 54%.

Regarding latency, hospital results consider the Hospital Santa Maria premises located in Lisbon. RU sites were selected in order to be in a maximum 10 km radius with Collocated RU-DU-CU, MEC deployment and 7.2 Splitting Option. As the minimum Max Latency of the services is 5ms and given by remote surgery, the latency results show that the network would be able to serve all hospital scenarios, with minimum network delay of 0.38 ms and maximum of 0.79 ms for the RU sites farther from the Hospital.

Afterwards, the Smart Factory sector is analysed with the same Private Network deployments as the Hospital one. Starting at the 3.5 GHz band with the reference mix and a Network Slicing Deployment, users start at 200 at receive increments of 50. One verifies that the network saturates at 400 users with all low priority Non-GBR users being delayed and 10% of AR/VR users delayed too. DelayC-GBR Services have best user's satisfaction until 250 users, with it lowering to medium-low at 300 users for the DelayC-GBR services with lowest priority and data rate. Non-GBR VNO has the highest capacity share, reaching 81% at its maximum, as it comprises of AR/VR service. When the network saturates, the Non-GBR share starts to decrease and the DelayC-GBR increases. This happens because Delay-GBR services are the last to be delayed. The next subsection approaches the Smart Factory Shared RAN Private Network Deployment using the same approach as in the Hospital Scenario. A different service mix with less GBR and Non-GBR services mixes is defined. At 10% of Shared capacity, GBR and DelayC-GBR services are served with minimum data rate while the network is saturated. At 20%, AR/VR starts to be served. At 50% all users are served with bad user's satisfaction. DelayC-GBR services reach their medium to good user's satisfaction at 70% of Shared RAN. To end the different Smart Factory scenarios, next one analyses the Smart Factory Standalone Private Network Deployment. Using the 5.9 GHz band and considering small cells, one compares the network performance between the 5.9 GHz and the 3.5 GHz band, with users fixed at 500. Again, one verifies the network saturated at the 3.5 GHz band, with 20% of AR/VR users delayed. At the 5.9 GHz band, all users are served and DelayC-GBR services have medium to good user's satisfaction. Finally, the Smart Factory latency results. For this scenario, the Auto Europa factory, in Lisbon, was chosen. Therefore, RU sites in a 10 km radius were selected. The network delay results show a minimum delay of 0.314 ms and maximum of 0.814 ms. The strictest service is Motion Control, and its latency range goes from 0.5 ms to 2 ms. This means that the sites closer to the factory could serve all users with best user's satisfaction latency wise, while the sites farther, close to the 10 km mark would serve Motion Control sensors with medium to high user's satisfaction.

Regarding the Mission Critical sector, one analyses the two sub-6GHz 5G frequencies, 700 MHz and 3.6 GHz. Both bands use 2 MIMO layers as input parameter and consider a Network Slicing Private Deployment. The 700 MHz deployment considers a Rural topology, and the Service Mix is adapted in a way that no AR/VR is present, as the 700 MHz band capacity is not enough to serve satisfactorily eMBB services. The user incrementation starts at 50 users and increases by 10 per simulation. The network saturates at 100 users with File Sharing users being delayed. At 50 users, the Mission Critical GBR users are served with good user's satisfaction, but it decreases to low at 70 users. Therefore, the

results show that the 700 MHz band would be enough for the mission critical services, but with low user's satisfaction if in an area with more than 70 concurrent users. Next, one analyses the 3.6 GHz band with a change on the service mix, increasing Video mix, adding AR/VR and decreasing MCPTT. Also, an urban topology is considered. In this scenario, the simulations start at 100 users with increments of 50 users. The network saturates at 200 users with all low priority users delayed and 20% the AR/VR ones. At 100 users, all Mission Critical GBR users are served with good user's satisfaction, it lowers to the minimum when the network saturates. Finally, one verifies the latency results for both scenarios considered. The Rural Scenario results consider cells close to the area of Pedrógão Grande located in the centre of Portugal. With an Independent RU and Collocated DU-CU architecture, the network delay results showed a minimum of 2.480 ms and a maximum of 3.362 ms. These results show that the network would be able to provide the best user's satisfaction latency wise for the Pedrogão Grande scenario. The Urban Scenario considers sites in the City of Porto and their network delay results are a minimum of 0.192 ms and a maximum of 2.673 ms. These results show that the urban sites would also serve well all users. The large difference between the maximum values of both scenarios is due to the existence of a Core node in the City of Porto urban scenario.

For future work, one could adapt the VRRM Optimisation model to a dynamic resource allocation and delay process, so that the services resources and user's delay can be treated according to the load of the network. With this, one could verify the latency of the service in real time, taking into account the TTI of each delay cycle. Apart from that, the knowledge of 5G traffic profiles would allow latency results with more accuracy. In addition to that, the security of Private Networks should be addressed in future works, as it is of great importance for the sector.

Annex A

5G QoS Characteristics

This annex summarises the QoS Characteristics of services for the 5G Network (Adapted from [3GPP18]).

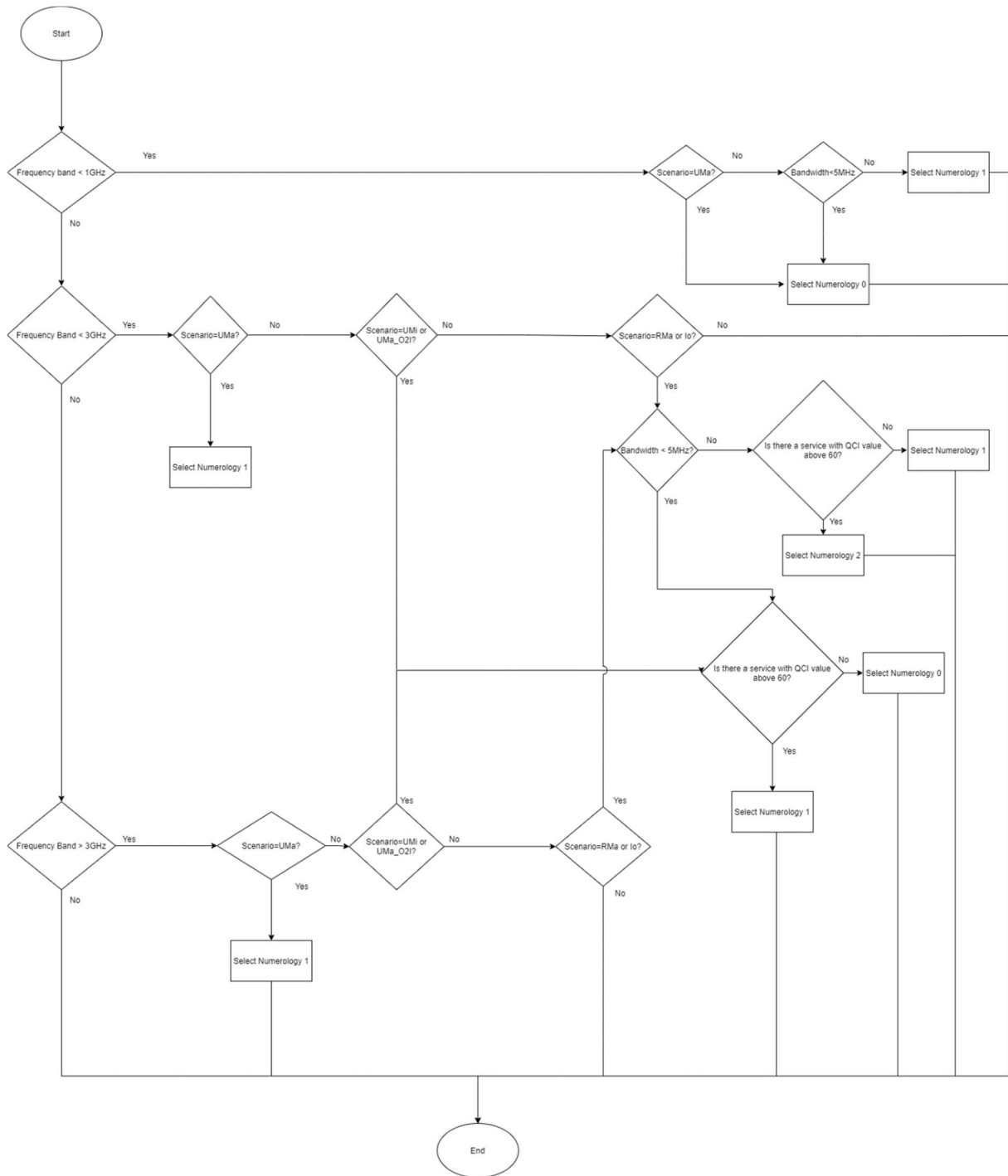
	QCI Value	Default Priority Level	Packet Delay Budget [ms]	Packet Error Rate	Example Services
GBR	1	20	100	10^{-2}	Conversational Voice
	2	40	150	10^{-3}	Conversational video
	3	30	50	10^{-3}	Real Time Gaming, V2X messages Electricity distribution – medium voltage, Process automation - monitoring
	4	50	300	10^{-6}	Non-Conversational Video (Buffered Streaming)
	65	7	75	10^{-2}	Mission Critical user plane Push To Talk voice (e.g., MCPTT
	66	20	100	10^{-2}	Non-Mission-Critical user plane Push To Talk voice
	67	15	100	10^{-3}	Mission Critical Video user plane
	75	25	50	10^{-2}	V2X messages
Non-GBR	5	10	100	10^{-6}	IMS Signalling
	6	60	300	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
	7	70	100	10^{-3}	Voice, Video (Live Streaming), Interactive Gaming
	8/9	80/90	300	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
	69	5	60	10^{-6}	Mission Critical delay sensitive signalling (e.g., MC-PTT signalling)
	70	55	200	10^{-6}	Mission Critical Data (e.g. example services are the

					same as QCI 6/8/9)
	79	65	50	10^{-2}	V2X messages
	80	68	10	10^{-6}	Low Latency eMBB applications Augmented Reality
Delay	82	19	10	10^{-4}	Discrete Automation
Critical GB	83	22	10	10^{-4}	Discrete Automation
	84	24	30	10^{-5}	Intelligent transport systems
	85	21	5	10^{-5}	Electricity Distribution high voltage

Annex B

Numerology Selection Algorithm

This annex shows the Numerology Selection Algorithm developed in the scope of this thesis.



Annex C

Military Applications Key Performance Indicators

This annex summarises the Key Performance Indicators for Military Applications (Adapted from [LiOu20]).

KPI	Value
Priority	High: Battlefield real-time confrontation tasks Middle: collaborative training tasks Low: logistics equipment support tasks
Delay	up to 1 ms
Reliability	Weapon strike: 99.999% Command and control: 99.9% Service support: 99%
Peak User Rate	20Gbps
Mobility	High: >200km/h Medium: 2-200km/h Low:<2km/h
Connection Density	High: >104/km ² Medium: 100~104/km ² Low:<100km ²
Security Classification	High: Classified Medium: Secretive Low: Unsecret
Energy Efficiency	High: Weapon sensor sensing Medium: Battlefield Situation Low: Remote control

Annex D

Scenarios Development

This annex summarises the possible scenarios for Private Networks.

Sector	Service	Application	Maximum Latency [ms]	Reliability	Availability	Data Rate	Connection Density	Failure Rate	Mobility [km/h]	Frequency Bands	References	QCI
	Remote Surgery	Audio/ Video feedback	100	-	-	1Gbps	-	10^{-7} (3.17 s of outage per year) (include text later)	-	Operator/ Unlicensed (Indoor)	[SFRD17]	67
		Haptic Feedback	25									80
		Interactive Live Holographic Feedback	5									86
Healthcare	Service Robots		5	-	-	450Mbps	-	10^{-7} (3.17 s of outage per year)	-	Operator	[GSMA18]	85
Energy	Smart Grid		5	>99.999% (include text later)	-	1 kbps per residential user	>1000 devices/km ²	-	-			
	Train-to-ground		50	-	>99.99%	<10Mbps	-	-	Up to 500		[GMVA17]	79

Railways	Train Communication Network		10	-	>99.999%	1-10Gbps	-	-	-	Specific		83
												6
	Passenger Communications		500	-	>95%	7.5-15 Gbps per train	1000/train	-	Up to 400			
Industry	Monitoring		50-100	>99.9%	-	0.1 Mbps – 0.5 Mbps	10 000 devices/km ²	-	-	Operator / Unlicensed (Indoor)	[Aija20]; [3GPP19]	3
	Safety Control		5-10	>99.999%	-	0.5 Mbps – 1 Mbps	1 000 devices/km ²	-	-			82
	Closed-Loop Control		2-10	>99.999%	-	1 Mbps – 5 Mbps	1 000 devices/km ²	-	-			
	Motion Control		0.5-2	>99.9999%	-	1 Mbps – 5 Mbps	1 000 devices/km	-	-			

							2					
Military		Service Support	1	99%	-	Up to 20Gbps	Low:<100 devices/km ²	-	Low:<2	Specific	[LiOu20]	70
		Command and Control		99.9%			Medium: 100~104/km ²		Medium: 2-200			85
		Weapon Strike		99.999%			High: >104/km ²		High: >200			86

Annex E

Additional Results

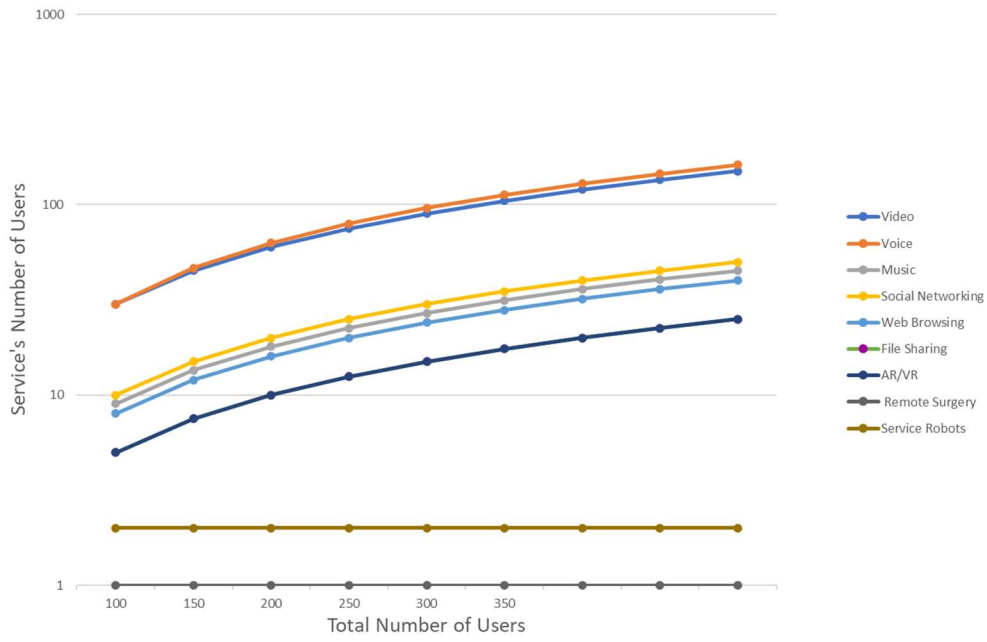


Figure E.1. Each service's number of users versus the total number of users for the Hospital Network Slicing Private Deployment.

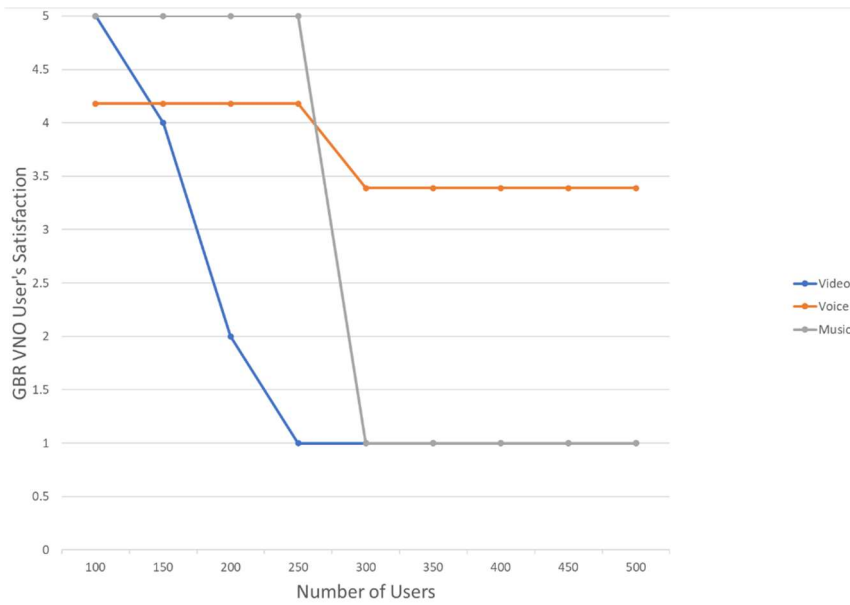


Figure E.2 GBR VNO User's Satisfaction versus the number of users for the Hospital Network Slicing Private Deployment.

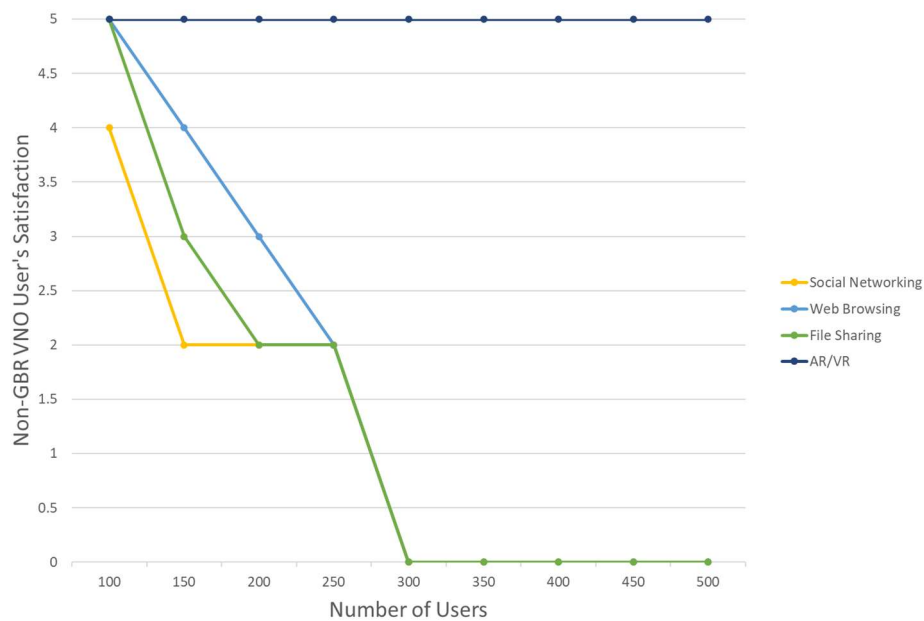


Figure E.3. Non-GBR VNO User's Satisfaction versus the Number of users for the Hospital Network Slicing Private Deployment.

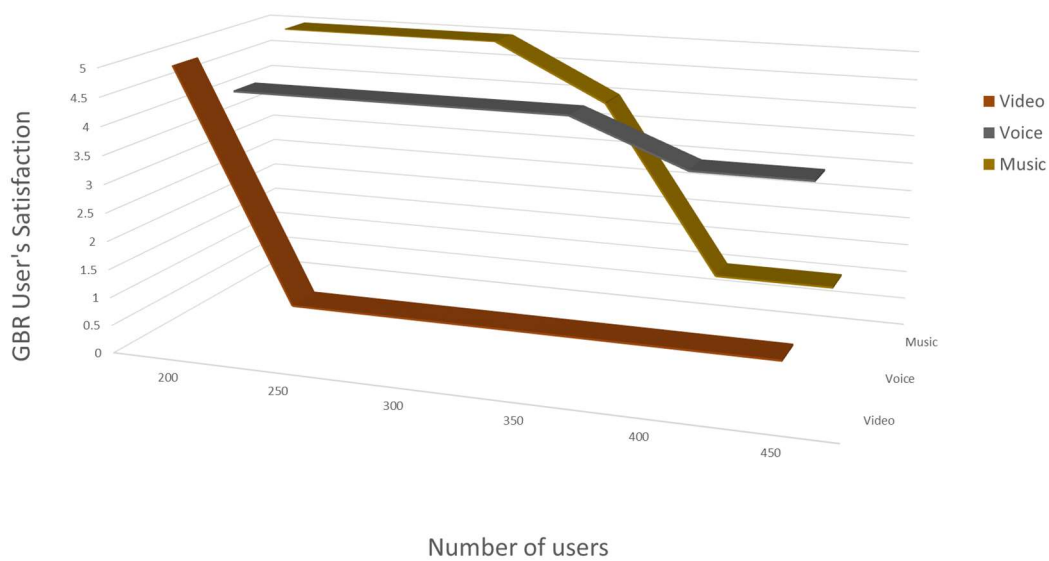


Figure E.4. GBR VNO User's Satisfaction for Smart Factory Network Slicing Private Deployment Scenario

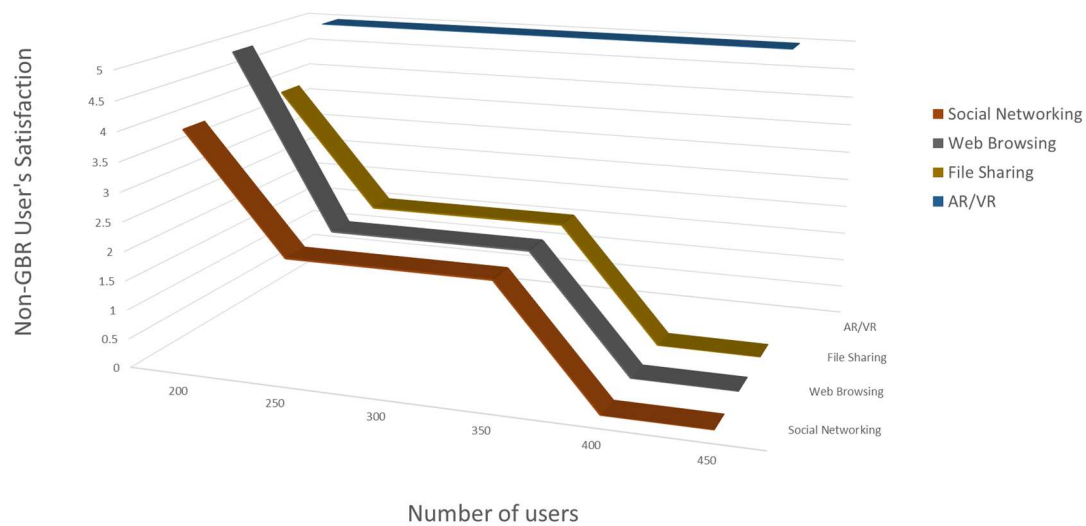


Figure E.5. Non-GBR VNO User's Satisfaction for Smart Factory Network Slicing Private Deployment Scenario

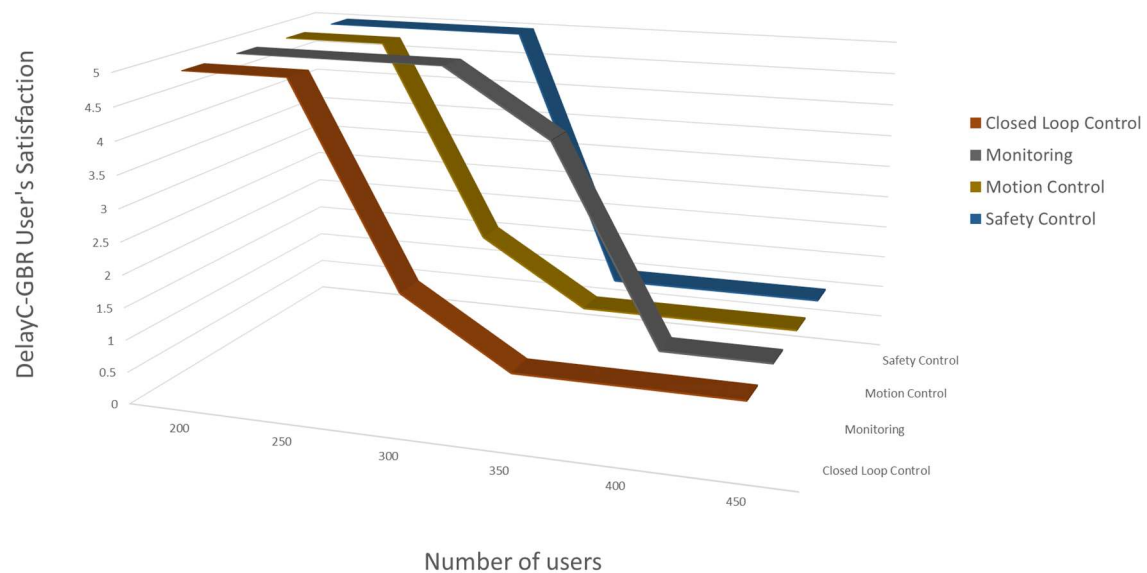


Figure E.6. DelayC-GBR User's Satisfaction for Smart Factory Network Slicing Private Deployment Scenario.

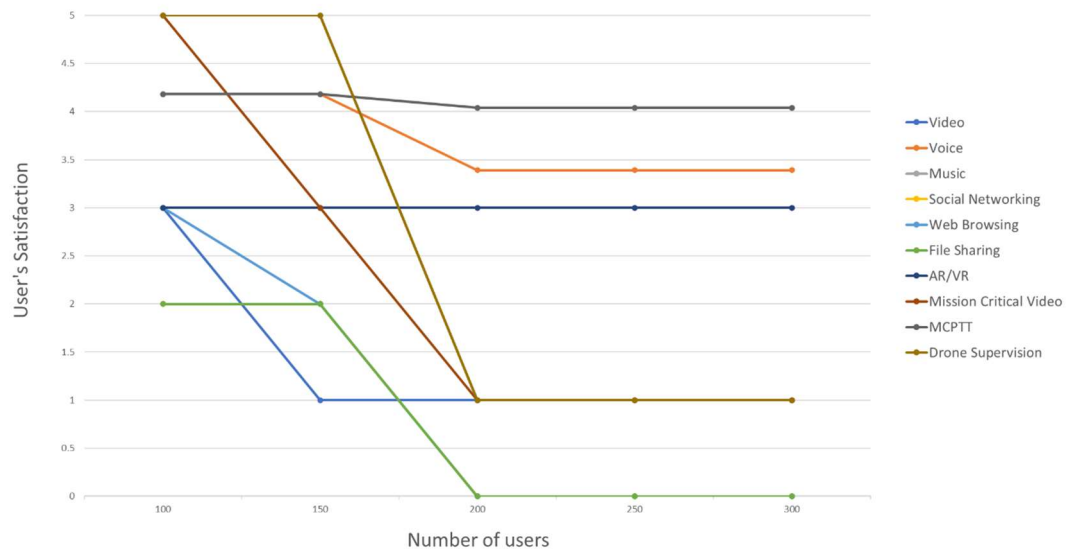


Figure E.7. User's Satisfaction versus the Number of Users for the 3.6GHz Mission Critical Scenario.

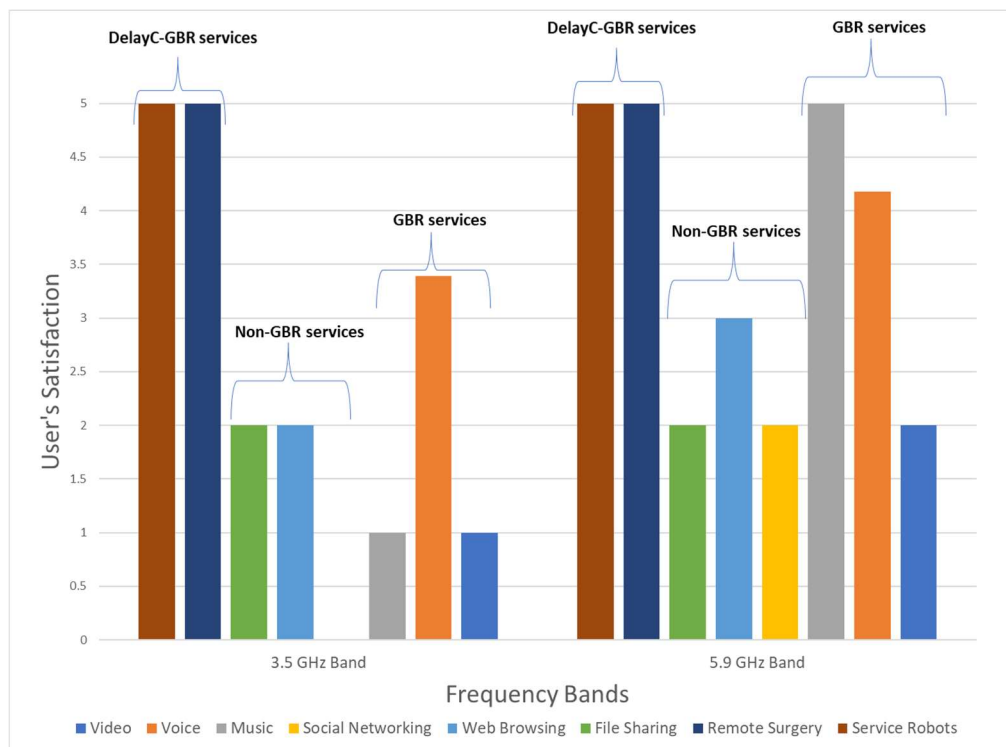


Figure E.8. User's Satisfaction for both 3.5 GHz and 5.9 GHz frequency bands.

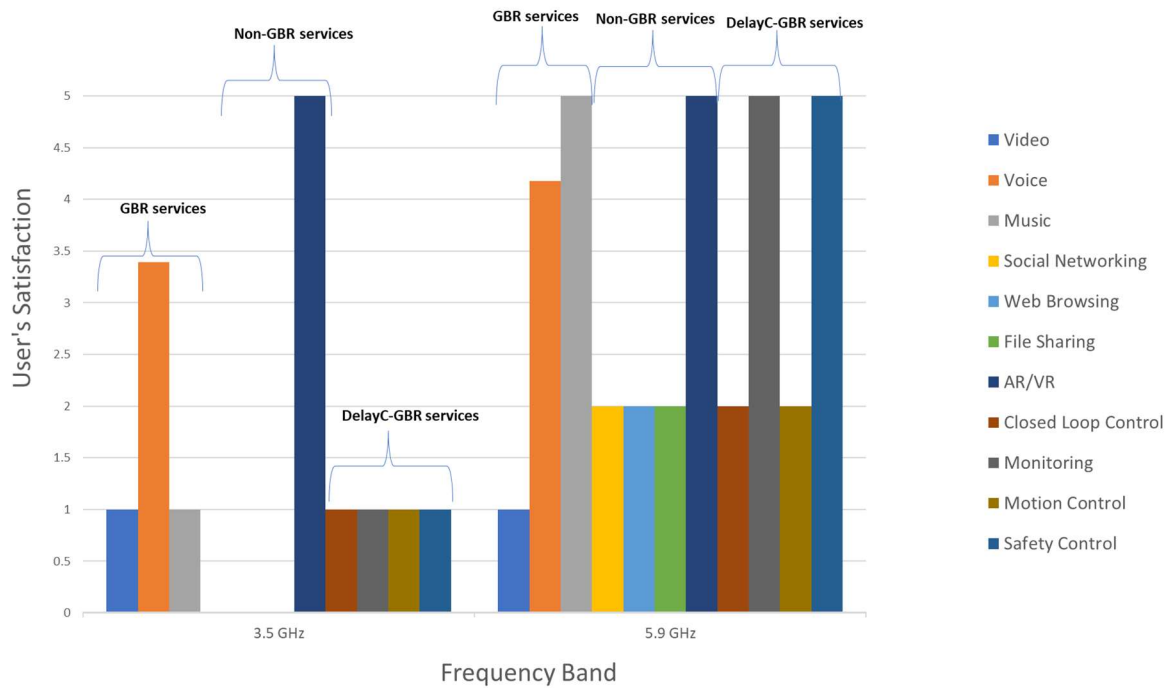


Figure E.9. User's Satisfaction for both 3.5 GHz and 5.9 GHz frequency bands.

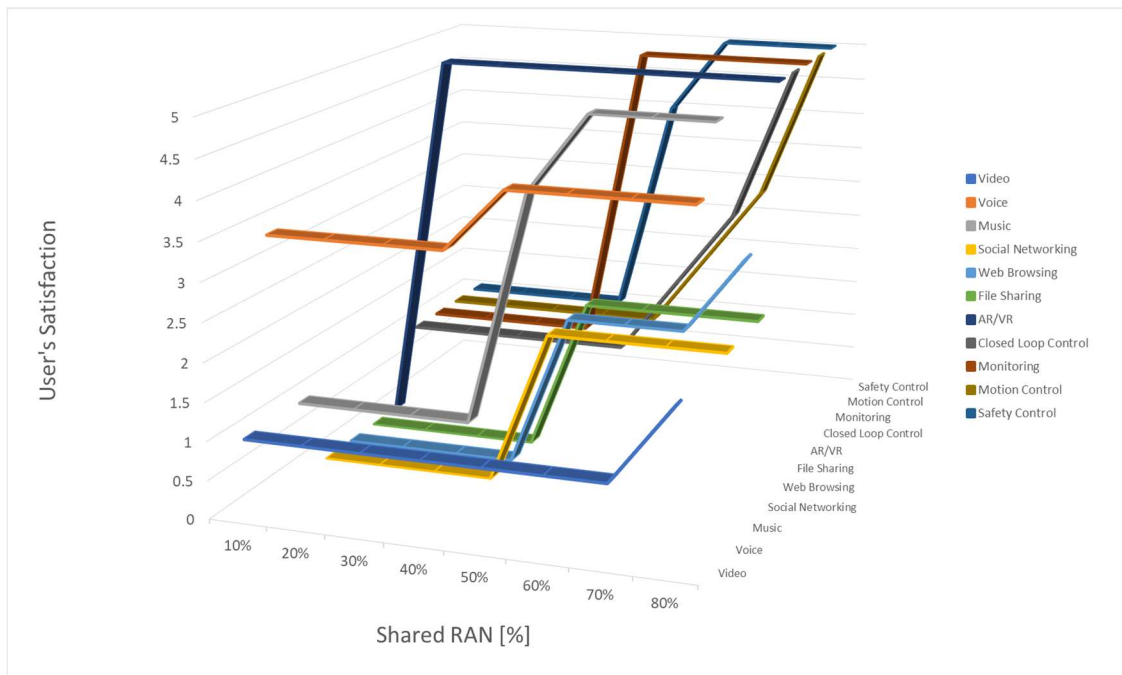


Figure E.10. User's Satisfaction versus Shared RAN percentage for the Smart Factory Shared RAN scenario.

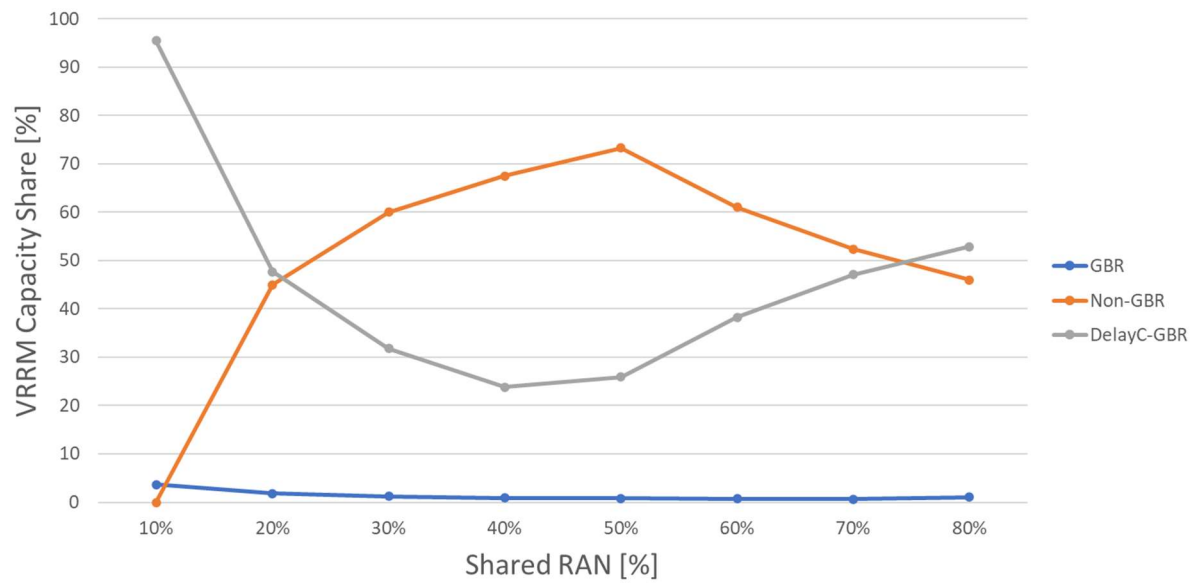


Figure E.11. VRRM Capacity Share versus Shared RAN percentage for the Smart Factory scenario.

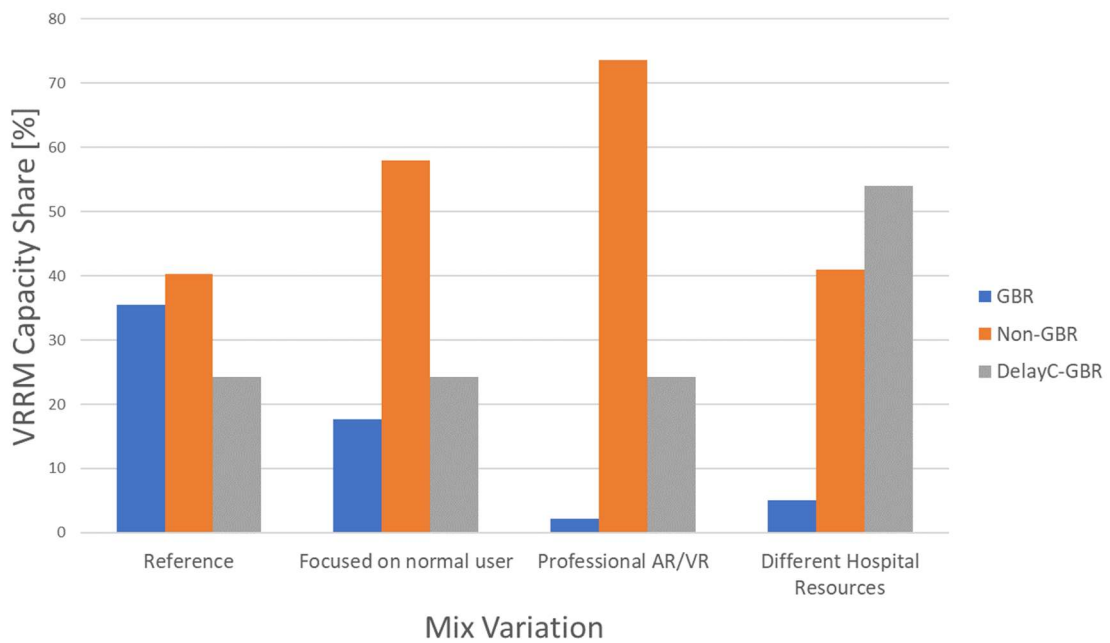


Figure E.12. VRRM Capacity Share versus Mix Variation for the Hospital Network Slicing Private Deployment.

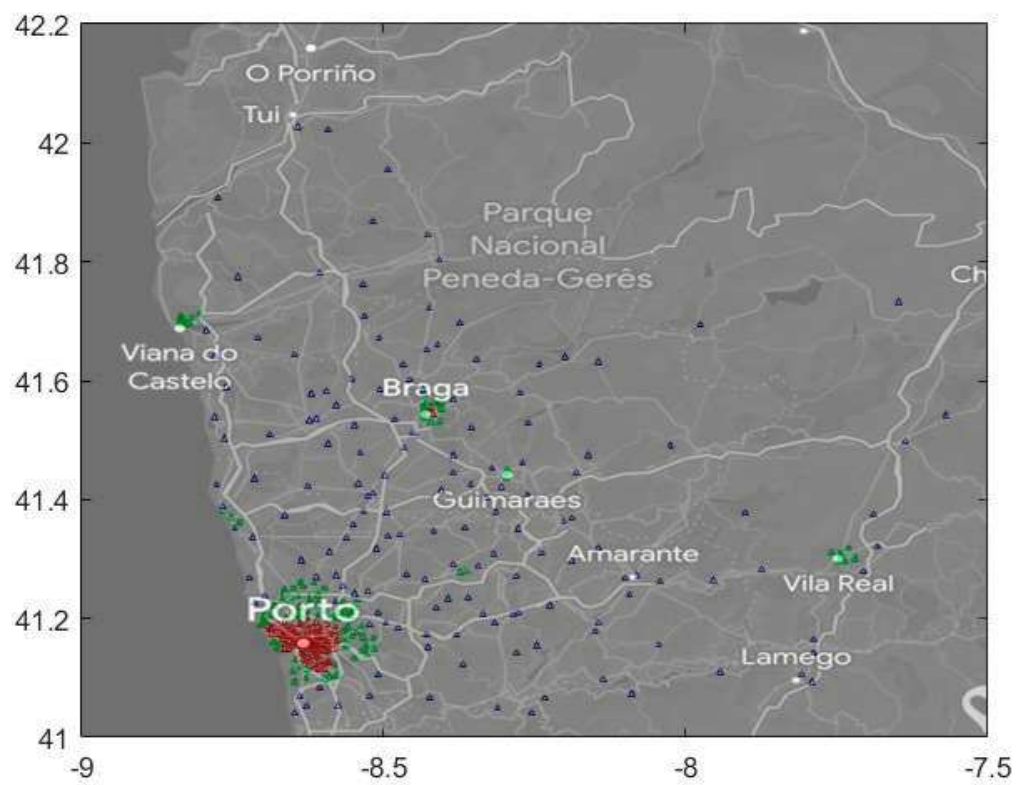


Figure E.13. RU sites provided by NOS for the 3600 MHz Mission Critical Scenario at city of Porto..

References

- [3GPP08] 3GPP, *LTE*, <https://www.3gpp.org/technologies/keywords-acronyms/98-lte>, Dec. 2008.
- [3GPP17] 3GPP, Technical Specification Group – User Equipment (UE) radio transmission and reception; TR38.101-1 V1.0.0 Release 15, Dec. 2017
- [3GPP18] 3GPP, *5G; System Architecture for the 5G System*, ETSI TS 23.501 version 15.3.0 Release 15, Sep. 2018.
- [3GPP19] 3GPP, *5G; Service requirements for next generation new services and markets*, ETSI TS 22.261 version 15.7.0 Release 15, Mar. 2019.
- [3GPP20] 3GPP, *User Equipment (UE) radio access capabilities*, TS 38.3063, Mar. 2020.
- [5GAm17] 5G Americas, *5G Services & Use Cases*, White Paper, Nov. 2017.
- [5GAm18] 5G Americas, *Wireless Technology Evolution Transition from 4G to 5G 3GPP releases 14 to 16*, White Paper, Oct. 2018.
- [5GAm19] 5G Americas, *5G Services Innovation*, White Paper, Nov. 2019.
- [5GAm20] 5G Americas, *The 5G Evolution: 3GPP Releases 16-17*, White Paper, Jan. 2020
- [Aija20] A. Aijaz, "Private 5G: The Future of Industrial Wireless," *IEEE Industrial Electronics Magazine*, vol. 14, no. 4, Dec. 2020, pp. 136-145.
- [AMRZ20] B. Ai, A.F. Molisch, M. Rupp and Z.-D. Zhong, "5G Key Technologies for Smart Railways," *Proceedings of the IEEE*, vol. 108, no. 6, June 2020, pp. 856-893.
- [BaCK20] L. Bastos, G. Capela and A. Koprulu, *Potential of 5G technologies for military application*, NATO Communications and Information Agency Working Paper, Sep. 2020. Available at: <http://www.mindev.gov.gr/wp-content/uploads/2020/11/Enclosure-2-Working-paper-Potential-of-5G-technologies-for-military-application.pdf>.
- [BoDF19] M.S. Bonfim, K.L. Dias and S.F. Fernandes, "Integrated NFV/SDN architectures: A systematic literature review," *ACM Computational Surveys*, vol. 51, no. 6, Feb. 2019, pp. 114.
- [BsSJ20] R. Bajracharya, R. Shrestha and H. Jung, "Future Is Unlicensed: Private 5G Unlicensed Network for Connecting Industries of Future", *Sensors*, vol. 20, no. 10, May 2020, pp. 2774.
- [Corr20] L.M. Correia, *Notes from Mobile Communication course*, Instituto Superior Técnico, Lisbon, Portugal, 2020.
- [DaPS18] E. Dahlman, S. Parkvall and J. Sköld, *5G NR: The Next Generation Wireless Access Technology*, Elsevier - Academic Press, London, UK, 2018.

- [DDLo15] B. Debaillie, C. Desset, and F. Louague, "Flexible and Future-Proof Power Model for Cellular Base Stations", Proc. of VTC Spring 2015 – 81st IEEE Vehicular Technology Conference, Glasgow, United Kingdom, May 2015.
- [Domi19] S. Domingues, *Analysis of the Performance of Multi-Access Edge Computing Network Slicing in 5G*, M.Sc. Thesis, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal, 2019.
- [FPEM17] X. Foukas, G. Patounas, A. Elmokashfi and M. K. Marina, "Network Slicing in 5G: Survey and Challenges," *IEEE Communications Magazine*, vol. 55, no. 5, May 2017, pp. 94-100.
- [GMBC19] A. Ghosh, A. Maeder, M. Baker and D. Chandramouli, "5G Evolution: A View on 5G Cellular Technology Beyond 3GPP Release 15", *IEEE Access*, vol. 7, Sep. 2019, pp. 127639-127651.
- [GMVA17] A. Gonzalez-Plaza, J. Moreno, I. Val, A. Arriola, P.M Rodriguez, F. Jimenez and C. Briso, "5G communications in high speed and metropolitan railways," *Proc. of 11th European Conference on Antennas and Propagation (EUCAP 2017)*, Paris, France, Mar. 2017.
- [GSMA18] GSMA, *Network Slicing Use Case Requirements*, White Paper, Apr 2018.
- [HaWW19] Z. Haibo, S. Weiping and Z. Weijun, "A Network Slicing Scheme for Private Communication in State Grid," *Proc. of 2nd IEEE International Conference on Information Communication and Signal Processing (ICICSP 2019)*, Weihai, China, Sep. 2019.
- [HaYW19] N. Hassan, K.A. Yau and C. Wu, "Edge Computing in 5G: A Review," *IEEE Access*, vol. 7, Aug. 2019, pp. 127276-127289.
- [HMML19] N. Huin, P. Medagliani, S. Martin, J. Leguay, L. Shi, S. Cai, J. Xu and H. Shi, "Hard-isolation for Network Slicing," *Proc of IEEE Conference on Computer Communications Workshops (INFOCOM 2019)*, Paris, France, May 2019.
- [HNHS19] M.A. Habibi, M. Nasimi, B. Han and H D. Schotten, "A Comprehensive Survey of RAN Architectures Toward 5G Mobile Communication System," *IEEE Access*, vol. 7, May 2019, pp. 70371-70421.
- [HUAW17] HUAWEI, *Service-Based Architecture for 5G Core Networks*, White Paper, Nov. 2017.
- [KYTH20] L.U. Khan, I. Yaqoob, N.H. Tran, Z. Han and C.S. Hong, "Network Slicing: Recent Advances, Taxonomy, Requirements, and Open Research Challenges," *IEEE Access*, vol. 8, Feb. 2020, pp. 36009-36028.
- [LiOu20] J. Liao and X. Ou, "5G Military Application Scenarios and Private Network Architectures," *Proc. of IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA 2020)*, Dalian, China, Aug. 2020.
- [Mari20] S. Marinheiro, *Analysis of the Implementation of Network Slicing in 5G Radio Networks*, M.Sc. Thesis, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal, 2020.

- [MZLZ20] N. Ma, X. Zhong, P. Liu and S. Zhou, "A SDN/NFV-based Core Network Slicing for Secure Mobile Communication," *Proc. of 91st IEEE Vehicular Technology Conference (VTC 2020-Spring)*, Antwerp, Belgium, May 2020.
- [OIFa17] S.O. Oladejo and O.E. Falowo, "5G network slicing: A multi-tenancy scenario," *Proc. of Global Wireless Summit (GWS 2017)*, Cape Town, Oct. 2017.
- [PFHP20] Q.-V. Pham, F. Fang, V.N. Ha, M.J. Piran, M. Le, L.B. Le, W.-J. Hwang and Z. Ding, "A Survey of Multi-Access Edge Computing in 5G and Beyond: Fundamentals, Technology Integration, and State-of-the-Art," *IEEE Access*, vol. 8, June 2020, pp. 116974-117017.
- [PMKL19] P. Porambage, Y. Miche, A. Kalliola, M. Liyanage and M. Ylianttila, "Secure Keying Scheme for Network Slicing in 5G Architecture," *Proc. of IEEE Conference on Standards for Communications and Networking (CSCN 2019)*, Granada, Spain, Oct. 2019.
- [PoOM20] W.Y. Poe, J. Ordonez-Lucena and K. Mahmood, "Provisioning Private 5G Networks by Means of Network Slicing: Architectures and Challenges," *Proc. of IEEE International Conference on Communications Workshops (ICC 2020)*, Dublin, Ireland, July 2020.
- [Rouz19] B. Rouzbehani, *On-demand RAN Slicing Techniques for SLA Assurance in Virtual Wireless Networks*, Ph.D. Thesis, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal, 2019.
- [SaKn18] Sanwork Knowledge, *LTE Network Architecture General overview Hindi*, Jan. 2018 Available at: <https://www.youtube.com/watch?v=6dt9xVMvtB8>.
- [SaMa19a] D. Sattar and A. Matrawy, "Towards Secure Slicing: Using Slice Isolation to Mitigate DDoS Attacks on 5G Core Network Slices," *Proc. of IEEE Conference on Communications and Network Security (CNS 2019)*, Washington, DC, USA, Jun. 2019.
- [SaMa19b] D. Sattar and A. Matrawy, "Optimal Slice Allocation in 5G Core Networks," *IEEE Networking Letters*, vol. 1, no. 2, June 2019, pp. 48-51.
- [SFRD17] D. Soldani, F. Fadini, H. Rasanen, J. Duran, T. Niemela, D. Chandramouli, T. Hoglund and K. Doppler, "5G Mobile Systems for Healthcare," *Proc. of 85th IEEE Vehicular Technology Conference (VTC 2017-Spring)*, Sydney, Australia, June 2017.
- [VaMA21] J.F. de Valgas, J.F. Monserrat and H. Arslan, "Flexible Numerology in 5G NR: Interference Quantification and Proper Selection Depending on the Scenario", *Mobile Information Systems*, vol. 2021, Article ID 6651326, 9 pages, 2021.
- [VeNN20] A. Vergutz, G. Noubir and M. Nogueira, "Reliability for Smart Healthcare: A Network Slicing Perspective", *IEEE Network*, vol. 34, no. 4, July/Aug. 2020, pp. 91-97.
- [YBSS17] F.Z. Yousaf, M. Bredel, S. Schaller and F. Schneider, "NFV and SDN—Key Technology Enablers for 5G Networks," *IEEE Journal on Selected Areas Communications*, vol. 35, no. 11, Nov. 2017, pp. 2468-2478.

- [YLWW19] X. Yang, Y. Liu, I.C. Wong, Y. Wang and L. Cuthbert, "Effective isolation in dynamic network slicing," *Proc. of IEEE Wireless Communications and Networking Conference (WCNC 2019)*, Marrakesh, Morocco, Apr. 2019.
- [YLWW20] X. Yang, Y. Liu, I. C. Wong, Y. Wang and L. Cuthbert, "Genetic Algorithm for Inter-Slice Resource Management in 5G Network with Isolation," *Proc. of International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2020)*, Split, Croatia, Sep. 2020.
- [ZhLZ18] Q. Zhang, J. Liu and G. Zhao, "*Towards 5G enabled tactile robotic telesurgery*", arXiv preprint arXiv:1803.03586, 2018.
- [ZLCL17] H. Zhang, N. Liu, X. Chu, K. Long, A. Aghvami and V. C. M. Leung, "Network Slicing Based 5G and Future Mobile Networks: Mobility, Resource Management, and Challenges," *IEEE Communications Magazine*, vol. 55, no. 8, Aug. 2017, pp. 138-145.