TÉCNICO LISBOA

Performance comparison of voice communications between VoLTE and UMTS/GSM.

Ricardo Manuel Pinto dos Santos Gameiro

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 Chanters Ribeiro da Cunha Sanguino

Supervisor: Prof. Luís Manuel de Jesus Sousa Correia

Members of Committee: Prof. António José Castelo Branco Rodrigues

Eng. Marco Serrazina

November 2015

To the ones I love

Acknowledgements

First of all, I would like to thank Prof. Luís M. Correia, for the opportunity to work under his guidance, for the suggestions and advices on how to perform a master thesis; without his guidance and present support, this master thesis would not have been possible. Under his guidance I could learn on working methods, and improve my performance, which allow me to perform this thesis.

To Vodafone Portugal, for allowing me to develop a work with a great relevance at business level, and in particular to Engs. Pedro Lourenço and Marco Serrazina for the time and effort spent on meetings, suggestions and critics during my work.

To all GROW members, for the meetings we had during this year, and specially to all my master's thesis colleagues and friends: Ana Claúdia Casto, Andrea Marotta, Carlos Martins, João Pires and Miguel Sá.

To my colleagues and friends that accompanied me throughout the journey in IST, in particular: Daniel Silvestre, and Pedro Ganço. To my colleagues and friends that accompanied me the journey in ISEL, in particular: Diogo Silva, Gilberto Conde, Filipe Palhinha, João Sêbo Tiago Inacio, and Valter Rodrigues, and the ones that accompanied me in both journeys: Tiago Felício, Ricardo Romão, and João Martinho.

To the friends that I met in Campimeco, with whom I have lived so many experiences, for more than 15 years of friendship.

To my mother, for the support she gave me, and my father that could not be present on my last academic challenge, that I know he would be proud to witness.

To Daniela, that has been my best friend and my love for almost a decade. For being patient, for the support, and the love, that she gave me during all the journeys together.

Abstract

The goal of this thesis was to analyse the performance of the voice communications over LTE, which are made over Internet, comparing the performance achieved, with the well-known performance of the mobile communications over circuit switching, and the study of the impact of the voice packets on the LTE network in the presence of data packets. This work consists of a study of the delay that the voice has between the instant that an interlocutor speaks, until the instant that the other interlocutor listens, in order to analyse the performance of the voice, and the study of the impact on data services due to the presence of the voice service. Two distinct scenarios were implemented on a simulator, both scenarios had a single base station, the first scenario had just voice communications in UMTS and LTE, both made over Internet, where the mean call duration, and the number of users have been changed, in order to analyse the performance of the network. The second scenario had voice and data communications in LTE, where the number of user performing the voice service have been changed, maintaining a fixed mean call duration. Different data services have been used, with different configurations. The results shows that the number of voice users tend to decrease with the increment of the mean call duration, and that the video conferencing is the service which has the higher impact on their performance, due to the priority of the voice service.

Keywords

LTE, UMTS, VoLTE, Voice, Delay.

Resumo

O objectivo desta tese foi o de analisar o desempenho das comunicações de voz sobre LTE, que são feitas sobre a Internet, comparando o desempenho alcançado nesta rede, com o desempenho conhecido das comunicações móveis sobre comutação de circuitos, e o estudo que os pacotes de voz têm sobre a rede LTE na presença de pacotes de dados. Este trabalho consiste no estudo do atraso que a voz tem entre o instante que um interlocutor fala, até ao instante que o outro interlocutor ouve, com o objectivo de analisar a performance da voz, e o estudo do impacto nos serviços de dados devido à presença do serviço de voz. Foram implementados dois cenários distintos num simulador, ambos os cenários tiveram uma estação base, o primeiro cenário teve apenas comunicações de voz em UMTS e LTE, onde a duração média de chamadas, e o número de utilizadores ligados à rede foram alterados, com o intuito de analisar o desempenho da rede. O segundo cenário teve comunicações de voz e dados em LTE, onde o número de utilizadores a realizar o serviço de voz foram alterados, mantendo uma duração média de chamada fixa. Foram utilizados diferentes serviços de dados, com configurações diferentes. Os resultados mostram que o número de utilizadores de voz tende a diminuir com o aumento da duração média de chamada, e que a videoconferência é o serviço que tem o maior impacto na sua performance, devido à prioridade do serviço de voz.

Palavras-chave

LTE, UMTS, VoLTE, Voz, Atraso.

Table of Contents

Acknowledgements v
Abstract vii
Resumo viii
Table of Contents ix
List of Figures xi
List of Tables xv
List of Acronyms xvi
List of Symbols xx
List of Software xxii
1 Introduction 1
1.1 Overview 2
1.2 Motivation and Contents 4
2 Basic Concepts 7
2.1 Network Architecture 8
2.1.1 GSM/UMTS 8
2.1.2 LTE 9
2.2 Radio Interfaces 10
2.3 Voice Service and Performance Parameters 14
2.3.1 Voice over GSM/UMTS 14
2.3.2 Voice over LTE 16
2.4 State of the Art 18
3 Theoretical Models and Simulator Description 23
3.1 Theoretical Models 24
3.2 OPNET Simulator 26
3.2.1 General Description 26
3.2.2 Traffic Generation 28
3.2.3 Metrics 29

3.3 Simulator Implementation 32
3.4 Simulator Assessment 38

4 Results Analysis... 41

4.1 Scenarios Description..42
4.2 Voice Analysis on UMTS...44
4.3 Voice Analysis on LTE 55
4.4 Data Analysis on LTE 65

5 Conclusions 75

References 105

List of Figures

Figure 1.1. Global total traffic in mobile networks, 2010-2014 (extracted from [Eric15]). 3
Figure 1.2. Mobile voice minutes of use, VoWi-Fi, VoLTE and VoIP (extracted from [Cisc15]). .. 3
Figure 2.1. GSM/UMTS network architecture (adapted from [Corr13]). 9
Figure 2.2. System architecture for 3GPP access networks (extracted from [HoTo09]). 10
Figure 2.3. Frame structure type 1 for DL (extracted from [Agil07])... 12
Figure 2.4. DL slot structure of bandwidths above 1.4 MHz (extracted from [HoTo09]). 13
Figure 2.5. Basic time-frequency resource structure of LTE with normal CP (extracted from [SeTo11]). 13
Figure 2.6. Main speech KPIs (extracted from [HaRo03]). 15
Figure 2.7. Dynamic and semi-persistent scheduling for voice (extracted from [PoHo12]). 17
Figure 3.1. Process domain... 26
Figure 3.2. Node domain. 27
Figure 3.3. Network domain. 27
Figure 3.4. Examples of attributes edition for the Application Definition and Profile Definition nodes. 29
Figure 3.5. Available statistics. 30
Figure 3.6. Attributes configuration for the Application Definition and Profile Definition for the FTP application. 33
Figure 3.7. UMTS network... 34
Figure 3.8. UMTS' attributes configuration for the UE and RNC. 35
Figure 3.9. LTE network. 35
Figure 3.10. Bearers' configuration on LTE Attributes and UE. 36
Figure 3.11. Packet Data Convergence Protocol configuration. 37
Figure 3.12. Physical definitions on LTE Attributes and eNodeB. 38
Figure 3.13. Simulation assessment – number of samples. 39
Figure 3.14. LTE assessment – progress of the average end to end delay in a 70 minutes simulation. 40
Figure 3.15. Simulation assessment – simulation time convergence analysis. 40
Figure 3.16. Simulation assessment – number of seeds convergence analysis. 40
Figure 4.1. End to end delay results for UMTS with mean call duration of 50 seconds. 45
Figure 4.2. End to end delay results for UMTS with mean call duration of 100 seconds. 45
Figure 4.3. End to end delay results for UMTS with mean call duration of 150 seconds. 46
Figure 4.4. All end to end delay for UMTS. 46
Figure 4.5. Number of users in UMTS as a function of the mean call duration. 47
Figure 4.6. Global MOS for UMTS with a mean call duration of 50 seconds. 48
Figure 4.7. End to end delay for UMTS with mean call duration of 50 seconds, 40 users (seed 1). 48
Figure 4.8. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 1) 49
Figure 4.9. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 1) 50
Figure 4.10. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed

1). 51
Figure 4.11. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed1). 51
Figure 4.12 PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed1). 52
Figure 4.13. Curve fitting example for UMTS, mean call duration of 50 s, 40 users (seed1). 53
Figure 4.14. Comparison between original sampling and core sampling, UMTS, 40 users. 54
Figure 4.15. Comparison between original sampling and core sampling, UMTS, 50 users. 54
Figure 4.16. End to end delay results for LTE with mean call duration of 50 seconds. 55
Figure 4.17. End to end delay results for LTE with mean call duration of 100 seconds. 56
Figure 4.18. End to end delay results for LTE with mean call duration of 150 seconds. 56
Figure 4.19. End to end delay results for LTE with mean call duration of 200 seconds. 57
Figure 4.20. All end to end delay for LTE. 57
Figure 4.21. Number of users in LTE as a function of the mean call duration. 58
Figure 4.22. Curve fitting for a rational equation. 59
Figure 4.23. Global MOS for LTE with a mean call duration of 50 s. 60
Figure 4.24. PDF of the end to end delay for LTE, mean call duration of 50 s, 400 users. 60
Figure 4.25. PDF of the end to end delay for LTE, mean call duration of 100 s, 400 users. 61
Figure 4.26. PDF of the end to end delay for LTE, mean call duration of 150 s, 400 users. 62
Figure 4.27. PDF of the end to end delay for LTE, mean call duration of 200 s, 400 users 63
Figure 4.28. Comparison between original sampling and core sampling, LTE, 400 users. 65
Figure 4.29. Packet end to end delay for the 1st configuration of the Video Conference service. 66
Figure 4.30. Throughput for the 1st configuration of the Video Conference service. 67
Figure 4.31. Download response time for the 1st configuration of the FTP service. 67
Figure 4.32. Upload response time for the 1st configuration of the FTP service. 68
Figure 4.33. Throughput for the 1st configuration of the FTP service. 68
Figure 4.34. Download response time for the 1st configuration of the Email service. 69
Figure 4.35. Upload response time for the 1st configuration of the Email service. 70
Figure 4.36. Throughput for the 1st configuration of the Email service. 70
Figure 4.37. Page response time for the 1st configuration of the HTTP service. 71
Figure 4.38. Throughput for the 1st configuration of the HTTP service. 71
Figure A.1. UMTS assessment – progress of the average end to end delay in a 70 minutes simulation. 80
Figure A.2. Global MOS for UMTS with a mean call duration of 100 seconds. 80
Figure A.3. Global MOS for UMTS with a mean call duration of 150 seconds. 81
Figure A.4. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 2). 81
Figure A.5. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 3). 81
Figure A.6. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 4). 82
Figure A.7. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 5). 82
Figure A.8. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 2). 83
Figure A.9. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 3). 84
Figure A.10. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 4). 84

Figure A.11. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 5). 84
Figure A.12. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed 2). 85
Figure A.13. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed 3). 86
Figure A.14. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed 4). 86
Figure A.15. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed 5). 86
Figure A.16. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed 2). 87
Figure A.17. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed 3). 88
Figure A.18. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed 4). 88
Figure A.19. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed 5). 88
Figure A.20. PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed 2). 89
Figure A.21. PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed 3). 90
Figure A.22. PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed 4). 90
Figure A.23. PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed 5). 90
Figure A.24. All end to end delay for LTE in a logarithmic scale. 93
Figure A.25. Global MOS for LTE with a mean call duration of 100 seconds. 93
Figure A.26. Global MOS for LTE with a mean call duration of 150 seconds. 94
Figure A.27. Global MOS for LTE with a mean call duration of 200 seconds. 94
Figure A.28. PDF of the end to end delay for LTE, mean call duration of 50 s, 500 users. 94
Figure A.29. PDF of the end to end delay for LTE, mean call duration of 100 s, 500 users. 95
Figure A.30 Comparison between original sampling and core sampling, LTE, 500 users. 96
Figure A.31 Curve fitting for the packet end to end delay of the Video Conference (1st Configuration). 96
Figure A.32 Curve fitting for the throughput of the Video Conference (1st Configuration). 97
Figure A.33 Curve fitting for the download response time of the FTP (1st Configuration). 97
Figure A.34 Curve fitting for the download response time of the Email (1st Configuration). 97
Figure A.35 Curve fitting for the upload response time of the Email (1st Configuration). 98
Figure A.36 Packet end to end delay for the 2nd configuration of the Video Conference service. 98
Figure A.37 Curve fitting for the packet end to end delay of the Video Conference (2nd Configuration). 98
Figure A.38 Throughput for the 2nd configuration of the Video Conference service. 99
Figure A.39 Curve fitting for the throughput of the Video Conference (2nd Configuration). 99
Figure A.40 Download response time for the 2nd configuration of the FTP service. 100
Figure A.41 Curve fitting for the download response time of the FTP (2nd Configuration). 100
Figure A.42 Upload response time for the 2nd configuration of the FTP service. 101
Figure A.43 Curve fitting for the upload response time of the FTP (2nd Configuration). 101
Figure A.44. Throughput for the 2nd configuration of the FTP service. 102
Figure A.45 Download response time for the 2nd configuration of the Email service. 102

Figure A.46 Curve fitting for the download response time of the Email (2nd Configuration)...... 102
Figure A.47 Upload response time for the 2nd configuration of the Email service..................... 103
Figure A.48 Throughput for the 2nd configuration of the Email service. 103
Figure A.49 Page response time for the 2nd configuration of the HTTP service. 104
Figure A.50 Throughput for the 2nd configuration of the HTTP service. 104

List of Tables

Table 2.1. National frequency allocation (adapted from [ANA12a], [ANA12b] and [ANA12c]). .. 12
Table 2.2. Speech codec Supported (extracted from [HaRo03]). 15
Table 2.3. Quality of service characteristics (extracted from [PoHo12]). 17
Table 2.4. Simulation results (extracted from [AnMo11]). 21
Table 3.1. Relation between R-value and user satisfaction (extracted from [ITUT14]). 31
Table 3.2. OPNET configurations. 37
Table 3.3. Data traffic mix (adapted from [Alme13]). 38
Table 4.1. FTP configurations. 43
Table 4.2. Email configurations. 43
Table 4.3. HTTP configurations. 43
Table 4.4. Video conference configurations. 44
Table 4.5. PDF values for UMTS, mean call duration of 50 s, 40 users (seed 1). 49
Table 4.6. PDF values for UMTS, mean call duration of 50 s, 50 users (seed 1). 50
Table 4.7. PDF values for UMTS, mean call duration of 100 s, 40 users (seed 1). 51
Table 4.8. PDF values for UMTS, mean call duration of 100 s, 50 users (seed 1). 52
Table 4.9 PDF values for UMTS, mean call duration of 150 s, 40 users (seed 1). 52
Table 4.10. Global curve fitting values for UMTS, for 40 users, and for seed number 1, with a mean call duration of 50 s. 53
Table 4.11. Global curve fitting values for UMTS, for 50 users. 54
Table 4.12. PDF values for LTE, mean call duration of 50s, 500 users. 61
Table 4.13. PDF values for LTE, mean call duration of 100s, 500 users. 62
Table 4.14. PDF values for LTE, mean call duration of 150s, 400 users. 63
Table 4.15. PDF values for LTE, mean call duration of 200s, 400 users. 64
Table 4.16. Curve fitting values for LTE. 64
Table A.1. Remaining PDF values for UMTS, mean call duration of 50 s, 40 users. 83
Table A.2. Remaining discarded values for UMTS, mean call duration of 50 s, 40 users. 83
Table A.3. Remaining PDF values for UMTS, mean call duration of 50 s, 50 users. 85
Table A.4. Remaining discarded values for UMTS, mean call duration of 50 s, 50 users. 85
Table A.5. Remaining PDF values for UMTS, mean call duration of 100 s, 40 users. 87
Table A.6. Remaining discarded values for UMTS, mean call duration of 100 s, 40 users. 87
Table A.7. Remaining PDF values for UMTS, mean call duration of 100 s, 50 users. 89
Table A.8. Remaining discarded values for UMTS, mean call duration of 100 s, 50 users. 89
Table A.9. Remaining PDF values for UMTS, mean call duration of 150 s, 40 users. 91
Table A.10. Remaining discarded values for UMTS, mean call duration of 150 s, 40 users. 91
Table A.11. Curve fitting values for UMTS, 40 users. 92
Table A.12. Curve fitting values for UMTS, 50 users. 92
Table A.13 PDF values for LTE, mean call duration of 50 s and 100 s, 500 users. 95
Table A.14 Discarded values for LTE, mean call duration of 50 s and 100 s, 500 users. 95
Table A.15 Curve fitting values for LTE, mean call duration of 50 s and 100s, 500 users. 96

List of Acronyms

3GPP	3rd Generation Partnership Project
AMC	Adaptive Modulation and Coding
AMR	Adaptive Multi-rate Codec
AMR-NB	AMR Narrowband
AMR-WB	AMR Wideband
ARP	Allocation and Retention Priority
BCCH	Broadcast Control Channel
BEP	Bit Error Probability
BER	Bit Error Ratio
BSC	Base Station Controller
BSS	Base Station Subsystem
BTS	Base Transceiver Station
CCCH	Common Control Channel
CN	Core Network
CP	Cyclic Prefix
CRC	Cyclic Redundancy Check
CS	Circuit Switching
CSFB	Circuit Switched Fallback
CSR	Call Success Rate
DCCH	Dedicated Control Channel
DCR	Dropped Call Rate
DTX	Discontinuous Transmission
DL	Downlink
DMRS	Demodulation Reference Signal
eNodeB	evolved Node B
EFR	Enhanced Full Rate
EPC	Evolved Packet Core Network
EPS	Evolved Packet System
E-UTRA	Evolved Universal Terrestrial Radio Access
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
FDMA	Frequency Division Multiple Access
FER	Frame Erasure Rate
FR	Full Rate

FSM	Finite State Machine
FTP	File Transfer protocol
GBR	Guarantied Bit Rate
GGSN	Gateway GPRS Support Node
GMSC	Gateway MSC
GMSK	Gaussian Minimum Shift Keying
GSM	Global System for Mobile Communications
HLR	Home Location Register
HR	Half Rate
HSS	Home Subscription Server
HSR	Handover Success Rate
HTTP	Hypertext Transfer Protocol
IMS	IP Multimedia Subsystem
IP	Internet Protocol
IPv6	Internet Protocol version 6
ISDN	Integrated Services Digital Network
KPI	Key Performance Indicators
LTE	Long Term Evolution
MBR	Maximum Bit Rate
MIMO	Multiple-Input and Multiple-Output
MME	Mobility Management Entity
MOS	Mean Opinion Score
MS	Mobile Station
MSC	Mobile Switching Centre
OFDMA	Orthogonal Frequency-Division Multiple Access
OVSF	Orthogonal Variable Spreading Factor
PBCH	Physical Broadcast Channel
PBPCH	Physical Broadcast Channel
PC	Personal Computer
PCC	Policy and Charging Control
PCRF	Policy and Changing Resource Function
PCU	Packet Control Unit
PCM	Pulse-Code Modulation
PDCCH	Physical Downlink Control Channel
PDF	Probability Density Function
PDSCH	Physical Downlink Shared Channel
P-GW	Packet Data Network Gateway
PRACH	Physical Random Access Channel
PS	Packet Switching
P-SCH	Primary Synchronization Channel

PSS	Primary Synchronisation Signal
PSTN	Public Switched Telephone Network
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel
QAM	Quadrature Amplitude Modulation
QCI	QoS Class Identifier
QoE	Quality of Experience
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RB	Resource Block
RNC	Radio Network Controller
RNS	Radio Network Subsystem
RoHC	Robust Header Compression
RS	Reference Signal
RSRP	Reference Symbol Received Power
RSRQ	Reference Signal Received Quality
RSSI	Received Signal Strength Indicator
RTP	Real-Time Protocol
SAE-GW	System Architecture Evolution Gateway
SC-FDMA	Single Carrier Frequency-Division Multiple Access
SF	Spreading Factor
SGSN	Serving GPRS Support Node
S-GW	Serving Gateway
SIP	Session Initiation Protocol
SRVCC	Single Radio Voice Call Continuity
S-SCH	Secondary Synchronization Channel
SSS	Secondary Synchronisation Signal
TCH	Traffic Channel
TCP	Transmission Control Protocol
TDD	Time Division Duplexing
TDMA	Time Division Multiple Access
TE	Terminal Equipment
TTI	Transmission Time Interval
UDP	User Datagram Protocol
UE	User Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunications System
UP	User Plane
UTRAN	UMTS Terrestrial Radio Access Network
VAD	Voice Activity Detector

VLR	Visited Location Register
VoIP	Voice over IP
VoLTE	Voice over LTE
VoWi-Fi	Voice over Wi-Fi
WCDMA	Wideband Code Division Multiple Access

List of Symbols

δ_g	Assessment global average end to end delay
δ_n	Assessment average end to end delay
μ	Average value
μ_s	Average end to end delay of each seed
μ_u	Average end to end delay for an user
μ_{UN}	Average end to end delay as a function of the users in the network
σ	Standard deviation
τ_{ack}	Acknowledge delay
τ_c	Compression delay
τ_d	Decoding delay
τ_{dc}	Decompression delay
τ_{dj}	Dejitter buffer delay
τ_{DL}	Download response time
τ_e	Encoding delay
τ_{ETE}	Packet end to end delay
τ_n	Network delay
τ_{pr}	Page response time
τ_{rq}	Request delay
τ_{rv}	Receive response packet delay
$\bar{\tau}$	Mean call duration
τ_{sd}	Send delay
τ_{sig}	Signalling delay
τ_{td}	Delay for setup and tear-down
τ_{UP}	Upload response time

A	Advantage factor (E-model)
b_{sr}	Bits successfully received
d_s	Minimum duration of the given service
F	File size
Id	All impairments due to delay and echo effects
Ie,eff	Effective equipment impairment factor
Is	All impairments that occur simultaneously with the voice signal
n	Number of samples
$N_{b/sym}$	Number of bits per symbol
N_{MIMO}	MIMO order
N_{RB}	Number of RBs
$N_{sub/RB}$	Number of sub-carriers per RB
$N_{sym/sub}$	Number of symbols per subcarrier
N_u	Number of active voice users
Nu_{max}	Maximum number of users
R	Transmission rating factor
R^2	Goodness of fit measure
$\overline{R^2}$	Average goodness of fit measure
R_b	Theoretical bit rate in the downlink
R_{bmax}	Maximum bit rate from the users' perspective
Ro	Basic signal-to-noise ratio
t_e	User request instant
Th	Throughput
t_r	End request instant
T_{RB}	Time duration of an RB
x_i	Given sample
y_i	Data sample
$\hat{y}_i$	Predicted data sample
$\bar{y}$	Mean of the predicted samples

List of Software

MATLAB R2015a	Numerical computing software
Microsoft Excel 2013	Calculation tool and tables processor
Microsoft Word 2013	Word processor
Riverbed OPNET Modeler 17.5	Software tool for network modelling and simulation

Chapter 1

Introduction

This chapter provides a brief overview of mobile voice communications, the main motivations in order to perform the study, the adopted methods, similar studies, and finalises with the description of the contents presented in each chapter of this master thesis.

1.1 Overview

Global System for Mobile Communications (GSM) was originally designed to carry voice traffic, using Circuit Switching (CS). Later on, data capability was added with the introduction of Packet Switching (PS). Data use has increased, but the traffic volume in GSM was clearly dominated by voice. Universal Mobile Telecommunications System (UMTS) boosted data use considerably, bringing high-speed radio capabilities. UMTS remained with CS for voice traffic and PS for data services.

There was a strong trend to make everything in Internet Protocol (IP), as it can be seen with the introduction of the Internet Protocol version 6 (IPv6) and the Internet of Things. Mobile communications followed this trend, and later on, the Long Term Evolution (LTE) was introduced.

LTE dropped CS, being able to use only PS with is all-IP network; it was designed to boost further data rates and capacities. LTE was designed to be able to provide a better performance than the other 3rd Generation Partnership Project (3GPP) systems, with a minimum peak user throughput of 100 Mbit/s in the downlink (DL) and 50 Mbit/s in the uplink (UL), reduced latency, and minimal terminal power consumption. LTE is also frequency flexible with allocations bandwidths from 1.4 MHz to 20 MHz.

Data volumes in mobile networks greatly exceed voice ones, as it can be seen in Figure 1.1, but LTE can also support voice efficiently. This support is not as trivial as in GSM or UMTS, since LTE is designed for PS connections only. The voice service in LTE uses the Voice over IP (VoIP) together with IP Multimedia Subsystem (IMS) and Session Initiation Protocol (SIP), which is called Voice over LTE (VoLTE).

There is also an alternative solution to support voice in the initial phase where the voice service runs on GSM/UMTS networks, which is called Circuit Switched Fallback (CSFB), and consists on moving voice calls from LTE to GSM/UMTS, hence, the call being made over CS. In order to support this solution, bi-directional handovers from LTE to GSM or UMTS must be supported. Each new 3GPP technology has been designed for interworking and coexistence with the existing systems, supporting bi-directional handovers between systems.

The concept of VoIP started in 1995, allowing users to call each other via their computers, allowing users to avoid long distance charges. This application/product only worked if both the caller and the receiver had the same software setup on their Personal Computers (PC). By 1998, some companies were able to offer PC-to-phone and phone-to-phone VoIP solutions. Since 2000, VoIP usage has expanded, the main consumers being business companies, nonetheless, nowadays there are a lot of non-business VoIP applications, such as Ventrilo, Teamspeak, Skype, Viber, etc. [VoHi15].

VoIP in general tends to increase, as seen in Figure 1.2. VoLTE and Voice over Wi-Fi (VoWi-Fi) tend to increase in comparison with VoIP provided by over the top solutions. VoIP is characterised by a solution over an IP architecture, but without software that can guarantee an acceptable maximum end to end delay, and gives a bad Quality of Experience (QoE) to the end user. The main reason to develop a

VoLTE solution is to mitigate the use of VoIP and to minimise the bad QoE of the end user.

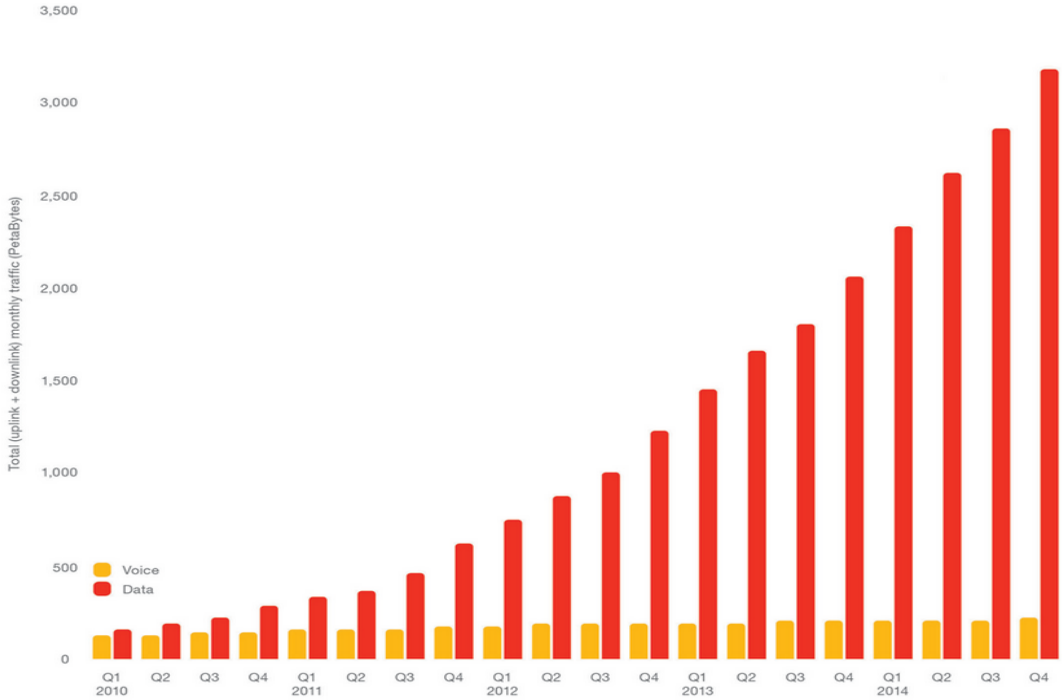


Figure 1.1. Global total traffic in mobile networks, 2010-2014 (extracted from [Eric15]).

The VoLTE solution has several advantages, it enables the possibility to make a call, and use the data services at the same time, without the drop of data service, due to the CSFB used before. Without the need of the CSFB, the call setup time reduced, once there are not the need of a handover in between systems.

As can be seen in Figure 1.2 The VoLTE and the VoWi-Fi tend to increase, this types of voice services combined have a strong benefit from the providers view point. When the users are calling inside a build or a house with Wi-Fi, this voice call can be provided via VoWi-Fi instead of VoLTE, which is helpful in terms of indoor coverage, reducing the effort of covering buildings and houses, by service providers. Once the users left the coverage area, it is made a handover from VoWiFi for VoLTE, [Apti15].

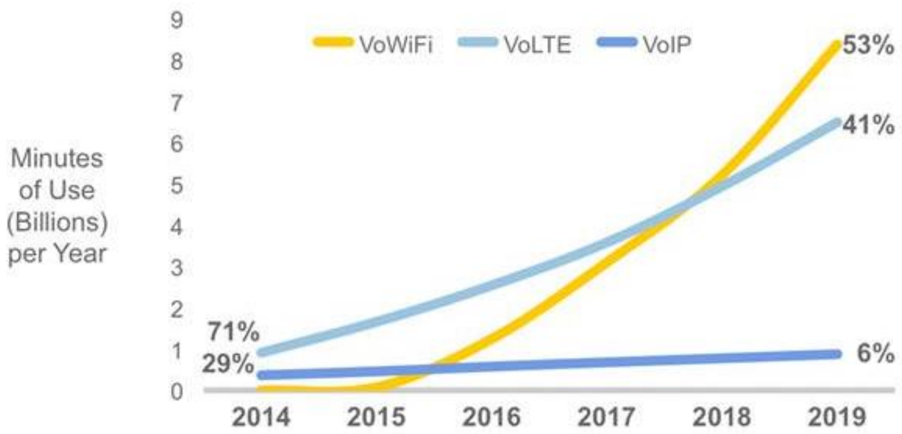


Figure 1.2. Mobile voice minutes of use, VoWi-Fi, VoLTE and VoIP (extracted from [Cisc15]).

1.2 Motivation and Contents

Due to the increasing number of mobile data providers, more users started the use of VoIP via UMTS and LTE, in order to avoid long distance charges. Since these software solutions were not customised to have priority over data, users often experience poor call quality on VoIP. This leads to the need of implementation and study of VoLTE performances, which is a solution that provides priority of voice packets over data ones.

Taking into account that voice communications over LTE should achieve the performance of CS networks, one of the main problems under study is the time interval in between the instant that the user speaks until the instant that the other user receives the signal, also known as mouth to ear delay, or end to end delay. Due to the priority of voice packets over data ones, some data packets can be delayed or even discarded, maximising the time interval of data transfer, reducing throughput.

The main method to evaluate the performance of VoIP over UMTS and VoLTE, is their implementation in these systems, in which users are only capable to perform voice calls. The goal is to analyse the circumstances in which the end to end delay reaches a certain value. The method adopted to evaluate the impact of VoLTE on data services is the implementation of a network in which users are capable to perform voice calls and data transfer, such as email, video conferring and file transferring.

Some studies have been made to study the end to end delay of VoIP. Some of these studies have included the evaluation of throughput, the Mean Opinion Score (MOS), jitter, and packet loss rate. Other studies evaluate VoIP capacity with different audio codecs, and the impact of voice traffic on bandwidth, with and without header compression. Regarding data services, studies have been done with the presence of voice services, however, evaluating the available capacity as a function of the number of available control channels, and of the scheduling algorithm.

The novel aspects of this work is the fact that the evaluation of the voice service in UMTS and LTE has been made as a function of the number of users in the network as well as of the mean calls duration. Another differentiating factor is the study of the impact on data services due to the priority that voice has over data in LTE.

The present chapter makes a brief overview of mobile wireless communication's history evolution, the evolution of voice communications and the software adopted for VoIP. This chapter also presents the motivation behind the thesis, as well as the problems under study, and the author's contributions to the study on performance comparison on voice communications.

Chapter 2 contains the main aspects of GMS, UMTS and LTE, in order to provide a perspective on the systems' differences, and the different approaches adopted by the systems for voice communications. The chapter ends with the state of the art, which presents studies on the performance evaluation for voice communications.

Chapter 3 presents the models, and the parameters adopted to evaluate the performance of voice in between the systems. This chapter also describes the software tool used to simulate the networks, as well as the parameters defined in the software tool. This chapter ends with an assessment on the software tool, in order to understand the duration of the simulations, number of necessary samples, and number of seeds per simulation.

The Chapter 4 presents the simulation scenarios, as well as the collected results and the analysis of results.

The conclusions of this thesis are presented in Chapter 5, together with suggestions for future work.

The Annexes present additional simulator's assessment and results.

Chapter 2

Basic Concepts

This chapter provides an overview of GSM, UMTS and LTE, mainly focussing on voice transmission aspects.

2.1 Network Architecture

This section provides an overview of the basic network architecture of GSM, UMTS and LTE, comparing the evolution of network architectures.

2.1.1 GSM/UMTS

This section describes GSM/UMTS's basic network architectures, based on [EbVö09] and [HoTo04]. The GSM/UMTS architecture supports CS and PS services. The fundamental components of a GSM/UMTS network are shown in Figure 2.1.

GSM's basic network is divided into three main high level domains: Mobile Station (MS), Base Station Subsystem (BSS) and Core Network (CN). BSS is constituted by the Base Transceiver Station (BTS) and the Base Station Controller (BSC), the CN is constituted by two subsystems, one responsible for CS services and other responsible for PS ones. UMTS's basic network architecture is built upon GSM's. The UMTS's basic network is divided into three main high level domains too: User Equipment (UE), Radio Network Subsystem (RNS) and CN. RNS is constituted by the Node B and the Radio Network Controller (RNC), forming the UMTS Terrestrial Radio Access Network (UTRAN).

Regarding GSM, MS is the device carried by the user, and the BTSs are responsible for signal and protocol processing. The essential control and protocol intelligence resides in the BSCs, for instance, BSCs are responsible for protocol functions for radio channel allocation, channel setup and management of handovers. Typically one BSC controls several BTSs, the connection between them being made by fixed lines through the interface called Abis.

The subsystem in the CN responsible for CS services is formed by the Mobile Switching Centre (MSC) that coexists with Visited Location Register (VLR), Home Location Register (HLR) and Public Switched Telephone Network (PSTN) or Integrated Services Digital Network (ISDN). The subsystem responsible for PS services is formed by Packet Control Unit (PCU), Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN).

The traffic from CS services is routed through the MSC, which is responsible for path search, data forwarding, service feature processing, allocation, administration of radio resources and be aware of the mobility of users for the handover, if needed. VLR and HLR are data bases that store mobile user data. Traffic is delivered to the PTSN/ISDN network by an MSC also known as Gateway MSC (GMSC).

PCU is an interface between BSC and SGSN. SGSN delivers data packets from and to the MS, and GGSN acts as an interface to external packet data networks, for instance Internet. Regarding the UMTS system, UE is the device carried by the user, Node B has the same functions as the BTS, and RNC is the equivalent of BSC in GSM architecture.

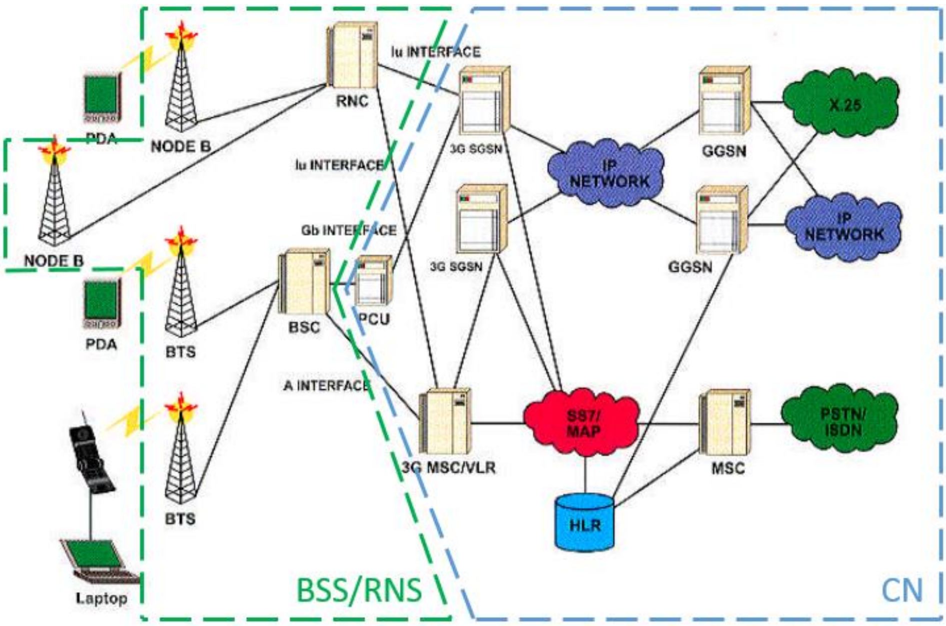


Figure 2.1. GSM/UMTS network architecture (adapted from [Corr13]).

2.1.2 LTE

This section describes LTE's basic network architecture based on [HoTo09]. LTE aims at optimisation for PS services in general, support for higher throughput and improvements in the packet delivery delays. LTE supports only PS services in opposition to GSM and UMTS (both enable PS services and CS services). In order to achieve this optimisation, a new flat network is used, through a simpler and more effective architecture. It was also considered the optimisation of the inter-working with other 3GPP systems.

LTE's basic network architecture is divided into four main high level domains: UE, Evolved UMTS Terrestrial Radio Access Network (E-UTRAN), Evolved Packet Core Network (EPC), and Services. EU, E-UTRAN and EPC form the IP Connectivity Layer, also called the Evolved Packet System (EPS). The maas a function of this layer is to provide a highly optimised IP based connectivity, without the need of CS nodes and interfaces seen in earlier 3GPP architectures. Figure 2.2 shows the co-located 3GPP architectures.

The E-UTRAN is at base station level, consisting of intelligent base stations called evolved Node B (eNodeB). E-UTRAN is a group of eNodeBs interconnected by the X2 interface. The eNodeBs functionality is to act as a layer bridge between UE and EPC. X2 interface carries the necessary signalling to enable exchange of information on the radio resource usage among the base stations, to provide handover command or to handle the data forwarding to another eNodeB, [PoHo12]. UEs, is the device that the user uses for communication, also called Terminal Equipment (TE). The UE interacts with eNodeB via LTE-Uu interface.

The EPC is formed by the following elements: Mobility Management Entity (MME), Serving Gateway (S-GW), Packet Data Network Gateway (P-GW), Policy and Changing Resource Function (PCRF), and

Home Subscription Server (HSS). The MME is responsible for authentication and security, mobility management, managing subscription profile and connectivity. The S-GW is responsible for User Plane (UP) tunnel management and switching. The P-GW performs traffic gating and filtering functions, usually acting as the IP point of attachment for the UE. The PCRF is the element that is responsible for Policy and Charging Control (PCC), it makes decisions on how to handle the services in terms of Quality of Service (QoS). The HSS is a subscription data repository for all permanent user data.

The connections between EPC and E-UTRAN are provided by the S1 interfaces. This connection can be made between eNodeBs to MME and eNodeBs to S-GW, the first interface is called S1-MME and the second interface is called S1-U. The combination of S-GW and P-GW is called System Architecture Evolution Gateway (SAE-GW), which is responsible for UP handling in the EPC. EPC and SAE-GW together are intended to handle the interoperability through other 3GPP systems.

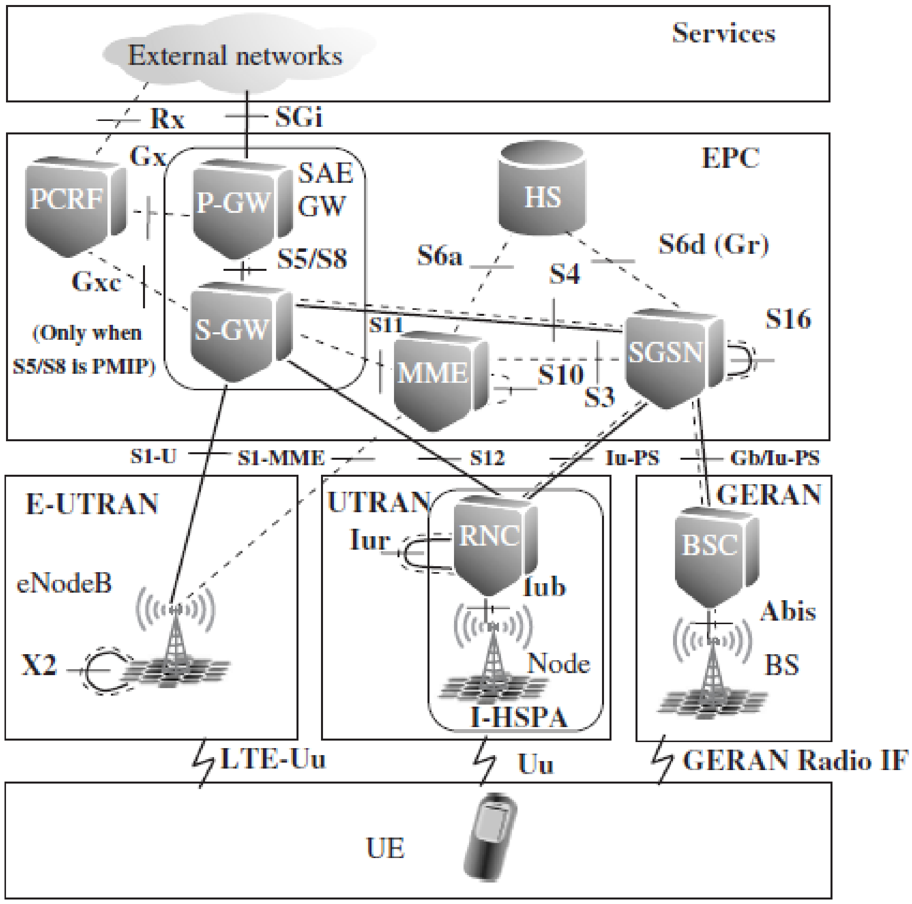


Figure 2.2. System architecture for 3GPP access networks (extracted from [HoTo09]).

2.2 Radio Interfaces

This section provides an overview of the basic radio interfaces of GSM, UMTS and LTE. GSM's basic radio interface is based on [EbVö09]. In GSM, the access technique used is a combination of Time

Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA). FDMA splits the available 25 MHz bandwidth into 124 single carrier channels of 200 kHz width, each of these channels being divided into 8 time-slots. This access technique enables up to 8 user channels per carrier.

Traffic Channel (TCH) is the channel for CS and PS services. The signalling channels are Broadcast Control Channel (BCCH), Common Control Channel (CCCH) that is a point-to-multipoint signalling channel, and Dedicated Control Channel (DCCH) that is a bidirectional point-to-point signalling channel. The modulation technique is Gaussian Minimum Shift Keying (GMSK).

UMTS's basic radio interface is based on [HoTo04]. The access technique is Wideband Code Division Multiple Access (WCDMA) with two types of codes, channelisation and scrambling. Scrambling is used to separate terminals or base stations from each other, not changing the signal bandwidth or the symbol rate. Channelisation codes are based on the Orthogonal Variable Spreading Factor (OVSF) technique, allowing the use of a Spreading Factor (SF) based on a code tree.

In UMTS, the voice service is done with an SF of 128, corresponding to 128 channels with a maximum user data rate of 15 kbit/s for UL, and 24 kbit/s for DL; the data rate for the voice service is 12.2 kbit/s. This allows up to 128 users, however, this number is reduced taking into account that part of the codes are reserved for the common channels, besides other constraints.

LTE's basic radio interface is based on [HoTo09] and [3GPP14]. Two access techniques are used, Orthogonal Frequency-Division Multiple Access (OFDMA) for DL and Single Carrier Frequency-Division Multiple Access (SC-FDMA) for UL, both with Cyclic Prefix (CP), the bandwidth being divided into various sub-carriers.

Currently, LTE uses Time Division Duplexing (TDD), a total of 17 FDD bands being available. ANACOM (the Portuguese national telecommunications authority) has made an auction for 800 MHz, 900 MHz, 1800 MHz, 2.1 GHz and 2.6 GHz frequency bands, [ANA12a]. Table 2.1 shows the frequency bands adopted by each system.

LTE has more than one bandwidth range option, from 1.4 to 20 MHz. Different bandwidth sizes enable a different number of Resource Blocks (RBs). It uses Quadrature Phase Shift Keying (QPSK), 16 Quadrature Amplitude Modulation (QAM) and 64QAM, like other 3GPP systems. It can take advantage of Adaptive Modulation and Coding (AMC), which is responsible for providing channel state information, so that LTE can use the channel variations for self-optimisation by changing the coding rate and the modulation. This optimisation improves the data rate.

The main physical channels for DL are the Physical Downlink Shared Channel (PDSCH), Physical Broadcast Channel (PBCH) and Physical Downlink Control Channel (PDCCH); these channels are used to carry information from the higher layers and to carry user data. The main physical signals are Reference Signal (RS), Primary Synchronisation Signal (PSS) and Secondary Synchronisation Signal (SSS); these signals are used for cell search and channel estimation purposes, and they are used by the physical layer but do not carry information originating from higher layers. The frame structure has 10 ms duration, and is composed of 20 slots of 0.5 ms duration each; it has two slots with 1 ms duration, called sub-frame, also designated as the Transmission Time Interval (TTI). Figure 2.3 shows the frame structure, with the allocation of physical channels and signals, the space for data transmission being shown in yellow.

Table 2.1. National frequency allocation (adapted from [ANA12a], [ANA12b] and [ANA12c]).

Frequency band [MHz]	System	Frequency interval (UL) [MHz]	Frequency interval (DL) [MHz]	Operators' bandwidth [MHz]
800	LTE	[832;862]	[791;821]	10
900	GSM	[890;914]	[935;959]	8
	UMTS	[880;890]	[925;935]	5
1800	GSM	[1710;1785]	[1805;1880]	6
	LTE			20
2100	UMTS	[1920;1980]	[2110;2170]	20
2600	LTE	[2500;2570]	[2620;2690]	20

The main physical channels for UL are the Physical Uplink Shared Channel (PUSCH) and Physical Uplink Control Channel (PUCCH). The main physical signals are the Demodulation Reference Signal (DMRS) and Physical Random Access Channel (PRACH). The UL frame structure is basically the same as DL regarding durations. CP has a direct impact on the number of symbols in the slot, as exemplified in Figure 2.4.

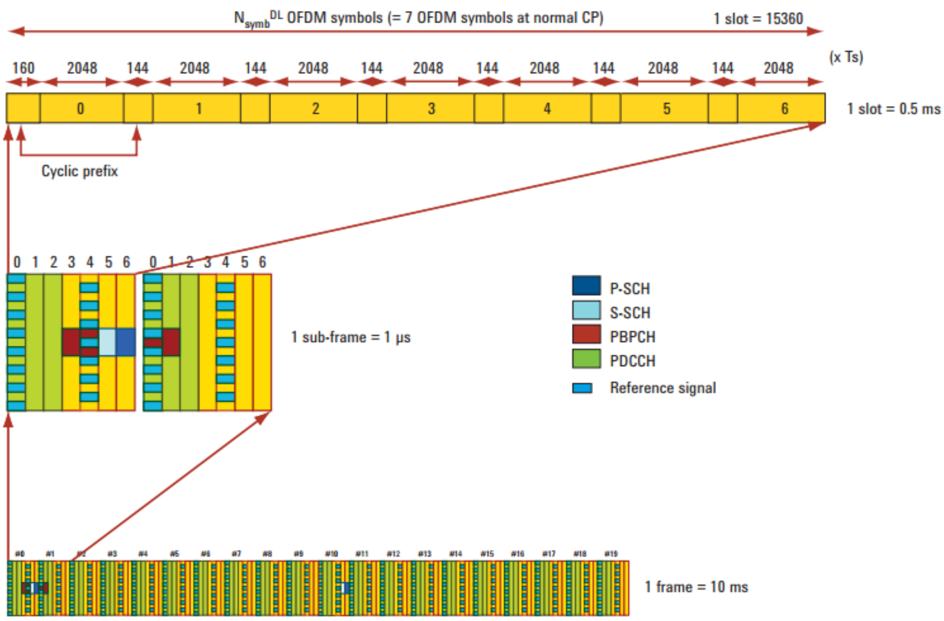


Figure 2.3. Frame structure type 1 for DL (extracted from [Agil07]).

RB is the basic unit when allocating data to the UE. The resource allocation in DL is made through a resource grid composed of RBs. An RB is an aggregation of 12 contiguous sub-carriers, each one with 15 kHz spacing and 7 symbols. The smallest unit possible is designated Resource Element, and is composed of one symbol in one sub-carrier. An RB has 84 Resource Elements, and keeps this size for all transmissions. Figure 2.5 shows an example of a resource structure.

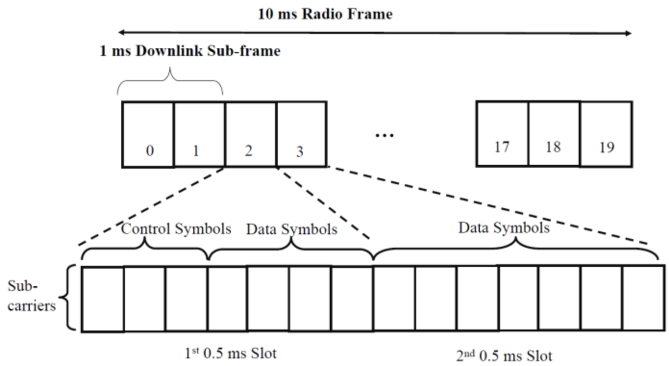


Figure 2.4. DL slot structure of bandwidths above 1.4 MHz (extracted from [HoTo09]).

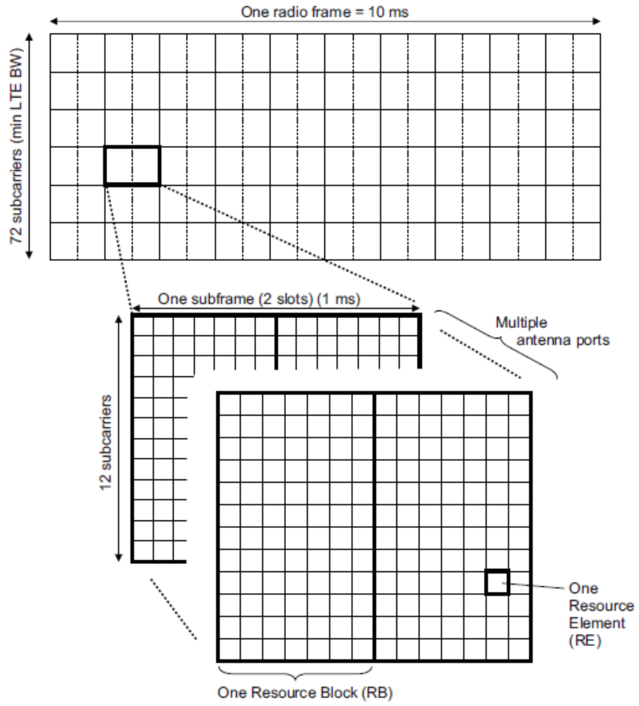


Figure 2.5. Basic time-frequency resource structure of LTE with normal CP (extracted from [SeTo11]).

LTE uses the concept of bearers to carry the data between UE and the core network, and scheduling to allocate the necessary RBs to deliver data. The default bearer is established to provide a logical connection between the UE and a P-GW for the purpose of delivering data. In order to differentiate the Guarantied Bit Rate (GBR) resource type such as real-time voice and video traffics, from the non-GBR resource type, LTE defines a set of QoS Class Identifier (QCI) with the intention of differentiation of QoS, priority, packet delay budget and the packet error loss rate, [PoHo12].

LTE introduces also the concept of a dedicated bearer. A dedicated bearer allows certain types of data traffic to be isolated from all other traffic. The voice bearer uses a dedicated bearer with a QCI of 1, [PoHo12].

In order to enable a high data rate, an optimisation is possible, by using multiple antennas with spatial multiplexing, i.e., Multiple-Input and Multiple-Output (MIMO).

2.3 Voice Service and Performance Parameters

This section provides an overview of the basic concepts on voice services and the main performance parameters of GSM, UMTS and LTE.

2.3.1 Voice over GSM/UMTS

This section describes GSM/UMTS's basic concepts on the voice service based on [HaRo03] and [HoTo04]. The voice traffic in GSM/UMTS is carried over CS, QoS being assessed by the blocking probability, normally around 1%.

The transmission and reception of a speech signal demands several successive operations, in order to convert speech into a radio signal and vice-versa. A transmitting and receiving chain with the following operations exists:

- Source Coding, converts the analogue signal into a digital one;
- Channel coding, adds extra bit to the flow, with the intention of detecting and correcting bit errors;
- Interleaving, separating bits to avoid adjacent bit errors;
- Ciphering, to avoid data to be intercepted;
- Burst formatting, adds synchronisation and equalisation information;
- Modulation, transforms the binary signal into an analogue one at a carrier frequency.

The receiver performs the reverse operations:

- Demodulation;
- Deciphering;
- Deinterleaving;
- Channel decoding;
- Source Decoding.

GSM and UMTS use proper voice codecs, which optimise speech signals, because they are dimensioned to human voice frequencies. GSM and UMTS support more than one codec modes, GSM Half Rate (HR), GSM Full Rate (FR), Enhanced Full Rate (EFR) and Adaptive Multi-rate Codec (AMR), which is divided into AMR Narrowband (AMR-NB) and AMR wideband (AMR-WB). Table 2.2 summarises speech codecs in use and the associated bit rates.

MS and BTS perform channel quality estimation of the received channel to estimate the best codec mode; the codec chosen in UL may be different from the one used in DL, but the channel mode (HR or FR) must be the same. HR can split a single full rate voice channel into two sub-channels that can maintain separate calls, with half of the bit rate and with quality reduction. GSM and UMTS have a feature to omit transmission when a silent period is detected during a speech connection, i.e., Discontinuous Transmission (DTX), which can reduce interference and is useful for power saving. Voice Activity Detector (VAD) distinguishes between voice and noise.

Table 2.2. Speech codec Supported (extracted from [HaRo03]).

Speech codec	Bit rate [kbit/s]	Bits per speech frame	System
GSM FR	22.8	260	GSM
GSM HR	11.4	112	GSM
AMR-NB	12.2	244	UMTS

GSM/UMTS's basic concepts on performance parameters are based on [HaRo03] and [HoTo04]. The performance of speech services can be measured using multiple different Key Performance Indicator (KPI). Bit Error Ratio (BER) and Dropped Call Rate (DCR) are used to quantify speech quality and the rate of the lost connections respectively, Call Success Rate (CSR) and Handover Success Rate (HSR) are used to measure the performance of the signalling channels associated with call originations and handovers. Figure 2.6 illustrates the different KPIs and the location where they are measured.

BER is a measurement of the raw bit error ratio in reception before the decoding process. DCR measures the percentage of connections lost; since a drop call has a very negative impact on the end-user-perceived QoS, the DCR should have a range from 1 to 2%. A well-performing network should have both the CSR and the HSR above 95%. AMR speech codec can tolerate about 1% of Frame Erasure Rate (FER) and the corresponding BER should be about 10^{-4}. Speech quality can be quantified using MOS, which ranges from 1 (bad) to 5 (excellent), being based on users' opinion. FER is a powerful KPI since is highly correlated with the final voice quality that the end user perceives. The Bit Error Probability (BEP) is a signal quality indicator, terminals reporting the mean BEP and its coefficient of variation (standard deviation/mean value).

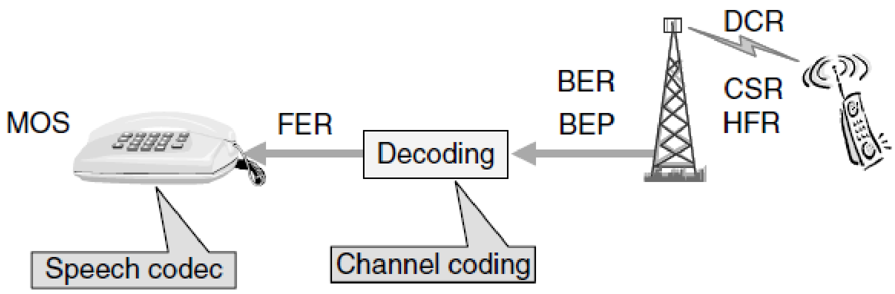


Figure 2.6. Main speech KPIs (extracted from [HaRo03]).

2.3.2 Voice over LTE

This section describes LTE's basic concepts on voice service based on [PoHo12]. The voice traffic in LTE is carried over PS. While VoLTE is not yet available, an interim solution to provide voice communications in LTE is CSFB, which handovers the call to GSM/UMTS, over CS. VoLTE uses VoIP together with IMS, based on the SIP protocol.

The traditional VoIP uses SIP signalling for call establishment, authentication, registration, presence maintenance and QoS requirements. During the session setup phase, the two UEs agree on the set of media they want to use for the session, and the codecs that will be used for the different media types. VoIP uses the Real Time Protocol (RTP), it defines how the audio data stream must be fragmented, adding in each fragment header the sequence and delivery time information. This type of data is transmitted over the User Datagram Protocol (UDP). Due to the high number of bits used in the headers, and due to the high data rate generated by this fact, it is necessary the use of header compression, such as Robust Header Compression (RoHC), [SeTo11]. In order to check errors, Cyclic Redundancy Check (CRC) is used.

IMS and SIP are key to deploy VoLTE. IMS provides interconnect and gateway functionalities that allow VoIP devices, such as VoLTE, to communicate with VoIP and non-VoIP devices. To avoid the high and variable delays, packet loss/discard and packets out of order, IMS provides end to end QoS and offers the UE the capability to negotiate parameters such as media type, media type bit rate, packet size, packet transport frequency, bandwidth adaption and direction of traffic. SIP signalling is used in VoLTE, as in the traditional VoIP, for the same purposes. LTE uses DTX in order to save power, like GSM and UMTS CS.

LTE defines the concept of bearers, which is a logical connection between the UE and a P-GW, also designated as EPS bearer, the default bearer being created when the UE attaches to LTE, and lasts as long as UE is attached to LTE. Additional default bearers may be created when simultaneous access to services are needed, the default bearers being always non-GBR bearers. LTE has dedicated bearers for QoS differentiation, which are triggered by the network or the UE; this type of bearers can be GBR and non-GBR. LTE defines a set of QCI to differentiate from GBR and non-GBR resource type. QCI is used for set the priority, packet delay budget and the packet error loss rate. Table 2.3 shows the set of QCIs defined in LTE.

For each EPS bearer (default and dedicated), QoS support is based on parameters such as: QCI, used as a reference to access node-specific parameters on bearers; GBR, indicating the bit rate that can be expected by a GBR bearer; Allocation and Retention Priority (ARP), to decide whether a bearer establishment/modification request can be accepted or not; and Maximum Bit Rate (MBR), limiting the bit rate that can be expected to be provided by a GBR bearer.

In LTE, there are no dedicated user-specific resources reserved in the air interface, in the typical case every single packet transmission is scheduled. Voice packets arrive every 20 ms, thus, packets need to be scheduled. The packet scheduling algorithm for voice can be fully dynamic or semi-persistent: the dynamic solution schedules every single voice packet that arrives with 20 ms periods, while the semi-persistent scheduling pre-allocates resources for voice packets every 20 ms. The second solution does not require control channel capacity unlike the first one, as shown in Figure 2.7. Voice packets can be transmitted in 1 ms (TTI), which allows the UE power amplifier to transmit only for a short time, making UL coverage a problem. The solution is to use TTI bundling where the same data is repeated in four consecutive TTIs, allowing the terminal to have a continuous transmission.

Table 2.3. Quality of service characteristics (extracted from [PoHo12]).

QCI	Resource type	Priority	Packet delay budget [ms]	Packet error loss rate	Example services
1	GBR	2	100	10^{-2}	Conversational voice
2		4	150	10^{-3}	Conversation video (live Streaming)
3		3	50	10^{-3}	Real time gaming
4		5	300	10^{-6}	Non-conversational video (buffered streaming)
5	Non-GBR	1	100	10^{-6}	IMS signalling
6		6	300	10^{-6}	Video (buffered streaming) Transmission Control Protocol (TCP)-based
7		7	100	10^{-3}	Voice, video (live streaming) interactive gaming
8		8	300	10^{-6}	Video (buffered streaming) TCP-based
9		9	300	10^{-6}	Sharing, progressive video

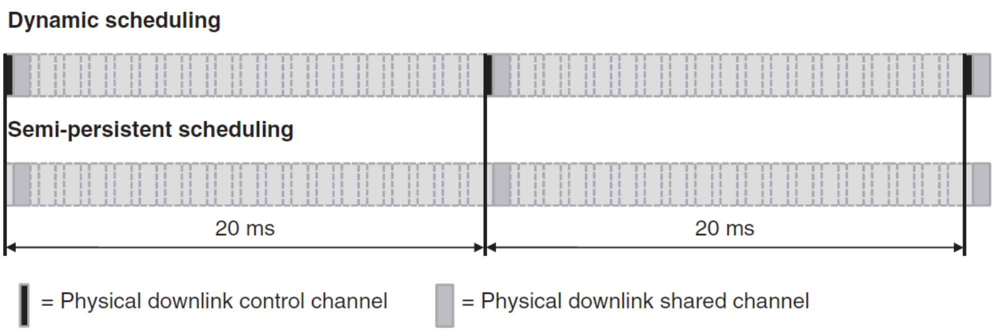


Figure 2.7. Dynamic and semi-persistent scheduling for voice (extracted from [PoHo12]).

LTE has a feature that enables secure service continuity when a VoLTE subscriber goes outside LTE coverage, but still has GSM/UMTS one. The Single Radio Voice Call Continuity (SRVCC) provides a handover from LTE to GSM/UMTS (from PS to CS), when a native IMS-based VoIP connection no longer can be maintained in LTE. The reverse SRVCC enables handover from GSM/UMTS to LTE. The

SRVCC is typically triggered when the signal level from the Reference Symbol Received Power (RSRP) or the signal quality from the Reference Signal Received Quality (RSRQ) drops below a predefined threshold.

LTE has an all-IP architecture, and the performance of speech services can be measured using multiple indicators, such as, end to end delay, jitter, RSRP, RSRQ and MOS. MOS is very subjective, being based on users' opinions as seen before. The end to end delay is measured from the input of the UE at the sender side to the output of the UE at the receiver side. The delay should preferably be below 200 ms, which is the value typically achieved in CS networks, [PoHo12]. In LTE, packets can be lost/discarded, so packet loss is an indicator to take into account, which should have a maximum of 2% packets unsuccessful delivery rate, so that the user is not in outage (not satisfied). The system capacity for VoIP can then be defined as the number of users present per cell when more than 95% of the users are satisfied; one group of network operators expressed a preference for the ability to support 60 satisfied VoIP sessions per MHz, [SeTo11].

Jitter is another performance indicator, VoIP applications being usually designed so that they can tolerate it in the order of 10 to 20 ms, [HoTo11]. The Block Error Rate (BLER) serves the purpose of radio link monitoring, and to set thresholds that are useful for synchronisation; the target BLER is 1% for the control channels and 10% for the data ones, [SeTo11]. The RSRP in a particular cell is the average of the power measured on the resource elements that contain cell-specific reference signals. The Received Signal Strength Indicator (RSSI) is the total received wideband power on a given frequency. The RSRQ is the ratio of the RSRP and the E-UTRAN carrier RSSI, for the reference signals. Both signals, RSRP and RSRQ are useful to know when is necessary to perform a SRVCC handover, [HoTo09].

The main difference in between VoLTE and the traditional CS is that in CS the available resources are allocated to users during the entire conversation, being available only for this particular communication. VoLTE allocates the resources every 20 ms, allocating only the necessary RBs to the users in need, thus, the resources are better distributed in the network. The main differences in between the traditional VoIP and VoLTE are the use of IMS, the bearers to provide priority of voice services over data services, the scheduling algorithms, and the TTI bundling in the VoLTE solution.

2.4 State of the Art

This thesis studies the performance evaluation of VoLTE and the comparison with GSM and UMTS. The main focus in most of the studies is the end to end delay and packet loss rate, which are important KPIs as shown in Section 2.3.2. However, there is no study comparing the performance of VoLTE against the GSM and UMTS, which is the main goal in this thesis.

In [TaGu14], the intention is to analyse performance parameters such as end to end delay, packet loss rate, and jitter. The maximum end to end delay is 150 ms, a minimum of 98% of packets successful delivery rate, and jitter should be 0, in a steady LTE network, to fulfil the requirements imposed by 3GPP and ITU-R. In this study different LTE bandwidths were simulated.

In order to design the LTE network, the authors have used the OPNET modeller from Riverbed Technologies Ltd, [Rive14]. This design includes a complete implementation for VoLTE. The baseline network consists of 7 eNodeBs and 21 UEs (3 in which eNodeB). Mobility was considered in this study, with a velocity of 5 m/s. Handovers between cells with the same frequency is the only type of handover likely to happen. The authors use as path loss model the free space model without any obstruction for the signal, and use two scenarios to evaluate the QoS factors, one with 20 MHz bandwidth to simulate the physical profile in the LTE network, and other with the same evaluations in 1.4 MHz and 5 MHz.

The authors concluded that the delay fulfils the 3GPP and ITU-R requirements in both scenarios, the average end to end delay is around 120 ms, and as expected the 20 MHz bandwidth has the lower end to end delay, and the 1.4 MHz bandwidth the higher one. The higher the bandwidth, the higher the data rate supported, as a result, the lower end to end delay. Regarding packet loss rate, it fulfils the requirements imposed, the average being around 0.005%. In both scenarios, jitter was almost 0 and fulfils the 3GPP and ITU-R. However, contrary to authors' expectations, the 20 MHz bandwidth has better jitter values than in the 5 MHz and 1.4 MHz ones.

In another study, [PaPa14], the intention was to analyse performance parameters such as one-way delay, throughput, packet loss and MOS with different audio codecs and different number of users. These performance parameters should fulfil the values previously indicated. The MOS value was computed trough a tool provided by the ITU-T.

The authors used a 3D modelling of a University campus, with a simulator called QualNet Developer, [QuDe15]. In this scenario, several types of propagation paths should be taken into consideration, so an adaptive propagation model was used to accurately predict the appropriate path loss. The authors used the 1800 MHz frequency band. In this study the 1.4 MHz, 5 MHz and 10 MHz bandwidths were used. Two simulations were made, the first one with 84 users with 1.4 MHz and 5 MHz and using different audio codecs, the G.711 and the G.723, with 64 kbit/s and 5.3 kbit/s respectively, and the second one with 300 users with 5 MHz and 10 MHz and the same audio codecs, both with a calls duration of 3 minutes.

In the first simulation, the authors concluded that with the G.711 audio codec and 1.4 MHz channel bandwidth, the average one-way delay exceeds the imposed value and the MOS value is very low. Changing the bandwidth to 5 MHz may improve network performances, but since there are only 84 users in the cell, the resources are not efficiently allocated. The packet loss is acceptable in both channel bandwidths. In the second simulation, the one-way delay and the MOS reach unacceptable values in the case of 5 MHz and using the G.711 audio codec; once again, increasing the channel bandwidth increases the network performance, however, the resources are not efficiently allocated. In the overall evaluation, the results satisfy the requirements imposed by the ITU-T, despite the inefficient resource allocation with the G.711 audio codec.

In [AaKj11], the intention was the study of VoIP capacity of an LTE network. In order to evaluate it, the dynamic and the semi-persistent scheduling algorithms were used. In this study, the authors also analyse capacity depending on channel bandwidth, modulation and coding scheme and different audio codecs. The audio codecs in use are G.711 (64 kbit/s) and G.723.1 (12.2 kbit/s), the bandwidths in use are 5 MHz and 10 MHz, and the modulation and coding scheme index are 9 and 15. An OPNET based simulation model is used, with one eNodeB and one EPC and a number of fixed UEs are randomly positioned in the system area with a 1 km radius. In this simulation there are no errors in the channel.

The authors concluded that the lower bit rate audio codec (G.723.1) provides much higher capacity, with the using of this audio codec a cell capacity can easily reach 150 users, where the end to end delay is just above 200 ms and the packet loss is about 0.5%, for the UL. The lower packet delay is achieved when 10 MHz is used. The VoIP capacity of 5 MHz with modulation and coding scheme index of 9 is 70 users, and with an index of 15 is 100 users, while for 10 MHz and index of 9 is 115 users, when the mean packet end to end delay is limited to 100 ms and the mean buffer overflow limit is 2%. The semi-persistent scheduling algorithm increases VoIP capacity, and decreases packet delays, but is not very efficient in terms of packet losses.

In [OzVa13], another type of performance evaluation was made. In this study, the authors' goal is to evaluate the performance of VoLTE in the presence of data traffic, and in heterogeneous networks, with macro and pico cells. Different simulations were performed, with different throughputs, different number of users per cell, limited number of PDCCH and unlimited PDCCH in a dynamic and a semi-persistent scheduling algorithms.

In the scenario that only exists the macro cell and unlimited PDCCH, the effect of data on VoLTE performance is only through inter-cell interference, because the scheduler always gives higher priority to VoLTE users. The impact of VoLTE on data performance is the reduced time and RB for data users, and the reduced multi-user diversity gain for data users. When there is a PDCCH limit, there is additional impact on data users such as, unused RB and reduced multi-user diversity due to the PDCCH limit. The use of semi-persistent scheduling algorithm is similar to the no PDCCH limit, because this algorithm needs less PDCCH. In the second scenario, with macro and pico cells, significant gains can be seen for VoLTE and data. This improvements came from an increase of available resources for macro cell users and the addition of small cells reduces the impact of PDCCH limitations on macro cells.

VoLTE capacity is always limited by the reduced number of PDCCH, except for semi-persistent scheduling. However VoLTE capacity can be increased significantly with dynamic scheduling when the PDCCH limitation is removed. VoLTE capacity would be higher if there were no data users in the system, even though the scheduler gives higher priority to VoLTE users within a cell, data users generate additional inter-cell interference. When the number of VoLTE users increases data throughput decreases almost linearly.

In another VoIP study, [AnMo11], the intention is to evaluate the MOS value for different audio codecs and different bandwidths, for a VoIP system. In this study, the authors used OPNET Modeler for environment simulation. Seven cells with five UEs per cell were configured, 1.4 MHz and 20 MHz bandwidths were used, and the audio codecs in study are G.711, G723.1 5.3K, G.729 A and GSM FR.

The authors concluded that using high bitrate codecs gives higher MOS values, however, the GSM FR codec performs quite well in the 20 MHz bandwidth. These results can be seen in Table 2.4.

In [JaIb12], the authors' goals are the study of the voice traffic and evaluate the impact on bandwidth. The research study was conducted in an urban area in a period of 3 weeks. The analyses have been made based on services groups, called speech call, video call and packet switch call. The study has been made in the busy-hour, with information derived from daily busy hour, weekly busy hour, monthly busy hour and busy hour of a cluster. The VoLTE service uses the RTP, and 8 AMR different modes in both, AMR and AMR-WB. The study has been made with and without RoHC and with VAD.

The authors concluded that if RoHC is used a better IP bandwidth consumption will be achieved. For the speech traffic in the busy hour, a cluster has a traffic of 4509.37 Erl. There are advantages of using RoHC and VAD compared with AMR without RoHC. The total bandwidth required is 91 964.98 kbytes in the busy-hour of the busiest day or 204.37 kbit/s for using AMR mode 7 codec rate with RoHC and VAD. The busiest hour is at 10 am with 29 423 Mbytes of traffic, which is equivalent to 65 Mbit/s of total bandwidth requirement in the IP network (on S1-U interface). The main conclusion is that the gross bandwidth throughput of VoLTE depends on the packet method and coding rate chosen.

Table 2.4. Simulation results (extracted from [AnMo11]).

Audio codec	**Bandwidth [MHz]**	
	1.4	**20**
GSM FR	2.51	3.49
G.729 A	3.02	3.03
G.723.1 5.3K	2.51	2.51
G.711	3.64	3.64

Chapter 3

Theoretical Models and Simulator Description

This chapter provides a description of the used simulator, as well as the implementation in the simulator, the metrics, and the theoretical models used. This chapter ends with a brief assessment of the used simulator, which is done in order to infer the minimum simulation conditions which provide realistic results.

3.1 Theoretical Models

This sections describes the theoretical models and metrics used in order to analyse the data collected from the simulator.

The theoretical bit rate in DL depends on some network conditions, e.g., the available bandwidth, the sub-carrier spacing, the associated CP, the coding scheme and the MIMO order used. The quality of the signal received by a UE depends on the channel quality from the serving sector, the level of interference from other sectors and the noise level. In order to enhance capacity and coverage, the transmitter tries to match the data rate for each user to the variations in received signal quality, which is referred to as link adaptation and typically based on AMC.

The modulation scheme and coding rate may be adapted according to the channel conditions, when using AMC. In LTE, the modulation and coding scheme is constant over the allocated frequency resources for a given UE. In terms of modulation schemes, a low-order modulation (e.g., QPSK, which uses 2 bits per symbol) is more robust and, as such, is able to tolerate higher levels of interference, although it provides a lower transmission bit rate, [Alme13].

In order to compute the theoretical bit rates in DL, the following expression is used, [Alme13]:

$$R_{b[\mathrm{Mbit/s}]} = \frac{N_{sub/RB} \cdot N_{sym/sub} \cdot N_{b/sym[\mathrm{bit}]} \cdot N_{RB} \cdot N_{MIMO}}{T_{RB[\mathrm{\mu s}]}} \tag{3.1}$$

where:

- $N_{sub/RB}$: number of sub-carriers per RB (12 when considering a 15 kHz sub-carrier spacing);
- $N_{sym/sub}$:number of symbols per subcarrier (7 when the normal CP is used);
- $N_{b/sym}$: number of bits per symbol, which depends on the modulation scheme and coding rate;
- N_{RB}: number of RBs;
- N_{MIMO}: MIMO order;
- T_{RB}: time duration of an RB, which is 500 μs;

The average value is given by:

$$\mu = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \tag{3.2}$$

where

- x_i: given samples;
- n: number of samples;

The standard deviation is given by:

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{n}} \tag{3.3}$$

In order to evaluate the maximum bit rate from the users' perspective, which could be qualified as a QoE metric, it was used the following:

$$R_{bmax[\mathrm{bit/s}]} = \frac{F_{[\mathrm{byte}]} \cdot 8}{S_{d[\mathrm{s}]}} \tag{3.4}$$

where:

- F: file size;
- S_d: minimum duration of the given service;

In order to evaluate the collected data and the goodness of fit, one used the Coefficient of Determination, R^2, which measures how well it fits a set of observations, [Matl15]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3.5}$$

where:

- y_i: data sample;
- $\widehat{y_i}$: predicted sample;
- $\bar{y}$: mean of the predicted samples;

In order to compute the mean of a given number of R^2, one uses [Corr13];

$$\overline{R^2} = \frac{1}{n}\sum_{i=1}^{n} R_i^2 \tag{3.6}$$

while for μ, it is given by;

$$\bar{\mu} = \frac{1}{n}\sum_{i=1}^{n} \mu_i \tag{3.7}$$

and for σ, it is given by;

$$\bar{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \sigma_i^2} \tag{3.8}$$

3.2 OPNET Simulator

This section provides a general description on the OPNET simulator, as well as a general description on how to generate traffic, and the metrics used to evaluate the performance of the voice services.

3.2.1 General Description

In order to use the models, the OPNET Modeler simulation tool [Rive13] was used, which is based on a hierarchical structure with different modelling domains, each domain having a given editor associated to it. The OPNET Modeler architecture is composed of three main domains: network; node; and process.

The lower level is the process domain and the associated process model editor is presented in Figure 3.1. Process models are used to specify the behaviour of the models that exist in the node domain. Process models are driven by events and interrupts, interrupts being generated when a given event occurs, allowing the process model to act in response to the event. In order to develop the process model, a combination of Finite State Machines (FSMs), libraries of kernel procedure and C/C++ programming language functions and variables are used, which is called Proto-C. The FSMs are used to graphically represent the progression of a process, being represented by states and transitions. The kernel procedures and the C/C++ code are used within the states and transitions to perform all the tasks related to events and interrupts.

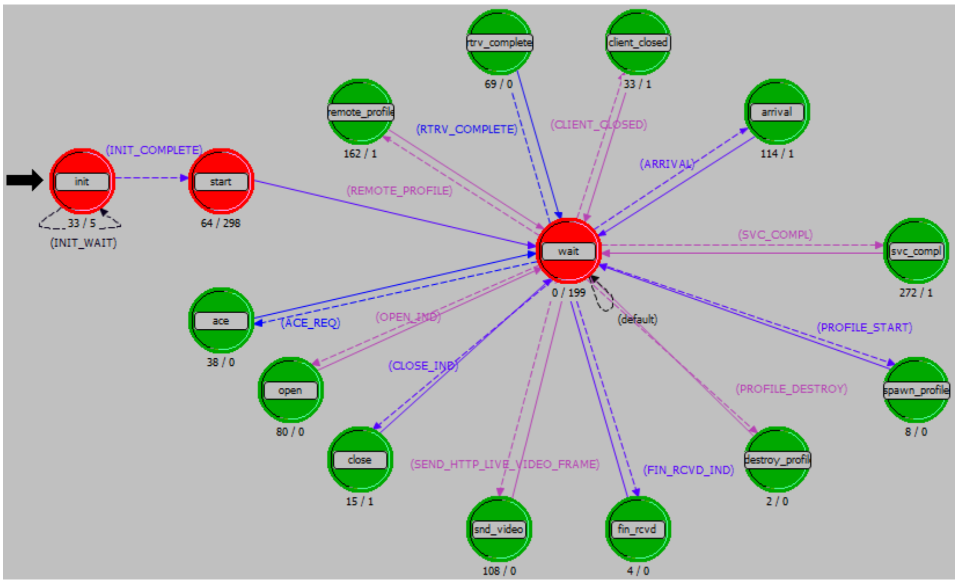


Figure 3.1. Process domain.

The next level is the node domain and the associated node editor, Figure 3.2. The node editor defines the behaviour of each network object, the behaviour being defined by using different modules and connections among them. Modules are used to represent, e.g., protocol layers and physical resources. Modules are connected by packet streams, statistic wires, and logical associations. Packet streams are a one-way pipes that are used to transmit packets between modules. Statistic wires are used to provide one-way connection between two modules. Logical associations are used to connect transmitters and receivers to indicate that they should be used together. Each node represented in Figure 3.2 has a process that specifies its behaviour.

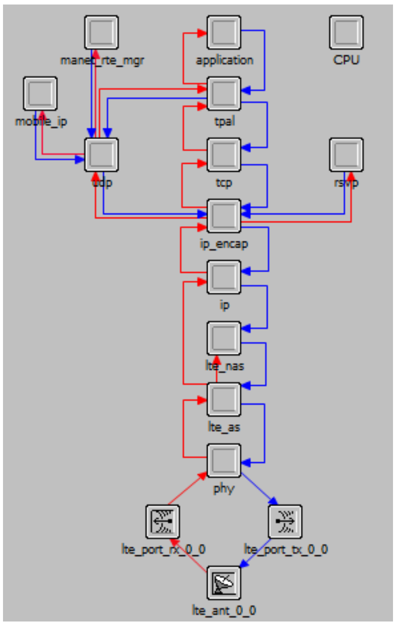


Figure 3.2. Node domain.

The top level is the network domain and the associated project editor, Figure 3.3. This editor is the main staging area, allowing one to graphically represent a topology of a communication network. The network is composed of node and link objects, which are instances of the lower level models. The network model also specifies the physical location, interconnections and configurations of the objects. Each node in Figure 3.3 is specified by the lower levels described before, the nodes being a UE, an eNodeB, and an EPC; the LTE Attributes node is used to specify settings, such as frequency and bandwidth.

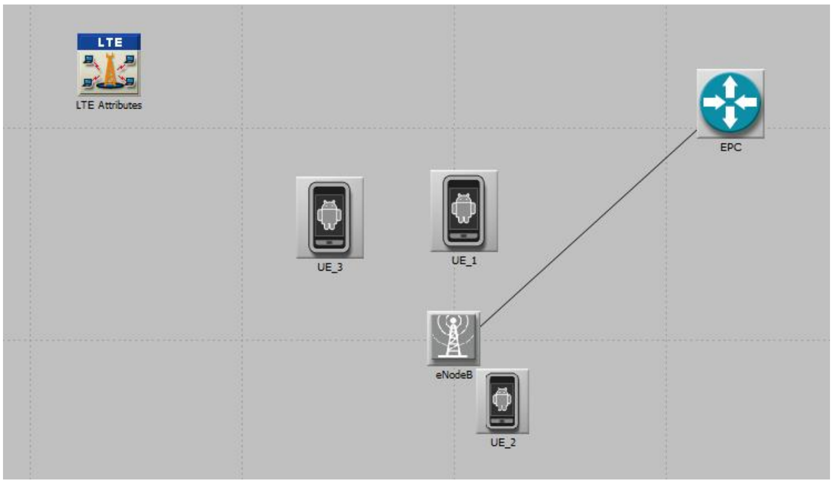


Figure 3.3. Network domain.

Within the process domain, there are extensions to expand the capabilities of the FSM, state variables, state executives, transition conditions and transition executives. State variables are private variables within a process used for the process actions, while state executives are actions performed by the process when a state is entered (enter executives) or left (exit executives), and transitions conditions and transition executives are used within the transitions between states, transitions conditions being expressions to decide if a transition can occur, and transitions executives actions that are performed while executing a given transition.

The three main domains that were briefly summarised represent the core of the modeller simulator engine, however, there are some other tools available to develop or to perform specific tasks. For example, there is the link model editor to create new types of link objects, the probe editor that allows to specify the statistics to be collected during simulation, and the packet format editor to design and specify new packet formats.

3.2.2 Traffic Generation

In order to simulate the behaviour of a given network and obtain performance results, traffic must be added. OPNET Modeler allows the addition of traffic to the network in two different ways: either manually, by setting the attributes from the applications, or automatically, by importing traffic from external files or programs. In order to import traffic from external sources, one can import end to end background traffic, which may represent the traffic from a real network, to infer the impact of changing the network configuration or of having traffic variations in a given network.

Regarding the manual traffic, two different types can be modelled: explicit and background. The background traffic in analytically modelled, this traffic affecting performance of explicit traffic by adding additional delays. Background traffic effects are modelled by calculating the increase in queue sizes, and additional delays because of queue lengths.

Regarding the explicit traffic, there are three general methods of generation. Packet generation, which is the most basic type, only supported for certain simplified node types; these nodes have the ability to generate stream of generic packets, using arguments such as inter-arrival time or the packet size. Application demands, which can be created to represent a flow between two nodes in the network, these demands being characterised in terms of the size and rate of the requests, and responses between the two nodes. Application traffic models, which are included in OPNET Modeler as a set of pre-defined applications, such as File Transfer Protocol (FTP), voice, and email, that can be customised to represent a more realistic application profile, however, there is also the possibility of defining custom applications.

OPNET Modeler has two node objects, defined to characterise applications traffic models, the Application Definitions, and the Profile Definitions nodes. Application Definition node, is responsible for defining the available applications in a given network. Profile Definition node, is responsible to define profiles that identify the behaviour of the available applications.

As seen in Figure 3.4, the Application Definition node allows the configuration of several applications, and the decision of which applications are available in simulations. Each application can be configured in great detail, such as size, compression delay, voice frames per packet, etc. The Profile Definition node allows to define duration, repeatability, and start time for the applications, and the repeatability of the profile itself.

Attribute	Value
⊟ Application Definitions	(...)
Number of Rows	1
⊟ Voice Application (GSM quality)	
Name	Voice Application (GSM quality)
⊟ Description	(...)
Custom	Off
Database	Off
Email	Off
Ftp	Off
Http	Off
Print	Off
Peer-to-peer File Sharing	Off
Remote Login	Off
Video Conferencing	Off
Video Streaming	Off
Voice	GSM Quality Speech

a) Application Definition node.

Attribute	Value
⊟ Profile Configuration	(...)
Number of Rows	1
⊟ Voice Application (GSM quality)	
Profile Name	Voice Application (GSM quality)
⊟ Applications	(...)
Number of Rows	1
⊟ Voice Application (GSM quality)	
Name	Voice Application (GSM quality)
Start Time Offset (seconds)	uniform (5,10)
Duration (seconds)	exponential (100)
⊟ Repeatability	(...)
Inter-repetition Time (secon...	exponential (300)
Number of Repetitions	Unlimited
Repetition Pattern	Serial
Operation Mode	Serial (Ordered)
Start Time (seconds)	uniform (100,110)
Duration (seconds)	End of Simulation
⊟ Repeatability	(...)
Inter-repetition Time (seconds)	constant (300)
Number of Repetitions	constant (0)
Repetition Pattern	Serial

b) Profile Definition node.

Figure 3.4. Examples of attributes edition for the Application Definition and Profile Definition nodes.

3.2.3 Metrics

In order to obtain results, OPNET Modeler has a number of statistics and metrics to be collected, as presented in Figure 3.5. Each of the modules at the node and link levels are a source of a significant number of available statistics. These statistics are divided into four different types: Global, Node, Module and Link. The selection of statistics is performed by specifying a list of probes, which perform the selection and control the flow of data from the selected statistics into output files.

After the collection of all desired data during a given simulation, OPNET Modeler gives the ability to analyse it in various ways, for example, traces of a given output can be presented, and these traces can be changed to show the average along time. These results can be shown separately in dedicated windows, or by combining several results in the same window. OPNET Modeler also gives the possibility to export data, for example, exporting the collected data to an Excel spreadsheet. In order to collect statistics, one need to specify the number of samples that the OPNET Modeler should collect during the simulation, the standard value being 100 samples.

Node Statistics and Global Statistics were selected in all simulations, the selected statistics in this thesis being: packet end to end delay, presented in (3.9); in order to compute the end to end delay as a function of the number of users, (3.10) and (3.11) were used; MOS, presented in (3.12), but since MOS is only available in the Global Statistics, one used (3.11) in order to compute the average MOS as a function of the number of users, but since it is impossible to access MOS for a particular user in the simulator, this expression gives the average MOS for the total number of users; download response time, presented in (3.13); upload response time, presented in (3.14); throughput, presented in (3.15); and page response time, presented in (3.16).

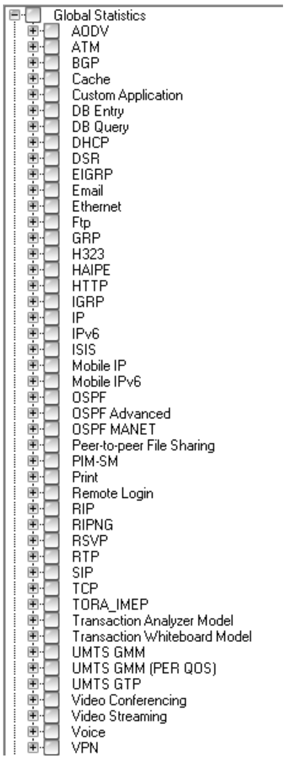


Figure 3.5. Available statistics.

The presented performance metrics are: packet end to end delay; MOS; download response time; upload response time; throughput; and page response time. The end to end delay for LTE should preferably be below 200 ms, which is the value typically achieved in CS. The packet end to end delay, is the total packet delay, called "analogue-to-analogue" or "mouth-to-ear" delay, [Rive13], being:

$$\tau_{ETE[\mathrm{s}]} = \tau_{n[\mathrm{s}]} + \tau_{e[\mathrm{s}]} + \tau_{d[\mathrm{s}]} + \tau_{c[\mathrm{s}]} + \tau_{dc[\mathrm{s}]} + \tau_{dj[\mathrm{s}]} \tag{3.9}$$

where:

- τ_n: network delay;
- τ_e: encoding delay;
- τ_d: decoding delay;
- τ_c: compression delay;
- τ_{dc}: decompression delay;
- τ_{dj}: dejitter buffer delay.

In order to compute the average end to end delay for the users, one should compute an average end to end delay for each user. The average end to end delay for all users is the average of the average values of each user, being given by:

$$\mu_{UN} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mu_{u_i} \tag{3.10}$$

where:

- μ_{UN}: average end to end delay as a function of the users in the network;
- μ_u: average end to end delay for an user;

Taking into account that each user has a given number of seeds per simulation, the average end to end delay for a single user is the average value of the average end to end delay for each seed, i.e., each seed has an average end to end delay, which is the average of all samples of this particular seed. The average end to end delay for the single user is the average value of the number of available seeds for this user, being given by:

$$\mu_u = \frac{1}{n} \cdot \sum_{i=1}^{n} \mu_{s_i} \tag{3.11}$$

where:

- μ_s: average end to end delay of each seed;

MOS is a very subjective KPI, however, the E-model [ITUT08] can provide a prediction of the expected voice quality, since it takes a wide range of telephony-band impairments into account, being applied to CS and PS. The primary output of the E-model is a scalar quality rating value, known as transmission rating factor, R, which can be transformed into other KPIs, such as MOS, being expressed by:

$$R = Ro - Is - Id - Ie,eff + A \tag{3.12}$$

where

- Ro: basic signal-to-noise ratio;
- Is: all impairments that occur simultaneously with the voice signal;
- Id: all impairments due to delay and echo effects;
- Ie,eff: effective equipment impairment factor;
- A: advantage factor;

The value of R ranges between 50 and 100, Table 3.1 showing lower limits of the respective MOS value.

Table 3.1. Relation between R-value and user satisfaction (extracted from [ITUT14]).

R-value	**MOS**	**User satisfaction**
90	4.34	Very satisfied
80	4.03	Satisfied
70	3.60	Some users dissatisfied
60	3.10	Many users dissatisfied
50	2.58	Nearly all users dissatisfied

Download response time is the time elapsed between sending a request and receiving the response packet from the server. According to [Rive13], it is given by:

$$\tau_{DL[\mathrm{s}]} = \tau_{rq[\mathrm{s}]} + \tau_{rv[\mathrm{s}]} + \tau_{sig[\mathrm{s}]} + \tau_{td[\mathrm{s}]} \tag{3.13}$$

where:

- τ_{rq}: request delay;
- τ_{rv}: receive response packet delay;
- τ_{sig}: signalling delay;
- τ_{td}: delay for setup and tear-down;

Upload response time is the time elapsed between sending a file to the server and receiving acknowledgements from it, [Rive13]:

$$\tau_{UP[\mathrm{s}]} = \tau_{sd[\mathrm{s}]} + \tau_{ack[\mathrm{s}]} + \tau_{sig[\mathrm{s}]} + \tau_{td[\mathrm{s}]} \tag{3.14}$$

where:

- τ_{sd}: send delay;
- τ_{ack}: acknowledge delay;

The throughput represents the average number of bits successfully received per second, [Rive13]:

$$Th_{[\mathrm{bit/s}]} = \frac{\sum b_{sr[\mathrm{bit}]}}{s} \tag{3.15}$$

where:

- b_{sr}: bits successfully received;

The page response time is the time elapsed in between the instant that the HTTP user request the page, until the instant that all the objects in the page, which are not pre-loaded, finish the load, being given by:

$$\tau_{pr[\mathrm{s}]} = t_e - t_r \tag{3.16}$$

where:

- t_r: user request instant;
- t_e: end request instant;

3.3 Simulator Implementation

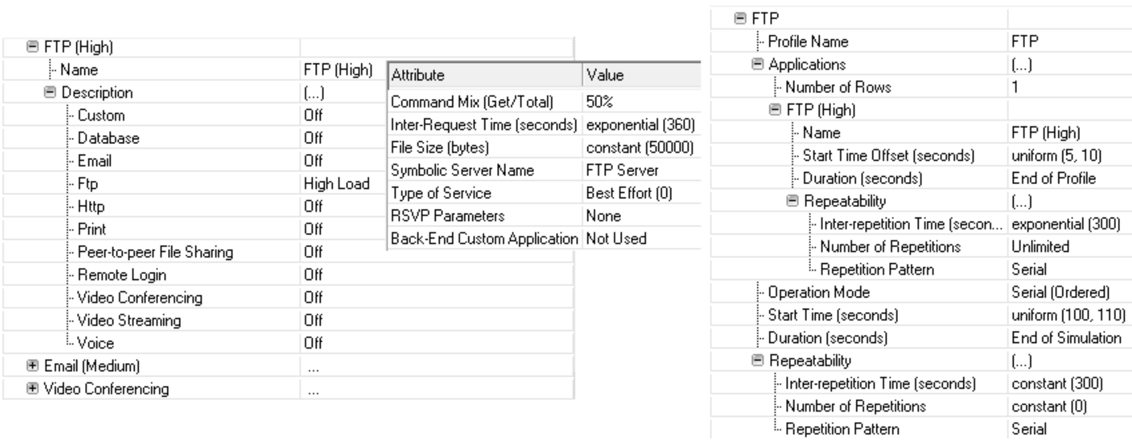
In this section, one focuses on the systems' implementation, from the definition of applications, base station's configurations, and system attributes' configurations, to user's configurations.

In the project manager, one can design a network or a simulation scenario using nodes and objects from the object palette, all networks having different objects according to the system in use; however, all systems have the Application Definition and the Profile Definition, which are used to define the applications available in each scenario. In the Application Definition, five applications were defined, voice, FTP, email, web browsing, from now on addressed as Hypertext Transfer Protocol (HTTP), and video conferencing. In the Profile Definition, one profile for each application was defined.

The duration of the voice application assumes an Exponential Distribution with a given mean value, the repetition pattern having an unlimited number of repetitions, and an inter-repetition time assuming an Exponential Distribution, with a mean value of 300 s, where the repeatability for the profile was set to 0, Figure 3.4. The type of service was set to Interactive Voice, the configuration for the voice service was set with a configuration named *GSM quality* for the VoIP service in the UMTS network, and for the LTE network, the VoLTE service was set with a configuration named Pulse-Code Modulation (PCM) quality, which has the G.711 has a standard speech codec.

The FTP application was set with a configuration named *FTP*, which has an inter-request time assuming an Exponential Distribution, and a given file size, Table 4.1. The duration was set until the end of the profile, the repetition pattern for the profile has an unlimited number of repetitions, and the inter-repetition time assumes an Exponential Distribution with a mean value of 300 s. The repeatability for the profile was set to 0, Figure 3.6.

The email application was set with a configuration named *Email*. The send and receive inter-arrival times assume an Exponential Distribution, these values, as well as the file size, being presented in Table 4.2. The HTTP application was set with a configuration named *HTTP*, which has a page inter-arrival time assuming an Exponential Distribution, and once again, these values, as well as the file size, are presented in Table 4.3. The duration, repetition pattern, inter-repetition time and repeatability for the email and HTTP profiles were set with the same values as the FTP application.



a) Application Definition node.

b) Profile Definition node.

Figure 3.6. Attributes configuration for the Application Definition and Profile Definition for the FTP application.

The video conferencing application was set with a configuration named *Video Conferencing*, which has a frame inter-arrival time information, and a frame size, these values being presented in Table 4.4. The duration assumes a Constant Distribution with a mean value of 100 s, the repetition pattern has an unlimited number of repetitions, and an inter-repetition time assuming an Exponential Distribution, with a mean value of 300 s, the repeatability for the profile was set to 0.

In the UE Attributes, one defines the supported profiles and the supported services, *i.e.*, if one user is set just to initiate calls (for voice or video services) the user is a “calling party” and must support the defined profile for this application, but if the user is a “called party” this user must support the given service. A user may be in both situations, hence, the user must support a service and a profile. In the UE Attributes, it is possible to specify which user is the “called party” for a given “calling party” by defining the Destination Preferences, rather than the “calling party” start a connection with a random user that supports the service. In the case of FTP, email, and HTTP, users must support the defined profile for the application.

The UMTS network is composed of the Application Attributes, Profile Attributes, NodeB, RNC, SGSN, GGSN, and UEs, Figure 3.7. The frequency configuration is done at the node domain, within the NodeB and UEs. The UMTS network does not support CS in this simulator, hence, the voice application supported has the GSM quality configuration, but the VoIP configuration was used. In order to support this configuration, some standard configurations were changed, Figure 3.8: the maximum bit rate for UL and DL was changed from 12.2 kbit/s to 64 kbit/s at the UE attributes, and the TTI for UL and DL transport channels was changed from 20 ms to 10 ms, which allow the use of VoIP in the UMTS network.

Regarding LTE, it is composed of LTE Attributes, Application Definition, Profile Definition, eNodeB, EPC, UEs, router, hub and application servers, Figure 3.9. The LTE Attributes object allows one to define the physical profiles (bandwidth and frequency) and the EPS bearers. Four EPS bearers were defined: two GBR and two Non-GBR. The GBR bearers are used for voice communications and video conferencing with QCIs 1 and 2, and the Non-GBR ones are used for signalling and FTP/email/HTTP with the QCIs 5 and 6. Inside the UE, the bearers are connected to the proper application, Figure 3.10.

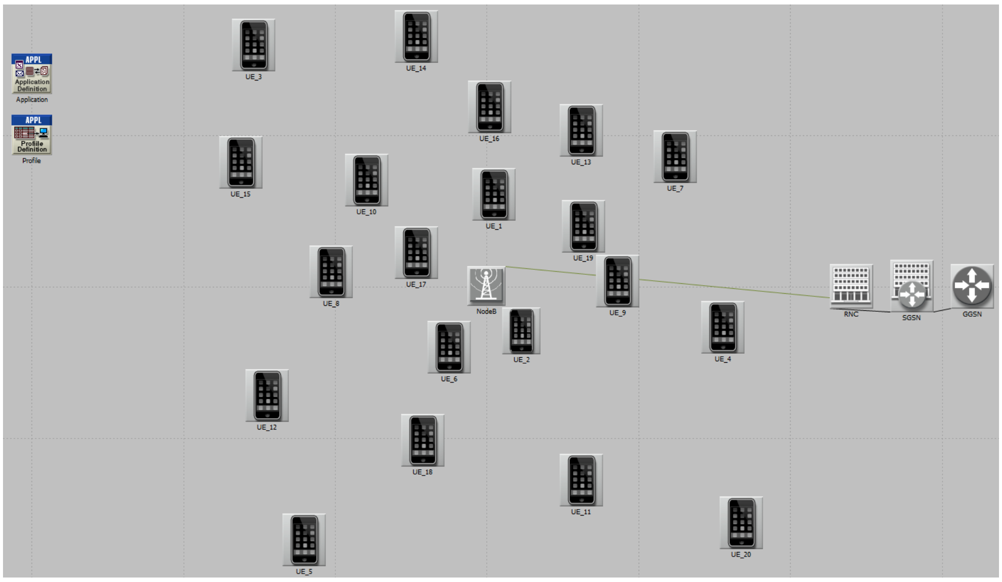


Figure 3.7. UMTS network.

Attribute	Value
UMTS	
UE IMSI	Auto Assigned
UE Serving SGSN ID	0
UMTS GMM Timers	Default
UMTS Logical Channel Configuration	Default
UMTS Logical Signaling Channel Conf...	Default
UMTS PDCP Compression	Disable
UMTS QoS Profile Configuration	(...)
Conversational	(...)
SDU Config	Default
Bit Rate Config	(...)
Maximum Bit Rate Uplink (kbps)	64
Maximum Bit Rate Downlink (...	64
Guaranteed Bit Rate Uplink (k...	Default
Guaranteed Bit Rate Downlin...	Default
Delivery Order	No
Maximum SDU Size (octets)	1500
Transfer Delay (ms)	65,535
Allocation/Retention Priority	Default
Mapped Logical Ch Queuing Sc...	Modified Weighted Round Robin
Streaming	Default
Interactive	Default
Background	Default
UMTS RACH QoS to ASC Mapping	Default
UMTS RLC Processing Time	0.0015
UMTS UE Cell State	CELL_DCH

Attribute	Value
UMTS RNC Parameters	
Admission Control Parameters	Default
Channel Configuration	(...)
Signaling Channel Config	Default
Data Channel Config (Per QoS)	(...)
Conversational	(...)
RLC Info	Default
RB Mapping Info	Default
UL TrChnl Info	(...)
Transmission Time Interval (msec)	10
Type of Channel Coding	Convolutional
Coding Rate	Rate 1/3
Rate Matching Attribute	256
CRC Size (bits)	16
DL TrChnl Info	(...)
Transmission Time Interval (ms)	10
Type of Channel Coding	Convolutional
Coding Rate	Rate 1/3
Rate Matching Attribute	256
CRC Size (bits)	16
RB Id	5
Streaming	Default
Interactive	Default
Background	Default

a) UE definition node. b) RNC definition node.

Figure 3.8. UMTS' attributes configuration for the UE and RNC.

The antennas for both networks are omnidirectional. The NodeB/eNodeB and UE nodes allow the configuration of gains, powers, sensitivities, heights and propagation models, which are shown in Table 3.2. The path loss propagation model used is the Free Space one. The height considered for the antennas, and the UE height are the average heights considered in voice and data communications, [Corr13]. Both systems have one transmission antenna. The used modulation and coding scheme has the index 9 for all users.

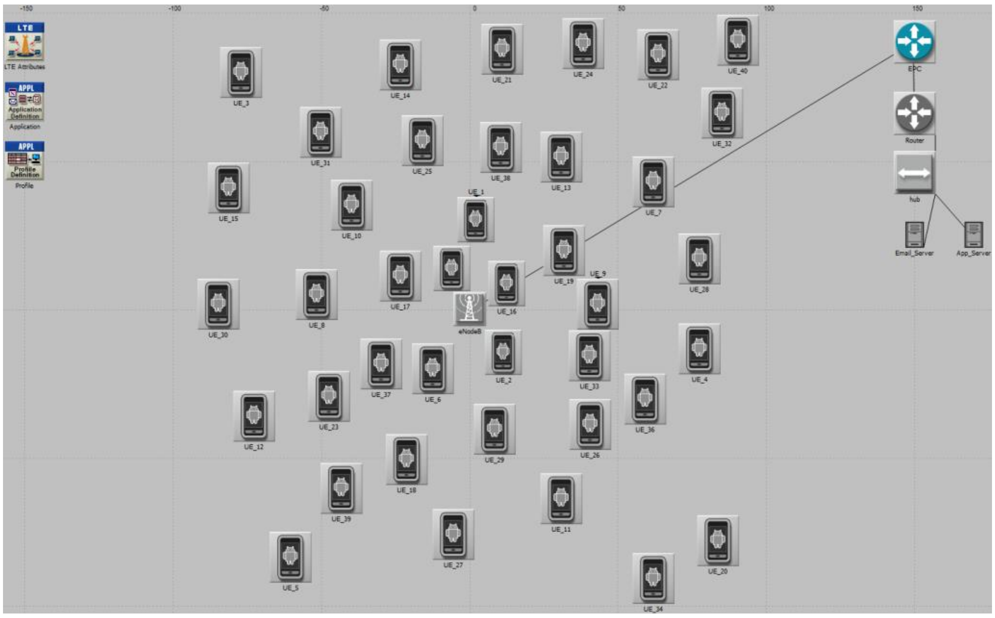


Figure 3.9. LTE network.

Attribute	Value
EPS Bearer Definitions	(...)
Number of Rows	4
Row 0	
Name	Platinum
QoS Class Identifier	5 (Non-GBR)
Allocation Retention Priority	5
Uplink Guaranteed Bit Rate (bps)	384 Kbps
Downlink Guaranteed Bit Rate (bps)	384 Kbps
Uplink Maximum Bit Rate (bps)	384 Kbps
Downlink Maximum Bit Rate (bps)	384 Kbps
Row 1	
Name	Gold
QoS Class Identifier	1 (GBR)
Allocation Retention Priority	1
Uplink Guaranteed Bit Rate (bps)	64 Kbps
Downlink Guaranteed Bit Rate (bps)	64 Kbps
Uplink Maximum Bit Rate (bps)	64 Kbps
Downlink Maximum Bit Rate (bps)	64 Kbps
Row 2	
Name	Silver
QoS Class Identifier	2 (GBR)
Allocation Retention Priority	2
Uplink Guaranteed Bit Rate (bps)	384 Kbps
Downlink Guaranteed Bit Rate (bps)	384 Kbps
Uplink Maximum Bit Rate (bps)	384 Kbps
Downlink Maximum Bit Rate (bps)	384 Kbps
Row 3	
Name	Bronze
QoS Class Identifier	6 (Non-GBR)
Allocation Retention Priority	5
Uplink Guaranteed Bit Rate (bps)	384 Kbps
Downlink Guaranteed Bit Rate (bps)	384 Kbps
Uplink Maximum Bit Rate (bps)	384 Kbps
Downlink Maximum Bit Rate (bps)	384 Kbps

a) LTE Attributes definition node.

Attribute	Value
EPS Bearer Configurations	(...)
Number of Rows	3
Row 0	
Bearer Name	Gold
TFT Packet Filters	(...)
Number of Rows	1
Row 0	
Match Property	IP ToS
Match Value	Interactive Voice (6)
Direction	Bidirectional
Radio Bearer RLC Configuration ...	Default
Action If Not Admitted	Discard Data
Radio Bearer PDCP Configuratio...	Default
Row 1	
Bearer Name	Silver
TFT Packet Filters	(...)
Number of Rows	1
Row 0	
Match Property	IP ToS
Match Value	Interactive Multimedia (5)
Direction	Bidirectional
Radio Bearer RLC Configuration ...	Default
Action If Not Admitted	Discard Data
Radio Bearer PDCP Configuratio...	Default
Row 2	
Bearer Name	Bronze
TFT Packet Filters	(...)
Number of Rows	1
Row 0	
Match Property	IP ToS
Match Value	Best Effort (0)
Direction	Bidirectional
Radio Bearer RLC Configuration ...	Default
Action If Not Admitted	Discard Data
Radio Bearer PDCP Configuratio...	Default

b) UE definition node.

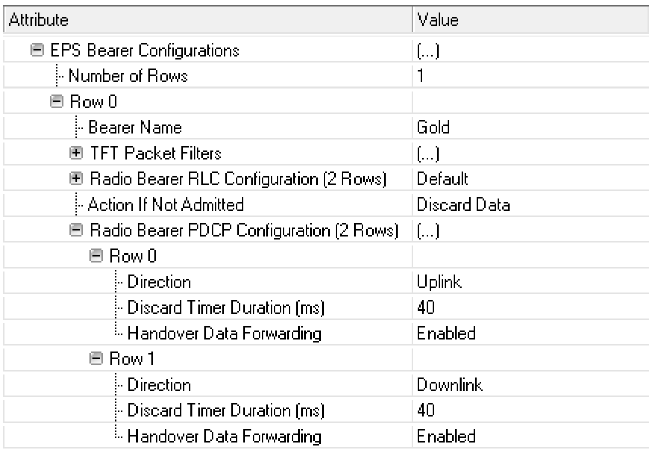
Figure 3.10. Bearers' configuration on LTE Attributes and UE.

The servers are defined to support the services, and since the user must support a profile, the server must support the service. One server must support the FTP service and the HTTP one, and the other must support the email one, and in this specific case the users that support this type of services must specify the Destination Preferences, assigning the designated server. Since there are three services that need servers, a router and a hub is needed to route the packets to the right server. In order to evaluate the impact of the voice service on data ones, data simulations started with 40 data users, these users being distributed as shown in Table 3.3. According to this distribution, there are 18 users performing video conference, 6 users performing FTP services, 4 users performing email services, and 12 users performing HTTP services.

In order to provide a good packet delay budget, the maximum time allowed for a packet in the buffers before being discarded was set to 40 ms, for both type of packets, voice and data ones, taking into account that the delay budget for voice has a maximum of 100 ms and for data of 300 ms, Table 2.3. The full configuration is shown in Figure 3.11. Figure 3.12 shows the configuration used in order to define the bandwidth for the system. In LTE Attributes, one selects the available bandwidth, as well as, the CP type, and UL and DL frequencies. In the eNodeB definitions, one selects the profile which contains the proper configurations.

Table 3.2. OPNET configurations.

Parameter	LTE	UMTS
Frequency [MHz]	1800	2100
Bandwidth [MHz]	5	5
Height of the base station antennas [m]	10	10
Base station gain [dBi]	14	14
Power of the base station [W]	40	40
Base station sensibility [dBm]	-125	-123.4
Height of the voice user [m]	1.5	1.5
Height of the data user [m]	1.2	-
User equipment gain [dBm]	0	0
Power of the user terminal [W]	0.2	0.2
User equipment sensibility [dBm]	-106.4	-106.4
Propagation model	Free space	Free space
Number of PDCCH symbols	1	-
Shadow fading		10

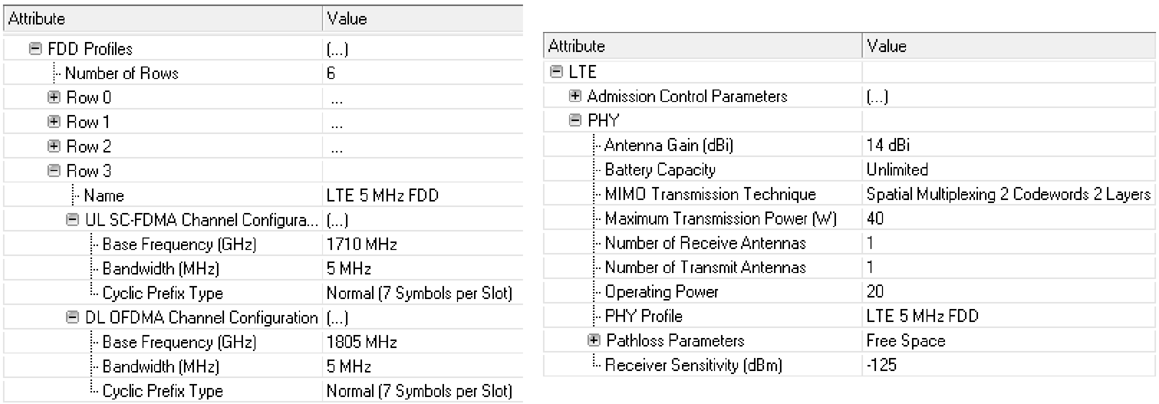


Attribute	Value
EPS Bearer Configurations	(...)
Number of Rows	1
Row 0	
Bearer Name	Gold
TFT Packet Filters	(...)
Radio Bearer RLC Configuration (2 Rows)	Default
Action If Not Admitted	Discard Data
Radio Bearer PDCP Configuration (2 Rows)	(...)
Row 0	
Direction	Uplink
Discard Timer Duration (ms)	40
Handover Data Forwarding	Enabled
Row 1	
Direction	Downlink
Discard Timer Duration (ms)	40
Handover Data Forwarding	Enabled

Figure 3.11. Packet Data Convergence Protocol configuration.

Table 3.3. Data traffic mix (adapted from [Alme13]).

Service	QCI	Penetration [%]
Video Streaming	2	45
FTP	6	15
E-mail	6	10
Web Browsing (HTTP)	6	30



Attribute	Value
FDD Profiles	(...)
Number of Rows	6
Row 0	...
Row 1	...
Row 2	...
Row 3	
Name	LTE 5 MHz FDD
UL SC-FDMA Channel Configura...	(...)
Base Frequency (GHz)	1710 MHz
Bandwidth (MHz)	5 MHz
Cyclic Prefix Type	Normal (7 Symbols per Slot)
DL OFDMA Channel Configuration	(...)
Base Frequency (GHz)	1805 MHz
Bandwidth (MHz)	5 MHz
Cyclic Prefix Type	Normal (7 Symbols per Slot)

Attribute	Value
LTE	
Admission Control Parameters	(...)
PHY	
Antenna Gain (dBi)	14 dBi
Battery Capacity	Unlimited
MIMO Transmission Technique	Spatial Multiplexing 2 Codewords 2 Layers
Maximum Transmission Power (W)	40
Number of Receive Antennas	1
Number of Transmit Antennas	1
Operating Power	20
PHY Profile	LTE 5 MHz FDD
Pathloss Parameters	Free Space
Receiver Sensitivity (dBm)	-125

a) LTE Attributes definition node. b) eNodeB definition node.

Figure 3.12. Physical definitions on LTE Attributes and eNodeB.

3.4 Simulator Assessment

Simulations with the scenario described in Section 3.3 were done to infer the number of necessary samples, the minimum simulation time, and the number of necessary seeds. Firstly, a simulation was done with 20 users, randomly distributed, and a 70 minutes time duration (4 200 s), using the voice service, with one sample per second (4 200 samples), one sample every 5 s (840 samples) and one sample every 10 s (420 samples).

The average value for the simulation was fixed as the global average end to end delay (δ_g), hence a measure of convergence, Δ, was computed as follows:

$$\Delta[\%] = \frac{\delta_n - \delta_g}{\delta_g} \cdot 100 \tag{3.17}$$

where:

- δ_n: assessment average end to end delay;
- δ_g: assessment global average end to end delay;

Figure 3.13 presents the error in between a simulation of 70 minutes with different number of samples, as can be seen with 420 samples (one sample every 10 seconds), the error is below 2.5%. The number of samples chosen was 420, which represents a sample every 10 seconds, despite having an associated error, this number of samples per second allow the reduction of necessary time of computational effort.

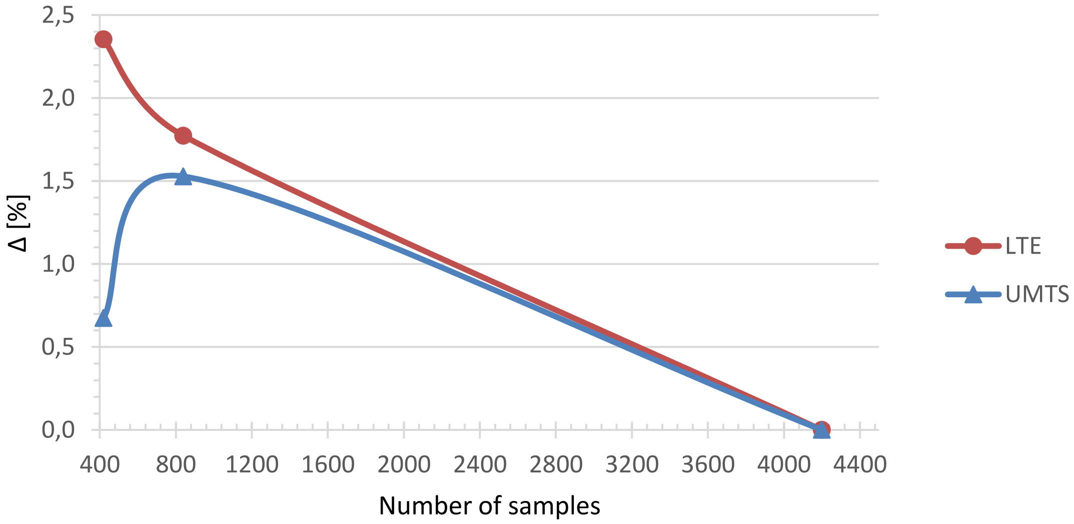


Figure 3.13. Simulation assessment – number of samples.

Secondly, simulations with the previous scenario were done for both systems, to infer the necessary simulation time. The first 10 minutes of simulation were discarded, since the values are not trust worthy, which can be seen in Figure 3.14; the average value begins to stabilise after the first 10 minutes, which can be seen in both systems, as shown in Annex 0. In order to obtain the minimum simulation time, the average end to end delay of the network in a 60 minutes simulation was computed. The average packet end to end delay is given in (3.2).

The average value for a simulation of 60 minutes was fixed as the global average end to end delay, hence, a measure of convergence, ΔT, was computed using (3.2) and (3.17). As one can be seen in Figure 3.15, 20 minutes duration for both systems are enough, hence, the end to end delay converges to less than 1%.

In order to obtain the minimum number of simulations, the scenario was simulated 20 minutes, 10 times, with different number of seeds. As can be seen in Figure 3.16, the maximum error is lower than 2.5% in both systems, for any number of seeds, however, from now on, 5 seeds will be used, to have a significant number of simulations, increasing the credibility.

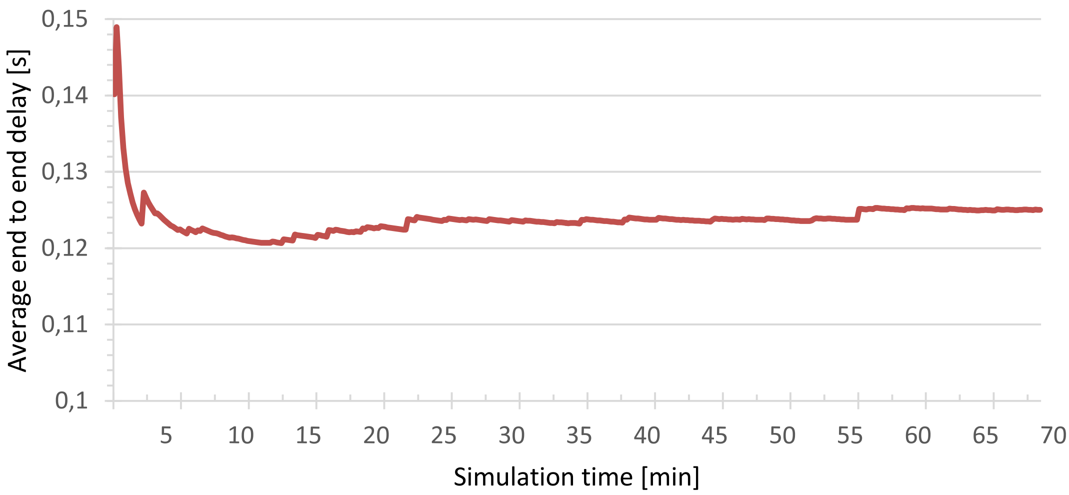


Figure 3.14. LTE assessment – progress of the average end to end delay in a 70 minutes simulation.

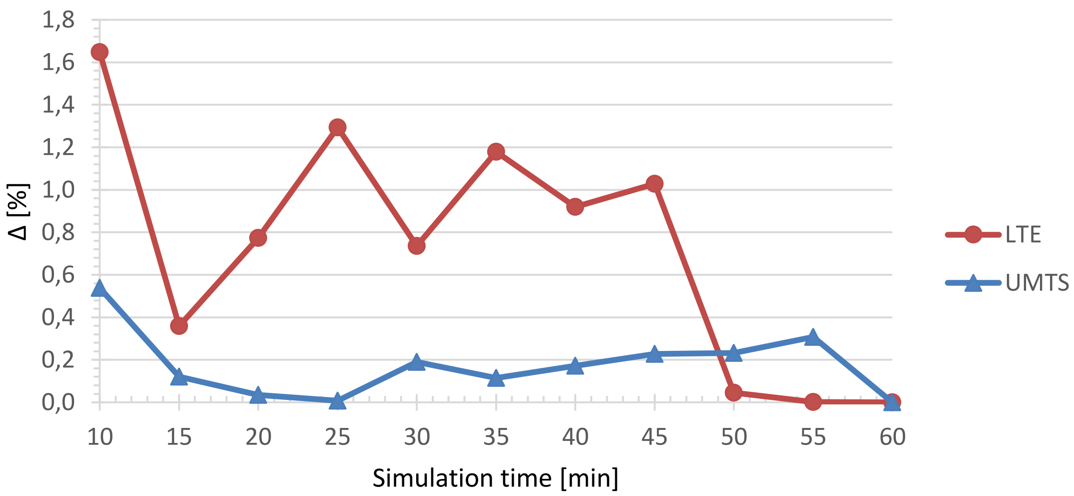


Figure 3.15. Simulation assessment – simulation time convergence analysis.

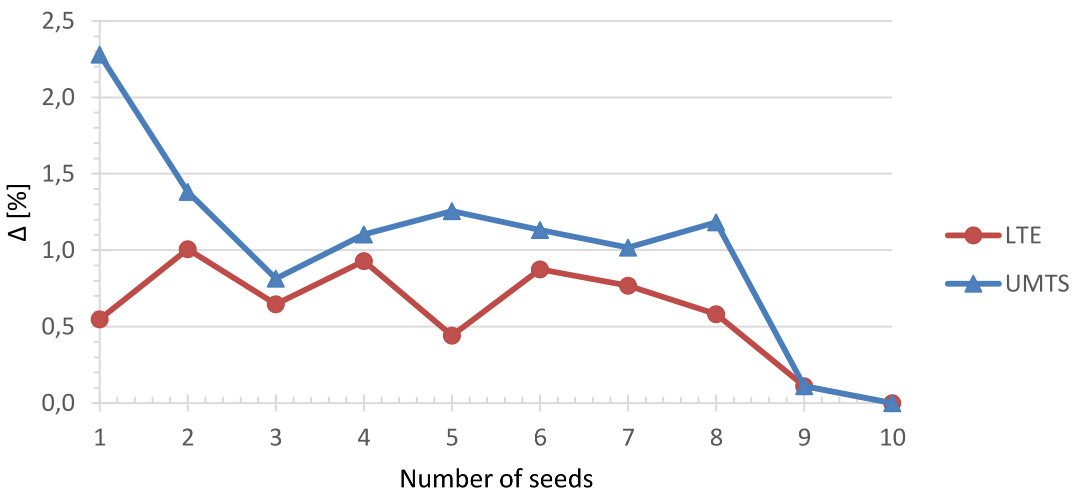


Figure 3.16. Simulation assessment – number of seeds convergence analysis.

Chapter 4

Results Analysis

This chapter contains the scenarios tested and the results obtained via simulation. The first section contains the description of the scenarios, and the second section contains the analysis of the obtained results.

4.1 Scenarios Description

The geographical scenario studied in this thesis is an area of 200×200 m^2, which contains the system's antennas, co-localised in the centre of the area. In this thesis, the main goal was the comparison between the voice performance of the 2 systems using VoIP and VoLTE solutions, with the well-known performance of CS, however, another study has been done. In the second study, the goal was to infer the impact that the voice service has on the data one in LTE, since the voice service has priority over the others, Table 2.3.

In the first part, the study of the voice comparison was done changing by incrementing the number of users and changing the mean call duration. In the second part, the study of the impact that voice has on data was done by incrementing the number of voice users and changing data sizes and request intervals.

In this thesis, 3 main scenarios were set: the first two scenarios correspond to the first part of the study and the third one to the second part. In the first scenario, a UMTS network has been set, in which users are only capable to perform VoIP calls; the mean call duration was changed within {50 s, 100 s, 150 s}. The number of users available in each simulation was 2, 10, 20, 30, 40, and 50 for the simulations with mean call duration of 50 s and 100 s, and 2, 10, 20, 30, and 40 for the one with 150 s.

In the second scenario, an LTE network has been set, in which users are only capable of performing VoLTE calls; the mean call duration had the values in {50 s, 100 s, 150 s, 200 s}. Regarding the number of users, the simulations with mean call duration of 50 s and 100 s had a number of users of 2, 100, 200, 300, 400, and 500, while the ones with 150 s and 200 s had 2, 100, 200, 300 and 400.

In the third scenario, an LTE network has been set, in which users are capable of performing VoLTE calls, FTP, HTTP, email, and video conference, in order to evaluate the impact of voice on data. Two different configurations have been set for data, in order to evaluate the impact on the same data type, but with different data sizes and request intervals. In both configurations, the mean call duration for voice is 150 s. The configurations for the attributes of FTP, email, HTTP, and the video conference services are presented in Table 4.1, Table 4.2, Table 4.3, Table 4.4, respectively. The number of users performing voice calls for the first configuration was 50, 100, 150, 200, 250, while for the second it was 50, 100, 150, and 200.

The FTP configuration allows one to define the Command mix, which is the percentage of file "get" commands to the total FTP commands, the remaining percent of the commands are FTP "put" transitions. This presents the ratio in between downloads and total transactions. The Email configuration allows one to define a specific group of emails in queue to be sent and received, and their sizes. The HTTP configuration allows one to define the number of objects per page, the number of images per page and their sizes, the sizes for the images assuming a Uniform Distribution. The type of service for the FTP, Email and HTTP services is Best Effort.

Table 4.1. FTP configurations.

FTP	1st configuration	2nd configuration
Name	Medium Load	High Load
Inter-request time [s]	300	300
Frame size [bytes]	50000	5000000
Command mix (get/total) [%]	50	50

Table 4.2. Email configurations.

E-mail	1st configuration	2nd configuration
Name	Medium Load	High Load
Send inter arrival time [s]	300	300
Queued e-mail to be sent	3	3
Received inter arrival time [s]	300	300
Queued e-mail to be received	3	3
E-mail size (bytes)	2000	10000

Table 4.3. HTTP configurations.

HTTP	1st configuration	2nd configuration
Name	Light searching	Medium searching
Specification	HTTP 1.1	HTTP 1.1
Page inter arrival time [s]	10	10
Object size [bytes]	1000	5000
Number of object per page	1	1
Image size [bytes]	[50; 2000]	[500; 2000]
Number of images per page	2	2

Table 4.4. Video conference configurations.

Video conference	1st configuration	2nd configuration
Name	Low resolution	Medium resolution
Incoming stream inter arrival time information [s]	0.1	0.1
Outgoing stream inter arrival time information [s]	0.1	0.1
Frame size information [bytes]	128×120 pixels	128×160 pixels

Due to the large capacity in the LTE network, more users were needed to have oscillations in the network, increasing the duration that each seed takes to finish the simulation. The simulation with 100 users takes about 4 hours for each seed to finish the simulation, and for 500 users it needs about 23 hours for each seed. Due to this fact, and taking into account that in LTE the error of using 1 seed, is less than 1%, Figure 3.16, the adopted number of seeds for LTE was 1, instead of 5.

4.2 Voice Analysis on UMTS

In this section, the results obtained from the simulations of UMTS' scenario are presented and analysed. The goal is the analysis of the impact that the increasing number of users has in the network, to understand the variations on the end to end delay.

Figure 4.1 presents the variations of the end to end delay as a function of the number of users, for the first scenario, with a mean call duration of 50 s. The number of users in the network increased from 2 up to 50. Simulations were done with 60 users, however, due to capacity constraints, 4 users did not managed to communicate, even with 5 seeds available for each user. Furthermore, with 60 users, the average number of seeds that each user could use were 2, the collected data from this simulation were discarded, since the values are not trust worthy.

The average end to end delay goes above 200 ms (represented by the green line), which is the acceptable value typically achieved in CS, when there are 40 or more users. The maximum and minimum values as a function of the number of users are represented in the figure, the red line associated with the maximum value and the purple line with the minimum one. The maximum value increases due to the increment of users, except for 50 users, which has a maximum value below the one for 40 users, however, this value is above the acceptable one.

Figure 4.2 presents the variations of the end to end delay as a function of the number of users, for the first scenario, with a mean call duration of 100 s. As above, simulations with 60 users were made, but the values are not trust worthy. The average end to end delay also goes above 200 ms when 40 or more users are connected to the network. The maximum value is increasing due to the increment of users, except for 50 users as in the previous situation.

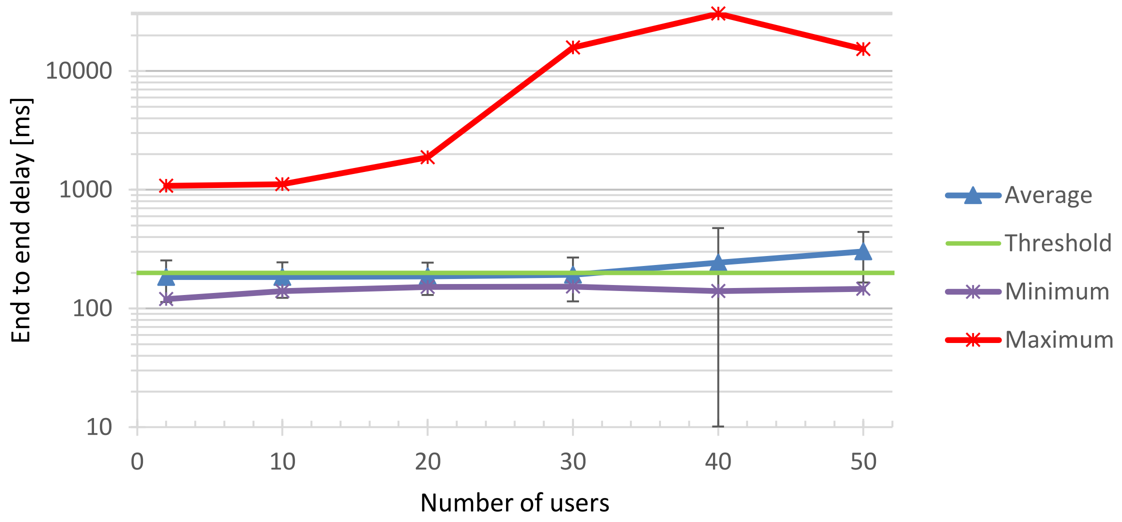


Figure 4.1. End to end delay results for UMTS with mean call duration of 50 seconds.

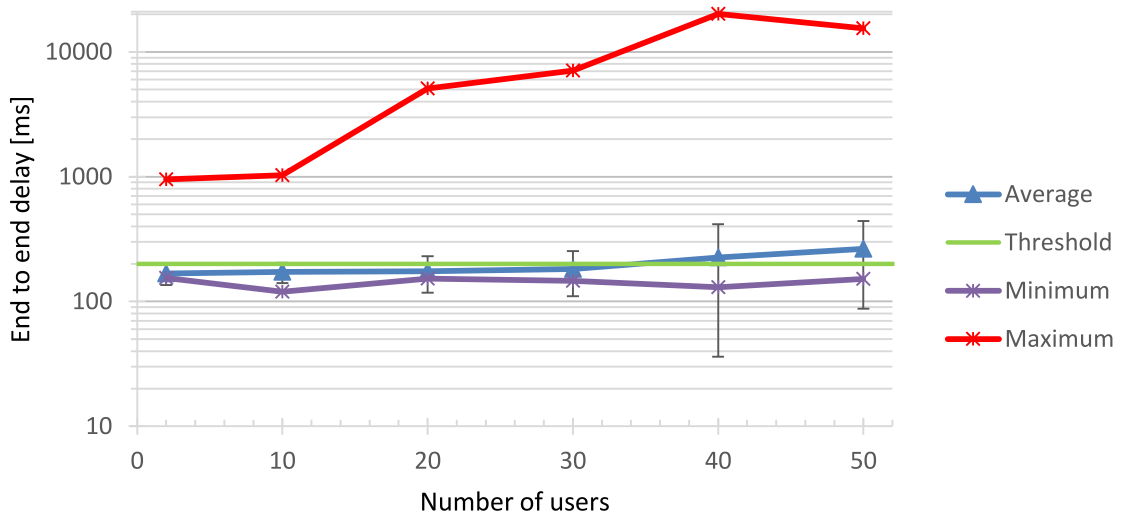


Figure 4.2. End to end delay results for UMTS with mean call duration of 100 seconds.

It can be seen that the increment of the mean call duration decreases the average end to end delay as a function of users, *i.e.*, the simulation with mean call duration of 50 s and 2 users has an average end to end delay higher than the simulation with mean call duration of 100 s with the same number of users, and similarly for 10 users and so on. The maximum end to end delay values are also higher in the simulations with a mean call duration of 50 s.

Figure 4.3 presents the variations of the end to end delay as a function of the number of users, for the first scenario, with a mean call duration of 150 s. In this case, unlike the previous ones, the simulations with 50 users were not trust worthy, so the results with 50 users were discarded. As previously stated, incrementing the mean call duration reduces the average end to end delay, and the maximum one.

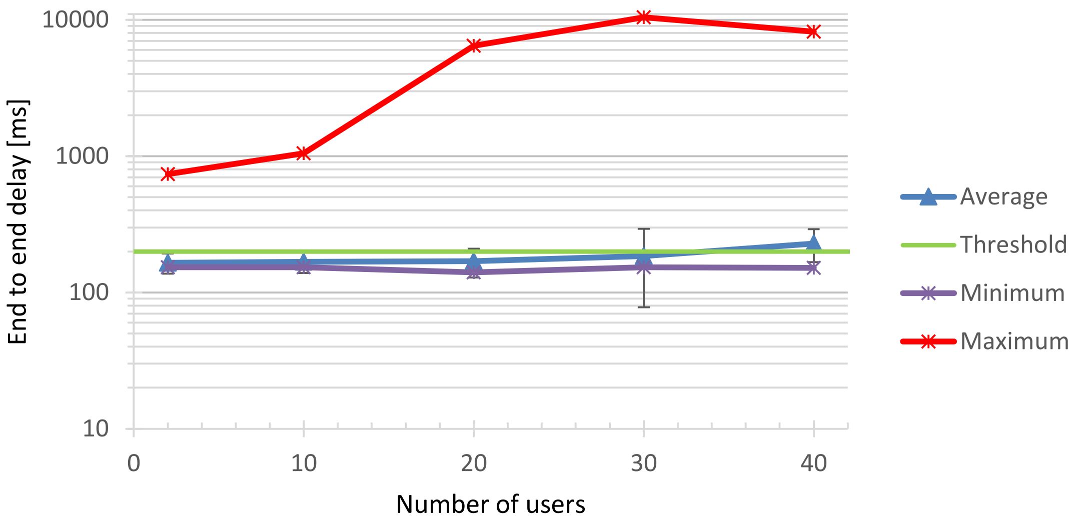


Figure 4.3. End to end delay results for UMTS with mean call duration of 150 seconds.

As seen in the previous figures, the maximum end to end delay values are decreasing with the increment of the mean call duration. Also, the standard deviation is decreasing, representing less variation in its values.

Figure 4.4 presents the differences in between the values set for the mean call duration in the UMTS scenario, in order to understand the impact that the average value and the standard deviation have in the network. The impact due to having a mean call duration of 50 s is higher than for 100 s or 150 s.

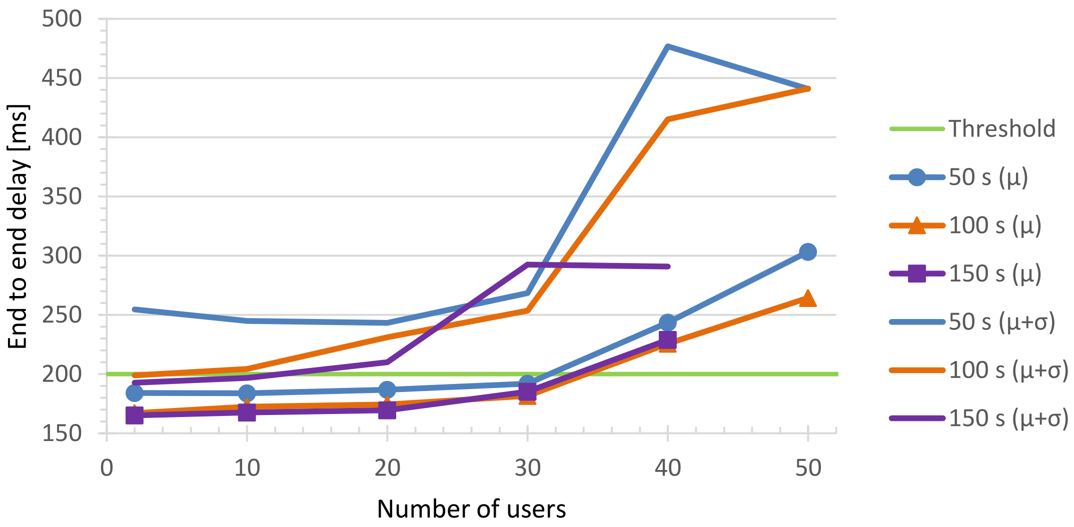


Figure 4.4. All end to end delay for UMTS.

With a mean call duration of 50 s, the average and standard deviation allow the VoIP call to reach values far above the defined threshold in every simulation, due to the fact that there is a large number of calls being started, because the duration of the call is small compared with the simulations with 100 s or 150 s, resulting in more variations in the network. The simulations with a mean call duration of 100 s, have a similar behaviour, however, the average end to end delay, as well as the standard deviation, are smaller, due to the smaller number of calls being started, reducing variations in the network.

The simulations with a mean call duration of 150 s have the smallest standard deviation, the behaviour of these simulations being better than the others, and the achieved values of the average and the standard deviation not being so high compared with the others; however, these values are above the defined threshold, except for the simulations with 2 and 10 users. The average value is lower than in the simulations with a mean call duration of 50 s and 100 s, except with 30 and 40 users; although this behaviour does not follow expectations, it can be explained, since the simulator was achieving its limits.

In order to understand the limitations due to the number of users, a linear extrapolation was performed, to estimate the number of users that reach the defined threshold. The extrapolation was performed using the values for the average end to end delay and the number of users immediately under and above the threshold. Figure 4.5 presents the number of users that reach the defined threshold (200 ms), as a function of the mean call duration. It was expected that in the simulation with a mean call duration of 50 s, the number of users is higher than in the others, however, this is not true, the simulations with a mean call duration of 100 s and 150 s support more 3 users and 2 users, respectively. One cannot take conclusion from this, since it is almost linear, and there are not significant changes due to the increment of the mean call duration.

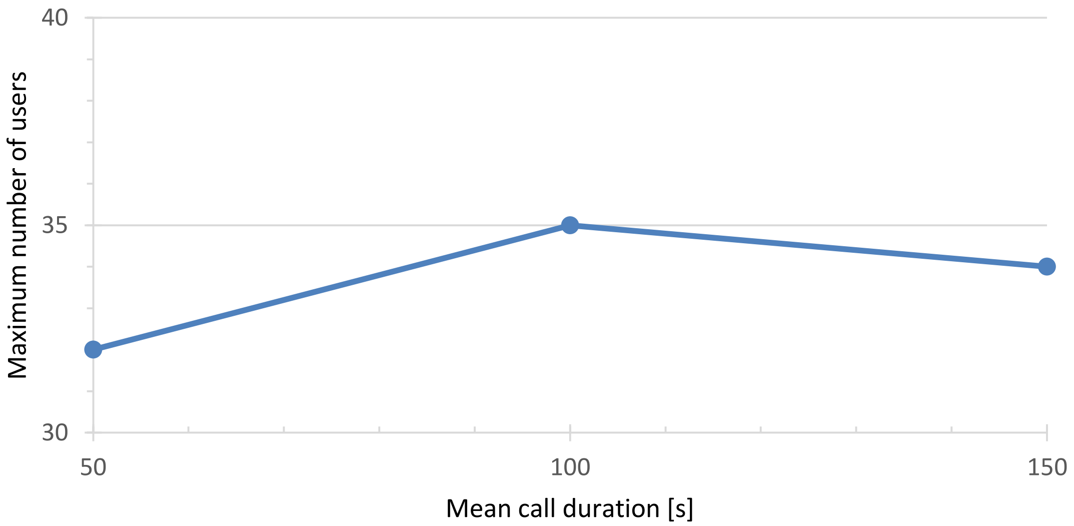


Figure 4.5. Number of users in UMTS as a function of the mean call duration.

Since MOS is not available on the Node Statistics, it was collected from the Global Statistics. It was computed as a function of the number of users, each seed with an average MOS, and the value as a function of the number of users, is the average of the averages from the 5 seeds. Figure 4.6 presents MOS as a function of the number of users, for the mean call duration of 50 s, where one can see that the average MOS and the maximum values have few oscillations, even with the increment of users. Increasing the mean call duration reduces the standard deviation of MOS, but the minimum values do not follow any pattern. The remaining MOS figures (for mean call duration of 100 s and 150 s) are displayed in Annex A.2, due to the very similar shapes and results. The average MOS is always in between 3.5 and 4, Table 3.1, representing a population in which some users could be dissatisfied.

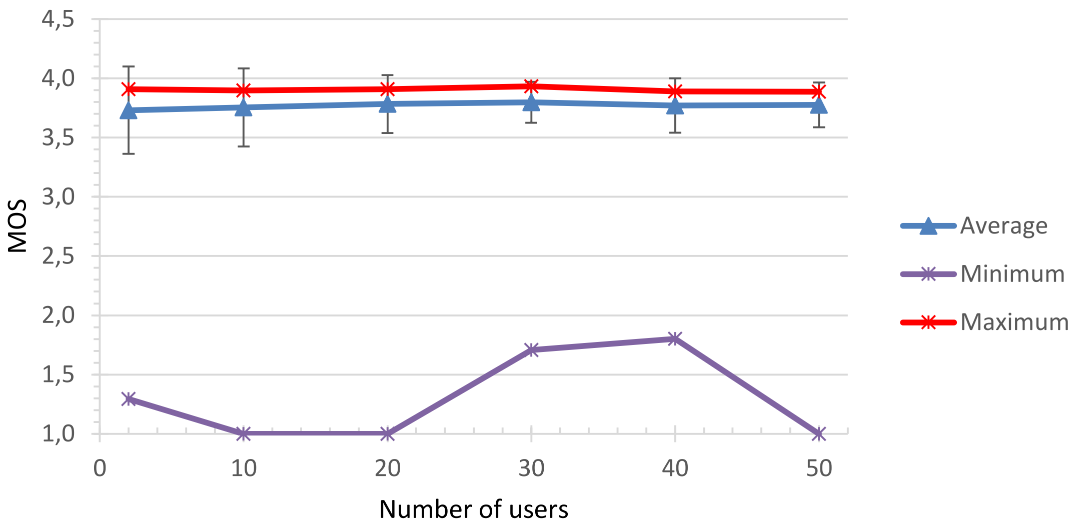


Figure 4.6. Global MOS for UMTS with a mean call duration of 50 seconds.

Taking into account that the simulations with 40 and 50 users have an average end to end delay above 200 ms, and that these simulations represent a high number of samples, these simulations can be evaluated in more depth. Figure 4.7 shows the end to end delay for the scenario with mean call duration of 50 s, with 40 users, for the first seed, *i.e.*, the samples were divided in 5 samplings, corresponding to each seed, without distinction of users. This figure shows the entire universe of samples collected for the first seed.

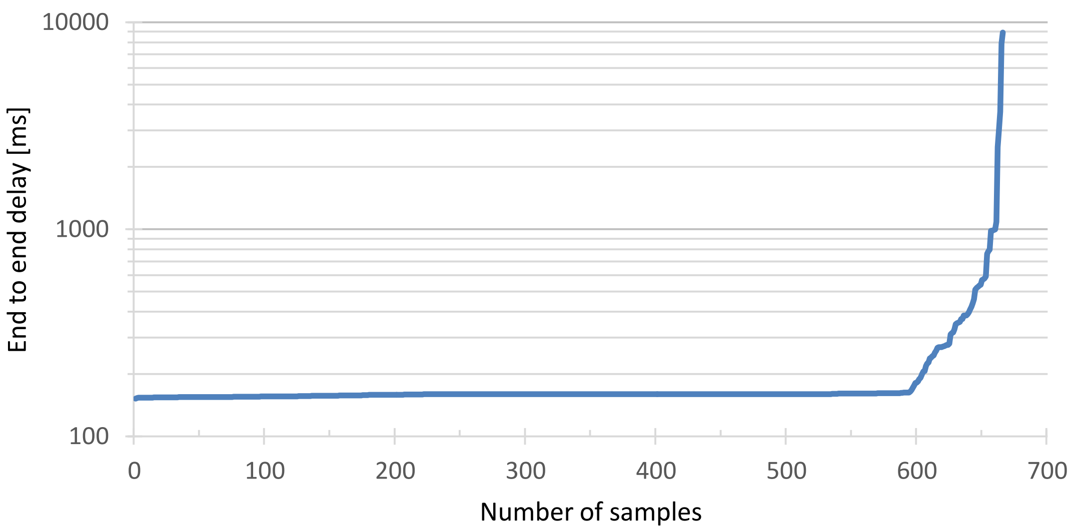


Figure 4.7. End to end delay for UMTS with mean call duration of 50 seconds, 40 users (seed 1).

In order to analyse all samples, it was divided into 3 main groups. The original, in which all samples are analysed; the core, in which there is the higher number of similar samples, which is between 153 ms and 163 ms, in almost every seed; and the acceptable, which are between the highest value of the core (163 ms) and defined threshold (200 ms), i.e., the typical value achieved in CS considered to be the quality threshold. As seen in Figure 4.7, there are a significant part of end to end delays that have acceptable values, however, there are unacceptable values, much higher than the defined threshold, with some samples reaching values in the order of seconds, as can be seen, this seed has a sample in the order of 9 s. These values were discarded in order to perform a Probability Density Function (PDF).

The PDF was performed for the simulations with an average end to end delay above 200 ms, for each seed, using the samples inside the “core” group, and the “original” group was analysed to compare the average, standard deviation, minimum, and maximum. Figure 4.8 represents the PDF for the simulation with 40 users and Figure 4.9 with 50 users, with a mean call duration of 50 s. Only the first seed of each simulation is presented in this section, the remaining seeds being displayed in Annex A.2.

The PDFs were performed using the boundaries where there is a higher concentration of samples; the end to end delay is centred in 160.5 ms in both scenarios, which is an acceptable value. There are no significant differences in the PDFs for 40 and 50 users. Table 4.5 and Table 4.6 shows the average value, the standard deviation, the minimum value, the maximum value, and the number of samples for the simulations with 40 and 50 users, respectively, for the original sampling and for the core one.

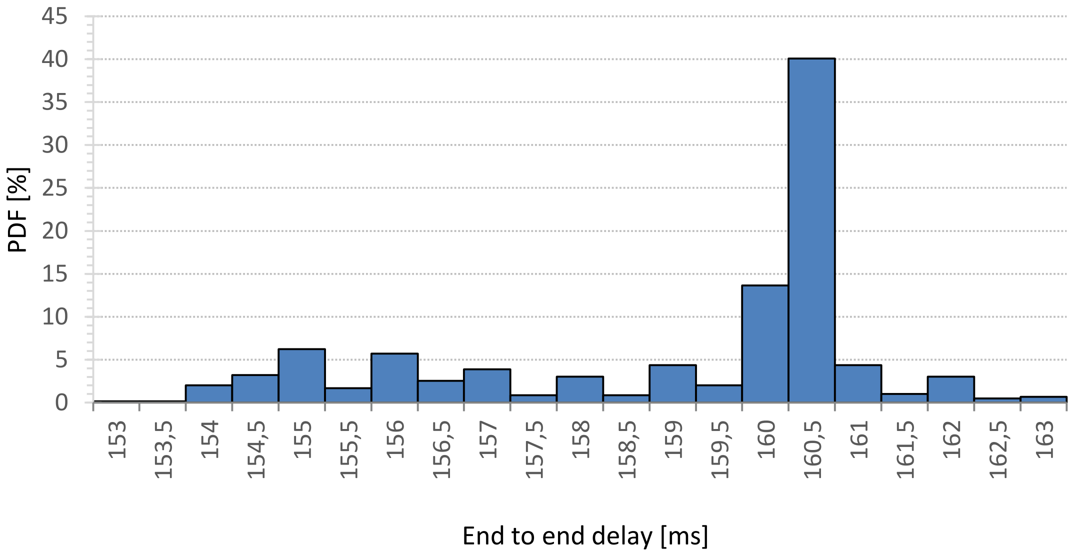


Figure 4.8. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 1)

Table 4.5. PDF values for UMTS, mean call duration of 50 s, 40 users (seed 1).

	Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
Original	220.52	502.19	152.02	8925.24	666
Core	158.75	2.18	152.02	162.68	594

The percentage of discarded values, which are above 163 ms for the simulation with 40 users, is 10.81%; the percentage of samples that are in between 163 ms and 200 ms is 1.50%, which are still acceptable values for voice communications; the remaining 9.31 % are above 200 ms. The difference in between the original number of samples and the core one, shows that the average value has an acceptable value, with an acceptable standard deviation, the core representing almost 90% of the simulation. Regarding the case with 50 users, 10.57% of the samples were discarded. The percentage of samples in between 163 ms and 200 ms is 2.35%, and the percentage of values above 200 ms is the remaining 8.22%. Once again, the samples used to perform the PDF represent almost 90% of the simulation.

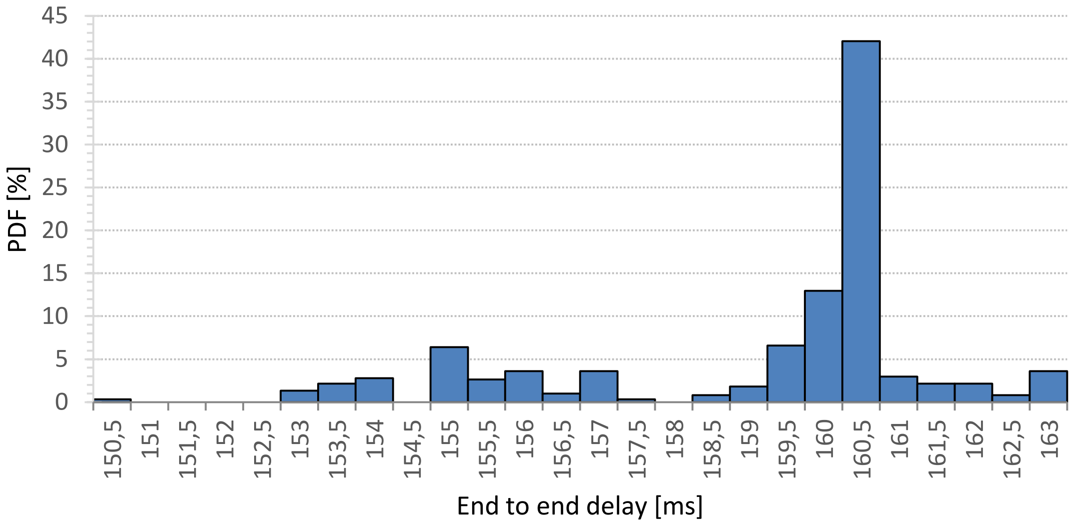


Figure 4.9. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 1)

Figure 4.10 and Table 4.7 present the PDF and the respective values for the first seed of the simulation with a mean call duration of 100 s and 40 users. The percentage of discarded values, which are above 162.5 ms, is 4.65%. The percentage of samples that are in between 162.5 ms and 200 ms is 1.60%, and the remaining 3.06 % are above the 200 ms.

Figure 4.11 and Table 4.8 present the simulation with 50 users and a mean call duration of 100 s. There are 3.06% of samples that were discarded in order to perform the PDF. The percentage of samples in between 162 ms and 200 ms is 0.87%, and the percentage of values above 200 ms is 2.19%. One can observe that with the increase of the mean call duration, the number of discarded samples, which are above the higher value of the core, is decreasing. The average value of the original group is closer to the average value of the core.

Table 4.6. PDF values for UMTS, mean call duration of 50 s, 50 users (seed 1).

	Average [ms]	**Standard Deviation [ms]**	**Minimum [ms]**	**Maximum [ms]**	**Number of Samples**
Original	185.25	126.59	150.00	1786.59	681
Core	158.88	2.44	150.00	162.68	609

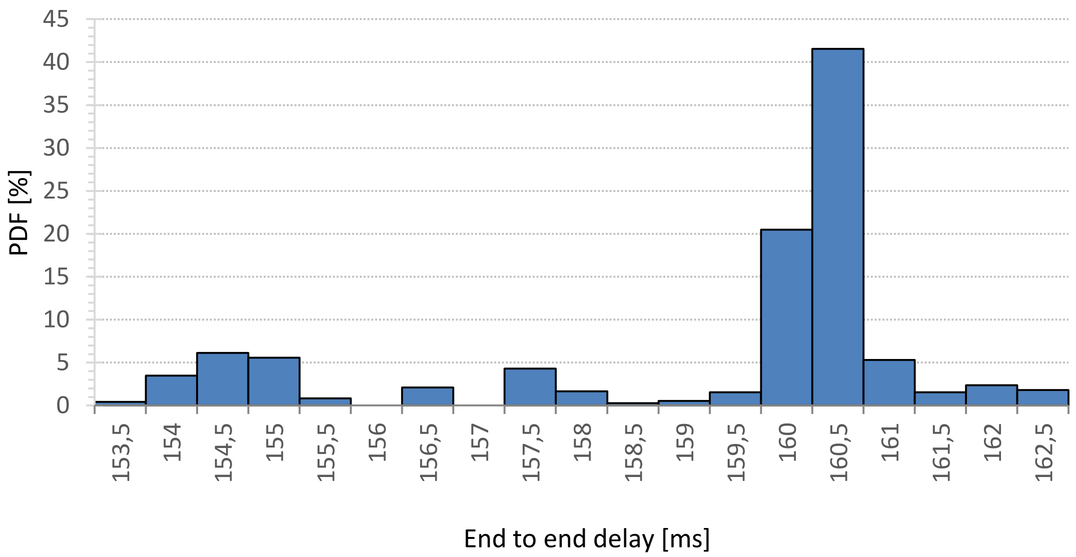


Figure 4.10. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed 1).

Table 4.7. PDF values for UMTS, mean call duration of 100 s, 40 users (seed 1).

	Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
Original	189.98	441.13	153.26	10902.26	752
Core	158.93	2.27	153.26	162	717

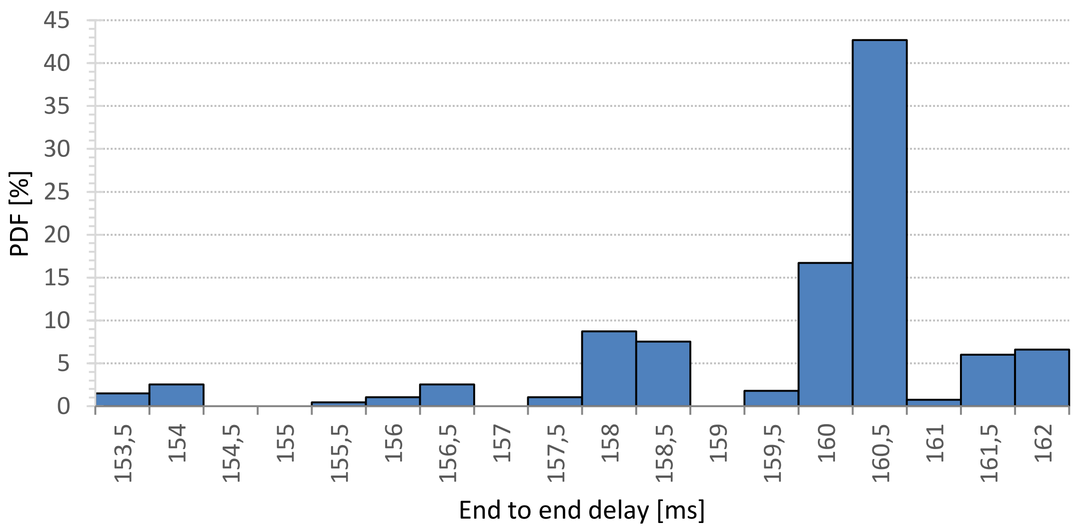


Figure 4.11. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed1).

Table 4.8. PDF values for UMTS, mean call duration of 100 s, 50 users (seed 1).

	Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
Original	184.64	270.70	153.37	4602.07	686
Core	159.46	1.73	153.37	161.93	665

Figure 4.12 and Table 4.9 present the PDF and the values for the simulation with mean call duration of 150 s and 40 users. The discarded percentage, which are above 160.5 ms, is 1.90%. The percentage of samples that are in between 160.5 ms and 200 ms is 0.71%, and the remaining 1.19 % are above 200 ms. Once again, one can observe that the discarded values are decreasing with the increment of the mean call duration.

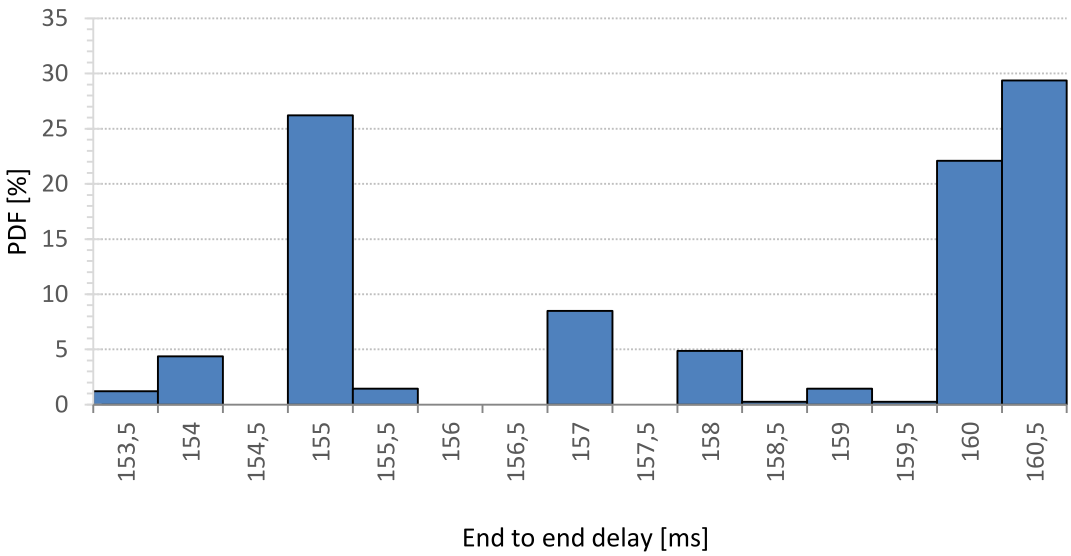


Figure 4.12 PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed1).

Table 4.9 PDF values for UMTS, mean call duration of 150 s, 40 users (seed 1).

	Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
Original	165.64	119.63	153.18	2565.25	420
Core	157.76	2.51	153.18	160.35	412

In order to evaluate the behaviour of the PDFs, the Curve Fitting Tool from Matlab was used, Figure

4.13 presenting the example of the fitting, for the PDF performed for the first seed of the simulation with a mean call duration of 50 s, and 40 users, Figure 4.8. All PDFs are very similar, except the one in Figure 4.12, following a Gaussian Distribution; the remaining figures obtained from the Curve Fitting Tool are not presented, due to the very similar shapes.

It is important to note that all Gaussian Distributions are locally limited, since the Gaussian Distribution itself exists from $-\infty$ to $+\infty$. Table 4.10 presents the relevant values obtained from the theoretical curve presented by the Curve Fitting Tool for the seed number 1, with a mean call duration of 50 s, and 40 users, showing also the mean values of R^2, μ, and σ for all seeds with 40 users.

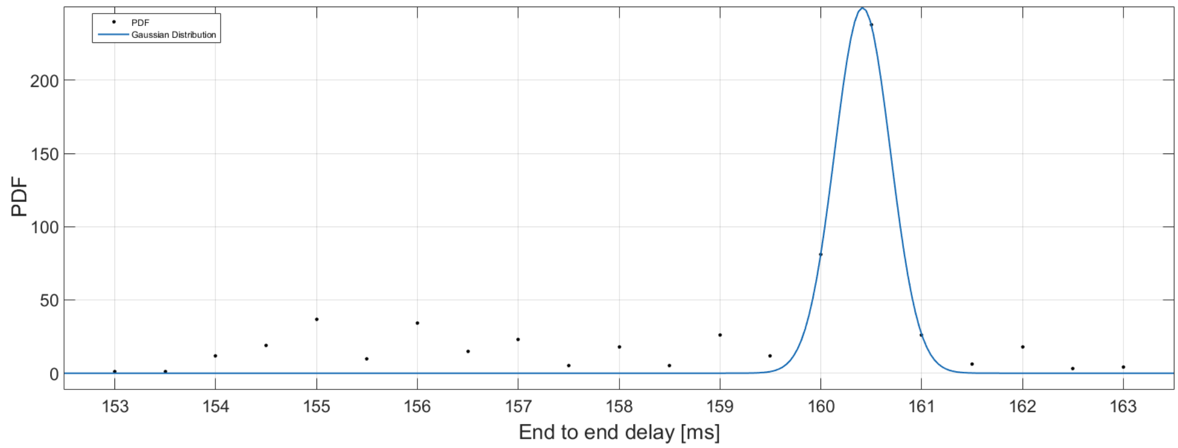


Figure 4.13. Curve fitting example for UMTS, mean call duration of 50 s, 40 users (seed1).

Table 4.11 presents the mean values of R^2, μ, and σ for all seeds with 50 users. The global values for R^2 were computed using (3.6), for μ with (3.7), and for σ with (3.8). The remaining relevant values for each seed are presented in A.2.

Figure 4.14 and Figure 4.15 present the comparison of the average end to end delay for the 5 seeds in between the original sampling and the core one, for 40 and 50 users, respectively. These average end to end delays are the average of each seed, computed for the original sampling and for the core one, using (3.11).

Table 4.10. Global curve fitting values for UMTS, for 40 users, and for seed number 1, with a mean call duration of 50 s.

Mean call duration [s]	**Seed**	**R^2**	**μ [ms]**	**σ [ms]**
50	**1**	0.8965	160.4	0.277
	Global	0.9128	160.36	0.250
100	**Global**	0.8695	160.4	0.239
150	**Global**	0.8683	160.4	0.275

Table 4.11. Global curve fitting values for UMTS, for 50 users.

Mean call duration [s]	Seed	R^2	μ [ms]	σ [ms]
50	**Global**	0.9121	160.4	0.263
100	**Global**	0.8943	160.34	0.213

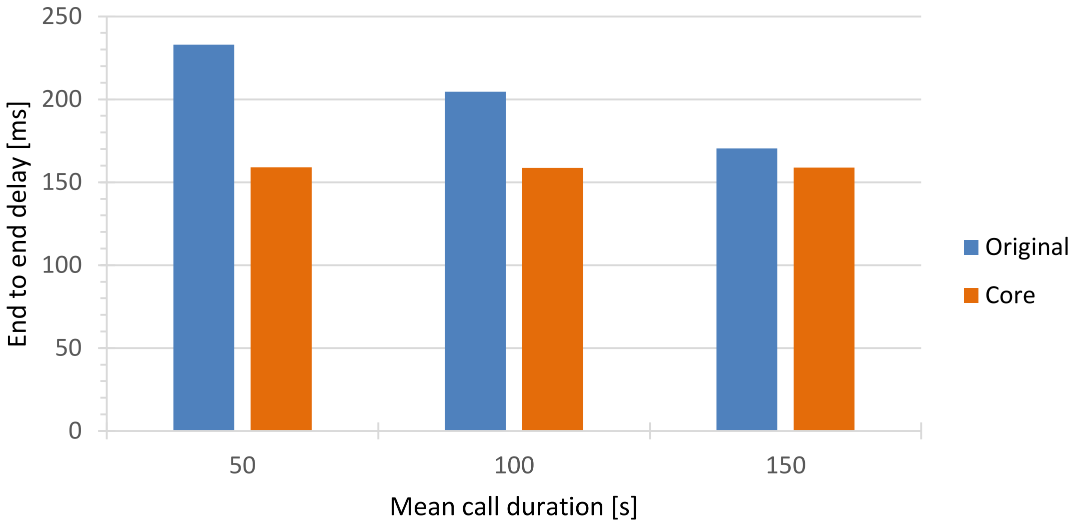


Figure 4.14. Comparison between original sampling and core sampling, UMTS, 40 users.

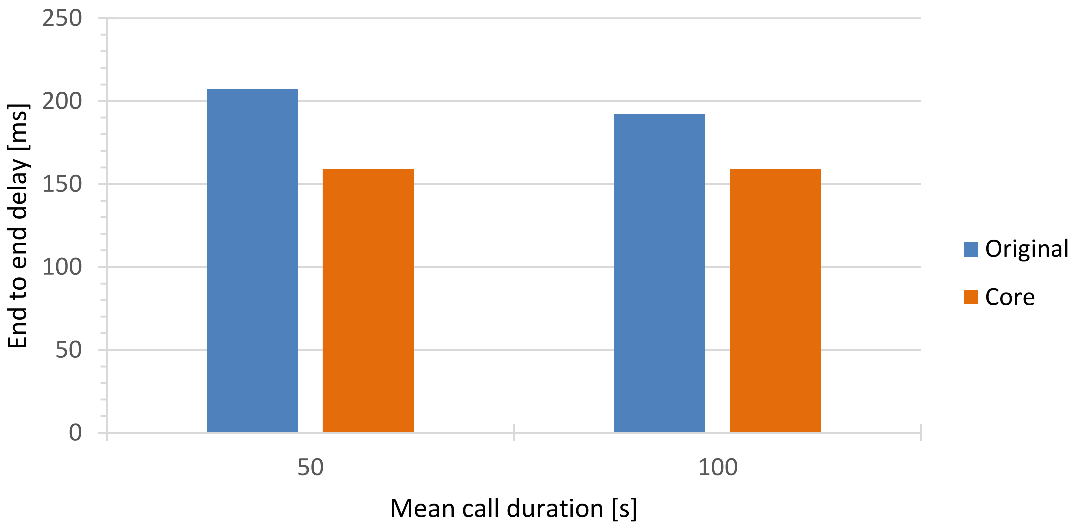


Figure 4.15. Comparison between original sampling and core sampling, UMTS, 50 users.

As previously stated, the difference in between the original samples and the core ones is decreasing with the increment of the mean call duration. In Figure 4.14, for a mean call duration of 50 s, the difference in between the original and the core is 74.04 ms, for mean call duration of 100 s, the difference is 45.78 ms, and for mean call duration of 150 s the difference is 11.74 ms. In Figure 4.15, the difference is 48.42 ms for a mean call duration of 50 s, and 33.20 ms for the mean call duration of 100 s. This decrease comes from the fact that the calls are longer, there are fewer calls being started, resulting in less variations in the network, as stated before.

4.3 Voice Analysis on LTE

In this section, the results obtained from the simulations of LTE's scenario, are presented and analysed. The goal is the analysis of the impact that the increasing number of users has in the network, to understand the variations on the end to end delay.

Figure 4.16 presents the variations of the end to end delay as a function of the number of users, for the second scenario, with a mean call duration of 50 s. In this particular simulation the defined threshold was not reached. As stated before, the LTE network has more capacity, requiring more users to evaluate its behaviour. Since the simulation with 500 users take 23 hours, and the forecast for the simulation with 600 users is about 44 hours, the number of users varied from 2 to 500.

Both the average and the minimum values remained stable, with small oscillations. The maximum value increases with the increment of users, and the standard deviation is stable, with small oscillations; however, it can be seen that the average value coupled with the standard deviation is above the defined threshold.

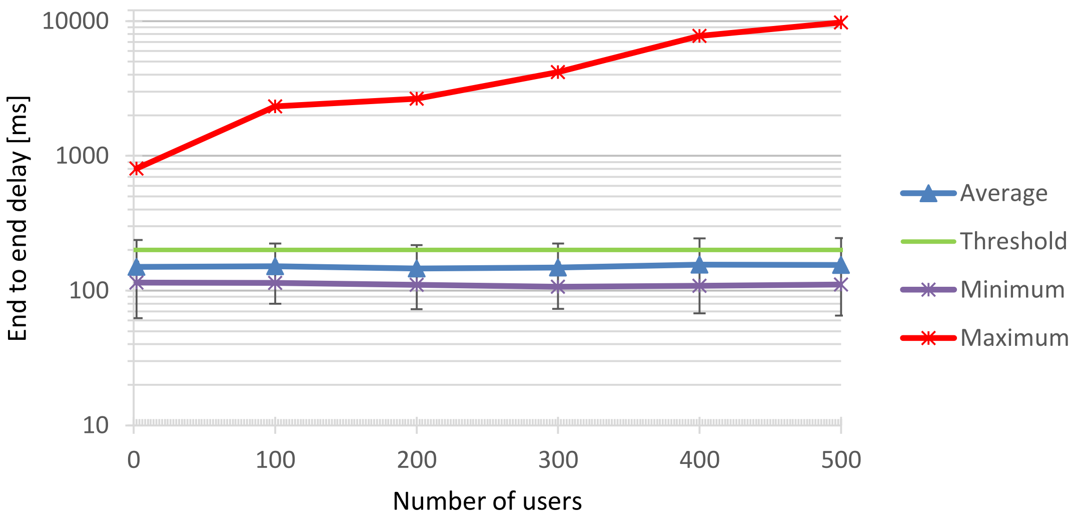


Figure 4.16. End to end delay results for LTE with mean call duration of 50 seconds.

Figure 4.17 presents the variations of the end to end delay for the second scenario, with a mean call duration of 100 s. The average end to end delay is lower than in the previous situation, and remains stable during all simulations until the simulation with 500 users, which is the simulation that reaches a value higher than the defined threshold. The minimum value follows the previous behaviour and the maximum one tends to increase with the increment of users. The standard deviation is smaller than the previous situation, except in the simulation with 500 users. In this simulation the average value coupled to the standard deviation does not reach the threshold in every simulation.

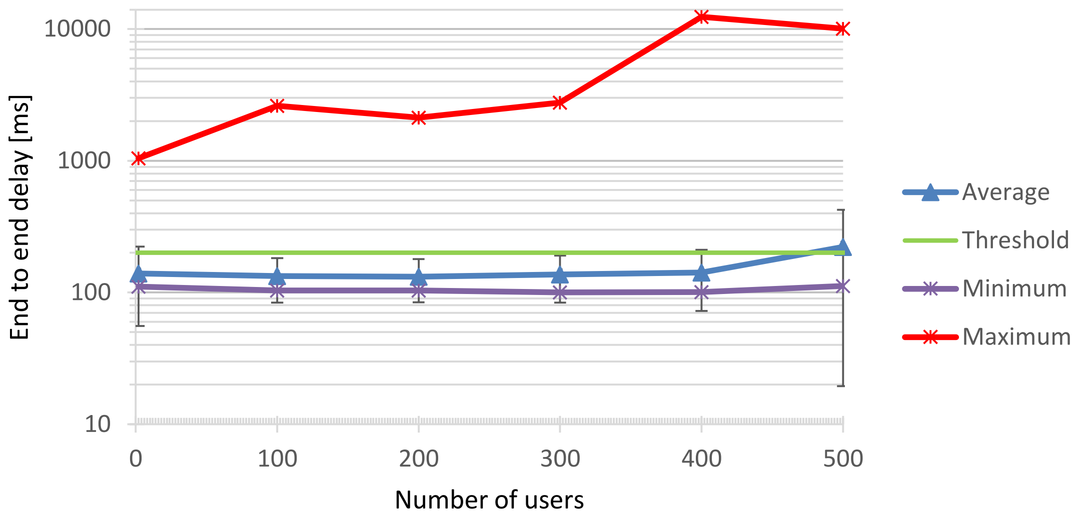


Figure 4.17. End to end delay results for LTE with mean call duration of 100 seconds.

Figure 4.18 presents the variations of the end to end delay for the second scenario, with a mean call duration of 150 s. The maximum value tends to increase as in the previous simulations, and the minimum value follows, once again, the behaviour of the previous simulation, reaching very similar values. The standard deviation is smaller than in the simulation with a mean call duration of 50 s, and the average value coupled to the standard deviation does not reach the threshold in every simulation. The average value is smaller than in the previous situations, due to the fact that with the increment of the mean call duration, there is a smaller number of calls being started, reducing once again variations in the network.

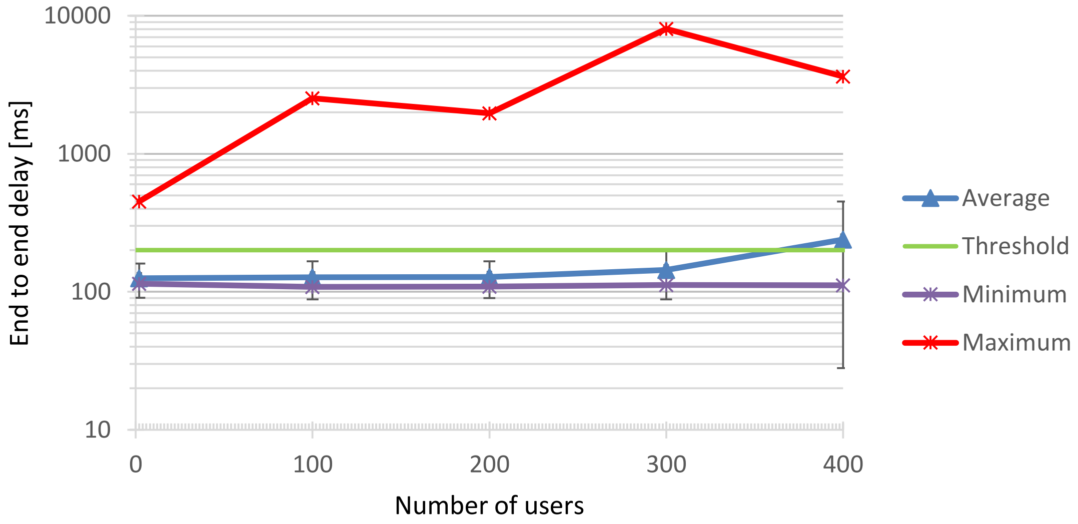


Figure 4.18. End to end delay results for LTE with mean call duration of 150 seconds.

Figure 4.19 presents the variations of the end to end delay for the second scenario, with a mean call duration of 200 s. The behaviour of this simulation is very alike the previous ones, the average, the maximum, and the standard deviation being smaller than the previous situations; however, in the simulation with 400 users, the maximum and the average reach much higher values than in the previous situations. The simulation with 400 users and a mean call duration of 200 s has a very different behaviour from the other simulations, which can come from the fact that for the given bandwidth, with this mean call duration, the limitations of the simulator were reached with a smaller number of users.

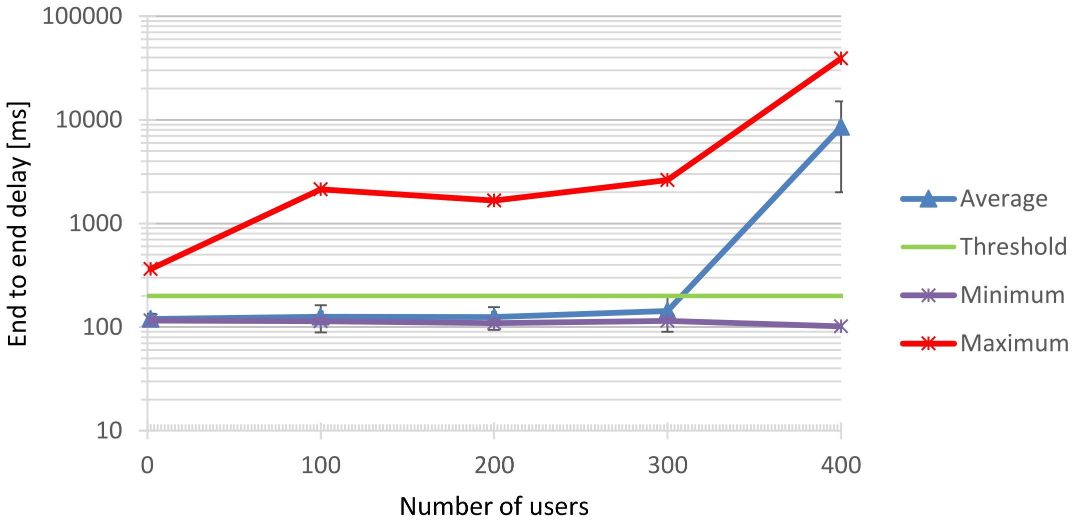


Figure 4.19. End to end delay results for LTE with mean call duration of 200 seconds.

Figure 4.20 presents the differences in between the values set for the mean call duration. The situation with a mean call duration of 50 s, has the highest average values, as well as a highest standard deviation, and as a consequence, the sum of these values is higher than the threshold in every simulation, nevertheless, this situation is the one that allows a higher number of users. The simulations with a mean call duration of 100 s, 150 s or 200 s have a lower end to end delay and standard deviation, and in most part of the simulations the sum of these values is under the threshold.

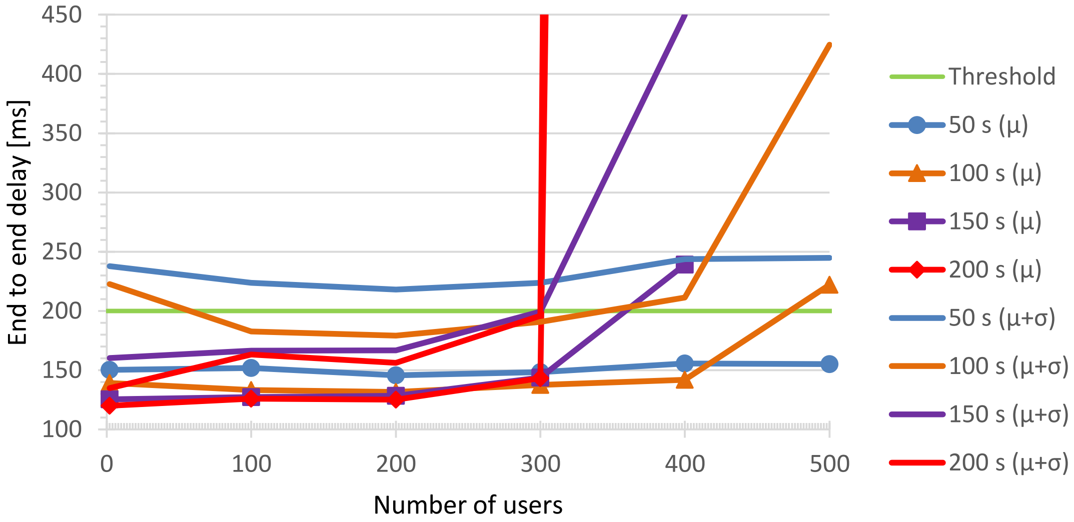


Figure 4.20. All end to end delay for LTE.

The increment of the mean call duration, reduces the average values, as well as the standard deviation, due to the reduced number of calls being started, originating less variations, however, the number of allowed users in the network is reducing with this increment. Since the end to end delay in the simulation with a mean call duration of 200 s and 400 users has a much higher value than the others, it can be seen in Figure 4.20 an abrupt increase for this simulation. In order to understand the variations under the defined threshold, Figure 4.20 was kept in the main body of the thesis, and the same figure in a logarithmic scale is presented in Annex A.3, to present the total behaviour of simulations.

In order to understand the limitations due to the number of users in the network, a linear extrapolation was performed once again. Figure 4.21 presents the number of users that reach the defined threshold as a function of the mean call duration. As expected, the number of users is decreasing with the increment of the mean call duration. As stated before, the simulation with a mean call duration of 200 s and 400 users has a different behaviour compared to the other ones; it can be seen that with 301 users the average value reaches the threshold. Increasing from an average of about 140 ms with 300 users, to 200 ms with 301 users, this fact shows that with 400 users, the simulator is beyond the limitations to the given bandwidth.

Using the Curve Fitting Tool, it is possible to define the equation for Figure 4.21, allowing the prediction of the maximum number of users that an LTE network with a bandwidth of 5 MHz can cover as a function of the mean call duration. The equation has R^2 = 0.9499, representing almost 95% of goodness of fit. This result is:

$$Nu_{max} = -1.96 \cdot \bar{\tau} + 680 \qquad (4.1)$$

where:

- Nu_{max}: maximum number of users;
- $\bar{\tau}$: mean call duration;

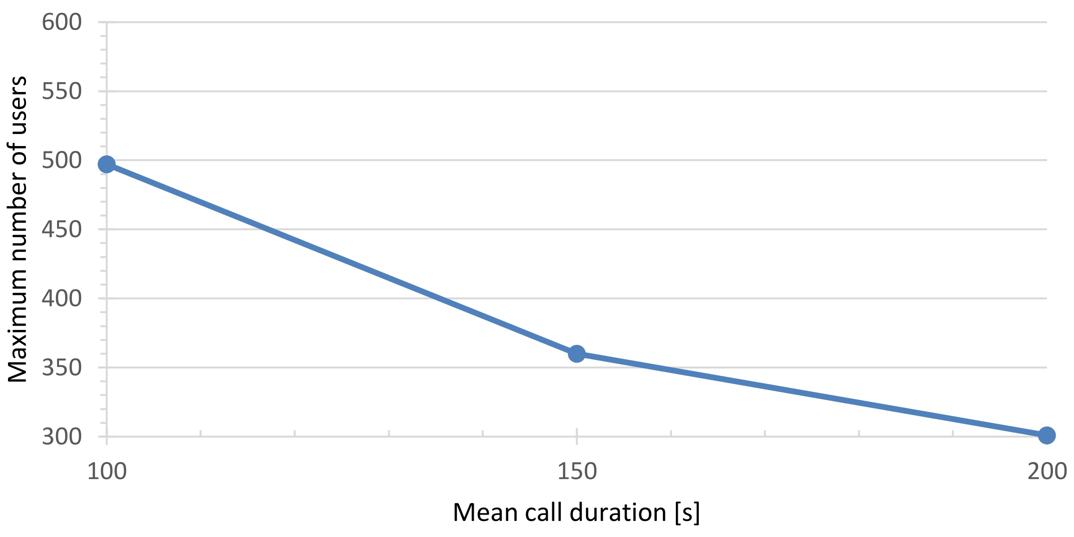


Figure 4.21. Number of users in LTE as a function of the mean call duration.

Equation (4.1) is represented by a linear equation, however, this equation is limited by the points presented in the figure. It is not possible to use this equation to represent the total behaviour of the network, since one would have a negative value for the number of users by increasing the mean call duration, which is not possible. Using the same tool, it is possible to define a rational function, Figure 4.22, which in the upper limits for the mean call duration saturates as it is supposed; however, in the lower limits it has an abrupt growth of users, which is not the expected behaviour for the network, nevertheless, this equation has a closer behaviour to the theoretically expected. This equation has R^2 = 0.9965, which represents a goodness of fit better than the previous equation, in addition to the closer behaviour to theory. This model is:

$$Nu_{max} = \frac{14990}{\bar{\tau}^{\,0.7406}} \tag{4.2}$$

In order to analyse the maximum bit rate allowed by the network in DL, (3.1) was used. Taking into account that the number of sub-carriers per RB is 12, which is a fixed parameter of the simulator, according to [Rive13], the CP in use is the normal, which has 7 symbols per sub-carrier, as seen in Figure 3.12, the number of bits per symbol is 2, due to the modulation and coding scheme index in use, which is 9, the number of available RB in a 5 MHz bandwidth is 25, the number of transmitting antennas is 1, and the time duration of an RB, which is 500 μs, gives a maximum of 8.4 Mbit/s.

It can be seen that the number of user connected to the network is around 500 when the mean call duration is 100 s. This is a very high number of users. Taking into account that the bit rate of the G.711 audio codec is 64 kbit/s, one can conclude that the theoretical maximum number of users that can perform the voice service at the same time, is 131 for a 5 MHz bandwidth. 500 users is far above the theoretical limits, however, this 500 users are not performing voice calls during the entire simulation, sometimes they are performing the voice service, with a given duration described by an Exponential Distribution, and a mean call duration of 100 s.

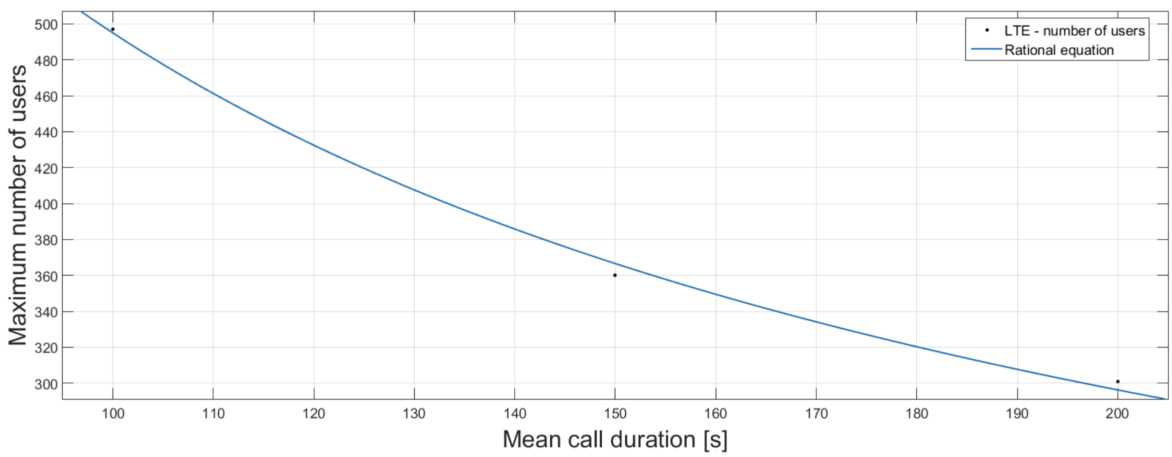


Figure 4.22. Curve fitting for a rational equation.

Figure 4.23 presents MOS for the second scenario, with a mean call duration of 50s. The average MOS has few oscillations, except in the simulations with the number of users that reach the threshold, which has a significant reduction. In these simulations, MOS reduces to a value below 1.5, Table 3.1 shows that with this value nearly all users are dissatisfied. The simulations with a mean call duration of 100 s and 150 s have a MOS in between 3.5 and 3, according to Table 3.1, with these values some users or many users are dissatisfied. The remaining MOS figures (for mean call duration of 100, 150 and 200 seconds) are displayed in A.3, due to the very similar shapes and results.

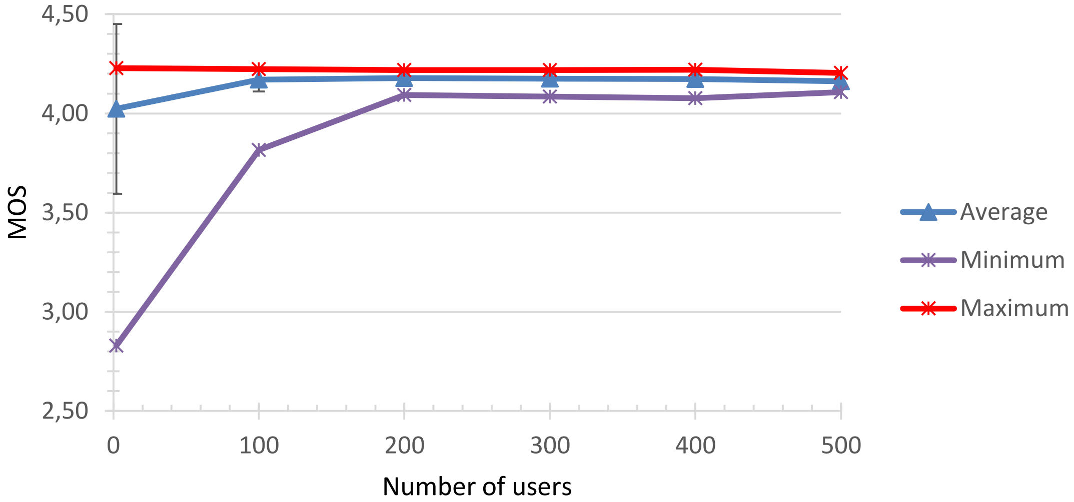


Figure 4.23. Global MOS for LTE with a mean call duration of 50 s.

The standard deviation has the same behaviour in every situation, with a large value with 2 users, stabilising with the increment of users. Regarding the maximum and minimum values, both have the same behaviour, starting with a lower value with 2 users, then stabilising, until the simulation with the number of users that reaches the threshold, where both values decrease. Figure 4.24 presents the PDF for the simulation with a mean call duration of 50 s and 400 users. Once again, the samples were divided into 3 main groups, the original group, which contains all samples, the core group, in which there is the higher number of similar samples, and the acceptable one, which is in between the highest value of the core group and defined threshold (200 ms).

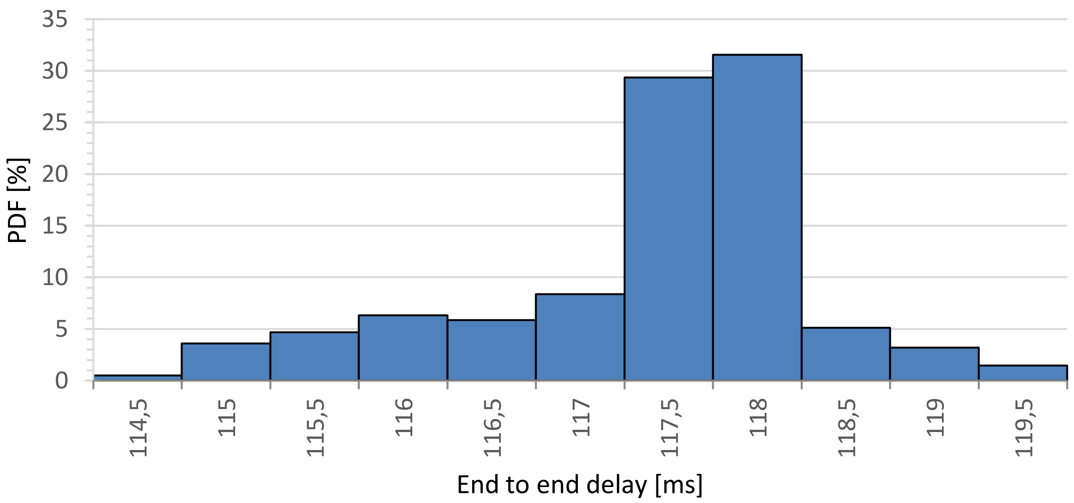


Figure 4.24. PDF of the end to end delay for LTE, mean call duration of 50 s, 400 users.

Table 4.12 shows the average, standard deviation, minimum and maximum values for the original sampling and the core one.The percentage of discarded values, which are above 119.5 ms, is 8.71%, the percentage of samples in between 119.5 ms and 200 ms is 2.36%, this value represents the percentage of acceptable values. The remaining 6.35% is above the define threshold. One should notice that the PDF performed for the simulation with mean call duration of 50 s and 400 users does not have a mean end to end delay higher than 200 ms. This fact led to a higher percentage of acceptable values and a lower percentage of samples above the threshold, than in the previous PDFs performed for UMTS.

Table 4.12. PDF values for LTE, mean call duration of 50s, 500 users.

	Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
Original	145.82	173.67	108.60	7754.80	7588
Core	117.12	0.94	114.30	119.45	6921

Figure 4.25 presents the PDF for the simulation with a mean call duration of 100 s and 400 users. Table 4.13 presents the relevant values for the original sampling and the core one. The percentage of discarded values (above the 130.5 ms) is 6.82%, the percentage of acceptable values is 3.38%, and the percentage of samples above the threshold is 3.44% for this simulation. This simulation has a lower percentage of samples above the threshold than in the previous one, which comes from the fact that with the increment of the mean call duration the standard deviation tends to be smaller, as seen from the comparison between Figure 4.16 and Figure 4.17.

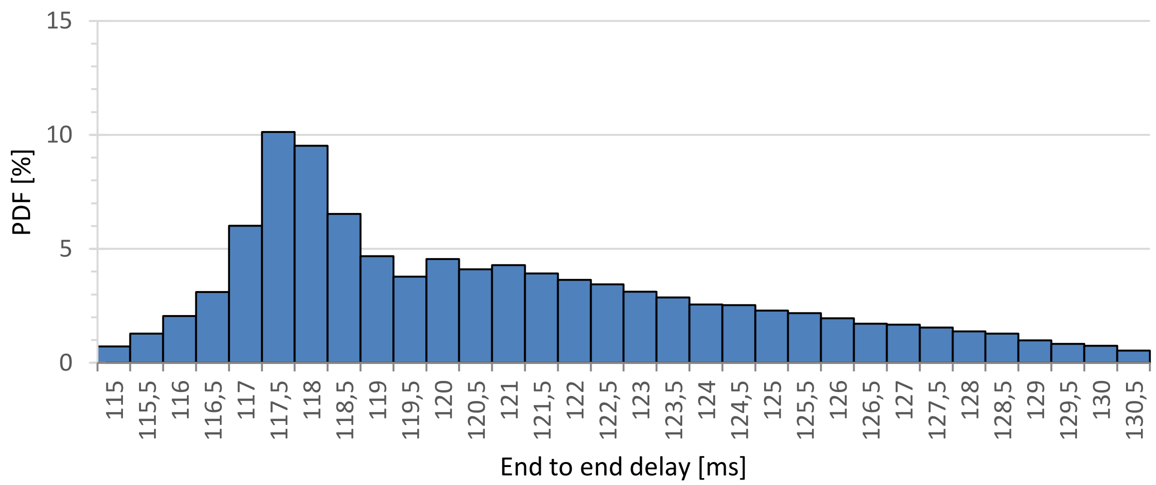


Figure 4.25. PDF of the end to end delay for LTE, mean call duration of 100 s, 400 users.

Table 4.13. PDF values for LTE, mean call duration of 100s, 500 users.

	Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
Original	136.83	155.96	101.00	12367.96	12389
Core	120.59	3.64	114.51	130.50	11538

Figure 4.26 presents the PDF performed for the simulation with mean call duration of 150 s, and 400 users. The percentage of discarded values (above the 180 ms) is 21.10%, the percentage of acceptable values is 6.90%, and the percentage of samples above the threshold is 14.20% for this simulation. Once again, there is a higher percentage of samples above the threshold than in the acceptable group. Table 4.14 presents the relevant values for the original sampling and the core one.

Figure 4.27 presents the PDF performed for the simulation with mean call duration of 200 s, and 400 users. This simulation has a mean end to end delay much higher than the previous ones, due to the fact that under these conditions, the simulator exceeded its limits. The PDF was performed with the higher number of similar samples, however, this sampling is more widespread than the others, which forced to expand the scope of the PDF. The PDF was performed in between the 141.5 ms and the 300 ms. The percentage of samples below the defined threshold is 14.05%, and the percentage of samples discarded (above 300 ms) is 76.35%. Table 4.15 presents the relevant values for the core sampling and the original sampling. It can be seen that the number of samples used to perform the PDF, which have less variations in between them, are much smaller than the total number of samples. The remaining samples have their occurrences very distant to each other.

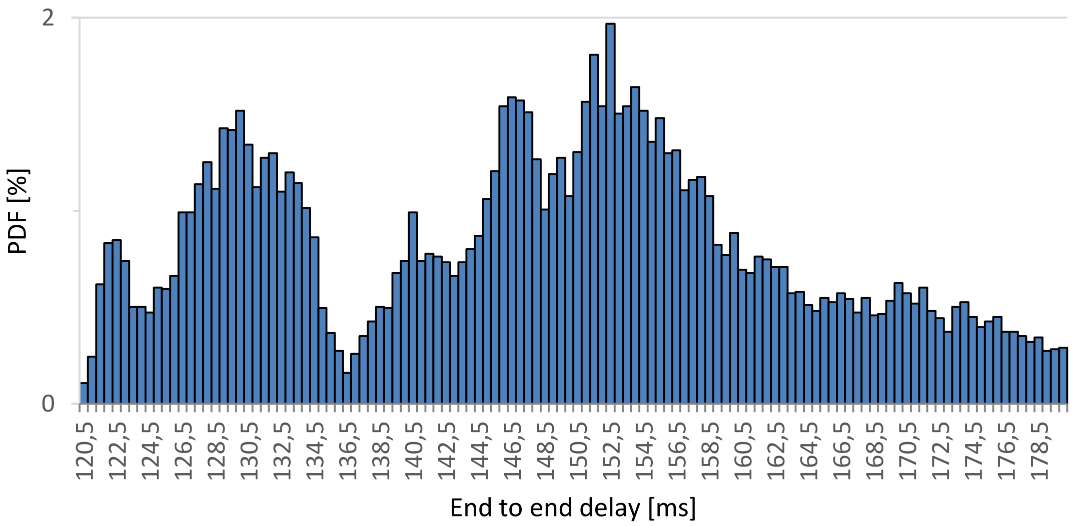


Figure 4.26. PDF of the end to end delay for LTE, mean call duration of 150 s, 400 users.

Table 4.14. PDF values for LTE, mean call duration of 150s, 400 users.

	Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
Original	219.81	257.67	111.75	3634.92	16619
Core	147.61	14.84	120.15	180	13104

One can conclude that with the increment of the mean call duration, the standard deviation for PDFs increases, which can be explained by the increase of the maximum value used to perform the PDFs, and that with the increment of the mean call duration, the percentage of samples that are above the define threshold is increasing.

The PDFs behaviour were evaluated to understand if they fits any distribution. The Curve Fitting Tool, as well as the Distribution Fitting Tool from Matlab were used to evaluate the fitting possibilities. The PDF presented in Figure 4.24 is described by a Gaussian Distribution (once again, one should note that all Gaussian Distributions are locally limited, since it exists from $-\infty$ to $+\infty$); Table 4.16 shows the relevant values for the curve fitting.

The remaining PDFs do not fit in any distribution, distributions for Figure 4.25 were tested, however, the maximum R^2 was 0.4129 for a Gaussian Distribution and 0.4374 for a Lognormal Distribution, which shows that the fittings are not appropriate for the collected data. As the mean call duration increases the PDFs tend to be wider, Figure 4.26 and Figure 4.27, which has a larger range for the core sampling; one should note that both simulations have an average end to end delay that exceeds the threshold, with 400 users, unlike the previous ones. One should note that the shape of the PDF should not change due to the changing of the mean call duration.

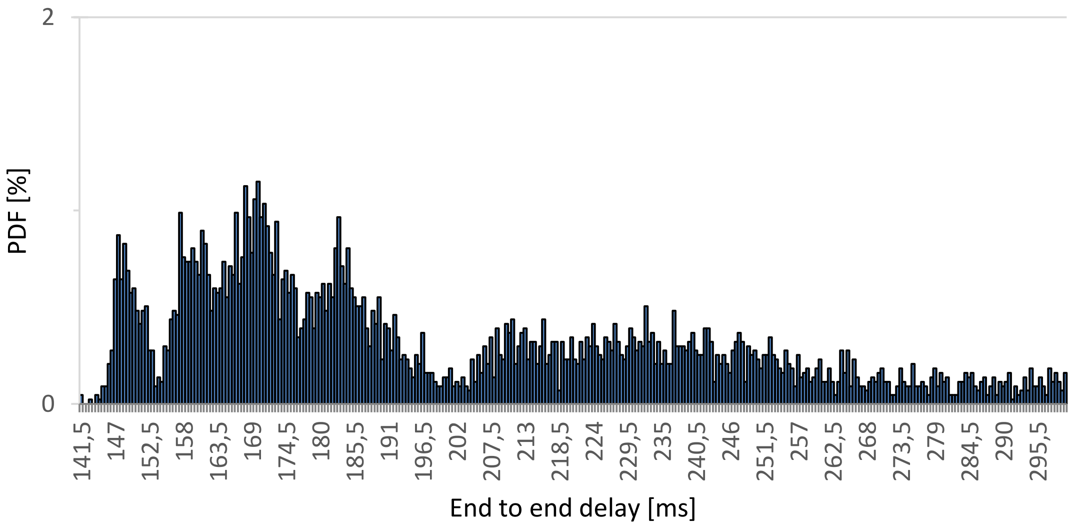


Figure 4.27. PDF of the end to end delay for LTE, mean call duration of 200 s, 400 users

Table 4.15. PDF values for LTE, mean call duration of 200s, 400 users.

	Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
Original	9472.39	10231.90	102	39131.28	18447
Core	198.94	39.88	141.05	299.91	4348

Regarding the PDF of Figure 4.25, the standard deviation is smaller compared to the PDF one of Figure 4.24, due to the percentage of samples inside or above the acceptable values; however, it has a wider range of core sampling, from the fact that the network has users connected in a larger period, leading to less oscillations on the end to end delay, because there is a lower number of calls being started, but also to a higher number of samples with a higher end to end delay, because the network has a higher load.

The PDF for the mean call duration of 50 s and 100 s, for 500 users, was also assessed by the Curve Fitting Tool as well as the Distribution Fitting Tool from Matlab, and compared the core sampling and the original one, the results being in Annex A.3. Figure 4.28 presents the difference in between the average end to end delay for the core sampling and the original one for a mean call duration of 50 s, 100 s, 150 s, and 200 s and for 400 users. Contrary to the similar analysis performed for UMTS, in the LTE one can see that the difference in between the original sampling and the core one is increasing with the increment of the mean call duration.

Table 4.16. Curve fitting values for LTE.

Mean call duration [s]	R^2	μ [ms]	σ [ms]
50	0.8949	117.7	0.396

In the UMTS analysis, it was concluded that with the increment of the mean call duration, calls are longer, reducing the number of calls being started, resulting in less variation in the network. In LTE, this fact is observed in Figure 4.20, analysing the standard deviation, which is decreasing with the increment of the mean call duration. However, the maximum value adopted to perform the PDF in UMTS tends to be constant, due to the concentration of similar samples, while in LTE, the maximum value for the core sampling is increasing, due to its lower concentration. This fact contributes to have an increment of the average end to end delay with the increment of the mean call duration. For the mean call duration of 50 s, the difference in between the original sampling is 28.70 ms, for the mean call duration of 100 s, the difference is 16.24 ms, in both simulations the average end to end delay has not yet reached the threshold, for the mean call duration of 150 s, the difference is 72.19 s, and for the mean call duration of 200 s, the difference is 9.27 s.

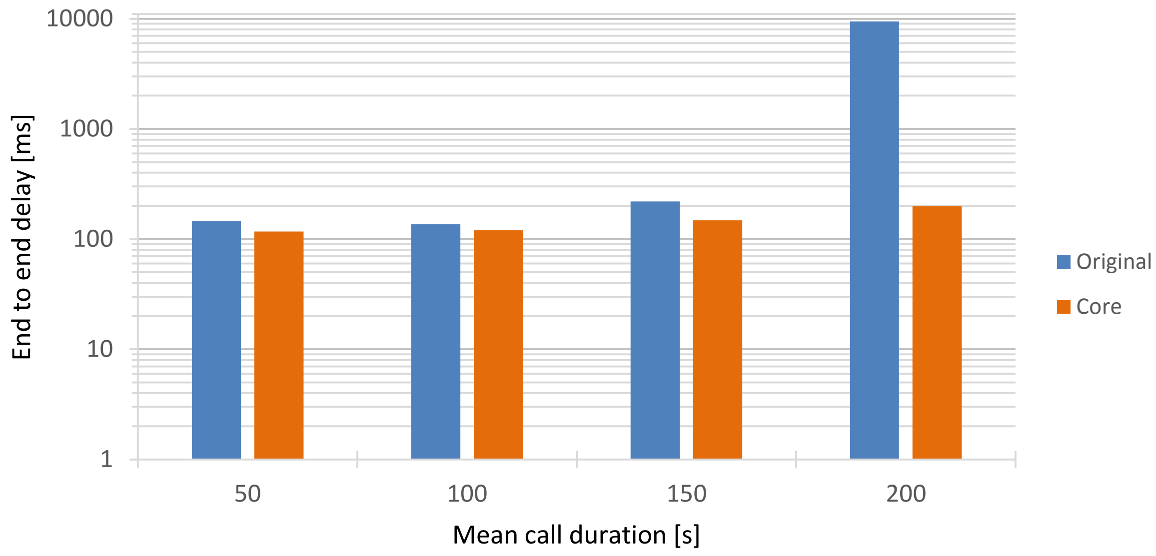


Figure 4.28. Comparison between original sampling and core sampling, LTE, 400 users.

4.4 Data Analysis on LTE

In the second study for the LTE network, a scenario with an eNodeB that can perform data and voice ones, and two different configurations was set. The first configuration has less traffic demand than the second one. For the first configuration, the number of users performing voice changed from 50 until 250, a simulation was done with 300 voice users, however, the collected data from this simulation were not trust worthy, since the number of users that can perform the services was significantly reduced, and the average values tend to decrease comparing to the simulations with less users. As in the voice analysis, the average values for each service are computed using the same approach, using (3.10).

The number of users performing voice with a mean call duration of 150 s, decreases from 360 to less than 300 when in presence of data services with the parameters of the first configuration. When in presence of data users, the average end to end delay for voice users never reaches the defined threshold, within the range of users stated before (from 50 until 250), however, this threshold is reached in the simulation with 300 users. This fact proves that the maximum number of users performing voice in presence of data, with the first configuration is between 250 and 300 users, thus, data services are analysed in between the minimum number of voice users, which is 50, until 250, once the data collected from the simulation of 300 present untrustworthy values.

Figure 4.29 presents the packet end to end delay for the Video Conference, for the first configuration. The packet end to end delay increases with the increment of the number of users performing voice services. The average value for the simulation with no voice users has values far from reality, around 35 ms, which are very low for this type of service. This service reaches very high average values with the increment of voice users, which can be explained by the high demand of traffic that this service has.

In order to evaluate the behaviour of the average value of the packet end to end delay for the Video conference, the Curve Fitting Tool from Matlab was used, to understand if it is possible to define a model for his behaviour. The fitting curves are all presented in Annex A.4. The packet end to end delay is represented by an exponential, with R^2=0.9685:

$$\tau_{ETE} = 0.5011 \cdot e^{0.01522 \cdot N_u} \tag{4.3}$$

Figure 4.30 presents the throughput for the same service, showing a great decrease with the increment of voice users. From Figure 4.29 and Figure 4.30, one can understand that video conference is highly affected with the increment of voice users, due to the priority of voice, even more than the remaining data services, which demand a lower throughput than video conference.

The throughput for Video Conference is decreasing, however, it is not decreasing as fast as a linear equation, the model that represents this decrease being a second degree polynomial, with R^2=0.9695:

$$Th = -19.603 \cdot {N_u}^2 + 1447.5 \cdot N_u + 1 \cdot 10^6 \tag{4.4}$$

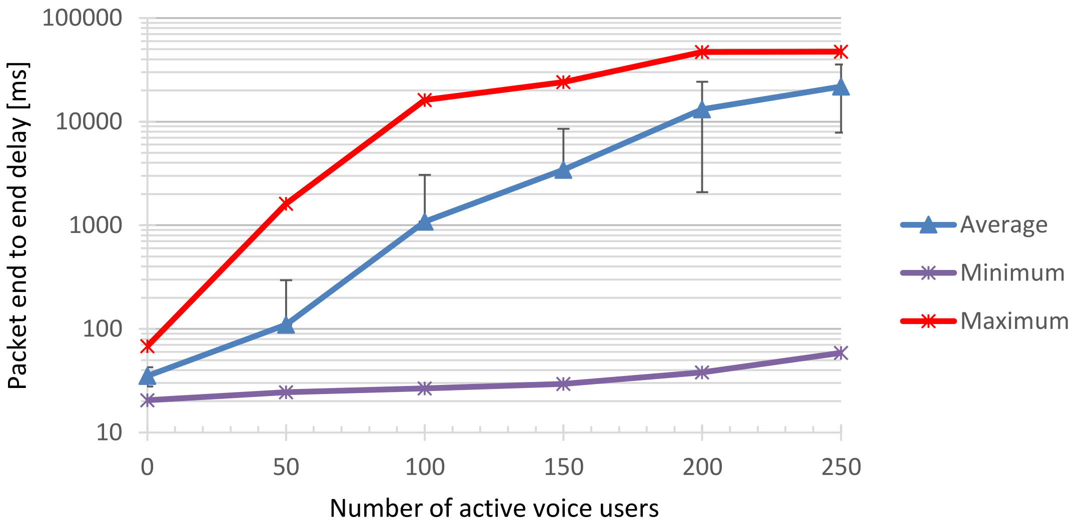


Figure 4.29. Packet end to end delay for the 1st configuration of the Video Conference service.

Taking into account the collected data from throughput, the maximum value for the first configuration of Video Conference service is 2.66 Mbit/s, which is below the maximum allowed by the network, however, using (3.4), with the data size of Video Conferencing, which is 17.28 kbytes, and the minimum packet end to end delay, which is 20.55 ms, from the users' perspective, the maximum bit rate is 6.93 Mbit/s. This maximum bit rate is below the maximum allowed by the network (8.4 Mbit/s), however, from the users' perspective this bit rate does not take into account the signalling traffic, increasing the QoE.

Figure 4.31 presents the download response time for the FTP service. The download response time, increases slightly with the increment of the number of voice users, however, FTP does not have a great impact on delays, the average value being around 470 ms, and the maximum value for the download response time being around 2.9 s when there are 250 voice users, which is an acceptable delay for the download of a 50 kbytes file.

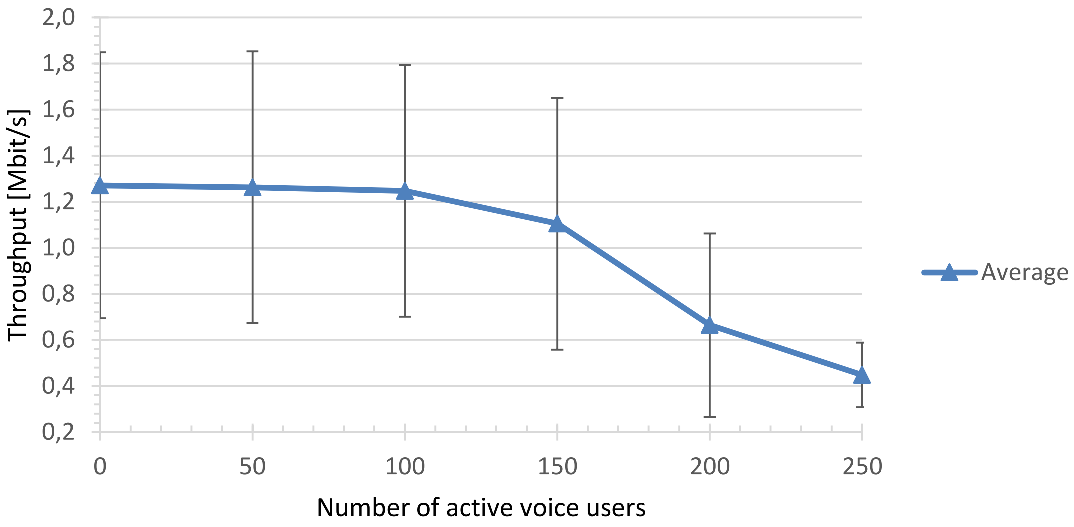


Figure 4.30. Throughput for the 1st configuration of the Video Conference service.

The download response time for FTP is increasing slightly, however, it is far from an exponential equation, the better model being a second degree polynomial, with R^2=0.9656:

$$\tau_{DL} = 7 \cdot 10^{-6} \cdot {N_u}^2 - 0.001 \cdot N_u + 0.2893 \qquad (4.5)$$

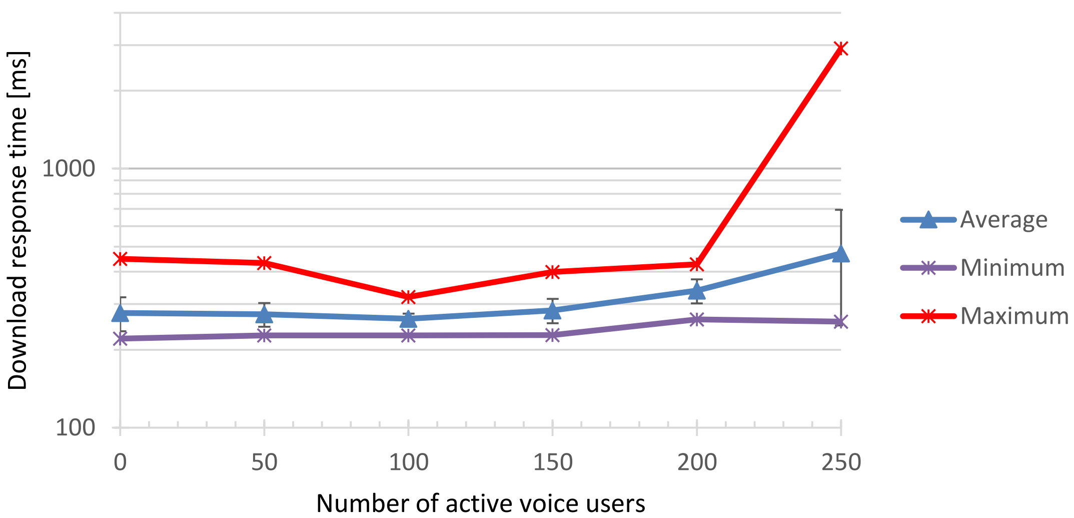


Figure 4.31. Download response time for the 1st configuration of the FTP service.

Figure 4.32 presents the upload response time for FTP. As expected, the UL response time has an average delay higher than the DL one, due to the higher capacity available in DL compared to UL. Like the DL response time, the UL one has a slightly increased with the increment of voice users. The average values are the expected ones for this type of service; however, it was expected that the maximum value would be always higher for UL compared with DL, but the maximum for UL is 631 ms, which is higher than all maximum ones in DL, except in the case of 250 voice users.

Figure 4.33 presents the throughput for the FTP service; the throughput has oscillations, but, it is not decreasing as expected. Neither the UL response time, Figure 4.32, neither the throughput, Figure 4.33, have a behaviour that can be represented by a simple model.

The maximum value of the throughput for the first configuration of the FTP service is 42.94 kbit/s, which is obviously below the maximum allowed by the network. Once again the maximum bit rate from the users' perspective is higher than the throughput collected; taking into account that the size of the data is 50 kbytes, and the minimum DL response time is 220.61 ms, from the users' perspective the maximum bit rate is 1.81 Mbit/s. For UL, the maximum bit rate observed by the user is 1.66 Mbit/s; as expected it is below the maximum bit rate for DL.

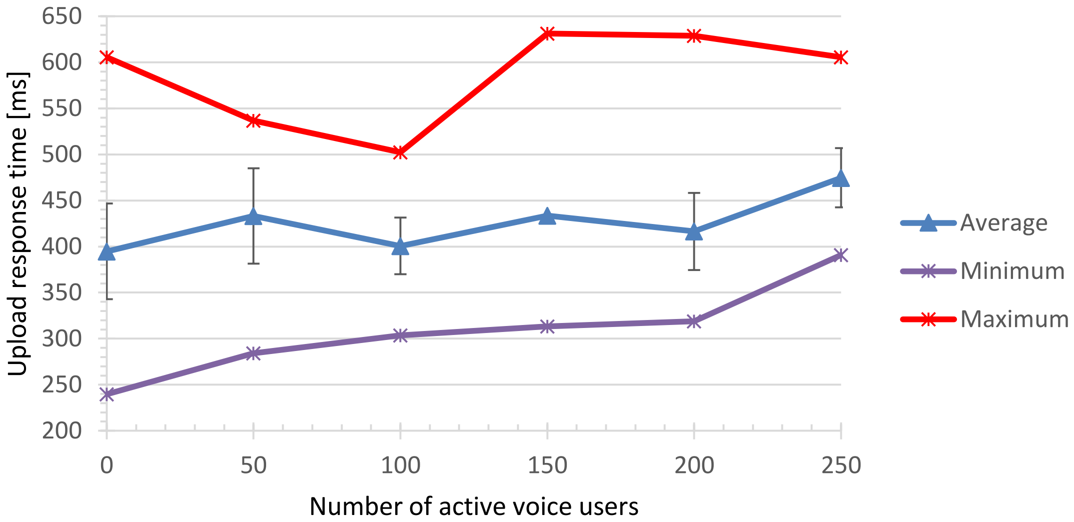


Figure 4.32. Upload response time for the 1st configuration of the FTP service.

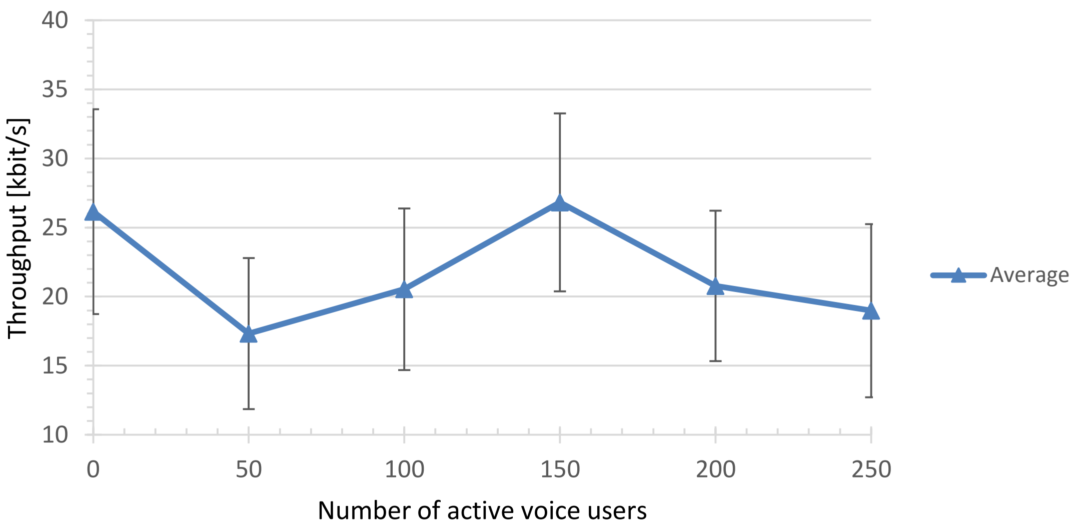


Figure 4.33. Throughput for the 1st configuration of the FTP service.

Figure 4.34 presents the DL response time for the Email service, and Figure 4.35 the UL one. Both averages have the expected behaviour, increasing with the increment of the number of voice users.

Both minimum and maximum values have the expected behaviour as well, and maxima for DL are always higher than the ones for UL. The average has acceptable values for DL and UL for a file with 2 kbytes for this type of service. Figure 4.36 presents the throughput for the Email service. Once more, the throughput does not have a great impact with the increment of the number of users.

The DL response time for Email has a behaviour similar to the one for FTP, which has a slightly increase, the better model for this behaviour being a second degree polynomial, with R^2=0.9510:

$$\tau_{DL} = 8 \cdot 10^{-7} \cdot {N_u}^2 - 7 \cdot 10^{-5} \cdot N_u + 0.1102 \tag{4.6}$$

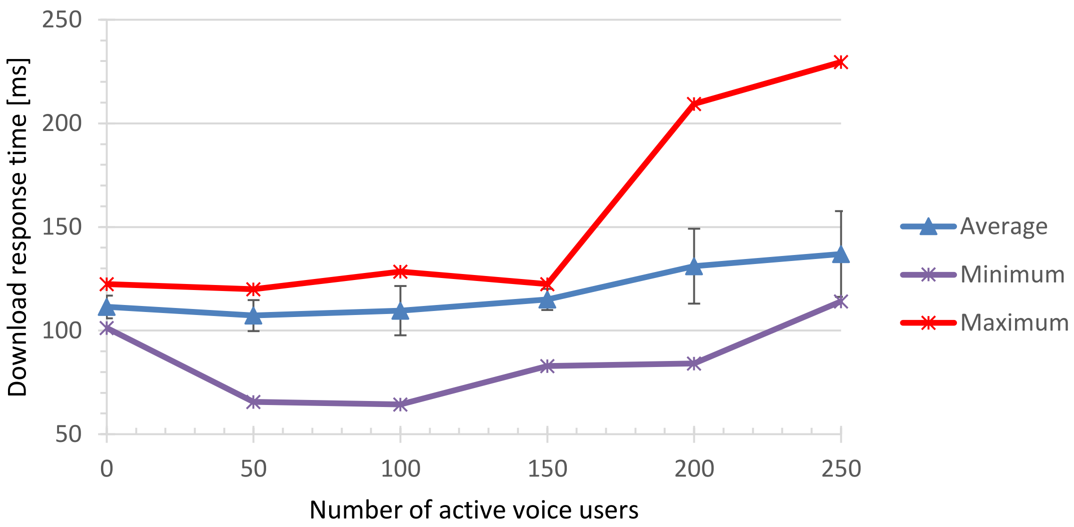


Figure 4.34. Download response time for the 1st configuration of the Email service.

The UL response time for Email has a slight increase, as the DL one, the better model being again a second degree polynomial, with R^2=0.7792:

$$\tau_{UL} = 1.019 \cdot 10^{-6} \cdot {N_u}^2 - 0.00012 \cdot N_u + 0.2475 \tag{4.7}$$

The maximum value of throughput for the first configuration of the Email service is 10.3 kbit/s, and the maximum bit rate from the users' perspective is 248.49 kbit/s for DL, and 111.64 kbit/s for UL. Taking into account that the size of the data is 2 kbytes, the minimum DL response time is 64.39 ms, and the UL is 143.32 ms; both values are below the maximum allowed in the network. The behaviour of throughput does not have a simple model that can represent its behaviour with a good value for the goodness of fit.

Figure 4.37 presents the page response time for the HTTP service. The average has almost the same value for the different number of active voice users. The increment of users has no impact on the HTTP service, except on the maximum value that has a slightly increment before stabilising. Figure 4.38 presents the throughput for the HTTP service, which is not affected by the increase of active voice users.

The maximum value of throughput for the first configuration of the HTTP service is 18.5 kbit/s, and the maximum bit rate from the users' perspective is 367.48 kbit/s. Taking into account that the maximum size of the data is 2 kbytes, and the minimum page response time is 43.54 ms, both values are below the maximum allowed in the network. Neither the page response time neither the throughput can be represented by a simple equation.

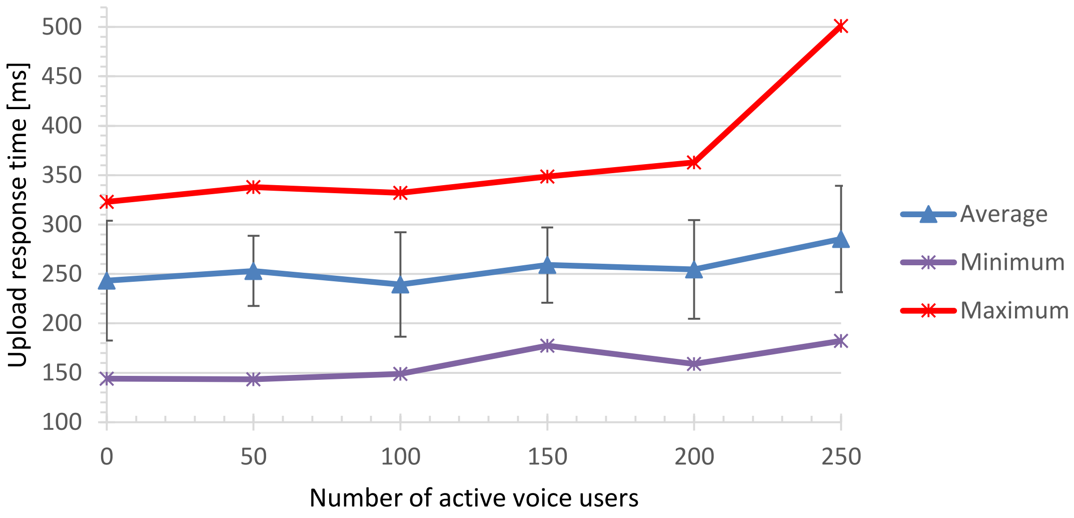


Figure 4.35. Upload response time for the 1st configuration of the Email service.

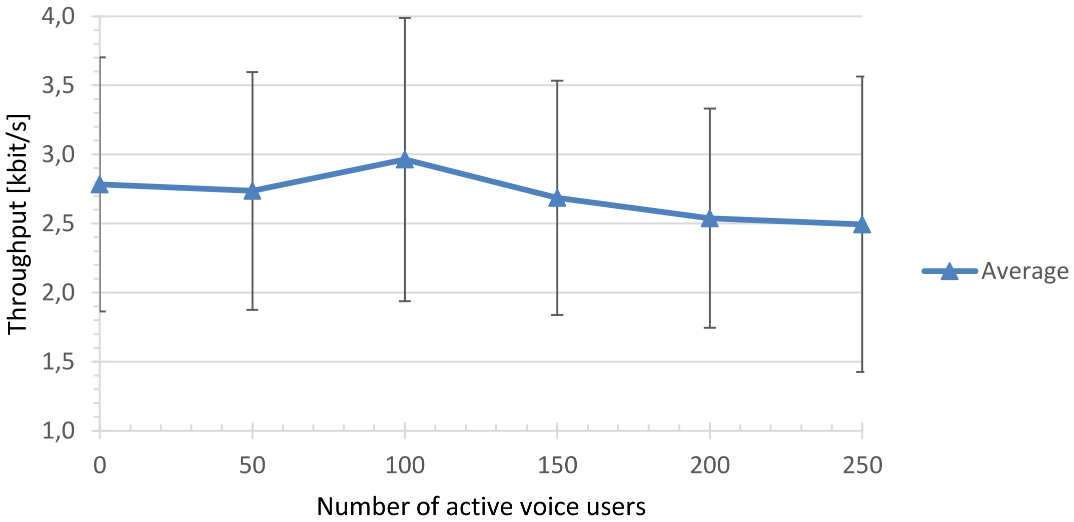


Figure 4.36. Throughput for the 1st configuration of the Email service.

In the second configuration, the number of active users performing voice changed from 50 until 200; a simulation was done with 250 voice users, however the collected data is not trust worthy. The average end to end delay for the voice users in the presence of data, with the second configuration, reached the defined threshold in the simulation with 150 active voice users. This fact shows that with a higher traffic demand for data, the number of voice users tends to decrease, even with priority over the data services, because the network has a higher load.

Since the shapes for the second configuration of the data services are very similar to the ones presented before, these figures are presented in Annex A.4, as well as the equations that represent the behaviour.

The packet end to end delay for Video Conference in the second configuration has a behaviour similar to the first one, so it is still modelled by an exponential. The throughput for this service is still represented by a second degree polynomial equation, due to the slight decrease.

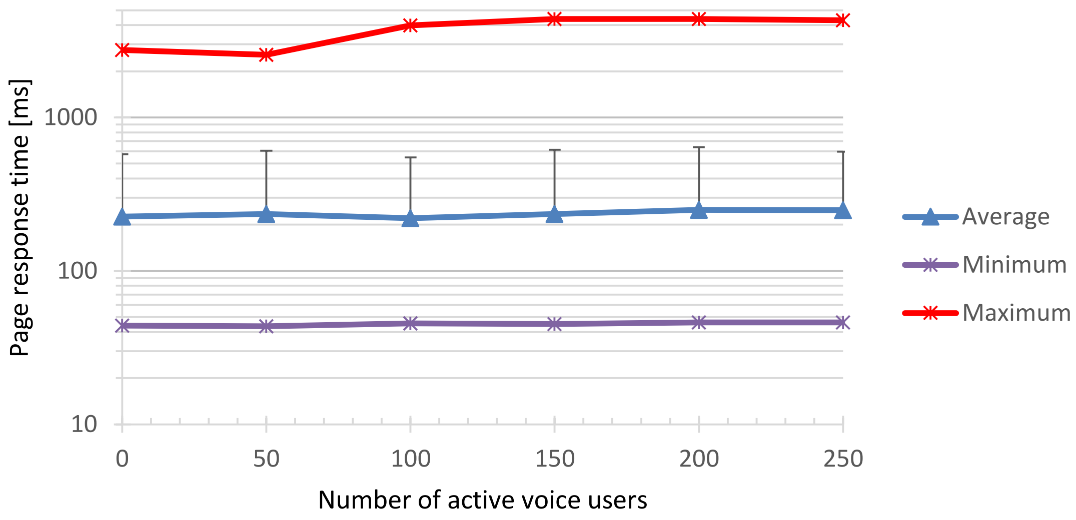


Figure 4.37. Page response time for the 1st configuration of the HTTP service.

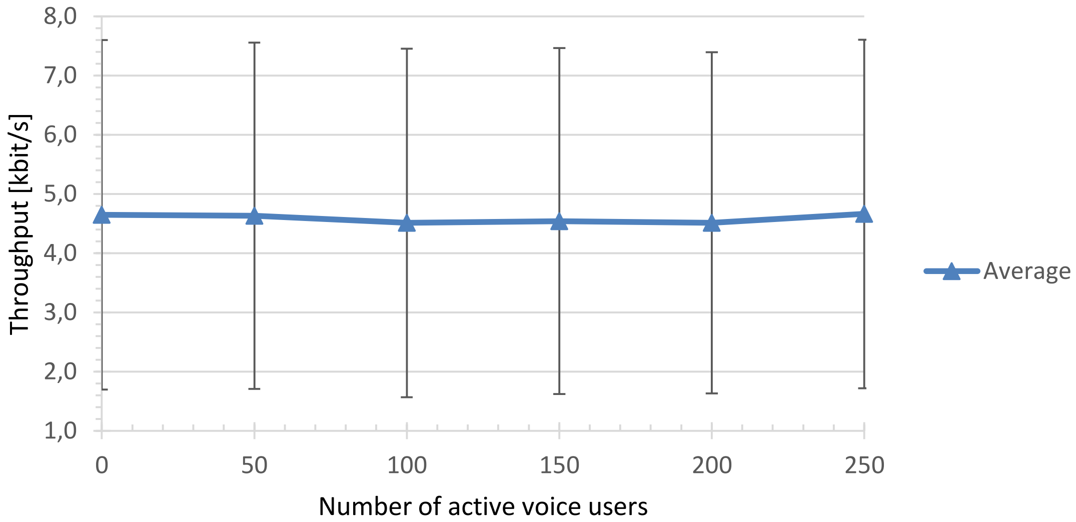


Figure 4.38. Throughput for the 1st configuration of the HTTP service.

The DL response time, as well as the UL one, for the second configuration of the FTP service, are represented by a second degree polynomial equation, due to their slightly increase. The DL response time of the Email service for the second configuration is represented, once again by a second degree polynomial equation. The throughput for the FTP, Email, HTTP services, and the page response time for the HTTP service, cannot be represented by an equation with a good value for the goodness of fit.

Figure A.36 presents the packet end to end delay for Video Conference, for the second configuration, which increases with the increment of the active voice users. This service is highly affected by the increment of voice users, as well as by the increment of the traffic demand from the service itself, and from the other services; even with no voice users, it is not possible to perform a Video Conference due to the high packet end to end delay. Figure A.38 presents the throughput for Video Conference, which is very affected by the increment of active voice users, similarly to the first configuration.

Taking into account the collected data from the throughput, the maximum value for the second configuration of the Video Conference service is 3.41 Mbit/s, which is still below the maximum allowed by the network, even with the increase of traffic demand. From the users' perspective, the maximum bit rate achieved is 8.48 Mbit/s; taking into account that in the second configuration the size of the data is 22.4 kbytes, and the minimum packet end to end delay is 21.13 ms, this maximum bit rate is above the maximum allowed by the network. From the users' perspective, this bit rate does not allow the network to have a decent connection for this type of service, with this traffic demand, when in presence of the other data services and the voice service; this is reflected in the average packet end to end delay, which is around 2.45 s in the simulation in which there are no voice users, which means that in the presence of voice, this average packet end to end delay will be higher, Figure A.36.

Figure A.40 presents the DL response time for FTP, for the second configuration. As expected all values increase compared to the first configuration for data services. Both values tend to increase with the increment of active voice users. The second configuration has an impact on the delays bigger than the first one; however, the delays for this service are acceptable for a DL of a file with 5 Mbytes. Figure A.42 presents the UL response time of the FTP service, for the same configuration; as expected, it has higher values than in the first configuration. It is expected that UL values are always higher than DL ones, however, this fact does not occur, which can be explained by the existence of HTTP services, which only have significant traffic in DL. The UL response time has very similar values to the DL one.

Figure A.44 presents the throughput for FTP, which, once again, has oscillations; however, this does not presents a trend, as expected. It was expected that the average throughput has a significant decrease with the increment of traffic demand, however, this fact does not occur. The maximum value of throughput for the second configuration of FTP is 4.14 Mbit/s, which is, once again, below the maximum allowed by the network. The maximum bit rate from the users' perspective is 4.91 Mbit/s, taking into account that the size of the data is 5 Mbytes, and the minimum DL response time is 8.14 s. Regarding the UL response time, the maximum bit rate observed by the user is also 4.91 Mbit/s, taking into account the minimum value for the UL response time, which is 8.15 s.

Figure A.45 and Figure A.47 present the DL and UL response times for the second configuration of the Email service, respectively. As expected, the average values are higher compared to the first configuration. The response times tend to increase with the increment of active voice users. The UL value continues to be higher than the DL one. The maximum average values for DL and UL are in accordance to the expected values for this type of service, when receiving or sending files with 10 kbytes.

Figure A.48 presents the throughput for the second configuration of the Email service. The throughput has higher values than in the first configuration, due to the higher traffic demand; however, the behaviour is very similar to the previous configuration, with oscillations, but nevertheless, it does not show a declining trend, as expected. The maximum value of throughput for the second configuration is 25.67

kbit/s, which is once again below the maximum allowed by the network. The maximum bit rate from the users' perspective is 684.58 kbit/s, taking into account that the size of the data is 10 kbytes, and the minimum DL response time is 116.86 ms. Regarding UL, the maximum bit rate observed by the user is 402.88 kbit/s, taking into account the minimum value for the UL response time is 198.57 ms.

Figure A.49 presents the page response time for the second configuration of the HTTP service, which has higher values than in the first configuration, due to the higher traffic demand; however, the behaviour is very similar to the previous configuration of this service, and there is an almost constant page response time for any number of active voice users.

Figure A.50 presents the throughput for this service, which is not affected due to the increment of active voice users, concluding that this is the least affected service. The maximum value of the throughput for the second configuration of the HTTP service is 45.47 kbit/s, and the maximum bit rate from the users' perspective is 559.91 kbit/s, taking into account that the maximum size of the data is 5 kbytes, and the minimum page response time is 71.44 ms. Both values are below the maximum allowed in the network.

In all services, the general response times increases due to the increment of traffic demand, as well as, the maximum throughput and the maximum bit rate from the users' perspective.

Chapter 5

Conclusions

This chapter outlines the main conclusions of this master thesis. It reviews the thesis goals, highlights the relevant results, and ends with suggestions for future work.

The main goals of this thesis was the compassion of the voice performance over PS in UMTS, also known as VoIP, and the voice performance over PS in LTE, known as VoLTE, with the well-known performance of voice over CS in GSM/UMTS networks, and the study of the impact that the voice service has on data ones in LTE. Both studies were performed in a similar scenario, which is an area of 200×200 m^2. The approach used in order to analyse voice performance was the increment of the number of users, with different mean call durations. The approach used in order to analyse the impact of the voice service on data ones was the increment of active voice users, with a fix number of users performing data services, and different traffic demanding.

In Chapter 1, one gives a brief historical overview of the evolution of voice services, from the very beginning, in which voice was performed by CS, to the present day where the voice service can have a very good performances on PS. The Chapter 1 presents the motivation for this studies and the contents of the thesis.

In Chapter 2, one presents the main relevant characteristic for each system regarding their network architecture differences and radio access technologies. Furthermore, this chapter also presents the main characteristics of the voice service over CS and PS, and their performance parameters.

The third chapter presents the theoretical models used to evaluate the collected data from the simulator and the network limitations, the description of the used simulator, the description of the parameters settings for the scenarios creation, and the simulator assessment. The description of the simulator is divided into three important sub-chapters: on how to implement a network, and which editors of the simulator are involved on the creation of a simulation; on how generate traffic in order to test the given network; and how to select the desired metrics to evaluate the network.

The simulator is divided into three main levels: the process domain and the associated process model editor; the node domain and the associated node editor; and the network domain and the associated project editor. The process domain is used to specify the behaviour of the models that exist in the node domain, the process models are driven by events and interrupts, the events and interrupts are developed by a combination of FSMs, libraries of kernel procedure, and C/C++ programming language, known as Proto-C. The node domain is responsible for defining the behaviour of each object, the behaviour is defined using different modules and connections among them. The network domain is responsible for graphically represent a topology of a communication network. The network is composed of node and link objects, which are instances of the previous levels presented before.

The traffic can be generated by two different ways: manually, by setting attributes from the applications, or by importing traffic from external files or programs. In order to generate traffic, the manually traffic generation was used, by setting certain applications, such as, voice, video, FTP, email, and HTTP.

Regarding the metrics, these are divided into four types: Global Statistics; Node Statistics; Module Statistics; and Link Statistics. The selection of statistics is performed by specifying a list of probes, which perform the selection and control the flow of data from the selected statistics.

The Chapter 3 also provides the description of the implementation of the networks in the simulator. In this particular chapter, one describes the nodes and settings used to implement a UMTS network and a

VoIP service, as well as for an LTE network and the VoLTE service, and the remaining data ones, which are video conference, FTP, email, and HTTP. The configuration of LTE bearers, and the priorities of the services are set in this sub-chapter, as well as the bandwidths, the antennas and users, the gains, the powers, and the propagation model selected in the simulator.

From the assessment, one concludes that collecting a sample every 10 s is enough due to the associated error, which is less than 2.5% in LTE and 1% in UMTS, compared to collecting one sample every second. Since the simulator is not stabilised from the very beginning of each simulation, one concluded that the first 10 minutes do not represent trust worthy values, hence, it is necessary to discard the first 10 minutes of each simulation. One concluded also that 20 minutes of simulation is enough to have trust worthy values, since compared with simulations with 60 minutes, the associated error is less than 0.8% in LTE and 0.2% in UMTS, so the duration of each simulation is 30 minutes, and the first 10 minutes were discarded. The number of seeds for each UMTS simulation is 5, in order to have a realistic scenario, but for LTE it is 1, due to the duration of each simulation, which can reach tens of hours.

Chapter 4 starts by presenting a description of the scenarios used in this thesis, followed by the mean call duration and the number of voice users used in both systems for the first study, in which exists only the voice service, followed by the data sizes, and number of voice users used in the second study, in which exist the voice service and data ones in an LTE network.

From the VoIP service in UMTS, one concludes that the number of voice users that can perform a VoIP call, tends to reduce when the mean call duration increases, which is the expect because the number of users using the available resources are higher and more constant; however, it does not have a strong trend that can be described by a function. This behaviour could come from the fact that only have been tested 3 different mean call durations. One can conclude that the standard deviation tends to be smaller when the mean call duration is higher, due to the fact that is a lower number of calls being started when the mean call duration is higher, because users are connected to the network for a longer duration, contributing to a more stable network than with a smaller mean call duration. When a communication is established the end to end delay is higher, and stabilises; if there is a low number of call being started, the end to end delay samples are stable during a longer period.

MOS for the VoIP connections is always in between the 3.5 and 4, which means that some users are dissatisfied; however, this metric does not present significant oscillations due to the increment of the mean call duration. In order to evaluated the dispersion of the end to end delay samples, PDFs were performed for each mean call duration, more precisely, 5 PDFs for each mean call duration, corresponding to each seed, for the number of active user of 40 and 50. The average end to end delay is centred in the 160.5 ms in the major part of the PDFs, and the PDFs are almost all represented by a limited Gaussian Distribution.

Regarding VoLTE, one concludes that the number of users tends to decrease with the increment of the mean call duration, as verified in UMTS. Contrary to UMTS, the trend of the number of users as a function of the mean call duration can be represented by a rational function, which tends to saturate as the mean call duration increases, reducing the number of users allowed in the network before reaching the defined threshold (200 ms); however, when the mean call duration is lower, the rational function tends to have a strong growth, which is not the expected behaviour; nevertheless, this function has a more realistic behaviour than a linear equation. Once again the standard deviation tends to decrease when the mean call duration increases, due to the fact that the users are connected to the network for a longer duration, reducing the number of call being started.

MOS for the VoLTE has a value higher than 4.2, which means that the users are satisfied, but this value tends to decrease when the number of users is increasing, as expected. The PDFs performed for LTE were very affected by the increment of the mean call duration, the lower and upper limits of the PDFs increasing. It is not easy to find distributions that fit the PDFs' behaviours; in fact, it was just found a distribution that fits the PDF with 50 s and 400 users, and two distributions which fit the PDFs with 50 s and 100 s, for 400 users. The increment of the mean call duration imposes a larger range for the PDFs, due to the fact that the samples are more scattered with a higher mean call duration. One can conclude that with the increment of the mean call duration, the percentage of the samples above the threshold (200 ms) tends to increase.

Regarding the analysis of the impact of the voice service on data ones, one can conclude that the number of users that can perform a VoLTE call reduced, as expected; even with the priority of the voice service over data ones, there is a lower number of RBs available for the voice service. One concludes also that the video conference is the data service with the higher impact due to the priority of the voice over the data services, the behaviour of the packet end to end delay is represented by an exponential equation as a function of the number of active voice users. The throughput of this service is the most affected, its decrease is represented by a second degree polynomial equation. The FTP is the second most affected service by the presence of the voice one, the increments of the download response time and the upload response time are represented by a second degree polynomial equation, the email service has a similar behaviour, however, it was not as affected as the FTP. The HTTP service is the least affected, the page response time and the throughput do not have a significant trend, which cannot be represented by an equation. The increment of the file sizes for the data services, reinforced the previous conclusions. However, the increment of the HTTP traffic demand, is not affected once again. The LTE scenarios were tested with a 5 MHz bandwidth, however, it is easy to extrapolate the conclusions, and the maximum number of users to a 10 or 20 MHz bandwidth.

For the future work, it could be considered to implement a VoLTE system that contains an IMS Model. The IMS Model provides connections that allows communications in between VoLTE and non-VoLTE devices, this feature is interesting to evaluate the performance of the VoLTE system, when it has to communicate with a non-VoLTE device. Another suggestion for a future study, is the evaluation of the performance of the VoLTE service with different audio codecs, and with silence suppression.

Annex A

Additional Simulator's Assessment Results and Simulation Results

This annex presents additional assessment result for UMTS, as well as, the additional results obtained in the simulations for UMTS and LTE.

A.1 Assessment Result for UMTS

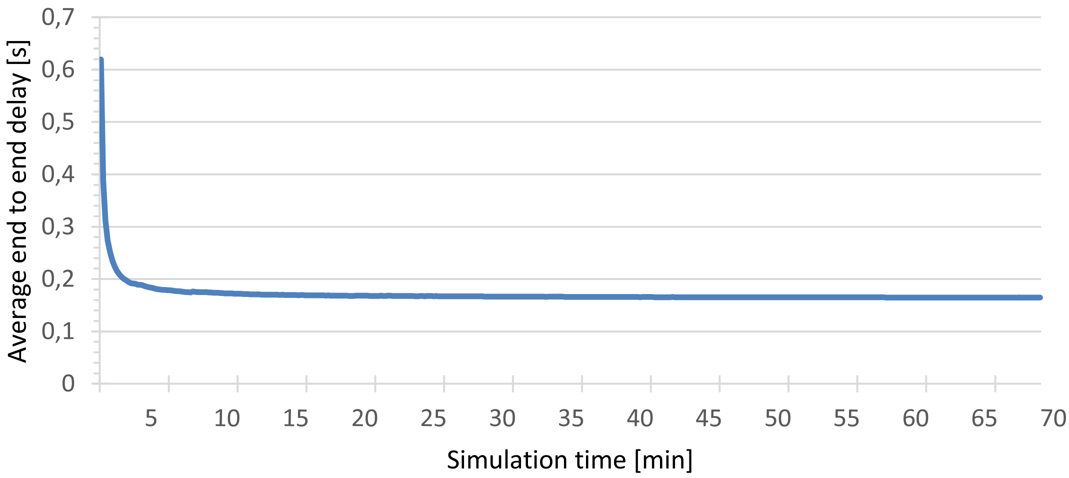


Figure A.1. UMTS assessment – progress of the average end to end delay in a 70 minutes simulation.

A.2 Additional UMTS Results

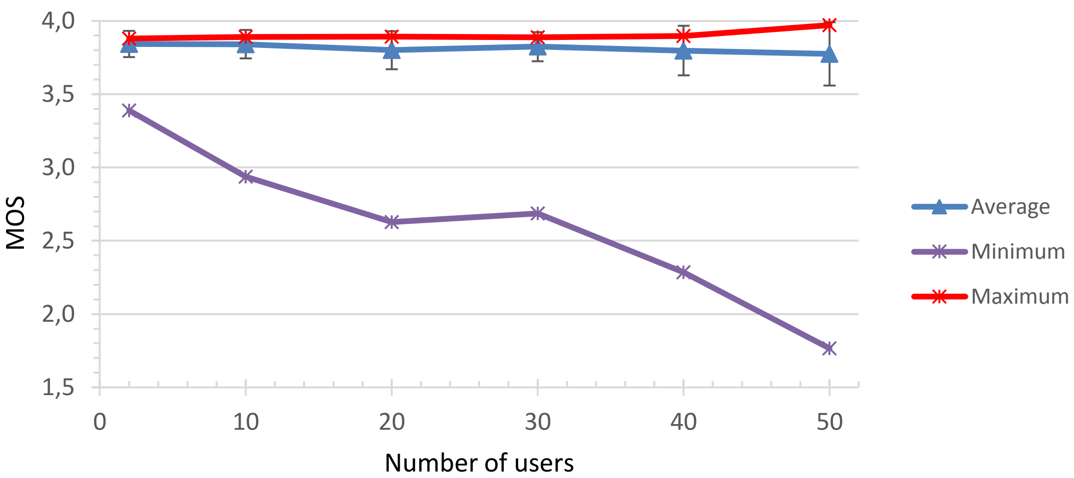


Figure A.2. Global MOS for UMTS with a mean call duration of 100 seconds.

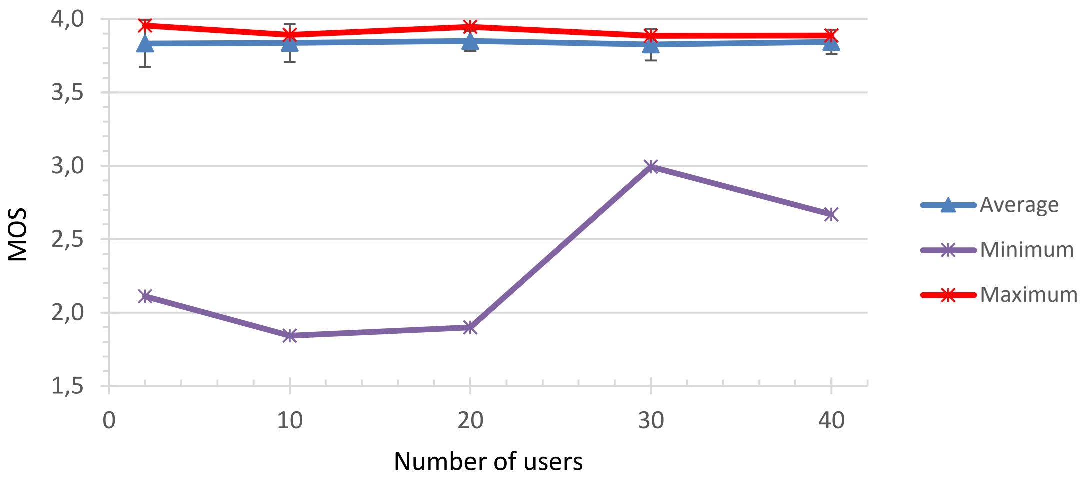


Figure A.3. Global MOS for UMTS with a mean call duration of 150 seconds.

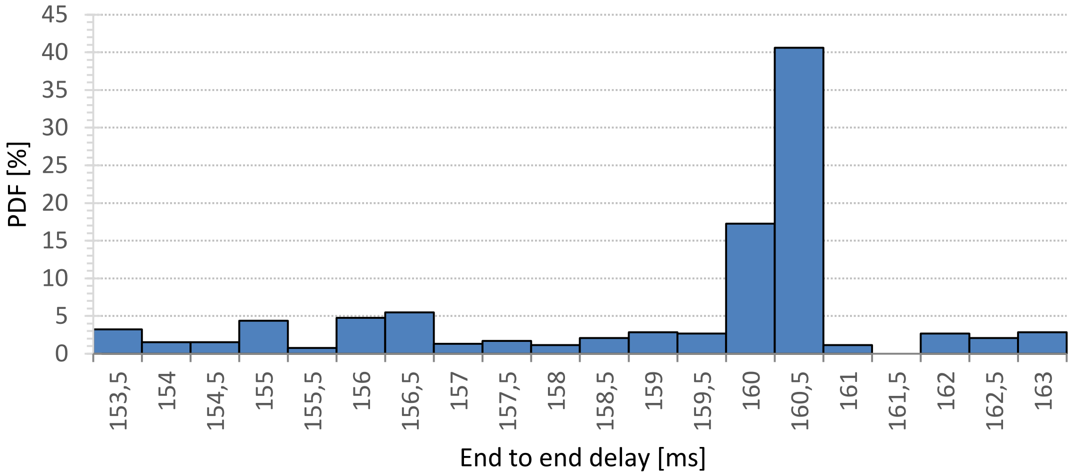


Figure A.4. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 2).

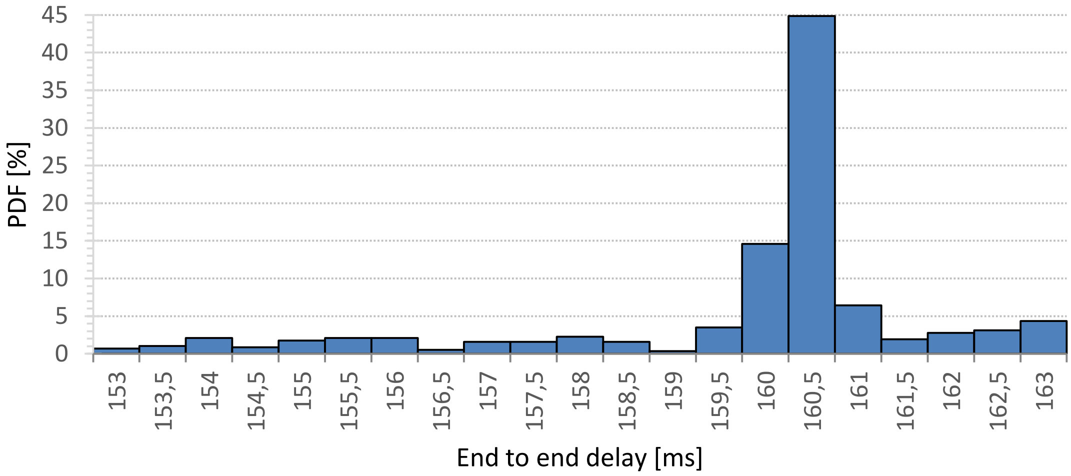


Figure A.5. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 3).

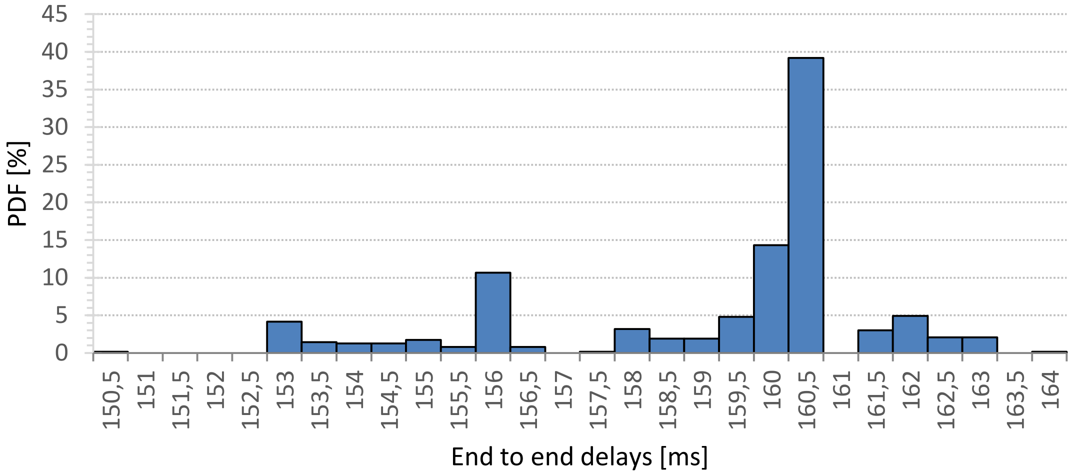


Figure A.6. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 4).

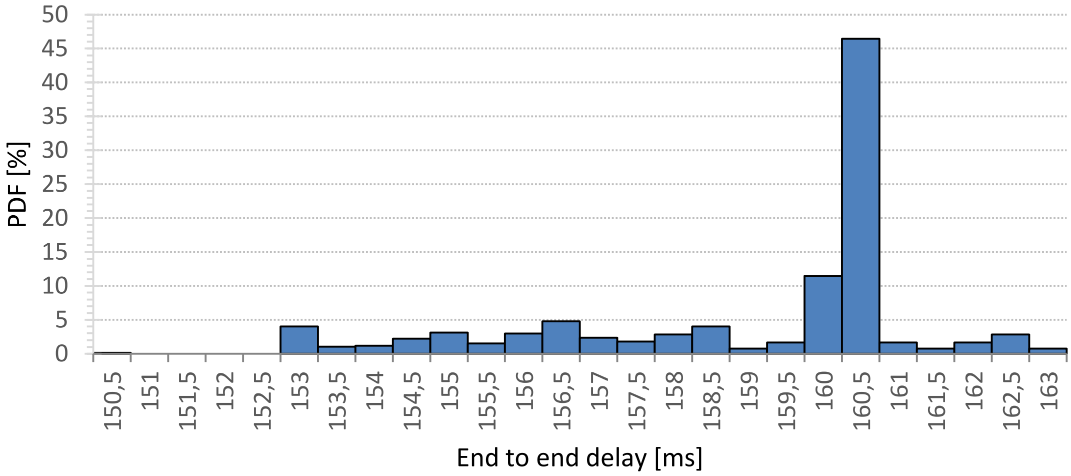


Figure A.7. PDF of the end to end delay for UMTS, mean call duration of 50 s, 40 users (seed 5).

Table A.1. Remaining PDF values for UMTS, mean call duration of 50 s, 40 users.

Seed		Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
2	Original	226.45	698.27	153.06	15896.83	597
	Core	158.89	2.25	153.06	162.71	527
3	Original	208.51	612.05	152.75	15320.56	633
	Core	159.53	2.04	152.75	162.67	575
4	Original	269.43	1296.37	150.01	24165.45	685
	Core	158.89	2.46	150.01	163.63	628
5	Original	240.04	1144.66	140.01	30488.52	745
	Core	158.70	2.46	150.03	162.56	671

Table A.2. Remaining discarded values for UMTS, mean call duration of 50 s, 40 users.

Seed	Above 163 ms [%]	In between 163 ms and 200 ms [%]	Above 200 ms [%]
2	11.73	1.17	10.55
3	9.16	1.11	8.06
4	8.32	1.46	6.86
5	9.80	1.34	8.46

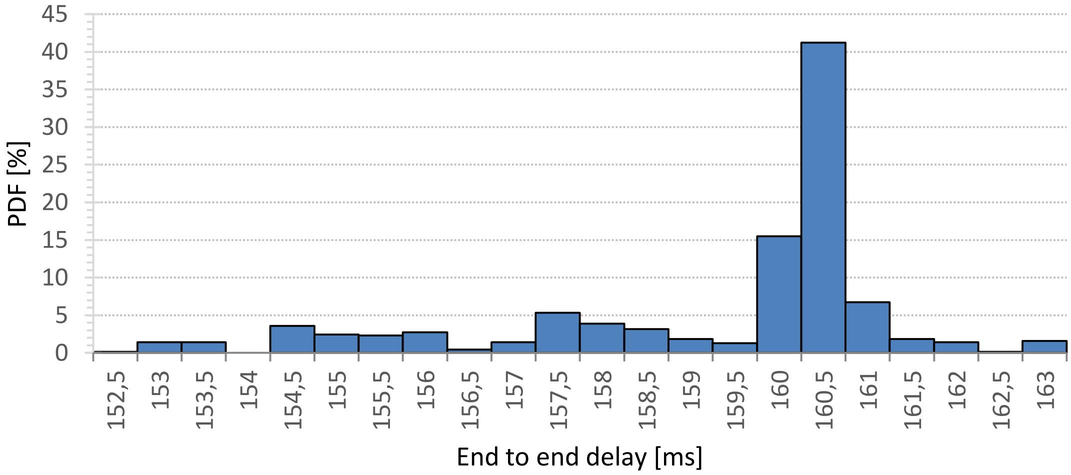


Figure A.8. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 2).

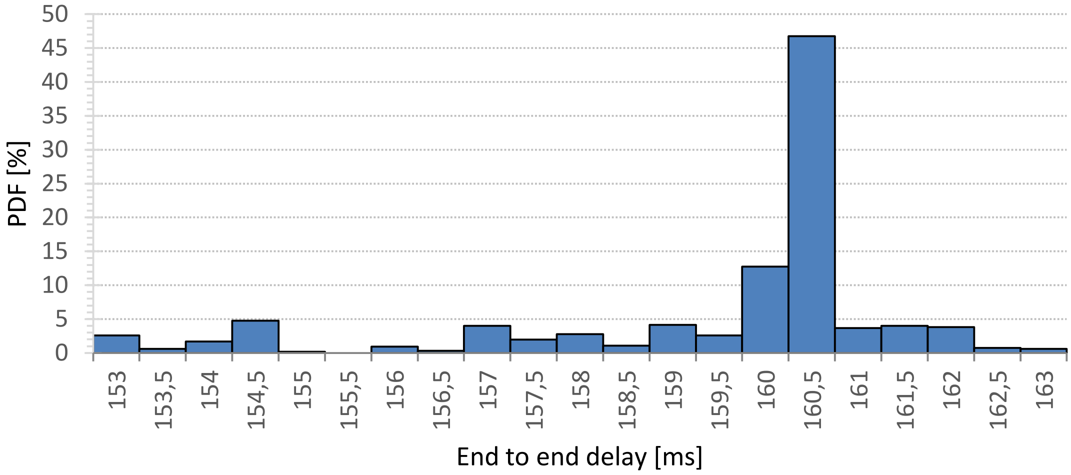


Figure A.9. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 3).

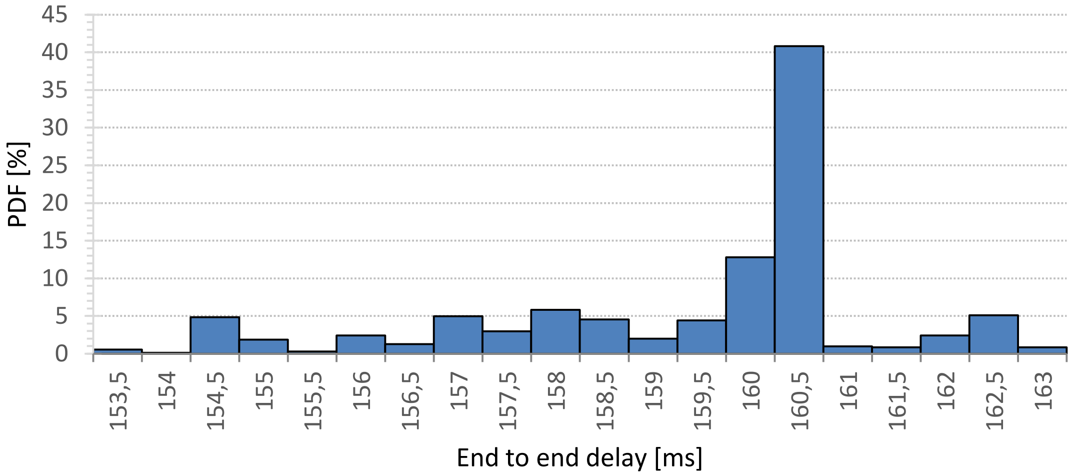


Figure A.10. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 4).

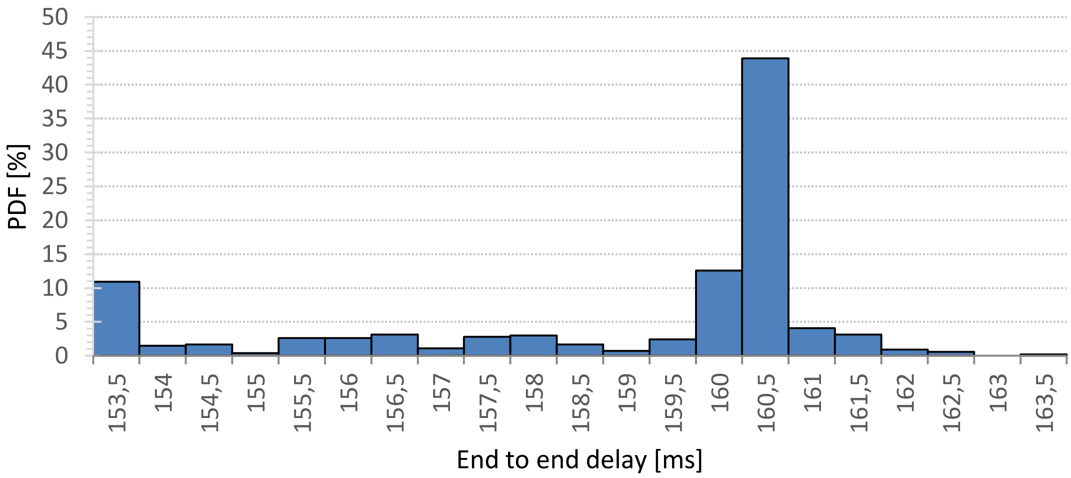


Figure A.11. PDF of the end to end delay for UMTS, mean call duration of 50 s, 50 users (seed 5).

Table A.3. Remaining PDF values for UMTS, mean call duration of 50 s, 50 users.

Seed		Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
2	**Original**	200.48	448.84	152.03	8949.40	766
	Core	158.99	2.12	152.03	162.72	696
3	**Original**	240.15	679.48	152.76	15305.86	717
	Core	159.18	2.11	152.76	162.54	652
4	**Original**	197.47	287.67	153.22	5354.90	771
	Core	159.08	2.03	153.22	162.56	703
5	**Original**	213.42	491.97	146.67	8808.20	594
	Core	158.53	2.59	152.88	163.06	539

Table A.4. Remaining discarded values for UMTS, mean call duration of 50 s, 50 users.

Seed	Above 163 ms [%]	In between 163 ms and 200 ms [%]	Above 200 ms [%]
2	9.14	1.83	7.31
3	9.07	1.12	7.95
4	8.82	1.95	6.87
5	9.09	1.68	7.41

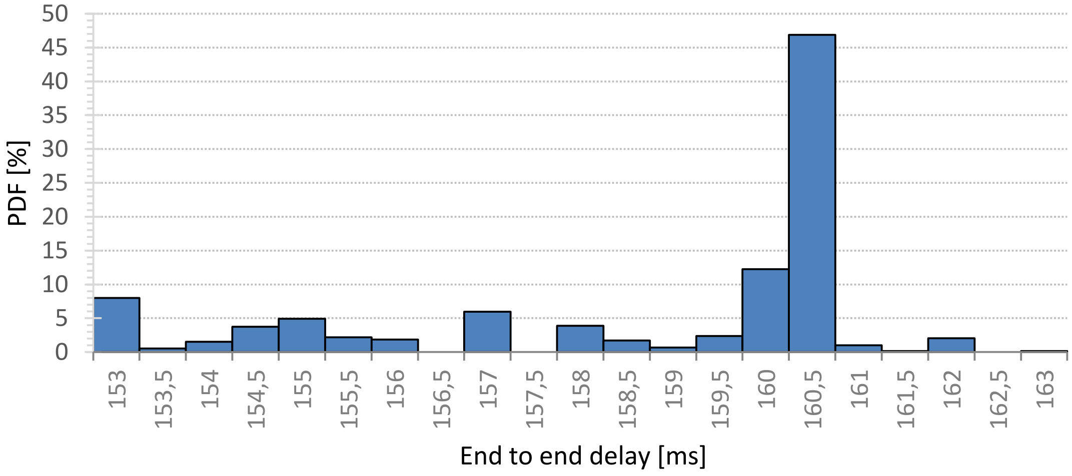


Figure A.12. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed 2).

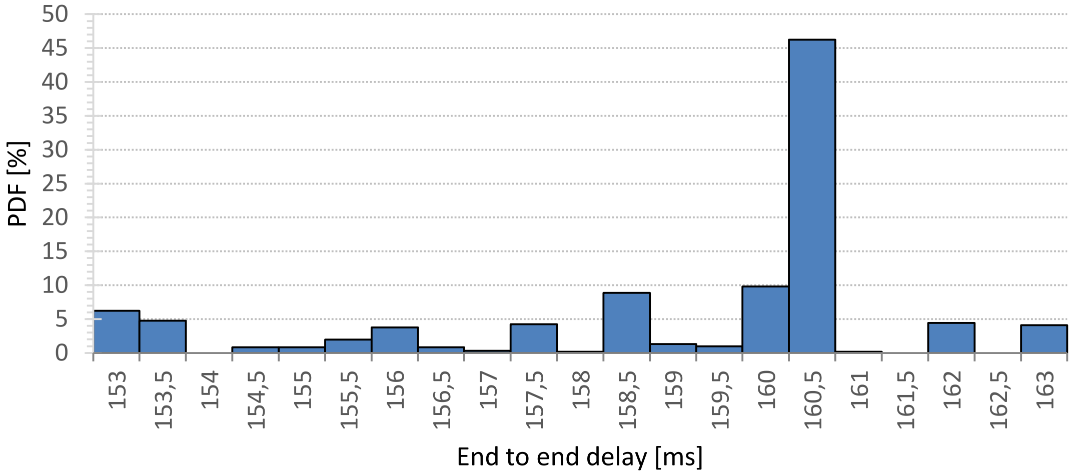


Figure A.13. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed 3).

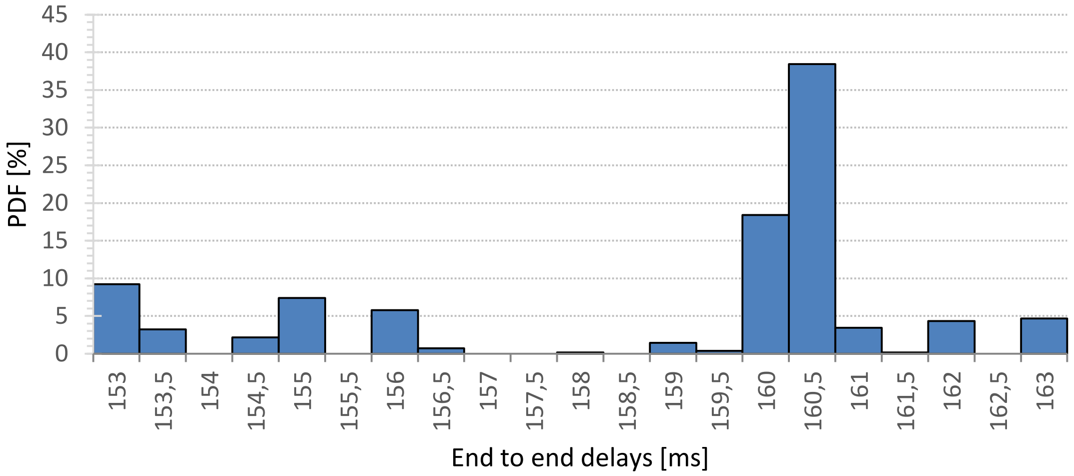


Figure A.14. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed 4).

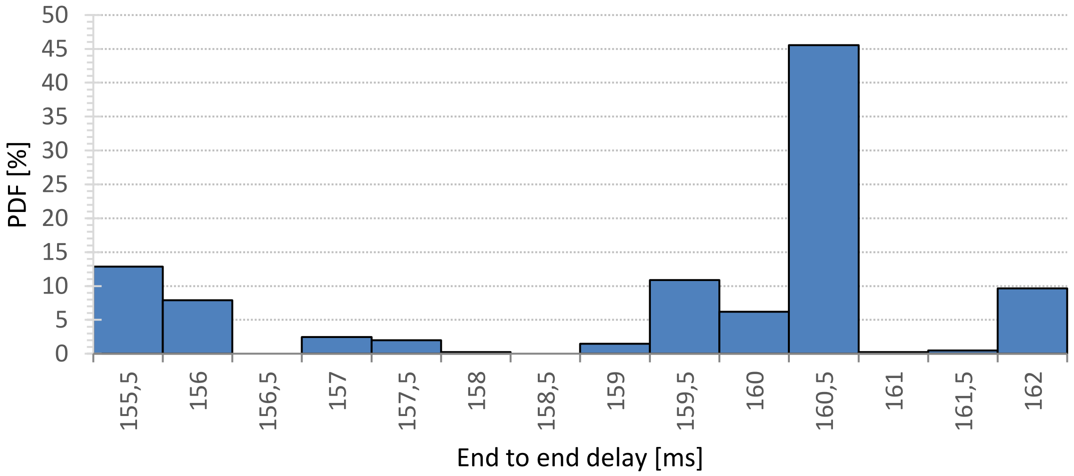


Figure A.15. PDF of the end to end delay for UMTS, mean call duration of 100 s, 40 users (seed 5).

Table A.5. Remaining PDF values for UMTS, mean call duration of 100 s, 40 users.

Seed		Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
2	**Original**	269.43	1204.78	140,01	2011.03	626
	Core	158.31	2.69	152.52	162.59	588
3	**Original**	218.95	742.92	130.00	15157.16	644
	Core	158.68	2.81	152.74	162.60	609
4	**Original**	162.37	38.23	152.82	858.84	572
	Core	158.54	2.93	152.82	162.64	554
5	**Original**	181.64	284.95	155.26	5576.86	417
	Core	159.01	2.02	155.26	161.81	404

Table A.6. Remaining discarded values for UMTS, mean call duration of 100 s, 40 users.

Seed	Above 163 ms [%]	In between 163 ms and 200 ms [%]	Above 200 ms [%]
2	5.91	0.96	4.95
3	5.28	1.40	3.88
4	3.15	1.57	1.57
5	3.12	0.96	2.16

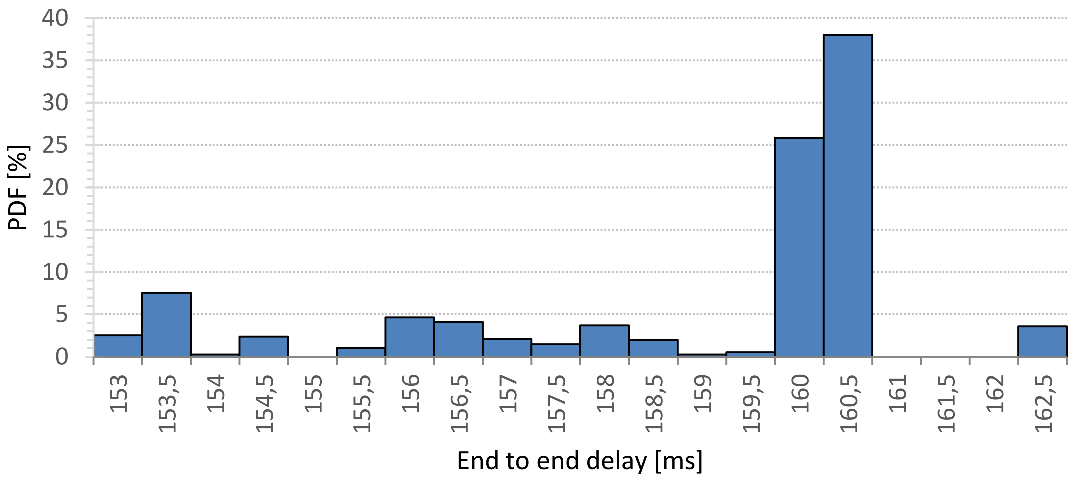


Figure A.16. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed 2).

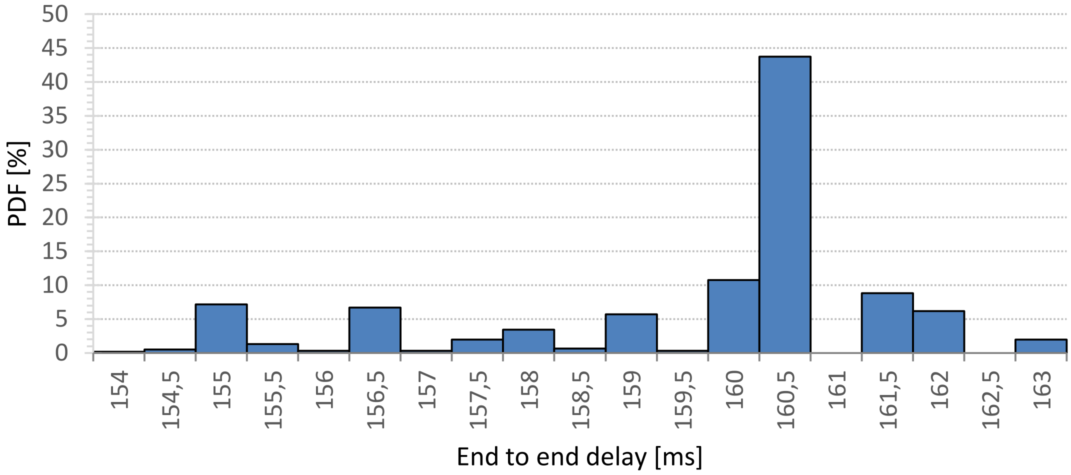


Figure A.17. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed 3).

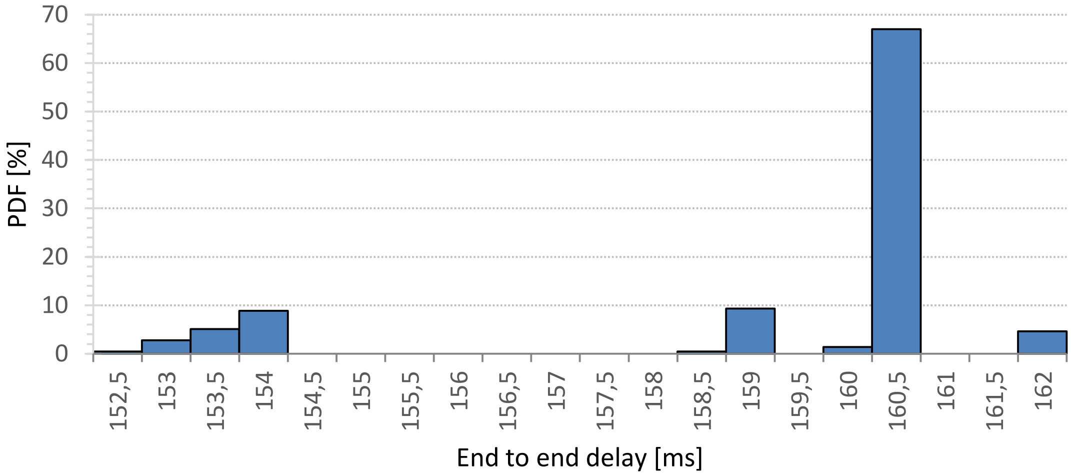


Figure A.18. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed 4).

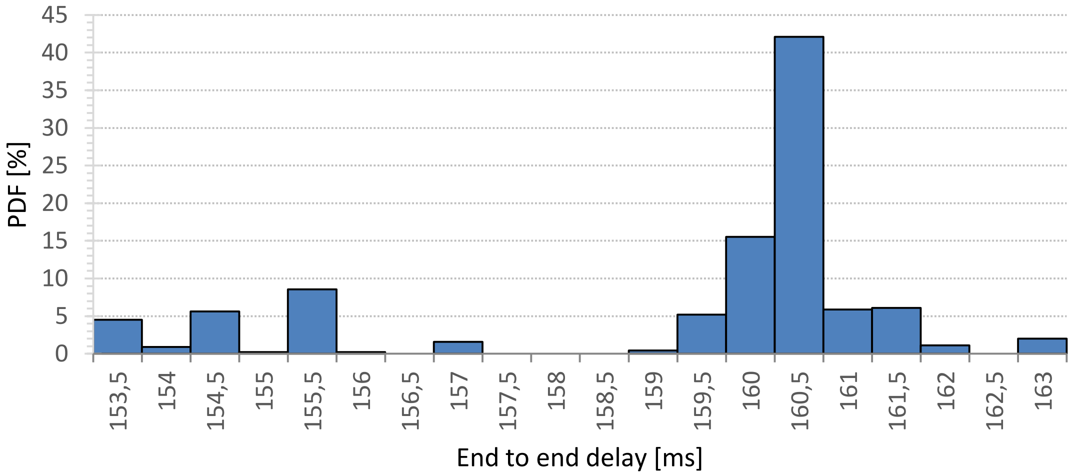


Figure A.19. PDF of the end to end delay for UMTS, mean call duration of 100 s, 50 users (seed 5).

Table A.7. Remaining PDF values for UMTS, mean call duration of 100 s, 50 users.

Seed		Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
2	**Original**	184.08	370.71	152.76	8289.07	783
	Core	158.60	2.45	152.76	162.22	755
3	**Original**	194.26	625.08	153.72	15424.98	634
	Core	159.29	1.99	153.72	162.64	613
4	**Original**	183.44	272.61	152.01	4147.16	222
	Core	158.85	2.56	152.01	161.76	215
5	**Original**	214.76	760.22	153.05	15322.04	468
	Core	158.99	2.37	153.05	162.61	444

Table A.8. Remaining discarded values for UMTS, mean call duration of 100 s, 50 users.

Seed	Above 163 ms [%]	In between 163 ms and 200 ms [%]	Above 200 ms [%]
2	3.58	1.28	2.30
3	3.31	1.42	1.89
4	3.15	0.90	2.25
5	5.13	2.56	2.56

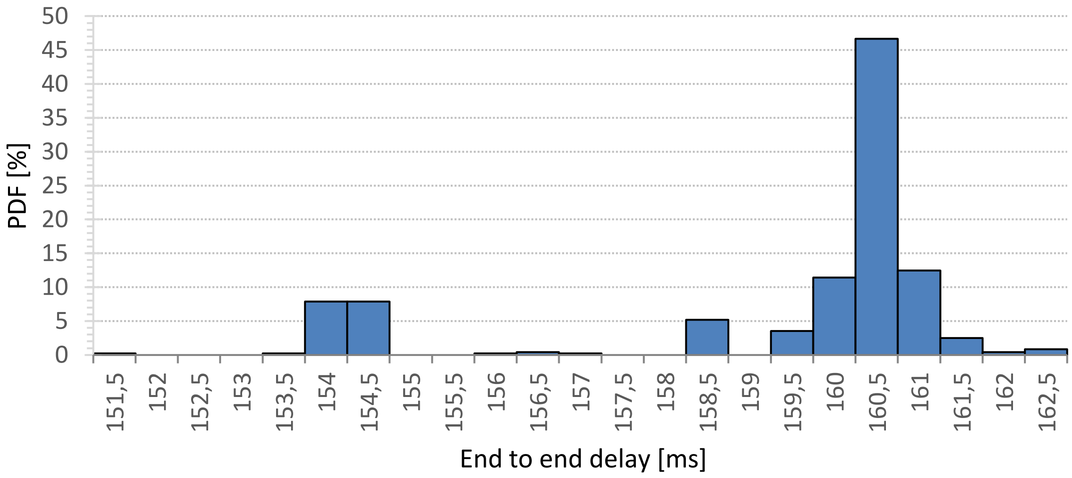


Figure A.20. PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed 2).

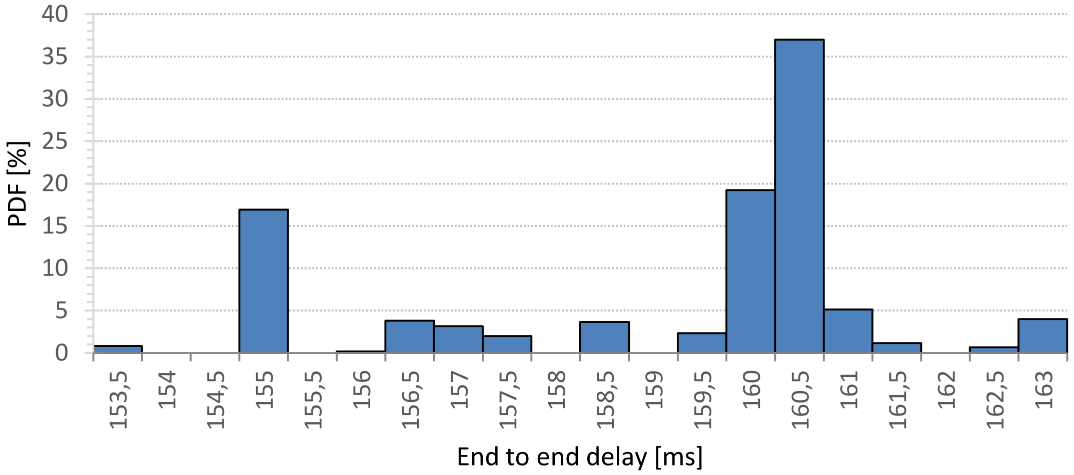


Figure A.21. PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed 3).

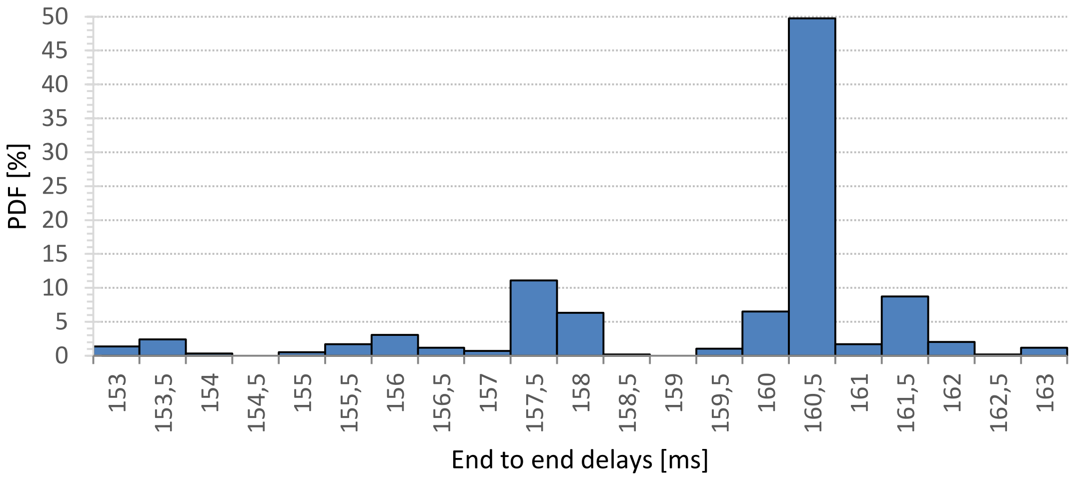


Figure A.22. PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed 4).

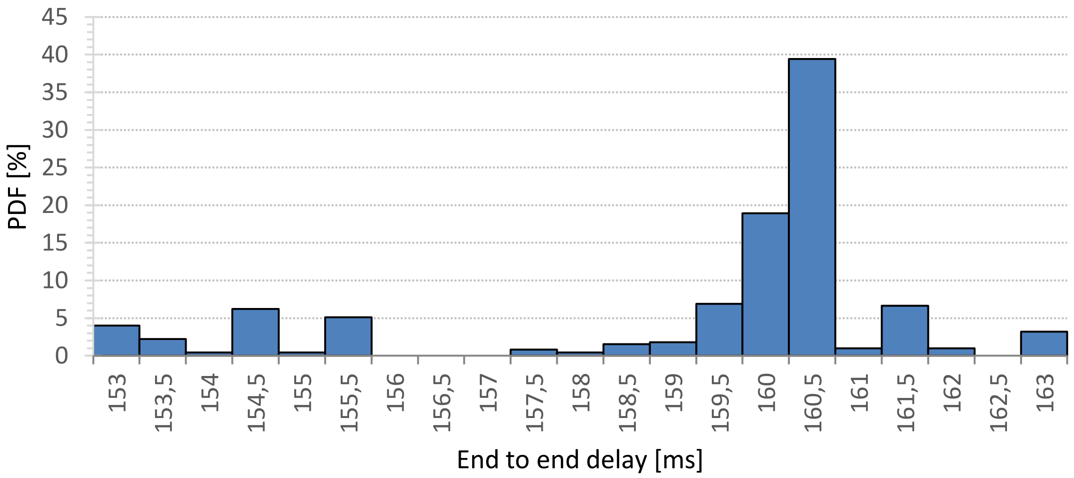


Figure A.23. PDF of the end to end delay for UMTS, mean call duration of 150 s, 40 users (seed 5).

Table A.9. Remaining PDF values for UMTS, mean call duration of 150 s, 40 users.

Seed		Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
2	**Original**	159.05	2.38	151.45	162.24	482
	Core	159.05	2.38	151.45	162.24	482
3	**Original**	167.69	181.55	153.21	4639.17	616
	Core	158.81	2.29	153.21	162.53	603
4	**Original**	186.32	362.23	152.81	8178.32	607
	Core	159.12	2.02	152.81	162.59	585
5	**Original**	173.22	216.22	152.79	5941.11	752
	Core	158.96	2.42	152.79	162.66	723

Table A.10. Remaining discarded values for UMTS, mean call duration of 150 s, 40 users.

Seed	Above 163 ms [%]	In between 163 ms and 200 ms [%]	Above 200 ms [%]
2	0	0	0
3	2.11	1.30	0.81
4	3.62	1.32	2.31
5	3.68	1.20	2.66

Table A.11. Curve fitting values for UMTS, 40 users.

Mean call duration [s]	Seed	R^2	μ [ms]	σ [ms]
50	**2**	0.9165	160.3	0.242
	3	0.9564	160.4	0.287
	4	0.8542	160.3	0.202
	5	0.9404	160.4	0.230
100	**1**	0.9323	160.4	0.295
	2	0.9138	160.4	0.221
	3	0.8901	160.4	0.187
	4	0.8558	160.4	0.282
	5	0.7554	160.4	0.188
150	**1**	-	-	-
	2	0.9233	160.5	0.305
	3	0.7843	160.4	0.312
	4	0.8839	160.4	0.215
	5	0.8815	160.3	0.256

Table A.12. Curve fitting values for UMTS, 50 users.

Mean call duration [s]	Seed	R^2	μ [ms]	σ [ms]
50	**1**	0.9032	160.4	0.258
	2	0.9368	160.4	0.301
	3	0.9415	160.4	0.256
	4	0.8810	160.4	0.230
	5	0.8978	160.4	0.263
100	**1**	0.8690	160.3	0.227
	2	0.9170	160.3	0.206
	3	0.8490	160.3	0.162
	4	0.9476	160.4	0.139
	5	0.8888	160.4	0.295

A.3 Additional Voice Results in LTE

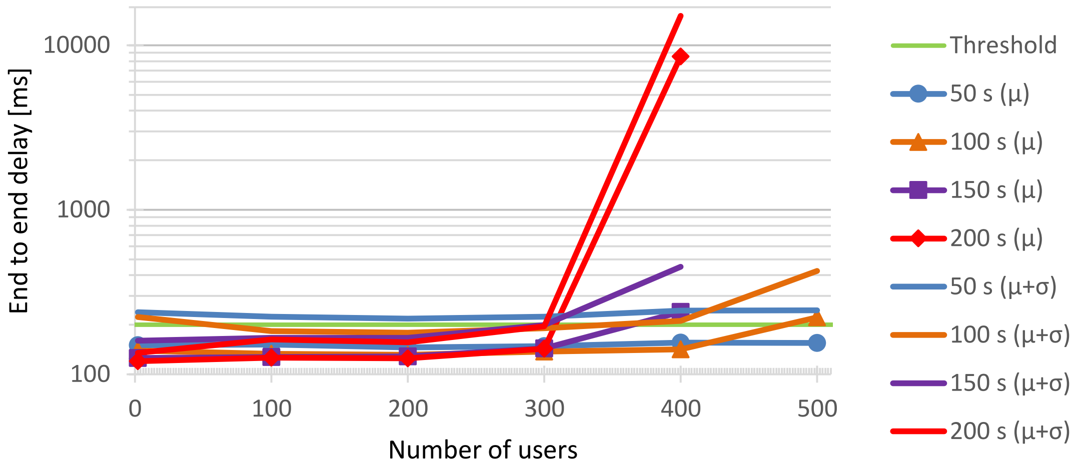


Figure A.24. All end to end delay for LTE in a logarithmic scale.

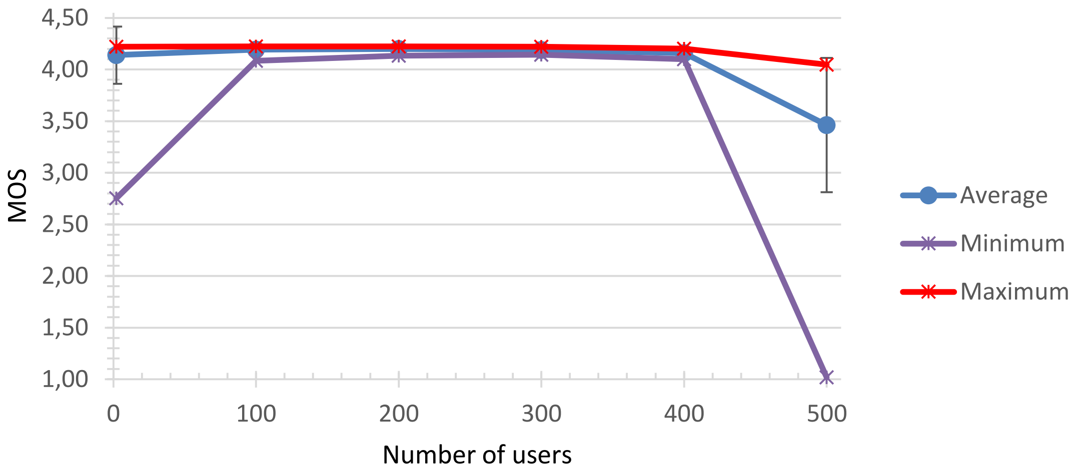


Figure A.25. Global MOS for LTE with a mean call duration of 100 seconds.

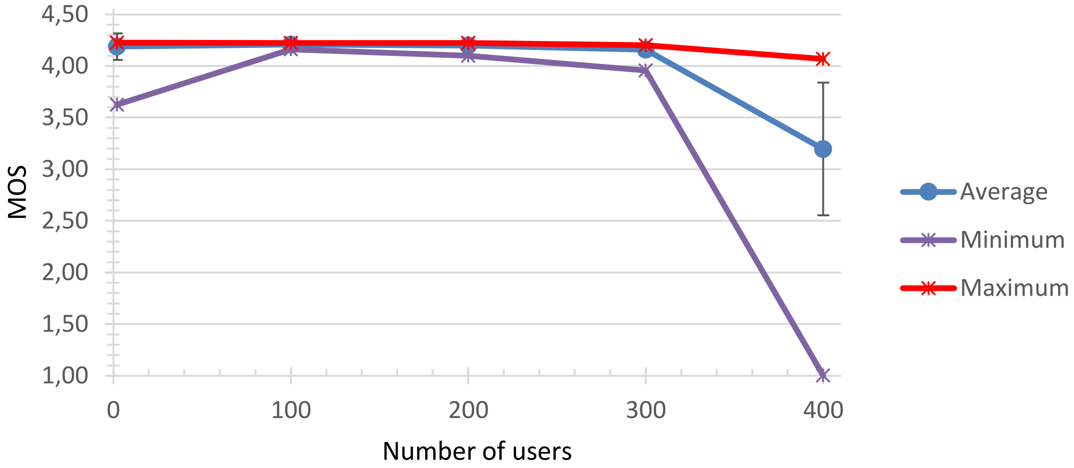


Figure A.26. Global MOS for LTE with a mean call duration of 150 seconds.

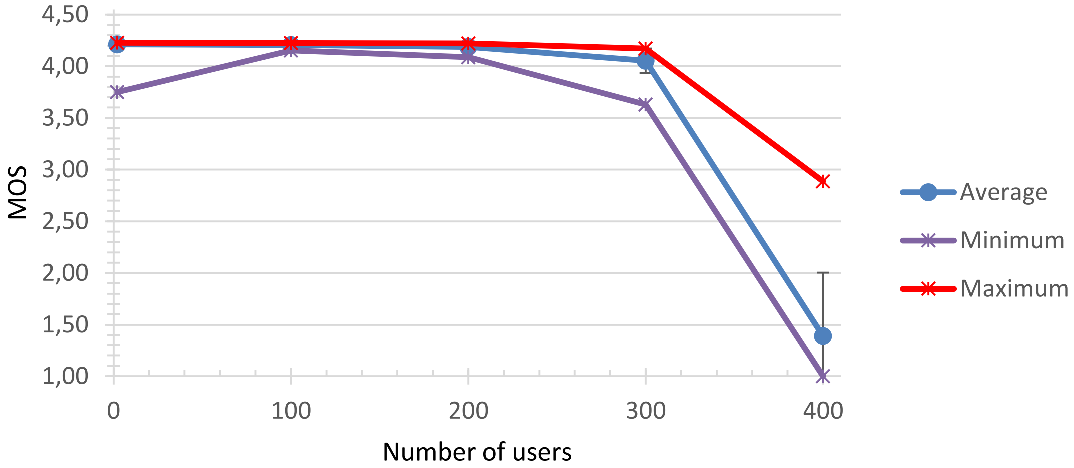


Figure A.27. Global MOS for LTE with a mean call duration of 200 seconds.

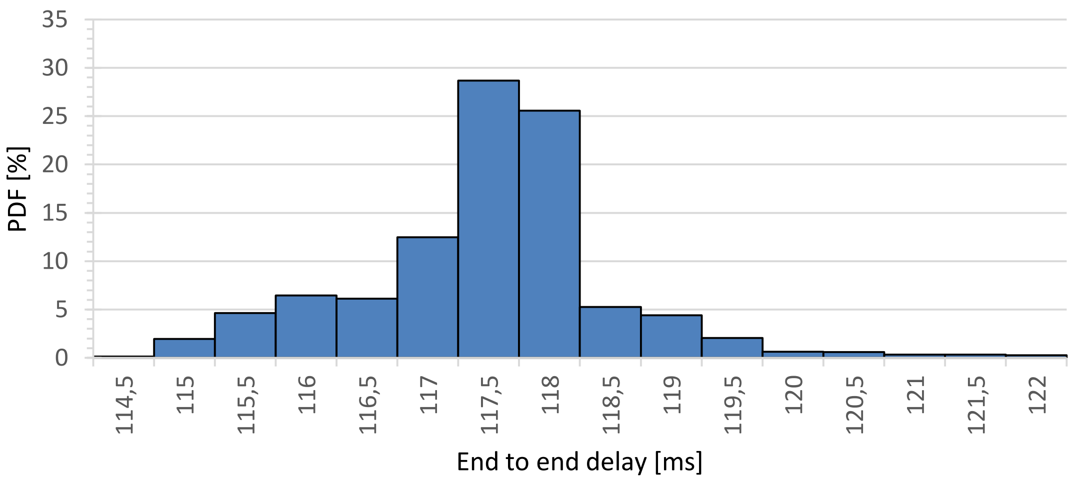


Figure A.28. PDF of the end to end delay for LTE, mean call duration of 50 s, 500 users.

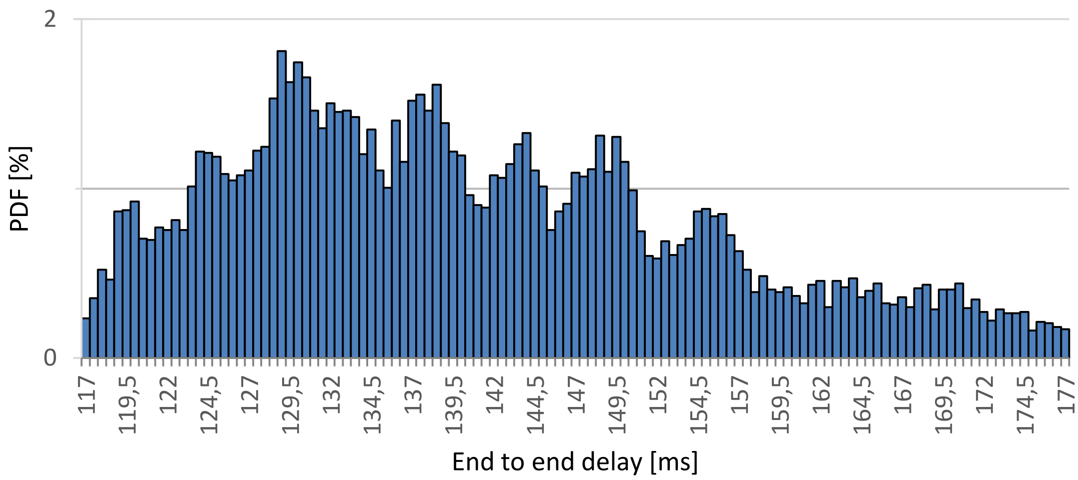


Figure A.29. PDF of the end to end delay for LTE, mean call duration of 100 s, 500 users.

Table A.13 PDF values for LTE, mean call duration of 50 s and 100 s, 500 users.

$\bar{\tau}$ [s]		Average [ms]	Standard Deviation [ms]	Minimum [ms]	Maximum [ms]	Number of Samples
50	Original	145.83	197.75	111.14	9766.12	9898
	Core	117.22	1.04	114.08	121.99	8602
100	Original	213.43	326.75	112.20	10069.69	15986
	Core	140.56	14.12	116.50	176.97	13635

Table A.14 Discarded values for LTE, mean call duration of 50 s and 100 s, 500 users.

$\bar{\tau}$ [s]	Above the core [%]	Acceptable values [%]	Above 200 ms [%]
50	13.01	7.22	5.79
100	14.61	2.69	11.92

Figure A.28 is represented by a limited Gaussian Distribution with an R^2 of 0.9102. Figure A.29 is represented by a Lognormal Distribution, with an R^2 of 0.7135, the Gaussian distribution was tested, however, the goodness of fit is lower, the R^2 for the Gaussian Distribution is 0.6307, 63%. The R^2 was computed using the equation (3.5). Table A.15 shows the relevant values for the curve fitting. The Lognormal Distribution has a relation with the Normal Distribution, the data from the Lognormal Distribution, μ and σ, can be analysed as a Normal Distribution if the dada were analysed as a logarithm of the values instead of the original values, so, in order to analyse the μ and σ from the Lognormal in miliseconds, one should convert the given value from the logarithm to the original value, [PoAc15]. The μ from the Lognormal is 4.94, and the σ is 0.099, the original values are presented in the Table A.15.

Table A.15 Curve fitting values for LTE, mean call duration of 50 s and 100s, 500 users.

Mean call duration [s]	**Distribution**	**R^2**	**μ [ms]**	**σ [ms]**
50	Normal	0.8949	117.7	0.396
100	Lognormal	0.7135	139.8	1.104

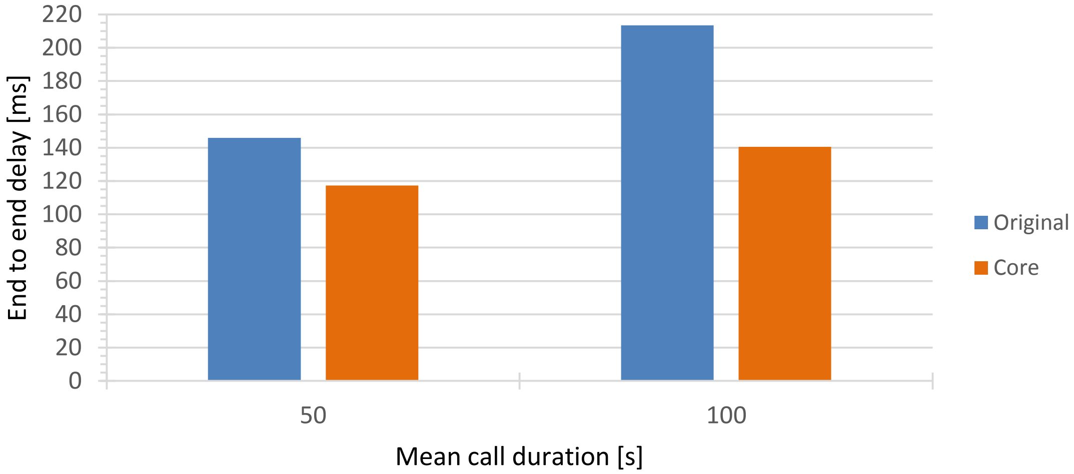


Figure A.30 Comparison between original sampling and core sampling, LTE, 500 users.

A.4 Additional Data Results in LTE

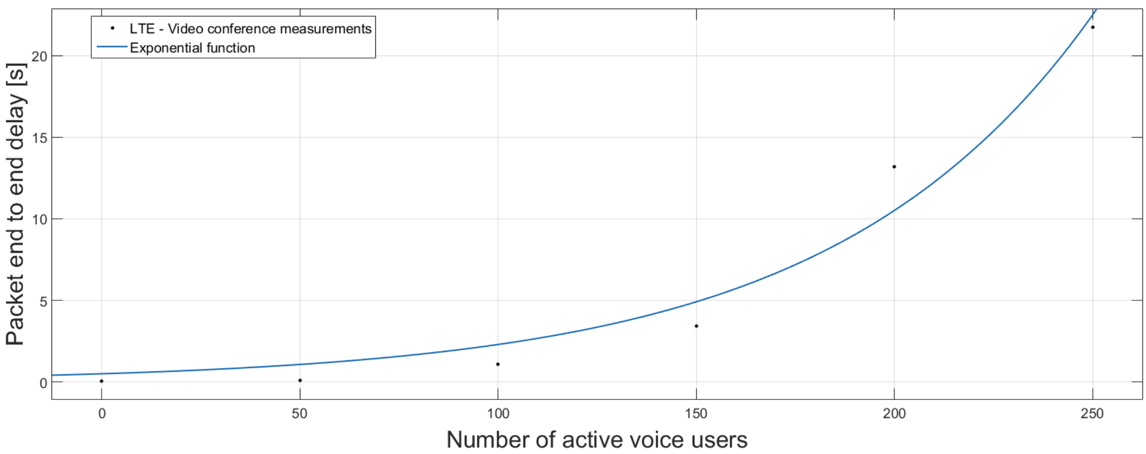


Figure A.31 Curve fitting for the packet end to end delay of the Video Conference (1st Configuration).

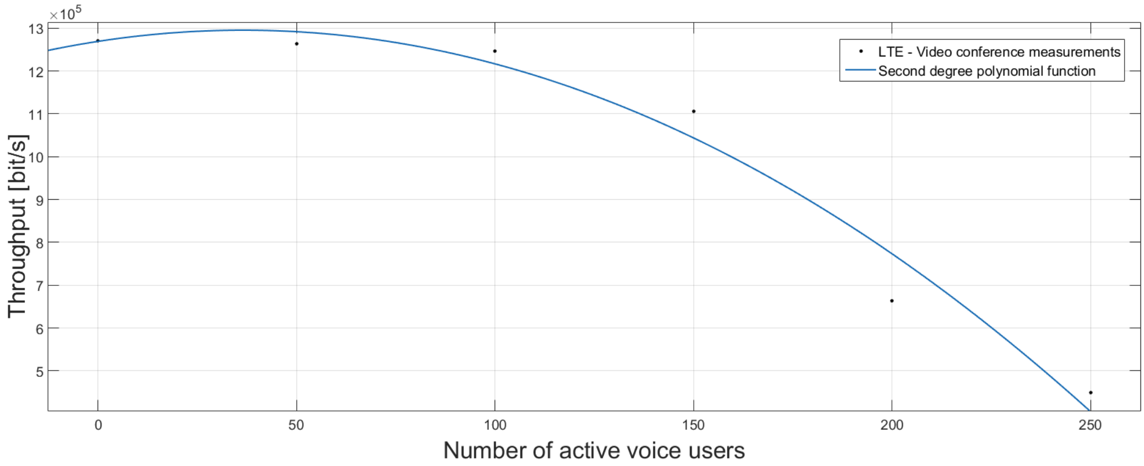


Figure A.32 Curve fitting for the throughput of the Video Conference (1st Configuration).

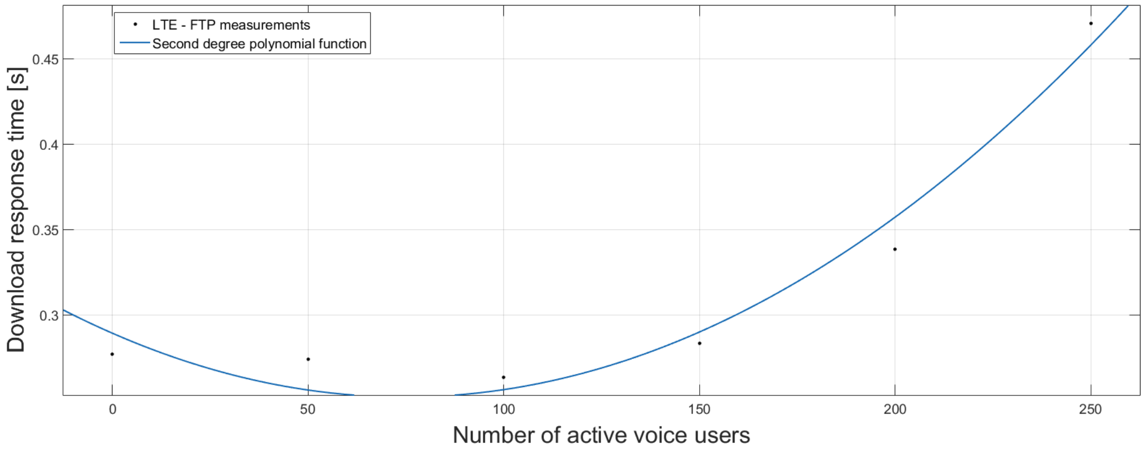


Figure A.33 Curve fitting for the download response time of the FTP (1st Configuration).

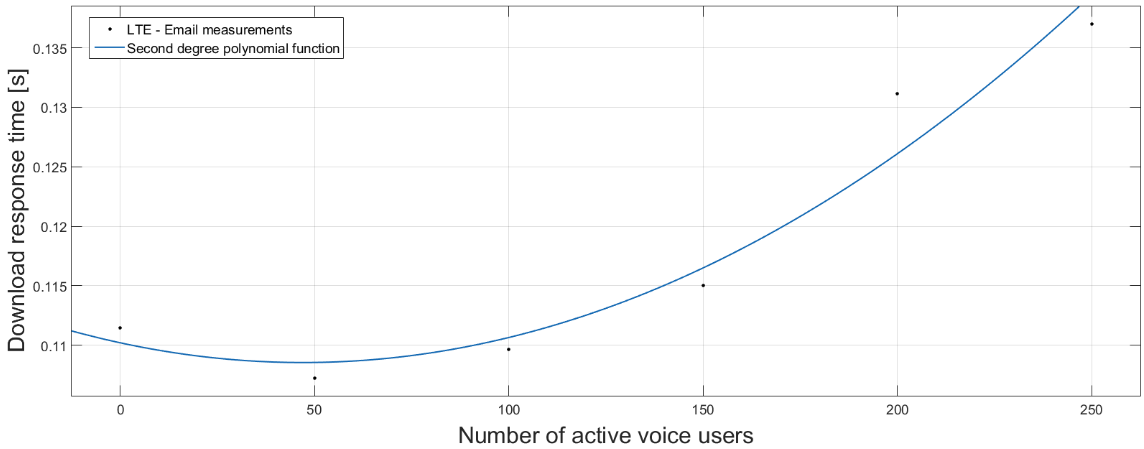


Figure A.34 Curve fitting for the download response time of the Email (1st Configuration).

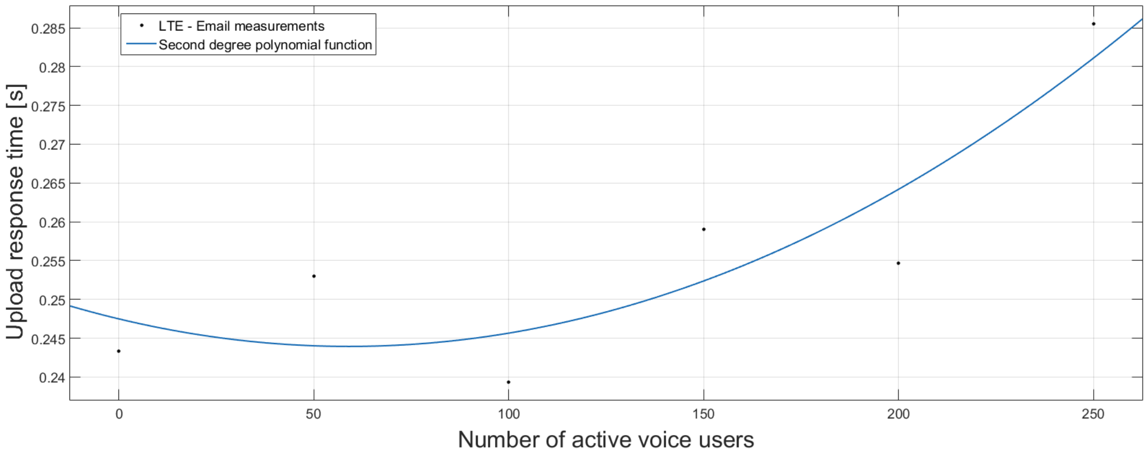


Figure A.35 Curve fitting for the upload response time of the Email (1st Configuration).

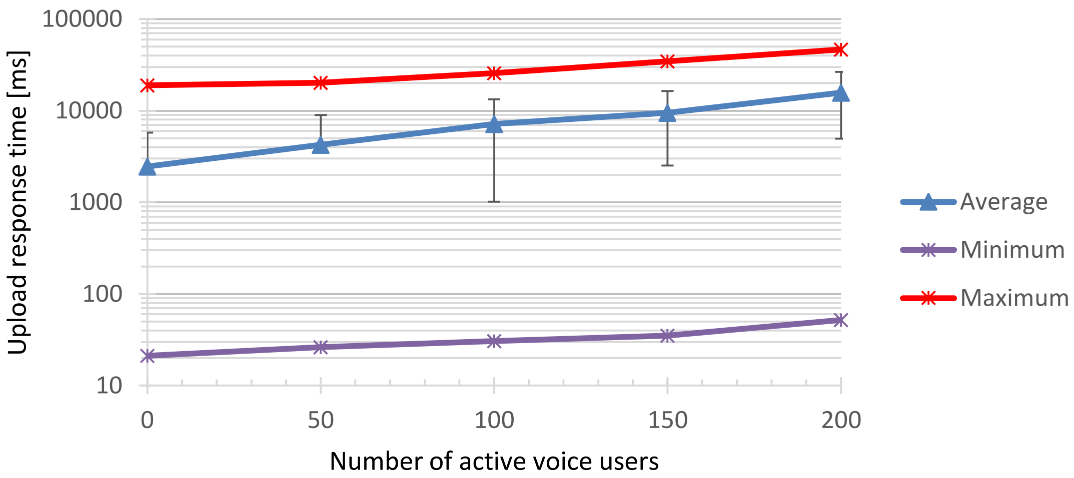


Figure A.36 Packet end to end delay for the 2nd configuration of the Video Conference service.

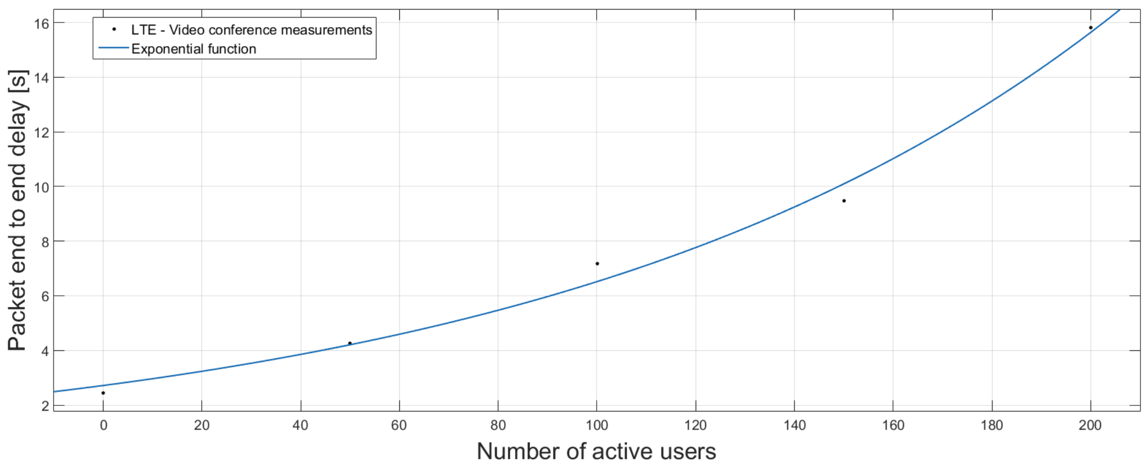


Figure A.37 Curve fitting for the packet end to end delay of the Video Conference (2nd Configuration).

$$\tau_{ETE} = 2.6131 \cdot e^{0.009 \cdot N_u} \quad \text{(A.1)}$$

R^2=0.9899

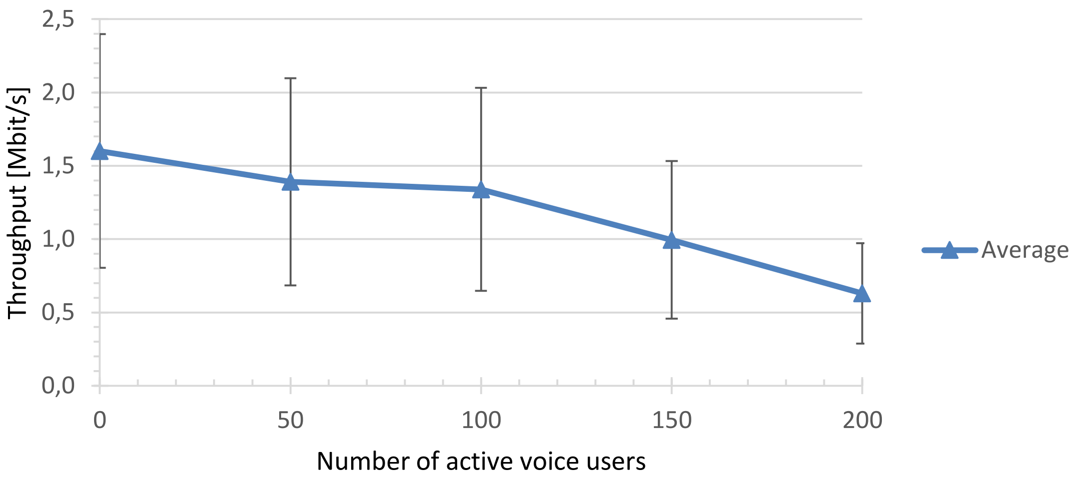


Figure A.38 Throughput for the 2nd configuration of the Video Conference service.

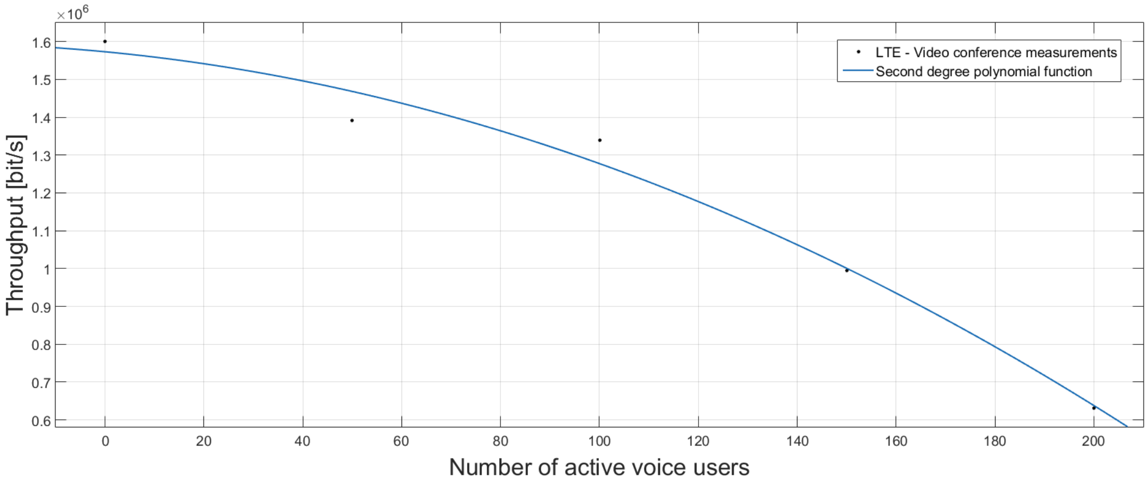


Figure A.39 Curve fitting for the throughput of the Video Conference (2nd Configuration).

$$Th = -17.198 \cdot {N_u}^2 - 1236.7 \cdot N_u + 2 \cdot 10^6 \tag{A.2}$$

R^2=0.9818

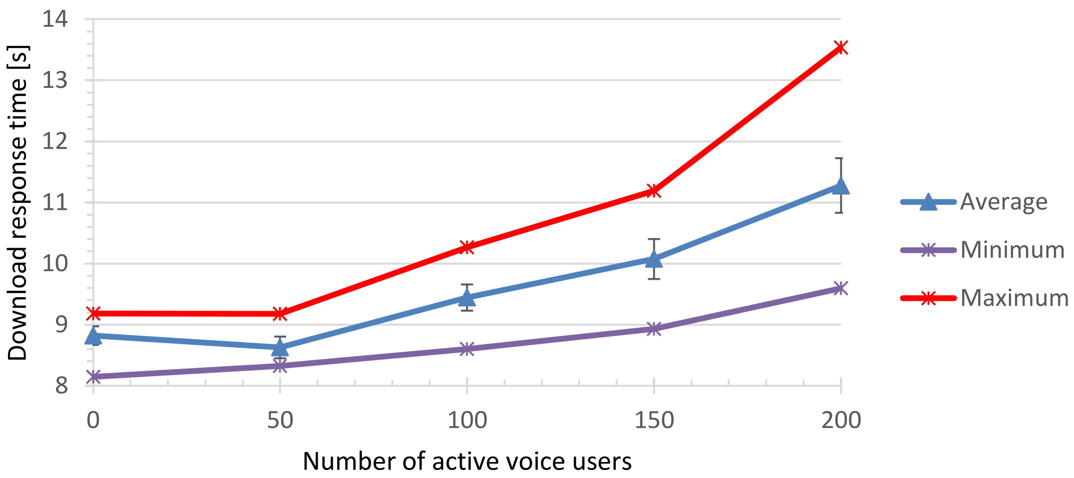


Figure A.40 Download response time for the 2nd configuration of the FTP service.

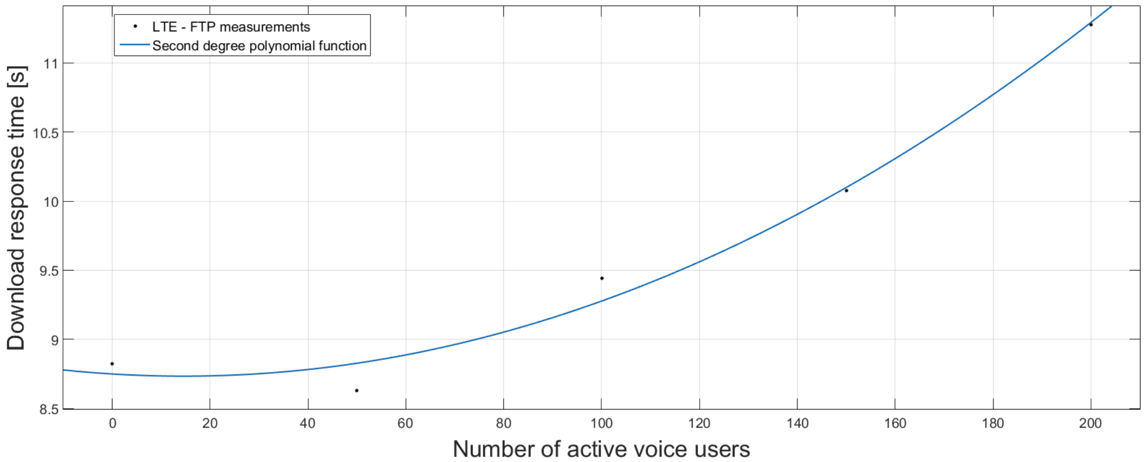


Figure A.41 Curve fitting for the download response time of the FTP (2nd Configuration).

$$\tau_{DL} = 7 \cdot 10^{-5} \cdot {N_u}^2 - 0.0022 \cdot N_u + 8.7496 \quad \text{(A.3)}$$

R^2=0.9841

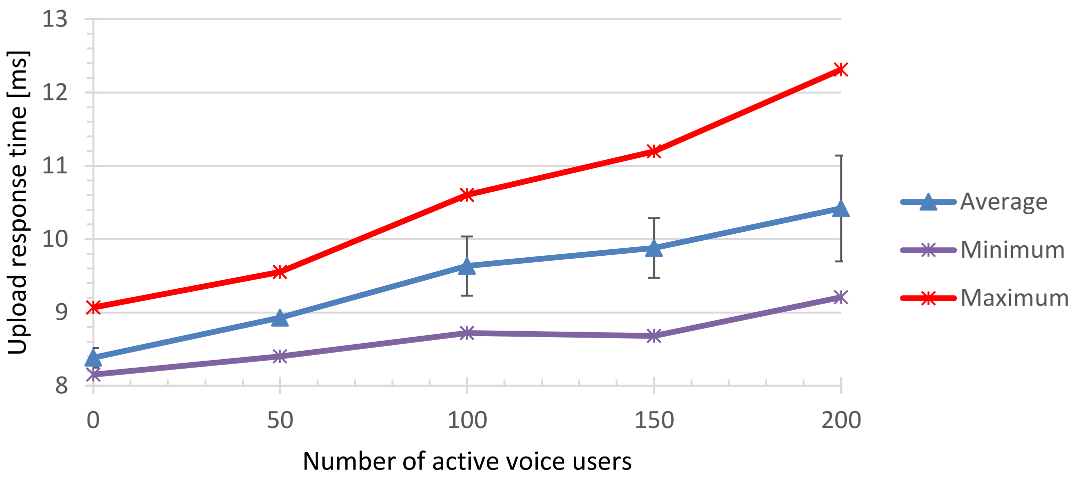


Figure A.42 Upload response time for the 2nd configuration of the FTP service.

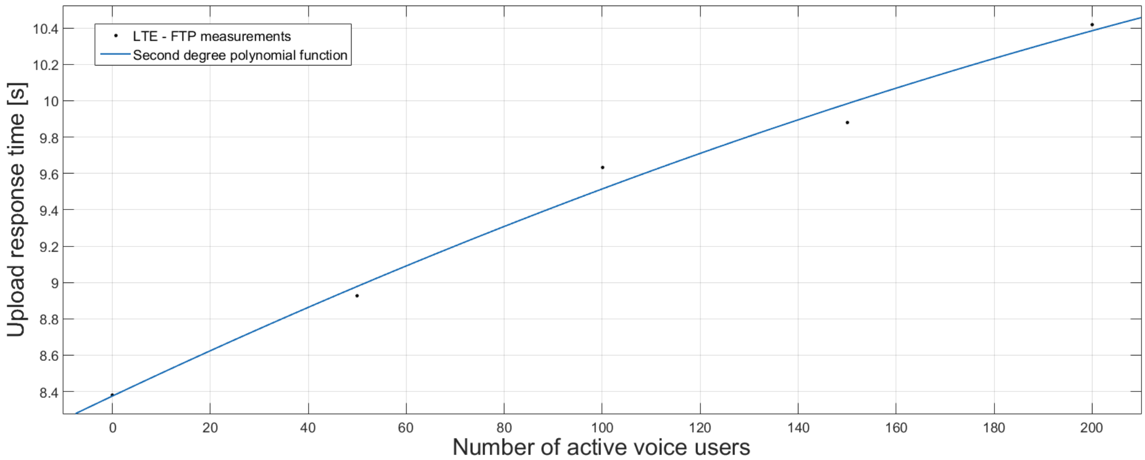


Figure A.43 Curve fitting for the upload response time of the FTP (2nd Configuration).

$$\tau_{UL} = -1.347 \cdot 10^{-5} \cdot {N_u}^2 - 0.0127 \cdot N_u + 8.375 \quad \text{(A.4)}$$

R^2=0.9888

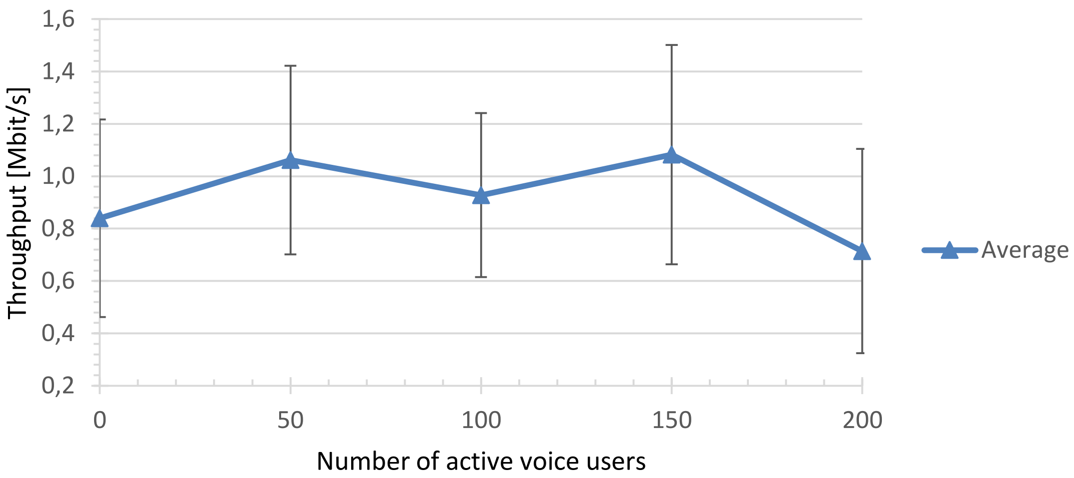


Figure A.44. Throughput for the 2nd configuration of the FTP service.

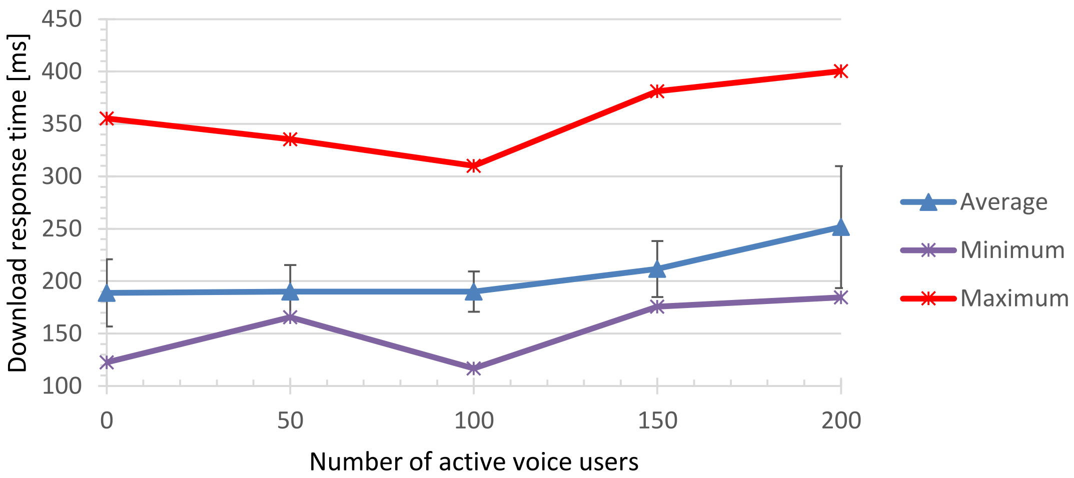


Figure A.45 Download response time for the 2nd configuration of the Email service.

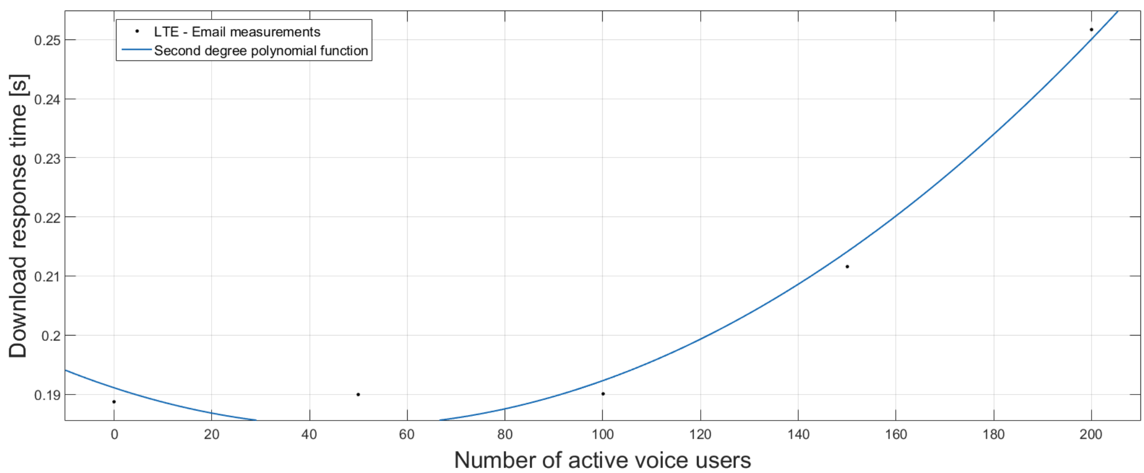


Figure A.46 Curve fitting for the download response time of the Email (2nd Configuration).

$$\tau_{DL} = 3 \cdot 10^{-6} \cdot {N_u}^2 - 0.0003 \cdot N_u + 0.1911 \qquad \text{(A.5)}$$

R^2=0.9837

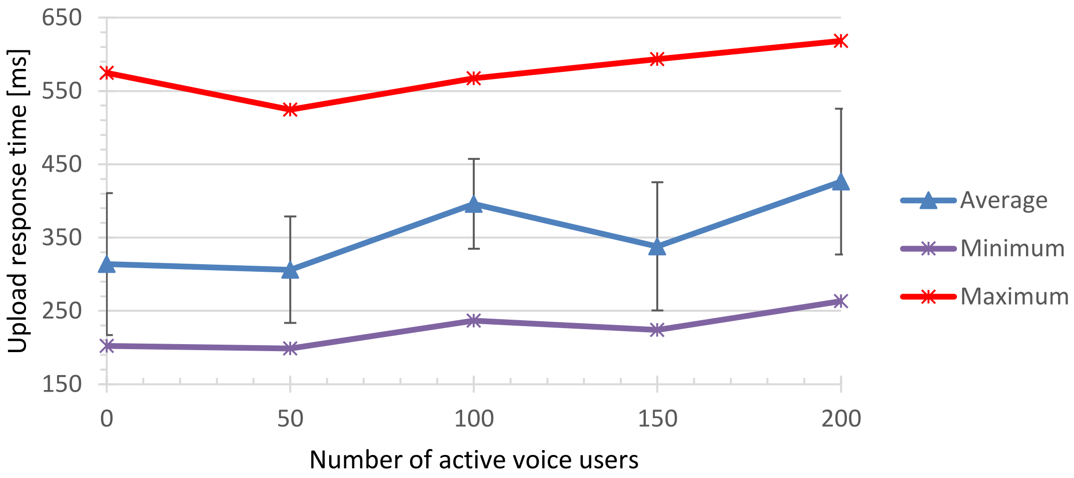


Figure A.47 Upload response time for the 2nd configuration of the Email service.

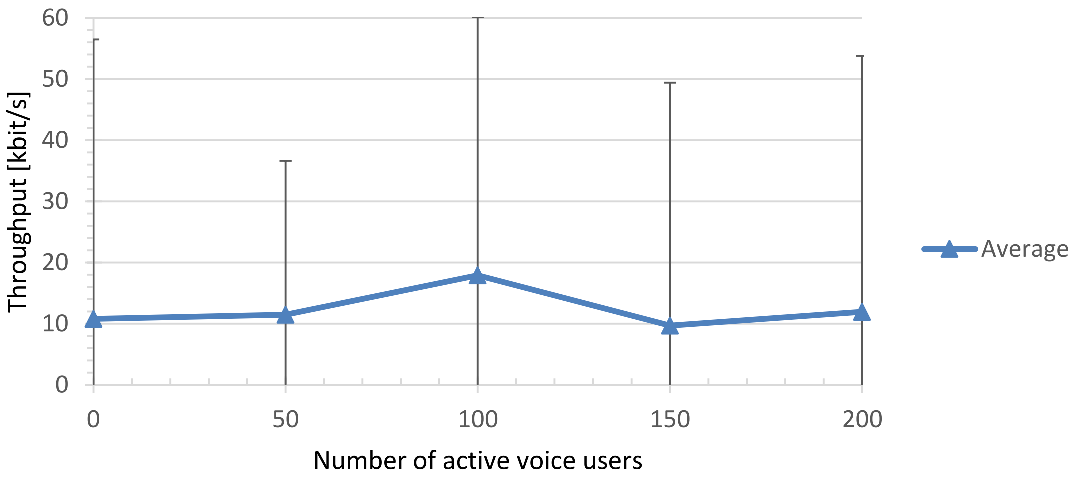


Figure A.48 Throughput for the 2nd configuration of the Email service.

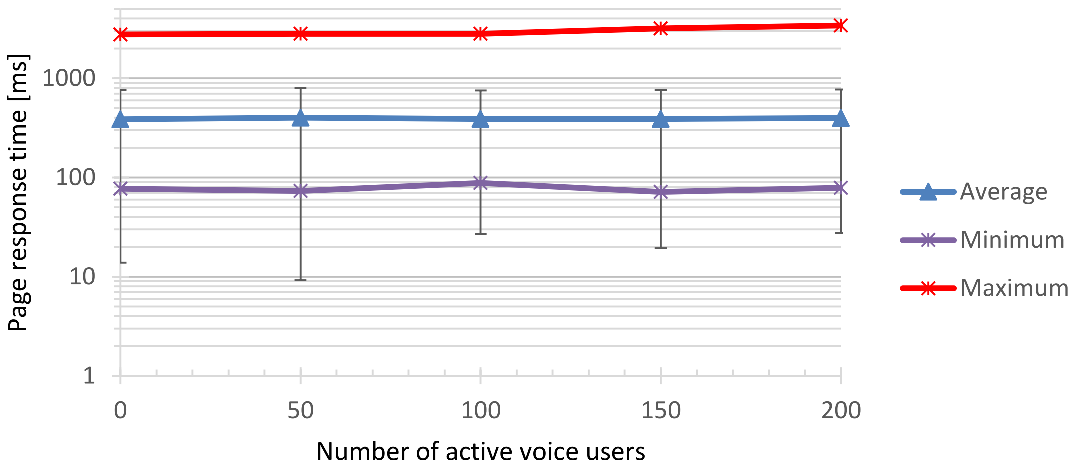


Figure A.49 Page response time for the 2nd configuration of the HTTP service.

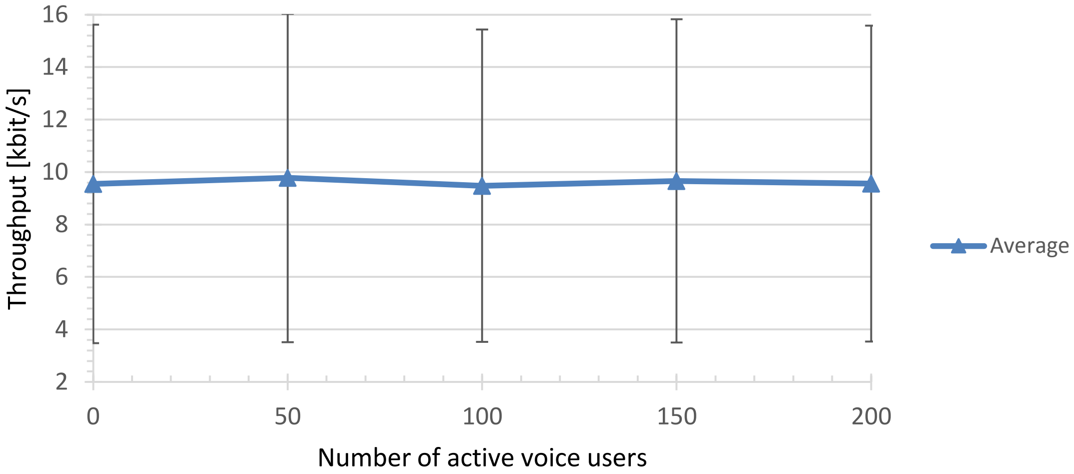


Figure A.50 Throughput for the 2nd configuration of the HTTP service.

References

[3GPP14] 3GPP, *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (Release 11),* TS 36.211, Ver. 11.5.0, January 2014 (http://www.3gpp.org).

[3GPP14a] 3GPP, *Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 11),* TS 36.101, Ver. 11.9.0, August 2014 (http://www.3gpp.org).

[AaKj11] Asheralieva, A., Khan, J. and Mahata, K., "Performance Analysis of VoIP Services on the LTE Network", in *Proc. of ATNAC'11 – IEEE Australasian Telecommunication Networks and Applications Conference*, Melbourne, Australia, November 2011.

[Agil07] Agilent Technologies, *Agilent Technologies Solutions for 3GPP LTE*, USA, September 2007 (http://cp.literature.agilent.com/litweb/pdf/5989-6331EN.pdf).

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

[ANA12a] ANACOM, *Final Report of the auction for the attribution of the frequency utilisation rights* (in Portuguese), Public Consultation, Lisbon, Portugal, January 2012 (http://www.anacom.pt/streaming/Relat_final_leilao_Jan2012.pdf?contentId=1112754&field=ATTACHED_FILE).

[ANA12b] ANACOM, *Table of National Frequency Allocation* (in Portuguese), Public Consultation, Lisbon, Portugal, June 2012 (http://www.anacom.pt/streaming/qnaf2010_2011_20junho2012.pdf?contentId=1129525&field=ATTACHED_FILE).

[ANA12c] ANACOM, *Decision to approve agreement on location of spectrum in the 1800 MHz band* (in Portuguese), Public Consultation, Lisbon, Portugal, March 2012 (http://www.anacom.pt/streaming/Decisao_ReshufflingMarco2012.pdf?contentId=1120288&field=ATTACHED_FILE).

[AnMo11] Andersson, K., Mostafa, S. and Ul-Islam, R., "Mobile VoIP User Experience in LTE", in *Proc. of LCN'11 - 36th IEEE Conference on Local Computer Networks*, Bonn, Germany, October 2011.

[Apti15] http://www.aptilo.com/wi-fi-calling/what-is-wi-fi-calling-vowifi, September 2015.

[Cisc15] Cisco, *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast*, Public Consultation, February 2015 (http://www.cisco.com/c/en/us/solutions/collateral/serviceprovider/visual-networking-index-vni/white_paper_c11-520862.html).

[Corr13] Correia, L., *Mobile Communications Systems*, Course notes, Instituto Superior Técnico, 2013.

[EbVö09] Eberspächer, J., Vögel, H., Bettstetter, C. and Hartmann, C., *GSM Architecture, Protocols and Services* (3rd Edition), John Wiley & Sons, Chichester, UK, 2009.

[Eric15] Ericsson, *Ericsson Mobility Report*, Public Consultation, Stockholm, Sweden, February 2015 (http://www.ericsson.com/ericsson-mobility-report).

[HaRo03] Holonen, T., Romero, J. and Melero, J., *GSM, GPRS and EDGE Performance Evolution Towards 3G/UMTS* (2nd Edition), John Wiley & Sons, Chichester, UK, 2003.

[HoTo04] Holma, H. and Toskala, A., *WCDMA for UMTS* (3rd Edition), John Wiley & Sons, Chichester, UK, 2004.

[HoTo09] Holma, H. and Toskala, A., *LTE for UMTS – OFDMA and SC-FDMA Based Radio Access* (1st Edition), John Wiley & Sons, Chichester, UK, 2009.

[HoTo11] Holma, H. and Toskala, A., *LTE for UMTS Evolution to LTE-Advanced* (2nd Edition), John Wiley & Sons, Chichester, UK, 2011.

[ITUT08] ITU-T, *E-model*, January 2015 (https://www.itu.int/ITU-T/studygroups/com12/emodelv1/tut.htm).

[ITUT14] ITU-T, *The E-model, a computational model for use in transmission planning*, Recommendation G.107, March 2014.

[JaIb12] Jailani, E., Ibrahim, M. and Rahman, R., "LTE Speech Traffic Estimation for Network Dimensioning", in *Proc of ISWTA'12 – IEEE Symposium on Wireless Technology and Applications*, Bandung, Indonesia, September 2012.

[MatI15] http://www.mathworks.com/help/curvefit/evaluating-goodness-of-fit.html, September 2015.

[Rive13] Riverbed Technology, *OPNET Modeler Documentation Set*, Version 17.5, 2013.

[Rive14] http://gb.riverbed.com/products/performance-management-control/opnet.html?redirect=opnet, January 2014.

[OzVa13] Ozturk, O. and Vajapeyam, M., "Performance of VoLTE and Data Traffic in LTE Heterogeneous Networks", in *Proc. of GLOBECOM'13 – IEEE Global Communications Conference*, Atlanta, USA, December 2013.

[PaPa14] Pastrav, A., Palade, T. and Puschita, E., "VoLTE Performances in a 3D Modeled Campus Area", in *Proc. of COMM'14 – 10th IEEE International Conference on Communications*, Bucharest, Romania, May 2014.

[PoAc15] http://www.portalaction.com.br/probabilidades/615-distribuicao-log-normal, *in Portuguese*, September 2015.

[PoHo12] Poikselkä, M., Holma, H., Hongisto, J., Kallio, J., and Toskala, A., *Voice over LTE (VoLTE)*

(1st Edition), John Wiley & Sons, Chichester, UK, 2012.

[QuDe15] QualNet Developer 6.1 Network Simulator, January 2015 (http://wwwscalable-networks.com).

[SeTo11] Sesia, S., Toufik, I. and Baker, M., *LTE – The UMTS Long Term Evolution (From Theory to Pratice)* (2nd Edition), John Wiley & Sons, Chichester, UK, 2011.

[TaGu14] Tabany, M. and Guy, C., "An End-to-End QoS Performance Evaluation of VoLTE in 4G E-UTRAN-based Wireless Networks", in *Proc. of ICWMC – 10th International Conference on Wireless and Mobile Communications*, Seville, Spain , June 2014.

[VoHi15] MainLine Communications, *The History of VoIP*, July 2015 (http://www.mainlinecom.com/voiphistory.html)