INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

Performance Analysis of UMTS/HSDPA Radio Resource Management Algorithms

João Lopes

Dissertation submitted for obtaining the degree of
Master in Electrical and Computer Engineering

Jury

Supervisor: Prof. Luis M. Correia

President: Prof. José Bioucas Dias

Members: Prof. António Rodrigues

October 2010

To all that I love...

Acknowledgements

This thesis is the final and visible work of more than a year of hard work. It was only possible to complete this work as desired, with the support of many people who encourage me, anytime I was down, or facing more difficulties. Those people kept me on the right way and made me believe that there was no reason to give up.

First, I would like to thank to Professor Luis M. Correia, for giving the opportunity to write this thesis and for the experience, knowledge and experience sharing that were fundamental throughout this work. I am thankful for the valuable time during our weekly meetings, as well as his advices and availability that always functioned as guidance to finish this work with the demanded and desired quality.

To Celfinet, that was the starting point of all this journey, to José Brito and Sérgio Pires, for giving their approval and for believing that my skills could be improved to became a better engineer in future challenges.

Next, I would like to thank to my colleagues, Nuno Rapaz and João Gaspar, for their patience, knowledge sharing, friendship and great wisdom and critical spirit. An enormous thanks to Luis Veiga for his unconditional patience, understanding and help; in many cases he supported me like a brother, absorbing my worries and difficulties, encourage me to look ahead.

Last but not of least importance, I would like to thank to my family, especially to my wife Ana, my dear mother and godmother, with their, care, understanding, love and believe throughout my 36 years of living; with them, I supported all the hardest times and problems to finish this journey.

To my wonderful 6 years old daughter Mariana, for understanding so clearly the reason why his daddy had so few time to play with her during the last months.

Abstract

The main purpose of this thesis was to study Radio Resource Management strategies in HSDPA in terms of network access control, having in mind a trade-off supporting the highest number of users, with the best possible QoS.

A simulator was developed in order to evaluate the performance in scenarios like variation of density users, changing the penetration rate of each service and different weight distribution between the 2 UMTS carriers involved, also changing the data packet volume applied.

The present network is managed by an algorithm based on a service/sub-system priority table, evaluating at the end, the average delay per user, the dependency of inputs parameters with the number of users employed in each scenario and the results of using different penetration rates for voice / data services, between both UMTS carriers.

One studied the average delay per user, for a reference scenario, this value being around 100 ms, which doubles if the packet data volume triplicates.

Data services bit rate, vary between 100 kbps in low demanding services, like WWW and E-Mail up to maximum values of 14.4 Mbps, in highest demanding services, like FTP.

Keywords

UMTS, HSDPA, RRM, Shared Carrier, Simulator.

Resumo

O principal objectivo desta tese consistiu no estudo de estratégias de gestão de recursos rádio na tecnologia HSDPA, em termos de controlo de acesso a uma rede móvel, estabelendo um compromisso entre qualidade de serviço e o número máximo de utilizadores activos.

Foi desenvolvido um simulador, como forma de avaliação do desempenho dos vários cenários, variando a densidade de utilizadores, a taxa de penetração de cada serviço, o seu peso e distribuição por cada uma das 2 portadoras UMTS utilizadas, tendo-se também alterado o tamanho dos pacotes transmitidos.

A presente rede é gerida por um algoritmo baseado num critério de prioritização de serviço por tabela de prioridades, avaliando-se no final o tempo de atraso médio por utilisador, a dependência dos parâmetros de entrada, com o número de utilisadores em cada cenário e o resultado da utilização de diferentes taxas de penetração para os serviços de voz e dados entre as 2 portadoras UMTS.

Foi feito um estudo do atraso médio associado a cada utilizador,cujo valor para o cenário de referência rondou os 100 ms; valor esse que é duplicado ao triplicar-se o volume dos pacotes a transmitir.

Os débitos médios obtidos, variaram entre valores mínimos de 100 kbps para serviços de baixo débito, como o web-browsing e o e-mail, até um máximo de 14.4 Mbps para os serviços mais exigentes como é o caso do FTP.

Palavras-chave

UMTS, HSDPA, Gestão de Recursos Rádio, Portadora Partilhada, Simulador.

Table of Contents

List of Figures

List of Tables

List of Acronyms

2G	Second Generation
3G	Third Generation
3GPP	3rd Generation Partnership Project
ACK	Acknowledge Message
AI	Air Interface
AMC	Adaptive Modulation and Coding
AMR	Adaptive Multi Rate
AROMA	Advanced Resource Management Solutions for future all IP heterogeneous Mobile Radio environments
ARP	Allocation Retention Priority
ARQ	Automatic Repeat Request
Asym	Assymetric
BHCA	Busy Hour Call Attempt
Bid	Bidirectional
BLER	Block Error Rate
BS	Base Station
CAC	Call Admission Control
CDMA	Code Division Multiple Access
CIR	Carrier Over Interference Ratio
CN	Core Network
CPU	Central Processing Unit
CQ-BMTA	Channel Quality – Based Minimum Throughput Assurance
CQI	Channel Quality Indicator
CRRM	Common Radio Resource Management
CS	Circuit Switch
DBS	Delay Based Scheduler
DC	Data Centric scenario
DCH	Dedicated Channel
DL	Downlink
DM	Data Mostly scenario
DS-CDMA	Direct Sequence – Code Division Multiple Access
DT	Discard Timer
DTX	Dual Transmission

E-DCH	Enhanced – Dedicated Channel
EDGE	Enhanced Data Rate for GSM Evolution
E-GPRS	Extended GPRS
ER	Exponential Rule
ETSI	European Telecommunications Standards Institute
FDD	Frequency Division Duplex
FIFO	First In First Out
FTP	File Transfer Protocol
GBR	Guaranteed Bit Rate
GGSN	Gateway GPRS Support Node
GMSC	Gateway Mobile Switching Center
GoS	Grade of Service
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HARQ	Hybrid Automatic Repeat Request
HLR	Home Location Register
HSDPA	High Speed Downlink Packet Access
HS-DPCCH	High Speed Dedicated Physical Control Channel
HS-DSCH	High Speed Downlink Shared Channel
HSPA	High Speed Packet Access
HSPA +	High Speed Packet Access Plus
HS-PDSCH	High Speed Physical Downlink Shared Channel
HS-SCCH	High Speed Shared Control Channel
HSUPA	High Speed Uplink Packet Access
HTTP	Hypertext Transfer Protocol
IEEE	Institute of electronics and electrical engineers
IMT-2000	International Mobile Telecommunications – 2000
IP	Internet Protocol
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
IPTV	TV over Internet Protocol
ISDN	Integrated Services Digital Network
ITU	International Telecommunication Union
KPI	Key Performance Indicator
MAC	Medium Access Control
MAC-hs	Medium Access Control for High Speed Data
Max CIR	Maximum Carrier over Interference Ratio
MC-CIS	Multi Carrier - Carrier Interdependence Scheduling
MC-CSS	Multi Carrier - Carrier Separately Scheduling

MC-TIS	Multi Carrier - Totally Integrated Service
ME	Mobile Equipment
M-LWDF	Modified – Largest Weighted Delay First
MPEG	Compression of Audio and Visual Digital Data
MR	Maximum Rate
MR_{delay}	Maximum Rate delay
MR_{min}	Maximum Rate minimum
MSC	Mobile Switching Center
NACK	Non Acknowledge Message
NBAP	Node B Application Part
NRT	Non Real Time
NTB	Non - Time Based
OCNS	Orthogonal Code Noise Simulator
O&M	Operation and Maintenance
OVSF	Orthogonal Variable Spreading Factor
PBR	Provided Bit Rate
P-CPICH	Primary Common Pilot Channel
PF	Proportional Fair
PLMN	Public Land Mobile Network
PS	Packet Switch
PSTN	Public Switch Mobile Network
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quaternary Phase Shift Keying
R5	UMTS Release 5
R7	UMTS Release 7
R99	UMTS Release 99
RAB	Radio Access Bearer
RAM	Random Access Memory
RAN	Radio Access Network
REF	Reference Scenario
RNC	Radio Network Controller
RNS	Radio Network Server
ROHC	Robust Header Compression
RR	Round Robin
RRC	Radio Resource Control
RRM	Radio Resource Management

RT	Real Time
RTP	Real Time Protocol
SC - PF	Single Carrier - Proportional Fair
SF	Spreading Factor
SGSN	Serving GPRS Support Node
SHO	Soft Handover
SIR	Signal to Interference Ratio
SMS	Short Message Service
SINR	Signal to Interference Noise Ratio
SPI	Scheduling Priority Indicator
SRB	Signalling Radio Bearer
Sym	Symmetric
TB	Time Based
TBS	Transport Block Size
TC	Traffic Class
TDD	Time Division Duplex
THP	Traffic Handling Priority
TTI	Time to Transmission Interval
UDP	User Data Protocol
UE	User Equipment
UICC	Universal Integrator Circuit Card
UL	Uplink
UMTS	Universal Mobile Telecommunication Systems
Und	Unidirectional
USIM	Universal Subscriber Identity Module
UTRA	UMTS Radio Access
UTRAN	UMTS Terrestrial Radio Access Network
VLR	Visitor Location Register
VoIP	Voice over IP
WCDMA	Wideband Code Division Multiple Access
WIFI	Wireless Fidelity
WIMAX	Worldwide Interoperability for Microwave Access
WWW	World Wide Web

List of Symbols

α_k	Packet duration of packet k in Pareto model
α_j	Average ortoghonality factor in the cell
β	Maximum Interference value
β_k	Weibull shape parameter
Δf_{BS}	Total available bandwidth
ΔP	Estimated power increase from increasing DCH traffic
η_{DL}	DL load factor
λ	Total amount of voice calls
λ_k	Weibull scale parameter
ρ	Signal to Interference Noise Ratio
σ	Measure of convergence of the ∂ function
$\overline{\tau_u}$	Average Delay per user
∂	Network performance function
Ψ	Street orientation angle
b	Building Separation
C_u	Interface between the UICC and the UE
d	Distance between BS and UE
$\overline{D_d}$	Mean value of time interval between two consecutive packets inside a packet call
$\overline{D_{pc}}$	Mean value of reading time between two consecutive packet call requests in a session
D_{Pv}	Data packet volume
D_r	Drop rate
E_b	Energy per bit / user
E_c	Energy per chip / user
$EIRP$	Equivalent Isotropic Radiated Power
F	Receiver Noise Figure
F_{a_j}	Activity factor of user j
f_{BS}	BS Frequency
G	Geometry Factor
G_{div}	Diversity Gain
G_P	Processing Gain

G_r	Receiving antenna Gain
G_{rdiv}	Gain of the receiving antenna plus diversity gain
G_t	Antenna gain transmission
G_{UD}	Geographic users distribution
H_B	Building height
H_{BS}	Base Station height
H_{UE}	UE height
i_j	Ratio of inter to intra-cell interference for user j.
I_u	Interface which links the RNC and the MSC or SGSN
I_{ub}	Interface between each Node B and the RNC
I_{ur}	Direct interface between 2 RNCs
k	Packet Size in Pareto model
J	State variable
L_{BS}	Base Station location
L_c	Cable Losses between transmitter and antenna
L_{int}	Indoor Penetration Loss
L_{msd}	Multi-screen Diffraction Loss
L_o	Free space attenuation
L_{ori}	Attenuation caused by main street orientation with respect to the direct radio path
L_P	Path Loss
$L_{P_{tot}}$	Total Path Loss (after total interference margin applied)
L_{rts}	Roof to street diffraction and scatter loss
M	Total considered margins to be considered
m_k	Mean Average packet duration in Weibull model
M_{FF}	Fast Fading Margin
M_{SF}	Slow Fading Margin
$M_{I_{\eta_{DL}}}$	Interference Margin in DL
N_0	Noise Spectral Density
N_{bc}	Number of voice blocked calls
N_{BS}	Number of Base Stations
$\overline{N_d}$	Mean value number of packets within a packet call
N_{dc}	Number of codes for data services
N_{df}	Number of delayed frames
N_{ds}	Total number of dropped sessions
N_{du}	Number of users connected in a data service
N_{mc}	BS Total available codes
$\overline{N_{pc}}$	Mean value of random packet call requests per session
N_{sc}	Number of channels used for signalling

N_t	Thermal noise density
N_{tc}	Total number of calls
N_{tot}	Number of total users
N_{ts}	Total amount of sessions
N_u	Number of active users
N_{uj}	Number of users in Node B j
$N_{u_{max}}^{NodeB}$	Number of users in the most populated Node B
N_{vc}	Number of codes for voice service
N_{vu}	Number of users connected in voice service
P_b	Blocking Probability for voice users
$P_{HS-DSCH}$	Received power of the HS-DSCH summing over all active HS-PDSCH codes
P_{inter}	Received inter-cell interference
P_{intra}	Received intra-cell interference
P_k	Packet probability of size S_k
P_{margin}	Power safety margin,for unpredictable variations of the non HSDPA power
P_{noise}	Received noise power
P_{Tx}^{Bs}	Total BS transmission power
$P_{Tx_{max}}$	Maximum Tx Power
$P_{T_{x}\,pil}^{\;BS}$	Base Station pilot Transmission Power
$P_{T_x sig_{\,ch}}$	Transmission power associated to signalling common channels
$P_{T_x sig_{R5}}$	Transmission power associated to signalling common channels within R5
$P_{T_x sig_{R99}}$	Transmission power associated to signalling common channels within R99
$P_{T_x traf_{\,ch}}$	Transmission power associated to traffic channels
$P_{T_x traf_{\,R5}}$	Transmission power associated to traffic channels within R5
$P_{T_x traf_{\,R5_{max}}}$	Maximum transmission power associated to traffic channels within R5
$P_{T_x traf_{\,R5_{min}}}$	Minimum transmission power associated to traffic channels within R5
$P_{T_x traf_{\,R99}}$	Transmission power associated to traffic channels within R99
$P_{T_x traf_{\,R99_{max}}}$	Maximum transmission power associated to traffic channels within R99
$P_{T_x traf_{\,R99_{min}}}$	Minimum transmission power associated to traffic channels within R99
P_r	Available receiving power at antenna port
P_{RX}	Received power in DL
$P_{RX_{min}}$	UMTS minimum receiver sensitivity
P_t	Power fed to the transmitting antenna
$P_{T_{x\,target}}$	Dynamic radio network parameter
PT	Services Priority table
R	Cell Radius

R_{b_j}	Bit rate associated to service of user j
$\overline{R_{b_G}}$	Global average bit rate
$\overline{R_{b_{BS}}}$	Average bit rate for each BS
$\overline{R_{b_S}}$	Average bit rate per service
R_c	Chip rate of WCDMA
s_1^2	Log-Normal distribution parameter (total volume of data per session)
s_2^2	Log-Normal distribution parameter (inter-arrival time)
S_c	Scenario choice
S_w	Street width
SF_{16}	HS-PDSCH Spreading Factor which value is 16
SP	Service Penetration
T	Simulation Time (1 hour duration)
T_f	Time frame duration
t_{OFF}	Mean Silent Phase in Pareto model
t_{ON}	Mean Activity Phase in Pareto model
T_r	Busy hour voice traffic
T_{vc}	Average voice calls duration
u_1	Log-Normal distribution parameter (total volume of data per session)
u_2	Log-Normal distribution parameter (inter-arrival time)
U_u	Radio interface between the UTRAN and the UE
v_j	Activity factor of user j at physical layer

List of Programmes

Microsoft Visual Studio	C++ Integrated Development environment
Microsoft Excel	Calculation tool
Microsoft Visio	Design tool (e.g flowcharts, diagrams, etc)
Microsoft Word	Text Editor tool

Chapter 1

Introduction

This chapter gives a brief overview of the work. The context in which this thesis fits in and the main motivations are brought-up; work targets and scope are established. At the end of the chapter, the work structure is provided.

1.1 Overview

The development of mobile communications has traditionally been viewed as a sequence of successive generations. It almost two decades since Global System for Mobile Communications (GSM) was first commercially available. In 1999, 3rd Generation Partnership Project (3GPP) launched Universal Mobile Telecommunications System (UMTS), a Third Generation (3G) first release. Although Enhanced Data Rate for GSM Evolution (EDGE) had already been accepted by the International Telecommunication Union (ITU) as a part of the International Mobile Telecommunications-2000 (IMT-2000) family of 3G standards, as it has a peak data rate above 384 kbps (ITUs 3G definition); deployment was about to start in Europe and Asia, including Japan and Korea, in the same frequency band, around 2 GHz. The first full set of specifications was completed at the end of 1999, called Release 99 (R99). In Japan, during 2001, the first commercial network was opened for use in key areas; in Europe, it was at the beginning of 2002 for the pre-commercial testing phase and for commercial use during 2003. UMTS was designed for coexistence with GSM most of them being deployed on top of the existing GSM ones [HoTo07].

UMTS uses Wideband Code Division Multiple Access (WCDMA) as access technique, and was designed from the beginning to offer multi-service applications [HoTo04], unlike previous Second Generation (2G) systems, designed originally for voice communications. The variable bit rate and the mixture of traffic on the air interface presents new possibilities for both operators and users, as streaming traffic and voice over Internet Protocol (IP). New challenges in network planning and optimisation, for instance the lower delays, lead to the introduction of Quality of Service (QoS). Since its introduction, 3G cellular technology has been heralded for its ability to deliver more voice channels and higher-bandwidth pipes. But, in reality, operators have started to realise that, while 3G allows for high-quality voice and media streaming services, it is a poor fit for high speed data.

The Internet has transformed the way to access information, communication and entertainment services at home and at work. Broadband connections have made the Internet experience richer for millions of people, and in the coming years, millions more will turn to wireless technology to deliver their broadband experience. Many new services are based on multimedia applications, such as Voice over Internet Protocol (VoIP), video conferencing, video on demand, massive on line games and Peer-to-Peer.

R99, in theory, enables 2 Mbps, but in practice gives 384 kbps [HoTo07], therefore, the exponential growth of data communications over mobile phones forced a further development of systems that would be capable of offering higher capacity, bit rate and enhanced multimedia services, available to consumers anywhere and anytime. 3GPP, specified important evolution steps on top of R99, for Down and Uplinks (DL and UL), these enhancements being commonly known as High Speed Packet Access (HSPA). High Speed Downlink Packet Access (HSDPA) was set as a standard Release 5 (R5) with the first specifications made available in March 2002. HSDPA was commercially deployed in 2005, with an initial peak data rate of 1.8 Mbps, increased to 3.6 Mbps during 2006, and by the end of 2007, 7.2 Mbps were available [HoTo06]. Nowadays, the maximum peak data rate of 14.4 Mbps is already provided by some operators. In December 2004, the UL counterpart was launched by 3GPP in Release 6 (R6) with the Enhanced Dedicated Channel (E-DCH), also known as High Speed Uplink Packet Access (HSUPA). Following the success accomplished by HSDPA, by the end of 2007, HSUPA started to be deployed, pushing at first the UL bit rates up to 1.45 Mbps, and up to 5.7 Mbps in later releases [DPSB07]. Further HSPA evolution is specified in 3GPP Release 7 (R7) and its commercial deployment was started in 2009. HSPA evolution is also known as HSPA+. HSPA evolution in R7 brings a maximum bit rate of 28 Mbps in DL and 11 Mbps in UL [HoTo07].

HSDPA promises to bridge the gap between 3G and the Internet, providing an overlay for the existing protocol stack that makes the delivery of high-speed data access to many users in a cell a reality.

In many ways, HSDPA greatest strength is not how progressive it is, but how well it interoperates with previous versions. HSDPA has been developed to be backwardly compatible with existing UMTS networks. But HSPDA is not a simple software upgrade to 3G systems. In many aspects, the change from R99 to HSDPA is as dramatic as that from voice-only GSM to EDGE was: changing both modulation and the way packets are processed. There are parts of the HSDPA standard that are relatively simple to implement using existing hardware; but, taken as a whole, HSDPA will simply break many deployed architectures and will require new hardware. Most base stations (BSs) also known as Node Bs, will need significant upgrades to cope with the increased data bit rate and the consequences of moving to a more complex protocol. One of the most sensible items is the Radio Resource Management (RRM) and what is the best way to prioritise the access of each user service to the network, particularly when the traffic values are high. This subject is the main question under consideration in the present thesis, when Call Admission Control (CAC) needs to create priorities on accessing the network.

From a technical point of view, WCDMA brings advanced capabilities, not just higher bit rates, but also QoS differentiation for higher efficiency of service delivery, simultaneous voice and data capability, low delays with packet round trip, and an essential inter-working with existing GSM/GPRS networks. Nevertheless, it also improves basic voice due to interference control mechanisms, like frequency reuse, fast power control, and soft handover.

UMTS/HSDPA is a technique that presents several enhancements compared to UMTS/R99, such as Adaptive Modulation and Coding (AMC), which adapts the data rate according to the conditions and quality of the channel, and Hybrid Automatic Request (HARQ), responsible for retransmitting packets at the physical layer.

HSDPA increases the DL data rate within a cell to a theoretical maximum of 14.4 Mbps, with 5.7 Mbps on the UL. However, it is not about delivering Ethernet bandwidth to one fortunate user. What is important is the ability to deliver reliably, many sessions of high-speed and burst data to a large number of users within that cell. The changes that HSDPA enables include better quality, more reliable and robust data services. In other words, while realistic data rates may only be a few megabits per second, the actual quality and number of users achieved will improve significantly. As can be seen in Figure 1-1 HSDPA/HSUPA are the real first releases which establish a new concept of wide band services, serving in the same way as precursors of incoming releases, even with higher bit rates and new potentialities.

HSDPA is deployed on the top of the R99 network, which means it is not expensive as a standalone data network. It uses the same carrier, or for higher capacity and higher bit rates, it can use other additional carriers.

It shares all R99 network elements, just needing new software packages and enhanced Node B and Radio Network Control (RNC) hardware.

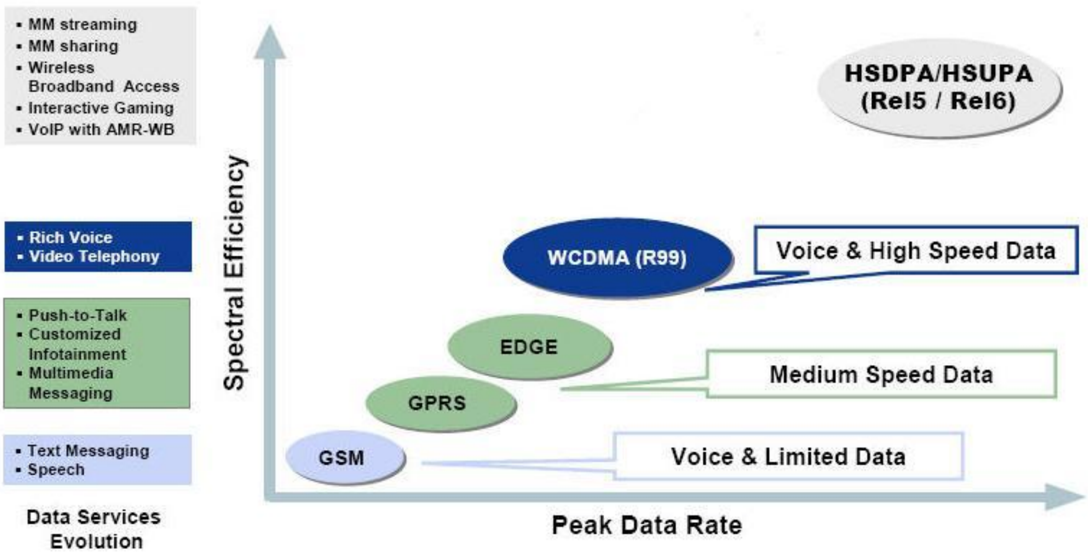

Figure 1-1 - Bit rates and spectral efficiency technology evolution (adapted from [QUAL04]).

1.2 Motivation and Contents

UMTS carriers are currently used by data services based on the HSDPA and HSUPA technologies, which can now deliver peak data rates up to 14.4 Mbps and 5.7 Mbps, respectively. The coverage of those services is unfortunately limited in the standard UMTS 2000 MHz band, even if by now some operators are developing some features like refarming to introduce UMTS in GSM band (called UMTS 900).

Traditional traffic modelling assumptions used for Circuit Switch (CS) voice traffic no longer hold true with the convergence of voice and data over Packet Switch (PS) infrastructures. Self-similar models need to be explored to appropriately account for the burstiness that packet traffic is expected to exhibit in all time scales. The task of demand characterisation must include an accurate description of the multiple user profiles, traffic source modelling, and service classes the network is expected to support, with their distinct geographical distribution, as well as forecasts of how the market should evolve over near and medium terms.

The appropriate assessment of QoS becomes a more complex issue, as new metrics and more intricate dependencies have to be considered when providing a varying range of services and applications that include voice, real-time (RT), and non-real time (NRT) data associated with development of complex RRM strategies and algorithms that makes modelling a complex but required task.

All those points have to be considered by the operator to obtain a proper dimensioning, resource allocation and rollout plan for system deployment. Additionally, any practical optimisation strategy has to rely on an accurate estimation of expected system performance in the presence of a specific multi-service profile. The assessment is proposed via performing a set of UMTS/HSDPA case studies using the proposed framework, in order to discuss validation and sensitivity of main network indicators to the multi-service and non-uniform aspects of the offered traffic.

From the operator point of view, a key interest is put in more accurate system performance and QoS assessment as a function of the multi-service and non-uniform traffic aspects. As more complex services are implemented over the UMTS infrastructure, more complex will be its modelling to accurately dimension network resources in order to meet the required QoS. PS and CS services have to be managed and dimensioned according to its main characteristics and specific needs.
The main purpose of this thesis is to study HSDPA functionalities and particularly its RRM features, in an environment with several users and service scenarios, taking into consideration a trade-off between coverage and capacity, to promote an efficient use of the available spectrum. The models were adapted to a multiple user simulator, allowing the analysis of HSDPA in a real network, by calculating several parameters, as the average bit rate, global and per service, user grade of service satisfaction and delay time for each user;

- analyse users impact in traffic variations;
- analyse number of users impact in specific scenarios corresponding to different available access technologies;
- comparison between different traffic profiles and associated service priorities.

The models were adapted to a multiple user simulation, allowing the analysis of HSDPA behaviour in real network, by calculating several parameters, as the user density and data volume packets to be transmitted. The simulator permits evaluating the influence of the variation of the system parameters in network performance, studying in each case the impact in resulting QoS.

The present work is focused on the dimensioning of an admission control model for voice and data users, following a service priority policy, according to network resources availability.

This thesis is composed of 5 chapters, including the present one and 3 annexes, identified from A to C. It is organised in the following way:

- In Chapter 2, one presents an introduction to UMTS and HSDPA. Some basic aspects are explained for both releases, which are supported in part by a common architecture. HSDPA new features, services and applications, capacity and coverage issues and RRM are described. At the end of the chapter the state of the art is present.

- In Chapter 3, the description of the multiple user models is presented. Afterwards, a detail description of the different simulator modules is performed, as well as the HSDPA RRM algorithms supporting the thesis. This chapter also addresses the characterisation of the functional blocks and input/output variables of the simulator. The way that real systems features are adapted to a simulation environment is explained, together with the simplifications and approximations that are necessary in a work of this type. In the last section of this chapter, the simulator assessment is evaluated.
- In Chapter 4, the results for each scenario in a UMTS network are processed for the presented simulation sets. Several different situations are performed, each of these consisting of characterising the response of the UMTS network (voice and HSDPA) to the variation of single input parameter. The results are interpreted, explained according to the theory, and compared to what should be expected considering theoretical principles.
- In Chapter 5, the simplifications and approximations on the design and building of the simulator are listed. Final conclusions and future research topics are drawn.

A set of annexes with auxiliary information is also included. In Annex A, the link budget used throughout this thesis is shown, whereas in Annex B, each traffic model service type characterisation is presented. In Annex C one describes the propagation model used in the present thesis.

Chapter 2

UMTS Basics

This chapter provides an overview of the UMTS Terrestrial Radio Access Network (UTRAN); first the system architecture is presented, followed by a description of services and applications within UMTS/HSDPA technologies. At the end a description is made about RRM state of the art.

2.1 Services and Applications

Multi-service traffic analysis with a given mixture of applications requires the definition of their main operation environments and respective deployment scenarios definition. An application is defined as a task that requires communication between two or more users, being characterised by service and also traffic and communications characteristics [VeCo02]. Services are firstly classified according to their characteristics: intrinsic time dependency or non time based, (TB or NTB), delivery requirements, RT or NRT, directionality (unidirectional or bidirectional, Und or Bid), symmetry of communications (symmetric or asymmetric, Sym or Asy), interactivity and number of intervenient players. In TB information, time is an integral part of the information to be communicated, typical examples being video, audio and animation, while NTB information includes images, graphics and text. In terms of delivery requirements, a RT application requires information delivery for immediate consumption; in contrast, NRT may be stored (perhaps temporarily) at the receiving points for later consumption. Services can be classified in the different categories according to given parameters, following 3GPP recommendations [3GPP05c] and [3GPP05d]. Due to a wide range of services and applications, the way to classify them, has to be distinguished for each application, QoS being necessary to accomplish.

The classification proposed by 3GPP, Table 2-1, consists of a separation in service classes depending the QoS they present: Conversational, Streaming, Interactive and Background.

Conversational services provide the means for Bid communications with RT (not store and forward) end-to-end information transfer between two users, or between a user and a service provider host. Messaging services offer user-to-user NRT communication between individual users via storage units with store and forward mailbox and/or message handling functions. Retrieval services provide users with the capacity to retrieve information stored in information centres (general available for public use). This information is sent to the user on demand only, with the possibility of being retrieved on an individual basis. Broadband services provide a continuous flow of information, which is distributed from a central source to an unlimited number of authorised receivers connected to the network. Each user can access this flow of information, but has no control over it. Cyclical services allow distributing information from a central source to a large number of users. However, the information is provided as a sequence of information entities (e.g., frames) with cyclical repetition. So the user has the ability of individual access to the cyclical distributed information, and can control start and order of presentation.

Many applications are interactive, even TV programs distribution, considered already with a return channel; conversational and messaging interactive services are one-to-one communications (except video conference, which is one-to-many); retrieval interactive services are one-to-one communications, whereas distribution services are one-to-many communications. It is also important to accurately describe the assumptions on latency/delay: absolute delay or latency is one of the key

QoS performance parameter. In order to provide interactive response to viewers, the response time between a user action and its effect should be less than 100 ms. To support network based video games, a response time of 50 ms or less is required to support twitch actions. This imposes minimum up and downstream bandwidth requirements. Latency only applies to RT applications; there is no latency requirement for NRT applications, although the associated delay is identified as a QoS issue [VeCo02].

Table 2-1 - UMTS QoS traffic classes main characteristics (adapted from [3GPP01] and [3GPP02a]).

Traffic Class	Conversational	Streaming	Interactive	Background
Fundamental characteristics	Preserve time relation (variation) between information entities of the stream Conversational pattern (stringent and low delay)	Preserve time relation (variation) between information entities of the stream	Request response pattern Preserve payload content	Destination is not expecting the data within a certain time Preserve payload content
Real time	Yes	Yes	No	No
Symmetry	Yes	No	No	No
Switching	Circuit switch	Circuit switch	Packet switch	Packet switch
Guaranteed rate	Yes	Yes	No	No
Bit rate [kbps]	12.2	12.2	64	64
Maximum Transfer Delay [ms]	80	250	-	-
Example of application	Voice	Streaming video	Web browsing	E-mails

Interactive classes and Background are mainly meant to be used by traditional Internet applications like WWW, Email, Telnet, FTP and News. Due to looser delay requirements, compare to Conversational and Streaming classes, both provide better error rate by means of channel coding and retransmission. The main difference between Interactive and Background class is that Interactive class is mainly used by interactive applications, e.g., Email or Web browsing, while Background is meant for background traffic, e.g., download of Emails or file downloading. The responsiveness of interactive applications is ensured by separating Interactive and Background ones. Traffic in the Interactive class has higher priority in scheduling than Background one, so Background applications use transmission resources only when Interactive applications do not need them. This is very important in wireless environments where the bandwidth is low compared to fixed networks [QUAL08]. Different environment can be seen as four types of market: residential, mixed, business and mass market. Mass market is driven by services like internet browsing (with huge variety of application and contents) and email. Residential traffic, in addition, includes a significant use of high quality TV services, with premium contents and several integrated services. A new Web generation is creating content, applications and business models. Broadband services allow users to share high amounts of data (videos and photos) and perform online games. In the meantime, VoIP has achieved mass market status and VoIP players are also providing new services – IPTV – where each user is free to choose its own favourite contents. Typical examples of IPTV are video on demand and Internet TV.

End users are increasingly accustomed to broadband web browsing performance through fixed Internet connections. In comparison, the browsing performance over early deployments of cellular systems supporting packet-switched services has suffered from limited bit rate and rather significant latencies [HoTo04].

Since the bit rates are one of the major issues in new releases development, in Table 2-2, one shows the association of such indicator per service. In those tables, one expresses some of the most common and most used wide band services, each one with its maximum and minimum bit rates.

Despite of their maximum and minimum bit rates, services are also characterised by their performance and feasibility to make applications work properly and final users confident in applications behaviour. From a technical view point, there are services that are sensorial, meaning that their final result is felt by the user a typical example being FTP, where a degradation on file transfer is immediately perceived by the service owner; on the other hand, there are services, where a variation on a parameter like bit rate is not felt by the final user, who keeps to proceed with its application; in such case, the best examples are voice and streaming which do not need a high bit rate to fulfil their requisites for normal functioning.

Streaming technologies are increasing with the Internet growth. It is a multimedia streaming technique, which uses a continuous stream transmission and buffers in receivers, enabling the client browser to access the data before the entire file is complete. So, the DL traffic is more important than the UL one, being an asymmetric technology.

The interactive class encompasses user data request from a remote appliance (e.g., Web browsing), it means a communications scheme that is characterised by the request response pattern of the end user. This class is one with very asymmetric traffic, being very tolerant in terms of delay. It is defined by requesting response patterns and preservation of payload contents. To be able to provide a good service, delay should be lower than 4 to 7s [3GPP03a].

Background class encompasses the services where transmission delays are not critical (e.g., Short Message Service (SMS), e-mail); it is the least delay-sensitive of all, since practically there are no delay requirements. The delay can range between a few seconds to several minutes. On contrary to interactive class the end user is not waiting for a response within a short time. This type of services only spends the network resources when they are not needed for application from the former classes. This class, like the interactive class, is intolerant to transmission errors. Values presented in Table 2-2, are typically used in several scenarios as references for different models and applications.

Table 2-2 - Maximum and minimum DL bit rates for each service (adapted from [Duar08]).

Services	Maximum Bit rate [Mbps]	Minimum Bit rate [Mbps]
Voice	0.048	0.012
Web	-	0.064
Streaming	5.600	0.300
E-mail	-	0.064
FTP	-	0.384
Chat	0.384	0.064
P2P	-	0.384

In Table 2-3, one shows some important aspects involved in the described services, with influence in their performance and in radio resource management.

Table 2-3 - Examples DL traffic models characterisation according to each service type (typical values) (extracted from [Duar08]).

Service	Parameter	Value
Voice	Average call duration [s]	120
	Average number of calls per user	0.825
Web	Average page size	300
	Average reading time [s]	40
	Average number of pages per session	10
FTP	Average file size [MB]	10
	Average number of files per session	1
P2P	Average file size [MB]	12.5
	Average session initiation time [s]	30
Chat	Average MSN message size [bytes]	50
	Average number of received messages during one session	25
E-mail	Average file size [kB]]	100
	Average number of e-mails per session	1
Streaming	Average video duration [s]	150
	Average video size [MB]	9.6
	Average number of videos per session	3

2.2 Architecture

The UMTS network supports two radio interfaces modes: Frequency Division Duplex (FDD) and Time Division Duplex (TDD). In this work, only FDD is considered, since TDD mode, which was designed to provide high data rates, was replaced by HSPA in FDD [3GPP07].

Three high level architecture modules compose the network [3GPP02a]: User Equipment (UE), UMTS Radio Access (UTRA) and Core Network (CN). The UE interacts with the user. UTRAN is responsible for radio interface, and allows connections to the CN, which is responsible for interaction with external networks, like other Public Land Mobile Networks (PLMN), Public Switched Telephone Networks (PSTN), Integrated Services Digital Networks (ISDN) and Internet. The structure of the network is shown in Figure 2-1.

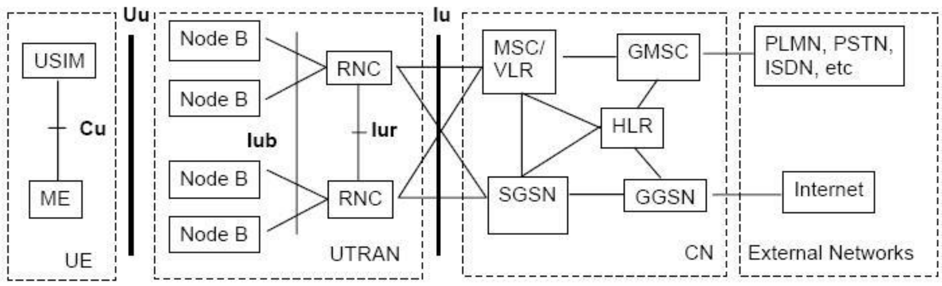

Figure 2-1 - UMTS Architecture (extracted from [HoTo04]).

The Node B performs physical layer (L1) functions, such as channel coding and interleaving, modulation, spreading, radio transmission, and reception [3GPP3b] and is also involved in Operation and Maintenance (O&M) functions [3GPP5a]. RNC is responsible for controlling the radio resources, i.e., frequencies, scrambling codes, spreading factors (SF), and power control. RNC controls Node Bs and performs RRM as well. The UTRAN architecture consists of one or several Radio Network Servers (RNS), which contain a RNC and one or several Node Bs. CN network elements are:

- Home Location Register (HLR) – database that stores user information.
- Mobile Switching Centre/Visitor Location Register (MSC/VLR): MSC is responsible for switching voice and data connections in the CS domain; VLR is a database containing all active network users.
- Gateway MSC (GMSC) – switch for connection to external networks in the CS domain;
- Serving GPRS Support Node (SGSN) – the same function as MSC/VLR but for the PS domain.
- Gateway GPRS Support Node (GGSN) – equivalent to GMSC in the PS domain.

UMTS architecture has not really gone through many changes to introduce HSDPA. Only, a new Medium Access Control (MAC) layer has been introduced in R5 to implement HSDPA, which improves capacity and spectral efficiency, being deployed together with R99, sharing all network elements. It requires software upgrade at the Node B and MSC on the network side, and a new UE on the user one.

2.3 Radio Interface

UMTS uses WCDMA, a wideband Direct-Sequence Code Division Multiple Access (DS-CDMA) spread spectrum air interface, with a chip rate of 3.84 Mcps leading to a radio channel of 4.4 MHz and separation of 5 MHz. The frequency bands are [3GPP05a] [1920,1980] MHz for UL and [2110,2170] MHz for DL. WCDMA capabilities include: high bit rates, low delay, QoS differentiation, smooth mobility for voice and packet data, RT services and high capacity capability.

Power control, soft and softer handover are key features in the air interface. It applies two types of power control: closed loop, to avoid the use of excessive power, and thus, an increase of interference (the near-far problem), and outer loop, to adjust the Signal to Interference Radio (SIR) of each individual link, instead of defining a SIR target for the worse case, resulting on a waste of capacity [Agil07].

HSDPA offers improved bandwidth to the end user, improved network capacity to the operator and improved interactivity for data applications, a new technology arrived. It is referred to as HSPA, and it is an upgrade from the initial version of WCDMA. Although designed for NRT traffic, simulations show that HSDPA provides enough capacity for low bit rate applications, as VoIP. HSDPA is backwards compatible with R99, so voice and data applications can still run on the upgraded networks, and the same radio channel will support R99 and HSDPA services simultaneously.

Furthermore, some logical functions traditionally performed in RNC, have been moved to the Node B. This new MAC layer is called Medium Access Control for HSDPA, (MAC-hs) being located in the Node B. The physical layer has been upgraded with new physical channels with shorter frame size (2 ms) and with link adaptation functions, such as adaptive modulation and coding. Link adaptation replaces two important features of the traditional system: variable spreading factor and fast power control. Another disabled feature in HSDPA is soft-handover [Agil07]. MAC-hs, includes four main functions [3GPP3c]:

- Flow control
- Scheduling and priority handling
- HARQ functionality handling
- Transport format and resource combination selection.

On the other hand, MAC-hs has slightly different functions at the UE side. These functions are HARQ handling, reordering and de-assembly. The output of the MAC-hs is the High Speed Downlink Shared Channel (HS-DSCH). Furthermore, a HS-DSCH channel is associated to DL and UL signalling channels. Besides these channels, HS-DSCH channel needs to have a Dedicated Channel (DCH) associated to carry user and system information in the UL [Guti03].

The result of adding HSDPA to UMTS is similar to that of adding Extended GPRS (E-GPRS) to GSM: improvement in peak data rates and the overall increase in system capacity, particular in small cells. To improve UMTS system performance, HSDPA makes a number of changes in the radio interface, mainly affecting the physical and transport layers [Agil07].

- Shorter radio frame.
- New high-speed DL channels.
- Use of 16 Quadrature Amplitude Modulation (QAM) in addition to Quadrature Phase Shift Keying (QPSK).
- Code multiplexing combined with time multiplexing.
- A new UL control channel.
- Fast link adaptation using Adaptive Modulation and Coding (AMC).
- MAC scheduling functions moved to Node B.

Higher data rates are accomplished through the use of a new higher order modulation, the 16QAM with 4 bits per symbol that can only be used under good radio channel conditions. QPSK is mainly used to maximise coverage and robustness. HSDPA introduces AMC, which adjusts the modulation and coding schemes to the radio channel conditions, and, together with 16QAM, allows higher data rates [HoTo06]. Moreover, due to the fact that scheduling, retransmissions, modulation and coding decisions are taken nearer the air interface, HSDPA also reduces significantly system delays. As a result, HSDPA shows to be the best solution to carry NRT, for instance Interactive Class or Background classes. HSDPA also aims to carry Streaming traffic in further releases; however, there may be certain limitations depending on the type of Streaming traffic that is sent.

HSDPA requires other link adaptation mechanisms (due to the exclusion of power control), to adapt the transmitted signal parameters to the continuously varying channel conditions.

As closed power control is not present, the channel quality variations must be minimised across the Time to Transmission Interval (TTI), which it is accomplished by reducing its duration from 10 ms in R99 down to 2 ms. The fast HARQ technique is added, which rapidly retransmits the missing transport blocks and combines the soft information from the original transmission with any subsequent retransmission before the decoding process [Guti03].

Some assumptions have to be taken in advance, such as: users with various traffic profiles, with a multi-service perspective, ranging from voice (in CS), to data at various bit rates (PS), non-uniformly distributed in the service area. For R99, coverage and capacity analysis is well defined by a link budget, which can estimate the total BS power requested by users, but for HSDPA this approach is no longer possible, since the channel is shared and the signal is adjusted to instantaneous channel conditions, which means that one cannot have a specific value for the well-known ratio, Energy per bit versus Noise spectral density, $E_b/_{N_0}$ for each service, and a different approach for the Signal to Interference Noise Ratio (SINR) value must be used. In the implementation of HSDPA, one can have two types of carriers, either sharing power between HSDPA and R99 users, or exclusively serving users connected to one of these versions. In the case of shared power, HSDPA can either be accommodated in a fixed percentage of the total BS power, or use the power remaining from R99 at that moment. In either case, one must take into consideration that a certain amount of power has to be served for signalling and control, and that HSDPA requires additional power for this purpose. In order to decide whether a service is using R99 or HSDPA, a threshold value (for example bit rate) has to be taken, i.e, data services requesting a bit rate lower than the threshold are allocated to R99, whether those with higher or equal values are served by HSDPA [LCCS06].

New channels are needed for HSDPA operation. HSDPA is always operated with the DCH from R99, DL packet access operation running in parallel, which can be used to carry CS services and the Signalling Radio Bearer (SRB). The new channels are [HoTo06]:

- High-Speed DL Shared Channel (HS-DSCH): transport channel carrying user data, being mapped onto the High Speed Physical DL Shared Channel (HS-PDSCH). For multi code operation, a fixed SF of 16 is used. When a user has no data to be transmitted, there is no transmission and the resource is allocated to another user during the 2 ms TTI.
- High-Speed Shared Control Channel (HS-SCCH): logical channel carrying time-critical signalling information, divided in two parts. The use of SF=128 enables carrying 40 bits per slot, being divided into 3 slots and having a 2 slots offset compared with the HS-DSCH. The first part contains the information that is needed before the reception of the HS-DSCH, as which codes to de-spread and modulation information, while the second part contains less urgent information
- High-Speed Dedicated Physical Control Channel (HS-DPCCH): physical channel containing control information for UL. It has a 3 slot structure, a fixed SF of 256 and is divided into two parts. The first one, consisting of one slot, carries the HARQ information. The second one, consisting of the remaining two slots, carries the Channel Quality Information (CQI) feedback.

Terminals supporting HSDPA are divided into 12 categories, which were specified by 3GPP, shown in Table 2-4.

Table 2-4 - HSDPA terminal capability categories, (extracted from [HoTo06]).

<table>
<tr><th>Category</th><th>Supported modulation</th><th>Maximum number of parallel codes per HS-DSCH</th><th>Minimum inter- TTI interval</th><th>Achievable maximum data rate [Mbps]</th></tr>
<tr><td>1</td><td rowspan="10">16 QAM</td><td rowspan="6">5</td><td rowspan="2">3</td><td rowspan="2">1.2</td></tr>
<tr><td>2</td></tr>
<tr><td>3</td><td rowspan="2">2</td><td rowspan="2">1.8</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td><td rowspan="6">1</td><td rowspan="2">3.6</td></tr>
<tr><td>6</td></tr>
<tr><td>7</td><td rowspan="2">10</td><td rowspan="2">7.2</td></tr>
<tr><td>8</td></tr>
<tr><td>9</td><td rowspan="2">15</td><td>10.2</td></tr>
<tr><td>10</td><td>14.4</td></tr>
<tr><td>11</td><td rowspan="2">QPSK</td><td rowspan="2">5</td><td>2</td><td>0.9</td></tr>
<tr><td>12</td><td>1</td><td>1.8</td></tr>
</table>

Those UE categories are differentiated from each other by the modulation, number of supported codes, minimum inter-TTI interval, and ARQ type at maximum data rate.

Although HSDPA uses a fixed SF of 16, from these 16 available codes, only 15 can be allocated to data transmission, as 1 code is needed for HS-SCCH transmission. From the BS view point, all the 15 codes can be allocated; however, for the UE, the allocated codes can vary within 5, 10 or 15 depending on the UE category.

2.4 Capacity and Coverage

Capacity and coverage are closely related in UMTS networks, and therefore both must be considered simultaneously. The coverage problem is directly related to the power availability, so power demand deriving from the system load level should be in accordance with the planned coverage. Therefore, the required transmitted power has to be lower than the maximum power available at the UE side. Coverage and capacity are user dependent. The number of users, its location and handling services define the interference on the radio interface, thus, coverage and available capacity. Three main parameters are limiting capacity: the number of available codes in DL, the BS transmission power and the system load [HoTo07]. The number of available codes in DL defines the maximum number of simultaneous active services handled by users. The number of channelisation codes is given by the SF: the higher the service bit rate is, the smallest the SF will be, leading to a decrease of the available SF, hence, of the allowed codes, since all the branch of the allocated code cannot be used due to the orthogonality. The use of different scrambling codes in the same sector is not recommended, because it would decrease the orthogonality of DL channels, and hence, increase interference. As the number of the available SF is only 512 (per carrier) this may be a limitative factor, mainly when the HSDPA shares the same frequency in one cell.

In a system where there is a carrier to share, the code management requires a higher accuracy: the number of codes employed for each service type, voice or data, depends on the number of active users for a particular service and has to be higher than the total amount of codes for a channelisation system type, otherwise the system is congested, (2.1).

$$N_{mc} \geq N_{vc} \times N_{vu} + N_{dc} \times N_{du} \quad (2.1)$$

where:

- N_{mc} : maximum number of codes for a channelisation system type.
- N_{vc} : number of codes for voice service.
- N_{dc} : number of codes for data services.
- N_{vu} : number of users connected in voice service.
- N_{du} : number of users connected in a data services.

Each service is characterised by a specific channelisation type according to the demanded radio resources, with direct impact in the final bit rate delivered by the network to the user; that is why for each data service, when there is a variation in the codes allocation, different SFs are employed, trying to minimise the difference between requested service bit rate and the effective bit rate processed by each final user; the correspondence between each service and the respective channelisation type is defined in Table 2-5.

Table 2-5 - Correspondence of each service and associated SF.

Service	Channelisation type
Voice	SF128
Data@384 kbps	SF8
HSDPA	SF16

As a shared channel used, end user performance depends on the number of active users. For DL, the load factor is defined by [HoTo04]:

$$\eta_{DL} = \sum_{j=1}^{Nu} v_j \left(\frac{\left({E_b}/{N_0} \right)_j}{{R_c}/{R_{b_j}}} \right) \left[\left(1 - \alpha_j \right) + i_j \right] \tag{2.2}$$

where:

- η_{DL}: DL load factor.
- N_u : number of active users.
- E_b: energy per bit / user.
- N_0 : noise spectral density.
- v_j : activity factor of user j at physical layer.
- R_c : WCDMA chip rate.
- R_{b_j}: bit rate associated to service of user j.
- α_j : average ortoghonality factor in the cell.
- i_j : ratio of inter to intra-cell interferences for user j.

The load equation predicts the amount of noise rise over thermal noise due to interference. The noise rise is given by the interference margin:

$$M_{I_{\eta_{DL}[dB]}} = -10\, log(1 - \eta_{DL}) \tag{2.3}$$

Shared transmission resources per cell are power and channelisation codes. Ideally, these transmission resources should be dynamically shared between HSDPA and R99 depending on the offered traffic and the QoS attributes for the services carried on these two domains. The resource allocation algorithms are an integral part of RRM algorithms, which includes both algorithms at the Node-B and at the centralized RNC [PeMi06]. The total transmit power is divided between common channels, RT DCH, NRT DCH and HSDPA. The power for common channels includes the transmit power for the primary common pilot channel (P-CPICH).

$$P_{Tx}^{Bs} = P_{T_x sig_{ch}} + P_{T_x traf_{ch}} \tag{2.4}$$

$$P_{T_x traf_{ch}} = P_{T_x traf_{R99}} + P_{T_x traf_{R5}} \tag{2.5}$$

$$P_{T_x sig_{ch}} = P_{T_x sig_{R99}} + P_{T_x sig_{R5}} \tag{2.6}$$

where

- P_{Tx}^{Bs} : total Base station transmission power.
- $P_{T_x sig_{ch}}$: transmission power associated to signalling common channels.
- $P_{T_x sig_{R99}}$: transmission power associated to signalling common channels within R99.
- $P_{T_x sig_{R5}}$: transmission power associated to signalling common channels within R5.
- $P_{T_x traf_{ch}}$: transmission power associated to traffic channels.
- $P_{T_x traf_{R99}}$: transmission power associated to traffic channels within R99.
- $P_{T_x traf_{R5}}$: transmission power associated to traffic channels within R5.

Usually $P_{T_x sig_{ch}}$ shall not exceed 30% of the total power available for both signalling and traffic channels in both Releases (R99 and R5). Such trade off has the purpose of not wasting too much capacity and manages the resources properly. The average RT DCH power is controlled by the Admission Control algorithm while the NRT DCH power is controlled by the PS algorithm in the RNC. The Node-B can be configured to report average measurements of the total transmit power and the non-HSDPA power.

$$P_{Tx}^{Bs} = P_{T_x traf_{R99}} + P_{T_x traf_{R5}} + P_{T_x sig_{R99}} + P_{T_x sig_{R5}} \tag{2.7}$$

$$P_{T_{x\,target}} > P_{T_x traf_{R99}} + P_{T_x traf_{R5}} + \Delta P \tag{2.8}$$

where

- $P_{T_{x\,target}}$: dynamic radio network parameter.
- ΔP : estimated power increase from increasing the DCH traffic.

Thus a capacity request for admitting new RT DCH connections or increasing the bit rate of NRT, has to take in account, not just the present active users, but also future incoming new users and associated signalling within, necessary to initiate a new call or data session.

DCH connections are only granted under (2.8) conditions. The RNC can dynamically allocate HSDPA power to the MAC-hs by sending messages to the Node B. If no messages with HSDPA power allocation are sent, the Node B is allowed to allocate all unused power to HSDPA. The non–HSDPA power for R99 channels is time-variant due to, e.g. fast closed loop power control. The RNC based HSDPA power allocation is not capable of tracking the fast R99 power variations, since the RNC only receives average power measurements from the Node-B. The RNC based HSDPA power allocation algorithm therefore has to reserve power control headroom before allocating HSDPA power. The fast Node-B based HSDPA power allocation is thus the preferred solution, as it results in more HSDPA transmission energy – improved HSDPA coverage. The fast Node-B based HSDPA power allocation algorithm adjusts the HSDPA power, every T seconds according to:

$$P_{T_x traf_{R5}} = P_{Tx}^{Bs} - \left(P_{T_x traf_{R99}} + P_{T_x sig_{ch}}\right) - P_{margin} \quad (2.9)$$

where:

- P_{margin} : power safety margin, needed to account for unpredictable variations of the non HSDPA power.

The RNC is in control of the overall power sharing between HSDPA and R99 via setting of $P_{T_{x_{target}}}$ which limits the maximum R99 power and therefore also explicitly determines the available power for HSDPA [PeMi06]. Another difference between HSDPA and R99 are the metrics used to access network performance. In R99, the most well known metric is, E_b/N_0, but this is not appropriate for HSDPA, since the bit rate may change every TTI. The metric used to evaluate HSDPA performance, fundamental for link budget planning and network dimensioning, is the average HS-DSCH SINR after HS-PDSCH de-spreading, being given by the average SINR:

$$\rho = SF_{16} \frac{P_{HS-DSCH}}{(1-\alpha_j)\, P_{intra} + P_{inter} + P_{noise}} \quad (2.10)$$

where:

- SF_{16} : HS-PDSCH SF (which value is 16).
- $P_{HS-DSCH}$: received power of the HS-DSCH summing over all active HS-PDSCH codes.
- P_{intra} : received intra-cell interference.
- P_{inter} : received inter-cell interference.
- P_{noise} : received noise power.

For performance analysis, the required instantaneous HS-DSCH ρ measurement is also used, defined as the SINR on the HS-DSCH channel to accomplish a specific Block Error Rate (BLER) target for the number of HS-PDSCH codes, modulation and coding scheme used [HoTo06].

Regarding coverage aspects, it is important to assure a minimum data rate at the cell edge. By using (2.9)**,** it is possible to express the Node B HS-DSCH transmit power as a function of the SINR, and consequently as a function of the desired data range at the cell edge:

$$P_{HS-DSCH} \geq \rho \left[1 - \alpha_j + G^{-1}\right] \frac{P_{Tx}^{Bs}}{SF_{16}} \quad (2.11)$$

where:

$$G = \frac{P_{Tx}^{Bs}}{P_{inter} + P_{noise}} \quad (2.12)$$

Given the allocated power and channelisation codes for HSDPA transmission, it is assumed that one scrambling code is assigned per cell with an associated set of orthogonal channelisation codes. The primary objective has been to minimise the blocking probability of R99 capacity request.If R99 capacity request is initially code blocked (due to lack of available channelisation codes) the required code is immediately made available by down-grading the number of HSDPA codes, if R99 has priority over HSDPA. RT DCH by default is given priority. HSDPA codes are only allocated if there are allocated calls on HSDPA in the cell [PeMi06]. In the Orthogonal Variable Spreading Factor (OVSF) code tree, the configuration provides up to 15 SF16 codes for HS-PDSCH, and up to 4 SF128 for HS-SCCH. All remaining OVSF codes can be used for non-HSDPA services (Speech, multi-RAB, etc).

2.5 Radio Resource Management

2.5.1 General scheduling algorithms

RRM algorithms are responsible for the efficient usage of the air interface resources being needed to maintain the planned coverage area, increase the capacity and end user performance, being also responsible for assuring network stability. The family of RRM algorithms can be divided into the resource allocation, admission control, congestion control, power control, handover control and packet scheduling functionalities. There are many packet scheduling algorithms, although all of them derive from the basic ones: the Round-Robin (RR), Proportional Fair (PF) and the Maximum Carrier-over-Interference (Max CIR) [HoTo04].

RR method is an algorithm that selects the user packets in a round robin fashion. In this method, the number of time slots allocated to each user can be chosen to be inversely proportional to the user data rates, so the same number of bits is transmitted for every user in a cycle. Obviously, this method is the “fairest” in the sense that the average delay and bit rate would be the same for all users (users are scheduled with an equal probability, independent of radio channel conditions). However, there are two disadvantages associated with the round-robin method. The first is that it disregards the conditions of the radio channel for each user, so users in poor radio conditions may experience low data rates, whereas users in good channel conditions may not even receive any data until the channel conditions turn poor again; it is obviously against the spirit of the HSDPA and it would lead to the lowest system bit rate. The second disadvantage of the RR scheduler is that there is no differentiation in the quality of services for different classes of users. One of the most well known scheduling algorithms in HSDPA is the Max CIR method: the scheduler attempts to take advantage of the variations in the radio channel conditions for different users to the maximum, and always chooses to serve the user experiencing the best channel condition, that is, the one with maximum CIR. Apparently, the Max CIR scheduler leads to the maximum system bit rate, but is the most unfair, as users in poor radio conditions may never get served or suffer from unacceptable delays. The maximum carrier to interference scheduler monopolises the cell resources for a small subset of users, and there may be a number of users at the cell edge that will never be scheduled.

One different proposal is the PF scheduling algorithm. This method takes into account both the short-term variation of the radio channel conditions and the long-term bit rate of each user. In this method, the user with the largest ratio data rate in the current time slot and the average data rate for the user in the past average window is served first. The size of the average window determines the maximum duration that a user can be starved from data, and as such it reflects the compromise between the maximum tolerable delay and the cell bit rate. According to this scheme, if a user is enjoying a very high average bit rate, its ratio will probably not be the highest. Then it may give way to other users with poor average bit rate and therefore high ratio in the next time slots, so the average bit rate of the latter can be improved. On the other hand, if the average bit rate of a user is low, the ratio could be high and it might be granted the right of transmission even if its current channel condition is not the best. It provides a trade-off between fairness and achievable HSDPA cell bit rate and provides a significant coverage extension [Shar06].

In the comparison of scheduling gains there is an advantage between PF scheduling over RR also called multi user diversity gain; multi user diversity gain comes from scheduling users when they experience relatively good SINR, while avoiding scheduling users that experience deep fades. Multi user diversity gain is only available at moderate UE speeds, where the packet scheduler is able to track fast fading radio channel variations. At high UE speeds, where the packet scheduler can no longer track radio channel variations by monitoring received CQI reports from UEs, multi user diversity gain is marginal. The same observation is true for stationary users where the channel is constant. Multi user diversity gain also depends on the number of simultaneous active HSDPA users in the cell [HoTo06].

2.5.2 RNC Algorithms

The RNC executes some important functions fundamental for a right service management. Some of the most important are: mobility, admission and congestion control, power control and packet retransmission data.

Power Control is needed to keep the interference levels at minimum in the air interface and to provide the required QoS. It is important that each transmission is carried out with the minimum required power to ensure the quality requirements for the considered service. Admission Control decides the admission or rejection of requests for set-up and reconfiguration of radio bearers. Since the maximum cell capacity is intrinsically connected to the amount of interference or, equivalently, the cell load level; the use of admission control algorithms is based on measurements and estimates the current network load situation as well as on the estimation of the load increase that the acceptance of the request would cause. A connection request can be admitted only after gaining permissions from the corresponding UL and DL algorithms. In the case of DL there is a hard limiting factor: the OVSF codes availability [RSAG05].

Congestion control in the radio interface is caused by excessive interference, thus algorithms need continuously to monitor the network status in order to correct overload situations.

RRM algorithms are responsible for mapping physical layer enhancements introduced by HSDPA to a capacity gain while providing attractive end user performance and system stability. With the arrival of HSDPA, RRM suffered some changes compared to R99. The most relevant is that RRM algorithms take place in the I_{ub} interface. Those functions have to be differentiated between RNC and Node B. In this context, HSDPA resource allocation refers to the functionality that allocates power and channelisation codes to the Node B for HSDPA transmission in each cell. HSDPA admission control is different from R99 DCH admission control algorithms, since HSDPA relies on a shared channel concept. Mobility management for HSDPA is another new functionality, since data are only transmitted from one cell to the UE at a time, and effective Node B buffer management is needed due to the distributed architecture [HoTo06].

Before the Node B can start transmitting data on the HS-DSCH, the controlling RNC needs to allocate channelisation codes and power for the transmission of HSDPA. Hence, the allocation of channelisation codes for HSDPA transmission only requires signalling between the RNC and the Node B. In general, it is advantageous to allocate as many HS-PDSCH channels as possible, since the spectral efficiency of the HS-DSCH is thereby improved. On the other hand, channelisation codes reserved for HS-PDSCH transmission cannot be simultaneously used for transmission of R99 channels, so allocation of many HS-PDSCH codes might result in call blocking of R99 users. Fortunately, if channelisation code congestion is detected the controlling RNC release some of the allocated HS-PDSCH codes to prevent R99 blocking, in voice or video connections.

HS-DSCH transmission to multiple users in parallel during a single TTI requires multiple HS-SCCH codes and multiple HS-PDSCH codes. Code multiplexing is found useful for scenarios where a Node B has more HS-PDSCH codes allocated than what is supported by HSDPA mobiles; the node B may support 10-15 HS-PDSCH codes, while the HSDPA terminal supports from 5 to 15 HS-PDSCH codes. In most cases, the scarcest DL transmission resource is power. The power for real time DCH is managed by RNC admission control, while non-real time DCHs are controlled by the RNC packet scheduler. The power for non-real time DCH is characterised as controllable power, since it can be adjusted via bit rate modifications, while the power for common channels and real time DCH is considered to be non-controllable. There are two main options for allocating HSDPA transmission power to each Node B cell:

- The controlling RNC allocates a fixed amount of HSDPA transmission power per cell. The Node B can afterwards use this power for transmission of HS-SCCH(s) and HS-PDSCH(s).
- The controlling RNC power may update HSDPA transmission power allocation any time later.

If the Node B does not explicitly allocate HSDPA transmission power to the Node B, the Node B is free to use any unused power in the cell for HSDPA transmission; as a result the Node B can adjust HSDPA transmission power, to equal in the limit, the maximum transmitted power minus the power used for transmission of non-HSDPA channels.

In the second scenario, the total available carrier transmission power can be better used, since the Node B can quickly adjust HSDPA transmission power based on short-term measurements of the current power used by all non-HSDPA channels. The state-of-the-art solution does therefore call for a dynamic algorithm at the RNC that can adjust power sharing between HSDPA and non-HSDPA channels based on the QoS attributes for ongoing calls on these two channel types. QoS for R99 DCH is conducted as a function of the user´s traffic class (TC), traffic handling priority (THP), allocation retention priority (ARP), and potentially also other bearer attributes. These QoS parameters from the I_u interface are not available in the Node B for MAC-hs packet scheduling. New QoS parameters have been defined for the I_{ub} interface between the RNC and the Node B. HSDPA QoS parameters in the I_{ub} are: Guaranteed bit rate (GBR), Scheduling priority indicator (SPI) and Discard timer (DT).

SPI takes values in the range [0...15], where a high number indicates high priority. The DT specifies the maximum time that a packet is allowed to be buffered in the Node B MAC-hs before it should be discarded. For the Conversational and Streaming traffic classes, the HSDPA GBR parameter can be set according to the bit rate requirement specified in the UMTS bearer attributes for this traffic class. A high SPI could be allocated to video streaming or other RT services, while Internet access applications could be assigned a low SPI value (SPI can also be adjusted dynamically during a packet call). The value of GBR and SPI for new HSDPA users who are requesting access can also be used in the HSDPA admission control decision [HoTo06].

Admission control determines whether a new user with HSDPA terminals is admitted in the cell, and what type of service, DCH or HSDPA, the user will get. Voice or CS video users are served with DCH. For PS and according to QoS requirements, the algorithm chooses the type of service based on Node B measurements like total power, non-HSDPA and HSDPA required power, as well as pilot measurements from the UE. The RNC has then to decide if the Node B has enough capacity to serve the new user with the required service, maintaining all other users with their specific QoS. If only best effort traffic with no strict QoS requirements are transmitted on HSDPA, then the admission control algorithm can be made fairly simple by only checking the availability of RNC and Node B hardware resources to serve a new HSDPA user. If more demanding services with stricter QoS requirements are considered for HSDPA, then a more advanced admission control algorithm is needed to ensure that QoS requirements for existing HSDPA users in the cell as well as the requirements of the new user can be fulfilled after the potential admission. Such algorithm can provide attractive capacity for high VoIP users. Mobility management is responsible for handovers. HSDPA does not support soft handover, since HS-DSCH and HS-SCCH transmission take place only from a single cell, called the serving HS-DSCH cell. In Figure 2-2 one shows the most important parameters associated to Admission Control.

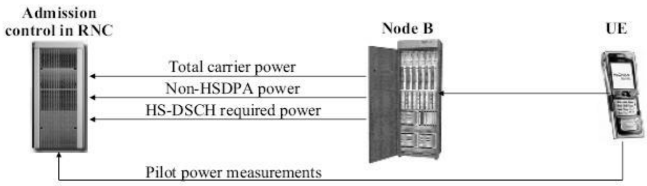

Figure 2-2 - Parameters associated with Admission Control (extracted from [HoTo06]).

2.5.3 Node B Algorithms

At the Node B, there are several functionalities, with impact in network performance and user quality of service.

In AMC, the modulation and coding rate are adapted to the instantaneous channel quality instead of adjusting the power; HARQ is a technique added, which rapidly retransmits the missing transport blocks and combines the soft information from the original transmission with any subsequent retransmission before the decoding process

Code Management is devoted to managing the OVSF code tree, used to allocate physical channel orthogonality among different users. The drawback is the limited number of such codes.

Packet scheduling algorithms are devoted to deciding the suitable radio transmission parameters for each connection, in a reduced time scale and in a very dynamic way. They operate on a TTI basis to take advantage of the short term variations in the interference level. The packet scheduling strategy may follow a time-scheduling approach (i.e, multiplex a low number of users simultaneously with relatively high bit rates), a code scheduling approach (i.e, multiplex a low number of users simultaneously with relatively low bit rates), and combination of both. Furthermore, prioritisation mechanisms can be considered in the scheduling algorithm.

The basic problem a packet scheduler has is how to share the available resources to the pool of users eligible to receive data [RSAG05].

The transmit power for HSDPA can be a large fraction of the total transmit power of a Node B. The continuous HSDPA power allocation scheme avoids large steps in power to prevent irregularities in the DL power control of dedicated channels. In contrast, the traffic-aware scheme switches the HSDPA power on only if data has to be transmitted, Figure 2-3.

The possibility to achieve very high bit rates using this structure depends primarily on two features of HSDPA: efficient power utilisation and time multiplexing rather than code multiplexing. This makes it possible to provide very high CIR values to a single user. In the absence of code multiplexing, the user does not have to share this high CIR with any other user, and hence achieve very high peak rates (the user perceived bit rate depends on how often the shared resource is available) [Wann06].

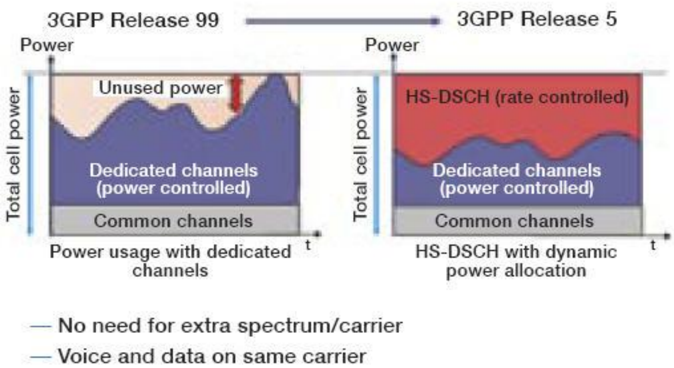

Figure 2-3 - HSDPA Power utilisation (extracted from [Wann06]).

Any increase in load (power consumption) will impact on the coverage on existing services. In a traffic scenario, it is expected that pilot coverage is good in low traffic situations; if traffic increases, the amount of DL interference raises and suddenly the pilot coverage is (almost) jeopardised in a few places. The HS-DSCH link adaptation algorithm at the Node B adjusts the transmit bit rate on the HS-DSCH every TTI, when a user is scheduled for transmission. The HS-DSCH transmit bit rate should be adjusted as a function of the per-TTI HS-DSCH SINR, experienced at user end. There are several situations that contribute to the variance of HS-DSCH – SINR, even if HS-DSCH transmission power is assumed to be constant. The total transmit power from the serving HS-DSCH cell is time variant due to the transmission of power controlled DCHs, the DL radio channel is time variant if the user is moving and the interference from the other end is also time-variant.

Under HS-DSCH link adaptation procedure, the UE periodically sends a CQI to the serving HS-DSCH cell on the UL (via HS-DPCCH). The CQI reported indicates with at least 90% probability, the maximum transport block size (TBS), number of codes and modulation from a set of reference ones that the UE is capable of supporting with a detection error no higher than 10% in the first transmission for a reference HS-PDSCH power. The RNC commands the UE to report the CQI with a certain periodicity from the set [2, 4, 6, 8, 10, 20, 40, 80,160], and can possible disable the report. In some cases it may exist a need to adjust the UE reported CQI by adding an offset, Figure 2-4.

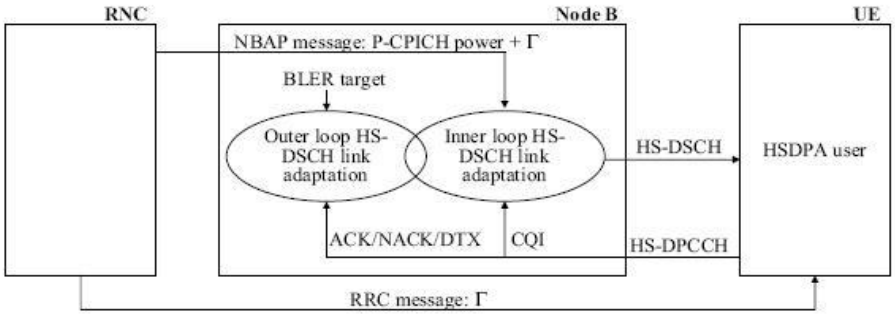

Figure 2-4 - HS-DSCH link adaptation algorithm at the Node B (extracted from [HoTo06]).

The conclusion from those studies indicate a need for an outer loop HS-DSCH link adaptation algorithm to further adjust the CQI index received from a user before applying it for adjustment of the HS-DSCH transmission format. The outer loop can be derived from ACK/NACK from previous transmissions. The algorithm adjusts the offset values to ascertain the average targeted retransmission probability. Too many retransmissions add an unnecessary delay, while too few indicate the transport block sizes used are not large enough, unnecessarily lowering bit rate. Reliable reception quality of the HS-SCCH is important since the transport block on the HS-DSCH can only be decoded if the HS-SCCH has first been correctly received. Therefore, sufficient power should be allocated to transmission of the HS-SCCH to ensure reliable reception. It is advantageous to reduce HS-SCCH transmission power to reduce interference levels in the network. Hence, it is generally recommended to have HS-SCCH power controlled every TTI, in which HS-SCCH transmit power is adjusted such that the desired user has a high probability of correctly decoding the channel. A large amount of HS-SCCH power is used for a mobile at the cell edge while a smaller amount can be used for another mobile close to the BS [HoTo06].

2.6 State of the art

In this section, the state of the art concerning this thesis main objective is presented. HSDPA scheduling schemes need to accommodate several traffic classes as well as maximise system bit rate while satisfying the QoS requirements of users. The packet scheduler determines to which user the shared channel transmission should be assigned at a given time. In todays networks, it is important to consider traffic mix, to analyse the performance of different scheduling algorithms, to simulate real networks, including different traffic classes. Services compete for time slots, power and codes, which is why it is crucial for users to have a time-shared and service differentiating system [WPKM05]. In a mixed best-effort and streaming traffic scenario, a reasonable fair scheduler can provide sufficient quality without service differentiation as long as the load is not high. At higher loads, service differentiation is however needed to protect the QoS of the streaming users. As previously stated, for RT services in particular, the packet delay has to be limited to a maximum tolerable value, otherwise packets are discarded; another important consideration is the fact that the PF scheduler does not take the delay into consideration (not suitable for real time services because of its severe packet loss).The average bit rate is continuously updated even if there is no data to be scheduled, increasing performance.

Both PF and Maximum Rate (MR) include the average bit rate in their metrics; MR scheduling presents two variations – MR_{min} and MR_{delay}. Users are otherwise ranked according to their instantaneous rate. When the delay budget is exceeded the priority of VoIP user´s packets is set based on their delay. Focusing on VoIP, the chosen schedulers represent different types of prioritisation functions. MR_{delay} drastically increases the priority of the VoIP flows when the delay reaches a certain value – strict prioritisation. The longer the user has waited since it last transmitted, the higher it will be prioritised by PF since the average bit rate drops. This increase in user prioritisation is gradual. For low data traffic the average cell bit rate is approximately the same for all schedulers [FLBW07]. In higher loads PF still increases the average cell bit rate, MR and MR_{delay} do not. About 10% of the flows get a higher bit rate with MR than PF for the same number of users. PF has fewer flows with low bit rate. The same observation holds for MR_{delay}. With MR_{min} the average cell bit rate drops when the number of users is increased. Pure MR is unable to serve the VoIP users, with larger delay budgets, the VoIP users are able to get a higher rate and can thereby compete better with the web users, but not well enough to reach the system quality constraint. When the load is increased it is foremost the VoIP users that suffer. This is because MR is biased towards web traffic. MR_{delay} on the other hand, prioritises VoIP traffic, higher than web traffic. Considering a mixture of RT and NRT services, there is a trade-off between the bit rate of NRT users and the delay requirement of RT users.

A more advanced scheduling algorithm, suitable for real time services known as the Modified-Largest Weighted Delay First (M-LWDF) incorporates the delay constraints of the TC along with the user´s instantaneous channel quality to determine their priority. Another modified version of the PF scheduling algorithm is known as the Exponential Rule, ER, scheme [HuNH07]. The scheduler uses the user´s priority as well as their queuing delay in the calculation of the priority function. Moreover, the schemes behaves like the PF algorithm when the difference in weighted delay is small and aims at providing better performance than the PF scheme when the variance in weighted delay is higher.

A Delay Based Scheduler (DBS) is proposed for HSDPA operated network that support real time applications [HuNH07]. The DBS considers the queuing delay along with the instantaneous channel condition of the user. Moreover the scheme accommodates multiple traffic classes by assigning priorities to users based on the delay requirements of their traffic types. Traffic classes with lower delay tolerance are given higher priority, whereas classes with a higher delay tolerance are given low priority. However the DBS does not provide guaranteed service to higher priority classes. Furthermore, on setting certain parameters, the DBS converges to a modified version of the Max CIR scheduler. At every TTI, the DBS collects and stores the channel condition of all users and the queuing delay of the head packet in their corresponding transmission queue. It then determines the maximum CQI value among all users to calculate the relative channel conditions. Users are then sorted in descending order of their scheduling function value. The DBS, M-LWDF and ER schemes maintain a lower percentage of packets dropped compared to the PF one which allows these schemes to enjoy a higher bit rate. However, DBS obtains a lower queuing delay, while maintaining a high bit rate, which is comparable to that of the M-LWDF and ER schemes, thus making the DBS a good candidate for scheduling RT traffic. Although, the DBS tries to maximise bit rate by giving users with good channel conditions a greater chance to be scheduled, it increases fairness by accommodating users with poor channel conditions [AKRS01].

Another fast packet scheduling scheme, aimed at providing a minimum bit rate guarantee for HSDPA users, consists of channel condition evaluation as well as their current average bit rates, hence giving higher priority to those with average bit rates below a certain threshold; such scheme is called Channel Quality – Based Minimum Bit rate Assurance (CQ-BMTA) [MaHa07]. Such algorithm gives more priority to those users with low average bit rates. Users are served based on their CQI as long as they are achieving high average bit rates. However if average bit rates start dropping below a threshold, their CQIs are multiplied by a weight factor that is inversely proportional to their bit rates. By introducing this weight, higher priorities are given to users with low average bit rate, which increases the degree of fairness in the system. The CQ-BMTA algorithm outperforms Max CIR and PF, because it increases the chance of those with low bit rate getting served by multiplying their priorities by a weight factor that increases as the difference between the minimum bit rate and their bit rate increases [GoPa05].

PF scheduling was designed for single carrier system requiring modifications in order to be suitable for a multi-carrier one. One of the most direct extensions of scheduling PF, Single Carrier PF (SC-PF), to multi-carrier system scheduler is combining the total carrier resource into one resource block. Multi-Carrier-Totally Integrated Scheduling (MC-TIS) is of this kind. Only one user is served at a time. It occupies all resource and got transmitted on *a* number of carriers simultaneously. Different from MC-TIS is Carrier Separately Scheduling (MC-CSS), which schedules carriers one by one. Under MC-CSS candidate scheduling user list was made into multiple copies, one for each carrier. For each carrier, the scheduling and resource allocation process are exactly the same. Proportional user fairness can be obtained on each carrier. Users may get resource of more than one carrier and get data transmitted on multiple carriers simultaneously. Because MC-CSS schedules carrier resource separately, it cannot obtain additional multi-user diversity gain in the frequency domain, i.e, it can only get diversity gain in the time domain as SC-PF. Similar to MC-TIS and MC-CSS, is Carriers Interdependence Scheduling (MC-CIS): it also makes candidate user list into multiple copies, one for each carrier. User scheduling and resource allocation are performed for each carrier separately. The difference between MC-CIS and MC-CSS is the average data rate updating process, i.e. the fairness component of the scheduler; in MC-CIS, it is in a user basis, by taking the achieved data rate on all carriers into account rather than on user/carrier basis in MC-CSS. Furthermore, because MC-CIS schedules user and allocate carrier resource interdependently, it can obtain additional frequency selectivity gain exploited from multi-user diversity in the frequency domain. In terms of performance, it can be seen that bit rate of multi-carrier is much higher than single carrier for the additional carrier resource. System bit rate under MC-CSS is exactly three times that of single carrier with PF scheduling, because of the additional multi-user diversity gain exploited in the frequency domain system bit rate under MC-CIS is higher than that of MC-CSS. Due to the impairment to multi-user diversity gain in both time and frequency domains MC-TIS scheduling gets the lowest bit rate performance (even lower than PF) [HuSA07].

The RNC follows the prioritisation of the traffic types when adjusting dynamic target threshold for NRT DCH packet scheduling. RT always has a higher priority to NRT DCH packet scheduling and HSDPA traffic types. Cell level priority for both HSDPA and NRT DCH traffic type is determined by the cell level weight values, which are dynamically adjusted by the RNC. The weight value is determined for each UE. In the case of multi-RAB, averaged weight is calculated from all interactive/background radio access bearers. The proposed algorithm description is based on packet rejection due to a full queue indication. In this model, first a packet is transferred to the main queue if the link is occupied and the queue has enough length otherwise it will be entered the retransmission queue. The queue used in this program is FIFO. This algorithm includes an internal and an external loop: the external loop is responsible for increasing the queue length and gathering the outputs. The internal loop is responsible for retransmitting the blocked packets without missing the following packets. This procedure can be repeated many times and can increase the system delay too much; its main advantage is to reduce blocking rate, but in voice it increases the voice delay. For data services it achieves improvements in delay and packet loss [ChAb07].

Several scheduling approaches can be employed: UE follows the best scheduling command (maximum allowed rate) where all other Node Bs may suffer from unexpected noise raise; in a second possibility, UE follows the worst scheduling command (the maximum allowed rate is the minimum among the signalled allowed rates), where the receiving quality would be better than expectations; in a third option, the UE can combine the scheduling commands by applying different weighting factors for each scheduling command. Another approach could be an UE that is allowed to transmit by more than one Node B, selecting the scheduling command with the maximum allowed data and lowest power control requirement from all received scheduling commands [SoBa04].

Chapter 3

Models and Simulator

This chapter presents an overview over multi-service and user applications management demands, to accomplish a good network performance, through the usage of a simulator; its constitutive blocks, inputs and outputs are presented. The chapter ends with the assessment of the simulator.

3.1 Models and Parameters

The main objective of the present thesis is to study service and applications management inherent to the RRM algorithms, considering different scheduling policies for the delivery of the usual services provided by cellular wireless networks, to get an optimal response, to a call or session request. Any optimisation study for HSDPA, at the level of the radio network, requires the introduction of service source models, in order to predict the user generated load in the network, based on multi-service traffic modelling. When a UE initiates an event (call or data session), BS RRM functionalities allocates and releases resources, changing power level, bit rate, number of codes, depending on application type and network capabilities, having in mind those users who finish their application and the remaining ones, who are continuing to transmit.

The number of carriers can assume several values, but in present thesis it is assumed a maximum of two carriers to support all voice and data services and within applications, Table 3-1. Following such criterion several scenarios can be built up, for different network traffic profiles, making variations in the predominance of each service per carrier; such changes allows to perform different studies on network behaviour for each different configuration.

Table 3-1 - Network characteristics and employed carriers.

Carrier 1	Carrier 2	Data services penetration rate	Voice services penetration rate	Network Classification	Scenarios
R99+R5	R99+R5	50%	50%	Reference (REF)	1
R99+R5	R5	60%	40%	Data Centric (DC)	2
R99+R5	R5	70%	30%	Data Mostly (DM)	3

The management algorithm is schematically described in Figure 3-1 and the system in Figure 3-2, including three main parameter types: user parameters, network parameters and algorithm parameters, the last one being responsible for rules creation to manage all users that are trying to access the network for a specific service.

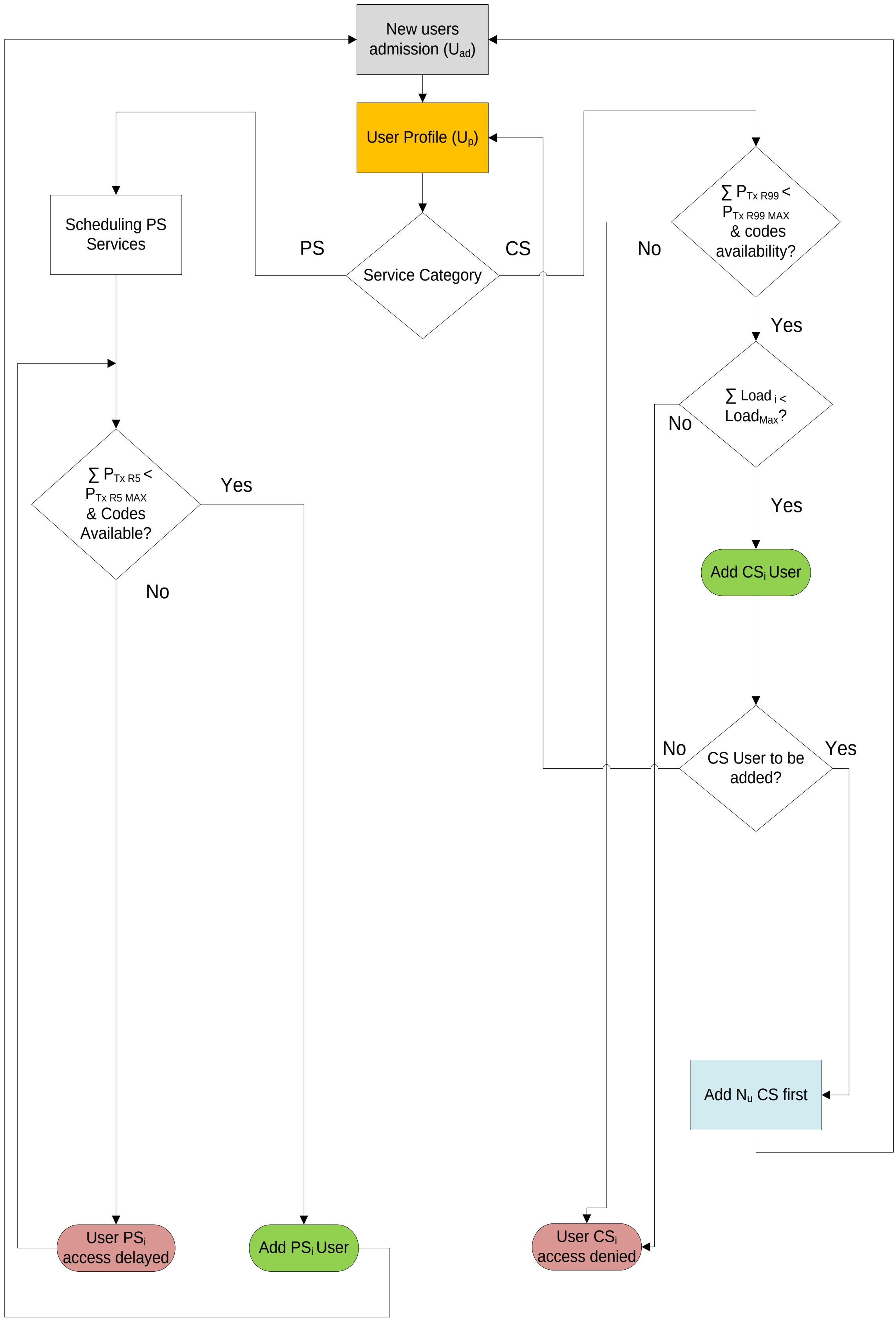

Figure 3-1 - RRM Global Model.

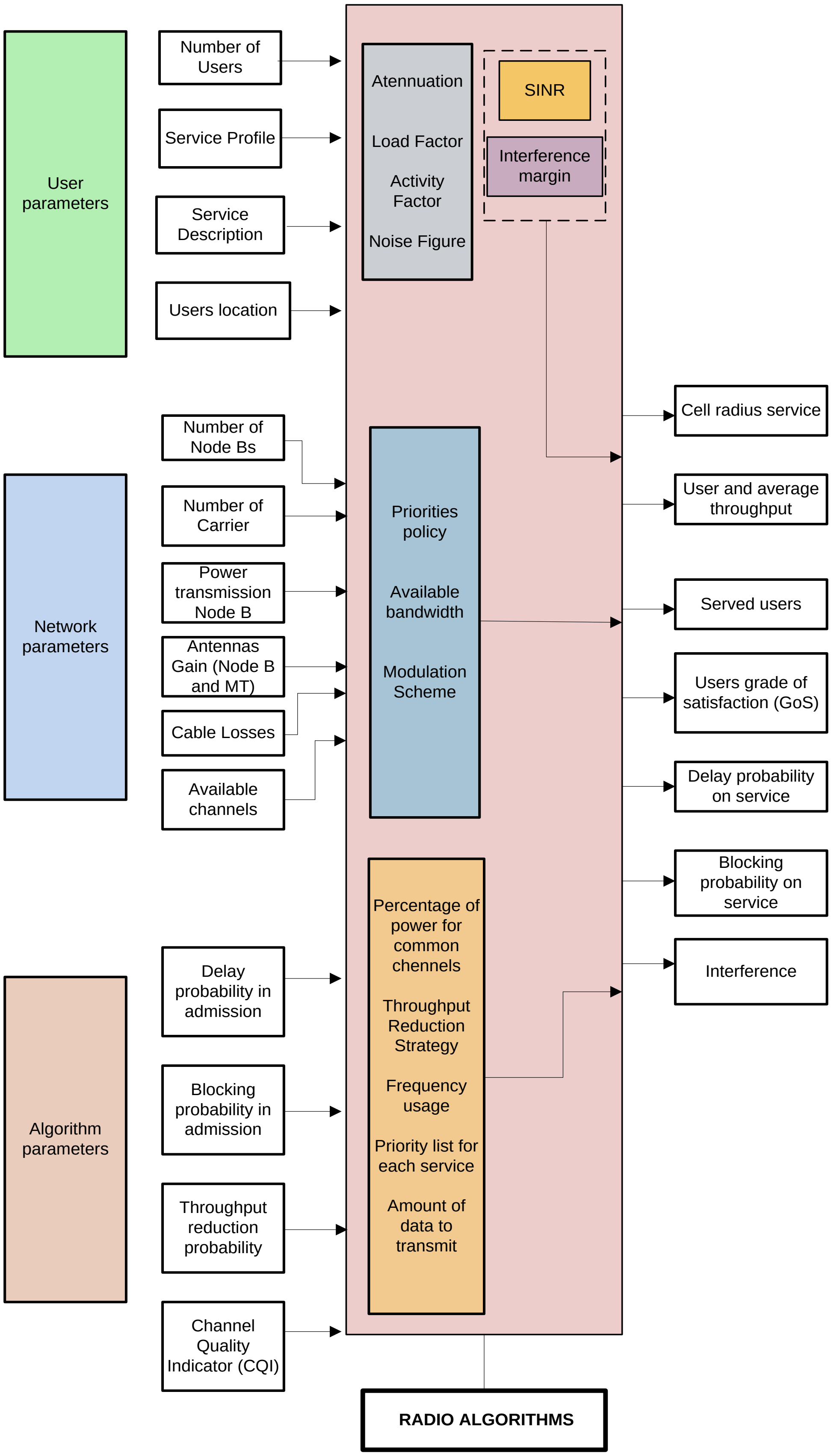

Figure 3-2 - Global parameters (inputs and outputs).

In Scenario 1, there is a traffic share for both voice and data services, in each carrier (50% for each). In Scenario 2, there is predominance in data services (60% service penetration): in carrier 1 in a mixed way and exclusively for data services in carrier 2. In the last scenario, the configuration is similar to scenario 2, even with higher unbalance between voice and data services (30% service penetration for voice).

When there is a user admission request, the first step performed by the algorithm is the user profile analysis: voice or data user; voice service has always precedence over data services, when accessing the network at the same time, voice users are firstly admitted and then (when all voice users are attached to the network) the remaining data users enter in the network while there are radio resources available (power and codes) and the maximum BS Load threshold η_{DL} is not reached - CAC model described in Figure 3-3. The purpose of the RRM global algorithm is to manage the different possible states: admitted, ended and delayed users which will be sooner or later re-admitted.

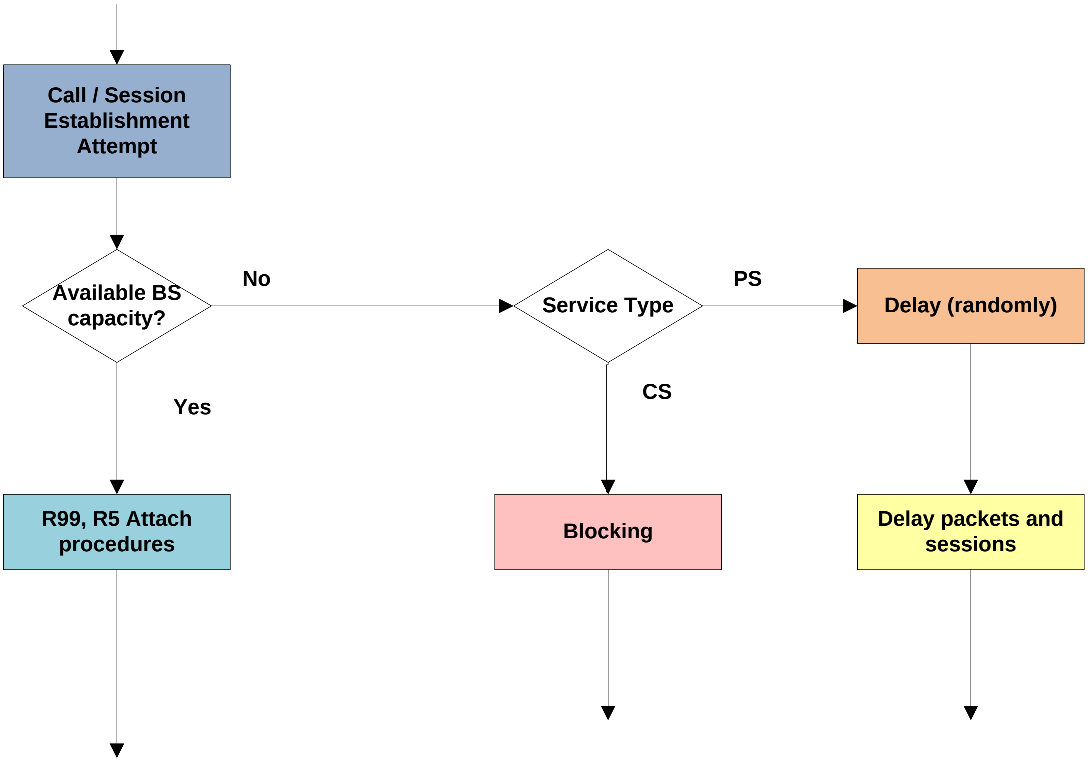

Figure 3-3 - RRM CAC Algorithm.

If the BS cannot accept a request, it automatically blocks the user in case of CS services and in case of PS the same user is delayed and driven to a queue list (delayed status). After entering in such a list, all users try again to re-enter in the network by order of priority, as soon as load and code availability thresholds are accomplished. In queued mobiles, where the service priority is the same, the access to the network is similar to FIFO: mobiles with higher delays get in the network first. Optimising resources to properly distribute traffic and bit rate per user begins with service category for CS and PS; voice has always precedence over data services whenever both compete for network accessibility. According to (2.7), the traffic power in the BS is dynamically distributed in two pieces, $P_{T_x traf}{}_{R99}$ and $P_{T_x traf}{}_{R5}$.For both services the system should guarantee, a minimum power for users who are connected in that moment, or even in such cases, where an additional user requests to access a specific service, i.e.:

$$\begin{cases} P_{T_x traf\,R99} \geq P_{T_x traf\,R99_{min}} \\ P_{T_x traf\,R5} \geq P_{T_x traf\,R5_{min}} \end{cases} \tag{3.1}$$

There is in time a dynamic variation in the traffic power associated with voice and data services, similar to a sliding window mechanism. In previous equations the traffic power is just applied to part of the total BS power, once the remaining has to be applied to signalling purposes. In Figure 3-4 the shared power according to service type is described according to (3.2). There is always a part of the transmission power reserved for signalling and control. The remaining traffic transmission power is distributed according the system needs between R99 and R5 users in a mechanism similar to a sliding window process. Despite of the power allocated for each release, there is always a maximum and minimum threshold to be followed, shown in Figure 3-4, where the minimum transmission power for traffic channels in R99 coincides with the maximum transmission power for traffic channels in R5 $(P_{T_x traf\,R99_{min}}, P_{T_x traf\,R5_{max}})$ and on the other hand, minimum traffic transmission power for R5 coincides with the maximum traffic transmission power for R99 $(P_{T_x traf\,R5_{min}}, P_{T_x traf\,R99_{max}})$.

$$\begin{cases} P_{T_x traf\,R99_{max}} = P_{TX}^{BS} - (P_{T_x traf\,R5_{min}} + P_{T_x sig\,ch}) \\ P_{T_x traf\,R5_{max}} = P_{TX}^{BS} - (P_{T_x traf\,R99_{min}} + P_{T_x sig\,ch}) \end{cases} \tag{3.2}$$

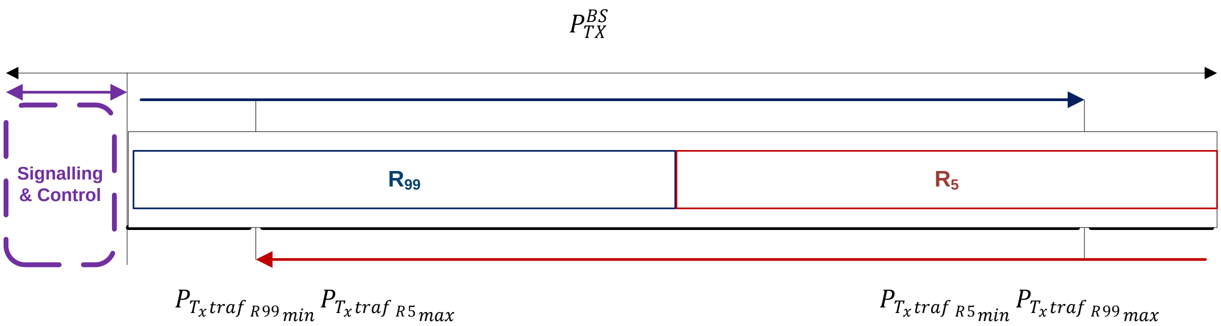

Figure 3-4 - RRM allocation in shared carrier.

The total BS power is employed in both traffic and signalling channels (2.4). The biggest amount of power (usually more than 60%) is employed in traffic services.

Depending on the scenario employed each new user admission or reactivation, has a different weight in the load of both carriers. If the scenario chosen is REF, the total power and the number of codes distribution per service is computed in both carriers like:

$$P_{Tx}^{Bs} = \frac{(P_{T_x traf\,R99} + P_{T_x traf\,R5})}{2} + \frac{(P_{T_x sig\,R5} + P_{T_x sig\,R99})}{2} \tag{3.3}$$

$$N_{mc} = \frac{N_{vc}}{2} + \frac{N_{dc}}{2} \tag{3.4}$$

For the DC scenario the same parameters are computed as follows:

$$P_{Tx}^{Bs} = \frac{P_{T_x traf\,R5}}{2} + P_{T_x traf\,R99} + \frac{P_{T_x sig\,R5}}{2} + P_{T_x sig\,R99} \tag{3.5}$$

$$N_{mc} = \frac{N_{vc}}{2} + N_{dc} \tag{3.6}$$

For the last considered scenario, DM, the equations are the same as in the DC scenario, with the variation applied in the service penetration (more 10% of PS services). This is the way chosen to simulate the existence of 2 UMTS carriers; with such structure, by dividing by a factor of 2, one duplicated the resources (codes and transmission power), for each service that is making use of both carriers. There is a linear relationship between the capacity of the R99 and HSDPA systems and the distribution of code resources. Therefore, a static distribution solution can be used to distribute the codes at the early stage of building a network, when there are no intense demands for either R99 voice services or HSDPA data services. When the network develops to a certain mature stage, dynamic distribution of the codes can be used to fully utilize the code resource in the system. With HSDPA, data transmission can be divided in time and codes (code multiplexing) to accommodate several users per TTI. Code multiplexing makes it possible to use all available codes per TTI, even if the codes are not all supported by a given end-user device: three users with a five-code device can be served simultaneously during the same TTI.

Delay-sensitive applications with short, burst data packages will benefit from code multiplexing when HSDPA usage increases in a network. The introduction of end-user devices that can handle 10 or even 15 codes increases maximum transmission rates to 14.4 Mbps. At the same time, however, it becomes necessary to allocate codes dynamically, because fixed allocation of 10 to 15 codes on a sector carrier for HSDPA seriously reduces the available codes for R99 traffic. During periods of high voice load, the codes should be assigned to voice and R99 data traffic; the rest of the time, they can be used for additional HSDPA traffic, especially when used in combination with code multiplexing.

There is a direct relation between the bit rate per service and the number of codes made available by the network, Table 3-2.

Table 3-2 - Bit rate and number of codes relation.

Services	Bit Rate [kbps]		Applicational Codes assigned per user	
	Min	Max	Min	Max
Voice	15		1	
Streaming	300	5 600	5	15
WWW	64	384	1	5
FTP	384	14 400	10	15
E-Mail	64	384	1	5

In a perspective of managing different services at the same time, establishing priorities and providing the best effort to reach the maximum number of active users, some approaches have to be taken, for each channelisation type (SF 128, SF 16 or SF 8):

- channel codes are stored for signalling and control (SF 128).
- 1 channel code is enough for each voice user (all channelisation types).
- 2 channel codes will be assigned to service data 64 kbps (SF 128).
- 5, 10 or 15 channel codes will be employed to HSDPA with different limits (from 0.384 Mbps till 14.4 Mbps) (SF8 or SF16). HS-PDSCH codes parameter is equal to 15 (possible maximum capacity). Such scenario will be employed for a dedicated carrier (data services) or shared carrier (voice and data services).
- No use of time or code multiplexing (all channelisation types).
- Presented results will be for static users (no associated mobility for all channelisation types).
- Each user processes just one service each time.

A correspondence between codes, services and number of users allowed can be a total of 126 channel codes for a 128 SF channelisation code type and 15 channel codes for a 16 SF channelisation code type, Table 3-3. Each service type (CS or PS) has a particular TTI, used for transmission and management between the users who are transmitting, those who are requesting access to a service, and the others who will no longer continue to transmit. For voice such time varies between 10 ms and 80 ms (10 ms is considered) and for data this value is 2 ms. The ability to mix HSDPA and non-HSDPA traffic reduces the need for early introduction of extra cell carriers, thereby increasing spectrum efficiency [DJMW06].

Introducing some rules related to radio conditions allows a satisfactory network performance, interference margins, service profile and THP; mechanisms like delay, introduce the concept of queuing lists, where each delayed user is kept (until it gets radio conditions to be re-admitted in the network).

Table 3-3 - Correspondence between services and code capacity request.

Service	Codes per user	Number of available codes	Number of users / service per SC or carrier
Only Voice	1 SF128	128	126
R99@384 kbps	1 SF8	128	7
HSDPA 3.6 Mbps	1 SF16	5	3
HSDPA 7.2 Mbps	1 SF16	10	1
HSDPA 14.4 Mbps	1 SF16	15	1

From a user view point, the most important output, when a data service is being requested, is bit rate, which means satisfaction in cases where there is a high demand and is reached a high value. Secondly, it is important to have low interference, representing low retransmission rate, higher integrity (low packet drops and error correction on bits transmission), low transmission delay and a better efficiency on employed modulation. Under such conditions, a better QoS can be reached, meaning that the service is provided with no interruptions, and keeping the thresholds of integrity and delays inside reasonable values. Another important parameter is latency, (service response time, end to end), relevant mainly for interactive services, but not directly dependent on RRM algorithms because it is influenced from other interfaces (not just the radio interface); that is why such parameter is out of scope of the present thesis. Other important parameters which influence the final grade of service (GoS) are: signalling efficiency, correction efficiency, modulation and rate of retransmissions. Despite of their importance, for the same reason of latency, they will not be analyzed in detail in the present thesis.

Capacity and coverage are the main issues in an HSDPA network. Analysis on a Node B basis, pointing out some important results on a perspective of user satisfaction and service performance, (for different scenarios and algorithms) is the objective to be accomplished.

The cell breathing process is a standard mechanism in UMTS R99, which implies a trade-off coverage and capacity. This RRM procedure has a preventive propose, in which the BS load is monitored. Should load level go beyond very high or very low levels, the power of the BS pilot channel will decrease or increase, respectively, inside a limited range. The result is a variation on the coverage area of the BS, which is computed after a pilot power change, the coverage matrix being updated accordingly.

3.2 Inputs and Outputs

In every model there is a set of parameters to be tuned, resulting in a more objective analysis for each studied scenario. The background behind the simulator to be employed in the present thesis is a urban scenario composed basically by streets and buildings, over which BSs and UEs are scattered.
The input parameters can be conceptually divided in two groups:

- Scenarios and UE users.
- Multi-RRM Algorithms.

The group “Scenarios and UE users” is composed of the following input parameters:

- N_{BS}: number of BS.
- N_u: number of active users (UEs).
- SP: service penetration (mix of voice, streaming, www, e-mail and FTP) and the corresponding individual source model parameters.
- G_{UD}: geographic users distribution.

The simulator presents a graphical MS-Windows based interface, which allows a significant level of flexibility, giving the possibility to a user of setting up a relatively large set of parameters and combinations with great simplicity and accuracy. The Multi-RRM Algorithms group defines the parameters required to define the BSs signal propagation conditions as well as some RRM initial parameters like the service priority list; a separation is performed between the parameters employed in the present thesis and the generic ones, which were not tuned during the present work. Global parameters are defined as a group that make part of any mobile network, in the present case including UMTS technology and by their definition, with the possibility to create a scope of work with several possible scenarios by changing each one separately or in a complementary way. Such parameters are:

- L_{BS} : BS location.
- N_{mc} : BS total available codes.
- G_t : antenna gains transmission.
- Δf_{BS} : available bandwidth.
- N_0 : noise spectral density.
- N_t : thermal noise density.
- f_{BS} : BS frequency.
- G_{UD} : geographic users distribution.

In the present thesis, there is a specific group of parameters that were tuned having in mind the final goal: serving the highest number of users, respecting determined QoS levels. Those parameters are directly or indirectly related to RRM algorithms behaviour when experiencing different scenarios; at the same time those adjusted items can be separated in system or user parameters:

- BS type (BSs equally configured with R99 / R5 capabilities).
- H_{BS} : BS height.
- P_{Tx}^{Bs} : total BS transmission power.
- N_u : number of active users.
- N_{sc} : number of channels used for signalling.
- N_{dc} : number of codes for data services.
- η_{DL} : load Factor Downlink.
- $M_{I_{\eta_{DL}}}$: interference Margin.
- ${P_{T_{x}}}_{pil}^{BS}$: BS pilot T_x power.
- PT : services priority table.
- $P_{Tx_{max}}$: maximum T_x power.
- S_c : scenario choice.
- D_{Pv} : data packet volume.
- T_{vc} : average voice calls duration.

Each of these parameters can be tuned to have different influence on the simulation results. The number of HS-PDSCH codes is a key parameter for the evaluation of HSDPA network performance and has direct impact for the bit rate increase at the end user. In the present thesis some compromises were established:

- Mobile categories supporting all possible number of HS-PDSCH codes,(maximum of 15).

- Multiple users influence is simulated by the introduction of the interference margin, specific for each Node B.
- All BS have the same technology and configuration, all supporting voice and data services.

The interference margin is a parameter to emulate the load in the cell, to be considered in the multiple service users scenarios. Due to the interference margin, path loss decreases, leading to lower bit rates.

The minimum and the maximum service bit rates can also be modified for HSDPA, one possible way to accomplish such situation, is a variation in the number of HS-PDSCH codes chosen.

The RRM simulator output block produces a set of results which are stored in a file. The simulator being structured in a modular architecture, can define itself the statistics and output parameters which are interesting and may easily implement its extraction, by setting up virtual scopes. Following output parameters have been defined and implemented, in order to reflect the network's overall performance in a RRM perspective. The most important output parameters to be generated are:

- Number of active users per service (N_u).
- Average bit rate, global and per service, $\overline{R_{b_G}}$ and $\overline{R_{b_s}}$.
- Codes usage for voice and data services (N_{vc} and N_{dc}).
- Rate of users in HSDPA (% N_{du}).
- Power distribution between HSDPA and voice users ($P_{T_x traf\ R99}$, $P_{T_x traf\ R5}$, $P_{T_x sig\ R99}$, $P_{T_x sig\ R5}$).
- Average Delay per user ($\overline{\tau_u}$).
- Blocking probability for voice service (P_b).
- BS Load factor computation (η_{DL}).

User satisfaction is the main goal for a service planning and development. Accomplishing such objective is a challenging task that can be reached through many ways. Different RRM strategies were implemented and compared in a perspective of evolution and results improvement, in order to reflect the network overall performance in a RRM perspective.

The RRM Blocking Probability, P_b, is a measure of blocked voice calls, is defined as:

$$P_b = \frac{N_{bc}}{N_{tc}} \times 100 \quad [\%] \tag{3.7}$$

where

- N_{bc}: number of blocked voice calls.
- N_{tc}: total number of voice calls.

The average delay, $\overline{\tau_u}$ is defined for all data services, as a measure of the delay affecting the transmission of packets, per user:

$$\overline{\tau_{u\,[s]}} = \frac{1}{N_{tot}} \sum_{i=0}^{T} N_{df} \times T_f \tag{3.8}$$

where:

- N_{df} : number of delayed frames.
- T : one hour simulation time.
- T_f : one frame duration time (10 ms).
- N_{tot} : total number of users.

The Drop rate, D_r, is defined according to the following expression:

$$D_r = \frac{N_{ds}}{N_{ts}} \times 100 \quad [\%] \tag{3.9}$$

where:

- N_{ds} : total number of dropped sessions.
- N_{ts} : total amount of sessions (successfully ended plus dropped).

The average bit rate by service is defined as:

$$\overline{R_{b_s}}_{[\text{Mbps}]} = \frac{1}{N_{BS}} \sum_{i=1}^{N_{BS}} \overline{R}_{b_s} \tag{3.10}$$

where:

- N_{BS}: number of BS used during the simulations.
- $\overline{R}_{b_s}$: average bit rate for each service (computation of all BS).

The average global bit rate, $\overline{R_{b_G}}$, is defined as the mean value of the average bit rate of the system computed for the only available RAN:

$$\overline{R_{b_G}}_{[\text{Mbps}]} = \frac{1}{N_{BS}} \sum_{i=1}^{N_{BS}} \overline{R}_{b_{BS}} \tag{3.11}$$

where:

- N_{BS}: number of BS used during the simulations.
- $\overline{R}_{b_{BS}}$: average bit rate for each BS.

The load and codes network was computed per service category, being both quantified for voice service and for data services.

The traffic voice for a Busy Hour call attempt is computed as:

$$T_{r\,[\text{Erl}]} = \sum_{i=1}^{T} T_{vc} \times \lambda \tag{3.12}$$

where

- T_r : busy hour voice traffic.
- λ : total number of voice calls.
- T_{vc} : average voice calls duration (mean holding time).
- T : one hour simulation time equivalent to a busy hour duration (3600 s).

Even if the focus of this thesis is not voice, this service is still important and is always present on capacity dimensioning.

3.3 Simulation Strategy

This thesis is based on the results produced by a RRM software simulation tool. It is a system level - time based simulator – with a resolution of 10 ms, which has been developed over Microsoft Visual Studio 2005 platform. In its original version [Serr02], the tool supported UMTS, HSDPA and WI-FI

RAN supervised by CRRM. For the current work, an adaptation of the original version was performed by the author in order to be focused in HSDPA RRM simulation capabilities. The simulator implements low-level system functionalities, like power control, link control, basic channel code management, radio bearer service, load control, access control, propagation estimation, interference estimation and generation. It is also possible to generate service traffic mixes through service source models parameterisation.

RRM functionalities are performed in HSDPA, mainly in common management of available resources and high level decisions, like call admission control. On the previous assumptions and to better understand the impact of scenarios and algorithms on RRM performance, two main approaches may be taken: one can make strategic changes on interesting scenario characteristics, keeping all the other parameters constant, namely the RRM algorithm/policy and then evaluate the RRM sensibility to scenario change, the other is to experience a coherent set of RRM algorithms/policies over a "static" scenario. Both strategies, enables the evaluation and identification of the relation cause/effect.

One of the relevant parameters that can be changed in a scenario is the user density, leading to a better understanding over the RRM outcome when dealing with higher, medium and low user population densities.

Originally the simulator had the structured described in Figure 3-5. The main difference between such model development algorithm and the present thesis, resides in the fact that in the first model structure, a combination of heterogeneous networks was developed, connected each other and supervised by a centralized structure – Common RRM - (CRRM), which allows inter-communication among those several radio networks and at the same time a right management resource and correspondence, between radio resources availability for each network and service type request. This study was performed in several scenarios, according to load variation, different mobility speed (from static to fast velocities) and different service policies and priorities. Users are distributed randomly in the network.

As shown in Figure 3-6, in the present algorithm simulation, there is just one network type (homogeneous network), supporting different service types; users are still distributed randomly; every BS presents similar configurations (all BS supporting the same technologies), where the main goal is to serve the maximum number of users, for each service requested inside pre-determined QoS thresholds. Besides intrinsic and fundamental system functionalities, like link control, code management, load control, access control, interference estimation and generation, the simulator is also capable of generating traffic mix, combining different scenarios and traffic services heterogeneity. The following service types are considered:

- Voice (conversational).
- Web browsing, Streaming and FTP (interactive).
- E-mail (background).

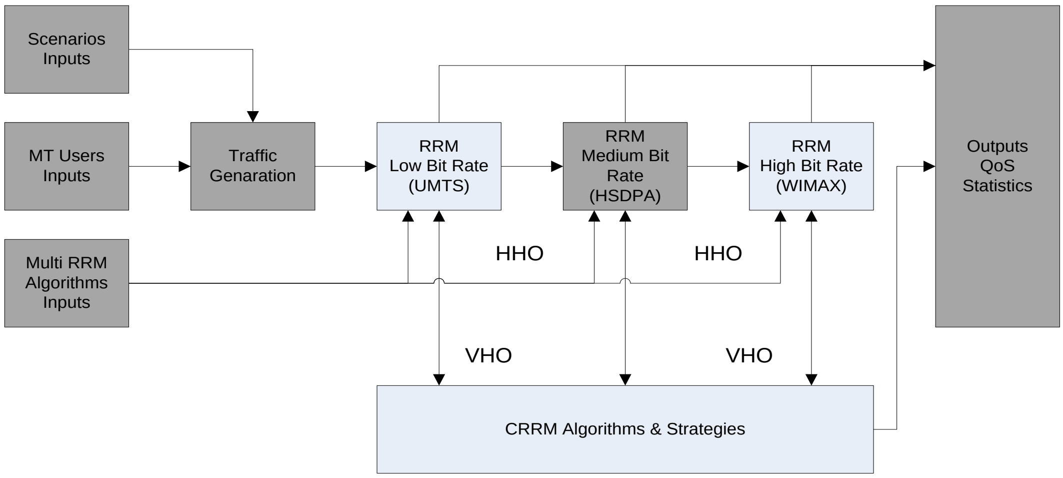

Figure 3-5 - Original Simulator Model.

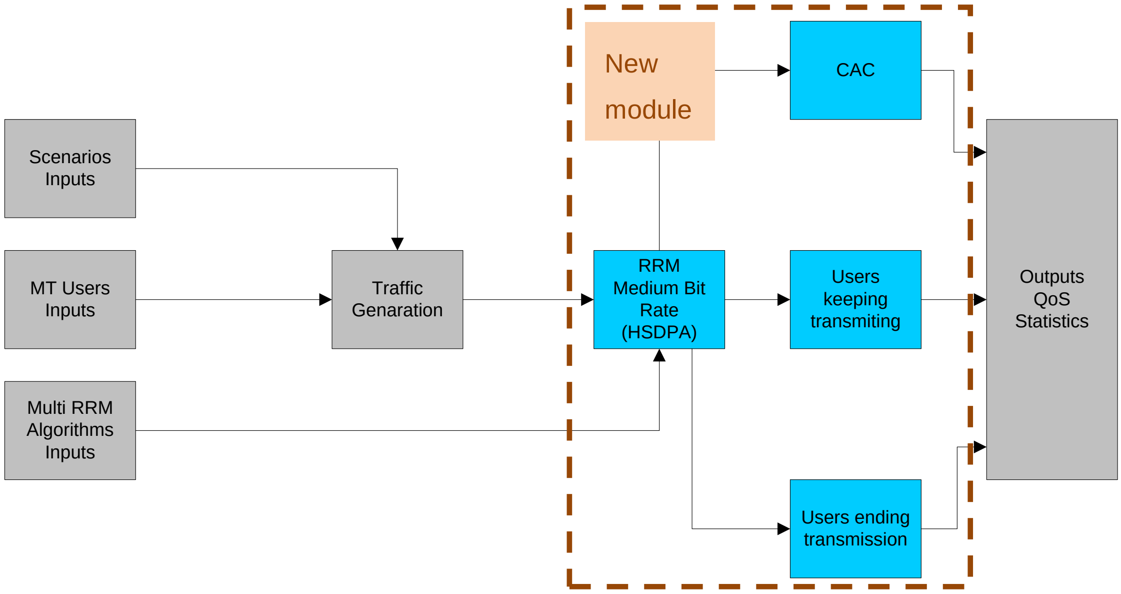

Figure 3-6 - RRM model structure in R5/HSDPA.

The simulator is composed for three main blocks, described in Figure 3-7:

- Users generation.
- Network management.
- Quality of service (statistical results).

For user generation it is necessary to take in account information concerning:

- Traffic and its information, related with each considered service and making use of real scenarios where is possible to perform a penetration rate for each service.
- Type of terrain. In the present simulator the employed database for simulations refers to an urban area of Lisbon.
- Scenarios characteristics, according to the type of service which is carried in both carriers.

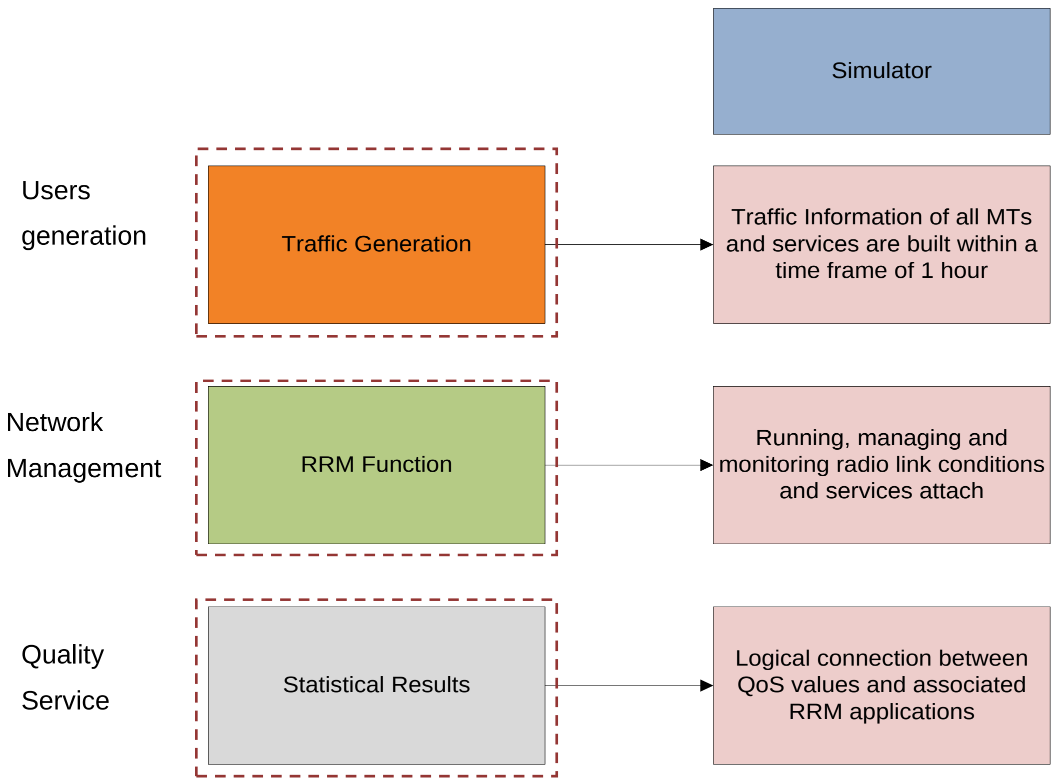

Figure 3-7 - Algorithm model for RRM policies and generation results.

3.4 Assessment

As stated before, the simulation tool used in this work consists of an upgrade of a previous version [Serr02], in which a re-adjustment in RRM models focused in UMTS (R99/R5) was made. The resulting (upgraded) version of the simulator was validated, and in particular a full debugging analysis of inputs and outputs of the new modules and functionalities was performed. The overall performance of the tool and its outputs were validated by hours of test simulations in a pre-production stage. The current simulator is a time-based, dynamic system, and as it happens with most simulators of this type, an instability period is expected at the beginning of the simulation period. After this period, a convergence to stability is observed. It is necessary to assess a starting point from which the simulation parameters are stable, after the initial oscillation period. The results obtained before the starting point should be naturally discarded.

The scenario for this simulation in the city of Lisbon included just one system (UMTS) and considered a user density of 10 000 [users km^{-2}]. Each user is active for a given service, characterised by a time variable bit rate, which value depends on the system and on the channel conditions. The present assessment result is based in the REF scenario, with 3 BS (2 carriers, both processing voice and data), equilibrating the penetration rate, for voice and data traffic (50% for both), according to Table 3-1. Regarding the results, the bit rate variation was chosen according to present network implementations (HSPA newest release – release 7) and strategies per service, making use of two carriers. The total number of available codes was dimensioned, considering that there are a number of codes exclusively for signalling; the services with highest bit rates, needs a higher number of codes; in Table 3-4, one defined the priority service associated to each service and the maximum and minimum bit rate thresholds, inside which is assigned randomly a final value, responding to each access to the network request, whenever radio resources, power and codes are available.

Table 3-4 - Assessment description for the REF scenario.

Service	Priority Service	Bit rate per service [Mbps] (theoretical values)		Maximum theoretical number of active users for each service – limitation factors: power and codes		Maximum number of active users per [s] (after 20 simulations)
		Min	Max	Power	Codes	
Voice	1	0.015		28	768	42
Streaming	2	0.300	5.600		76	11
WWW	3	0.064	0.384		152	9
FTP	4	0.384	14.400		51	7
E-MAIL	5	0.064	0.384		152	13

In the present thesis there is not any bit rate strategy reduction: every time there is a request from a UE with a specific service to enter in the network, the CAC algorithm stage evaluates the request, compares it with the available resources, and if the access rules are fulfilled the access is performed otherwise, the user is blocked if it is a voice user, or is delayed until there is an opportunity to re-enter in the network if it is a data user.

Load is a limiting factor, where there is a theoretical power transmission value, calculated with the assumption of a permanent scenario of 2 carriers, assuming that each new active user introduces whatever the service, an extra power of 0.05 watts, which results in a maximum of 28 users per omni BS (14 for each carrier, to accomplish a load maximum criteria $\eta_{DL} < 0.7$). When the limiting factor criterion is associated to the number of codes, the maximum number of available codes is different according to each service distribution weight at each time, having Table 3-3 as a reference for each service channelisation type. Each maximum value for codes in Table 3-4 is in line with the premise that all the resources will be used exclusively by that service. Otherwise with so many possible combinations of active users for each service per second, it would not be possible to establish a maximum theoretical value for such indicator. The most demanding services are FTP and Streaming, which is why, with good radio conditions, these services improve the associated bit rates (higher bit rates imply higher codes allocation within thresholds), resulting in higher code demands, and of course in less users accessing the network at the same time, without reaching any congestion. The proposed value of the current thesis resides basically on two main aspects. First, a simulation tool for a homogenous mobile network gathering one RAN technology, enhancing voice and high speed data, is developed, and second, a heavy series of simulations is processed in order to characterise the described network like in any work involving heavy simulation workload, decisions had to be taken in order to conciliate statistically relevant results with a feasible (limited) time frame. The maximisation of the number of simulation results by on the one hand, and the minimisation of time/CPU resources spent in simulation work on the other, configures a non-trivial optimisation problem. Bit rate is a decisive aspect to define the grade of success for a high speed data network; it can be defined per service or in a global way. To access the valid simulation period, a simulation of 1h 10min was performed. In Figure 3-8, one shows the evolution of the average bit rate of the whole system along 1h 10 min of simulation; such bit rate in a given second is computed as the average value of the average bit rates of the unique RAN considered. A close observation to Figure 3-8 shows that stability is achieved around the 7th minute (420 s) of simulation; consequently the analysis in all simulations is taken for one hour simulation (excluding this initial 10 min).

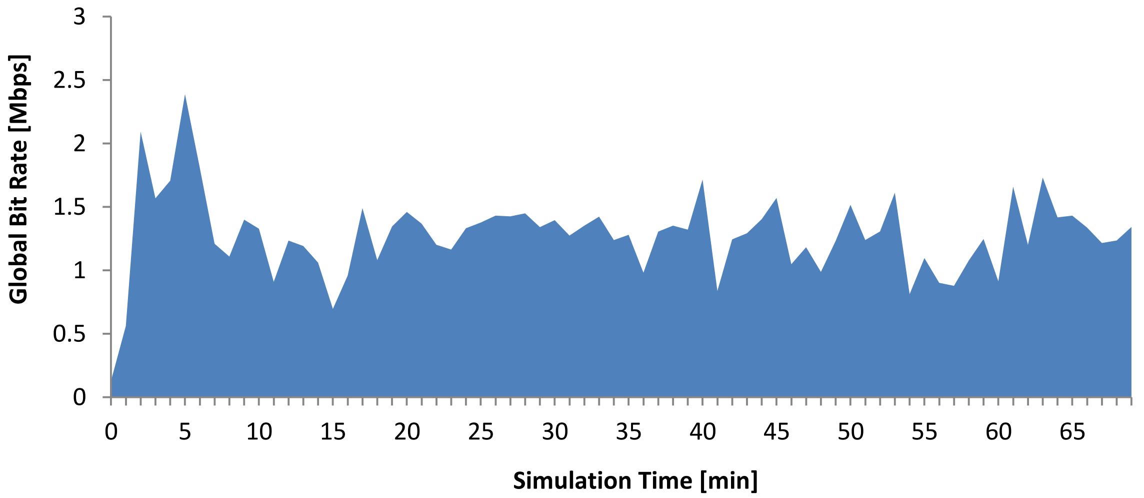

Figure 3-8 - Average Bit Rate for a one hour simulation period.

In order to better describe the strategy to characterise the network integrating one RAN technology, one defines the network performance function (∂) as a column matrix, which entries correspond to the output parameters defined in Section 3.2:

$$\partial = \begin{bmatrix} \overline{N_u} \\ \eta_{DL} \\ \overline{R_{b_S}} \\ \overline{R_{b_G}} \\ P_b \\ \overline{\tau_u} \\ S_c \end{bmatrix} \tag{3.13}$$

Consequently, it can be assumed that ∂ is a function of the simulator's inputs:

$$\partial \; = f\ (N_u, T_{vc}, D_{Pv}, G_{UD}, SP, PT, S_c) \tag{3.14}$$

The REF scenario can be described as a set of fixed values that are assigned to the input parameters of the simulator. The characterisation of the ∂ function is given by the characterisation of its individual functions. In practical terms, this is achieved by defining a given number of scenarios, where all the parameters assume the fixed values defined for the REF scenario, except for one, that floats inside a given range of values.

It has been seen that the choice of an optimal number of simulations must account for the simulation time/CPU resources and number of parameters of the ∂ function. It is also important to get a convergence in the values for the ∂ function, obtained by running 20 simulations and by defining a measure of convergence σ:

$$\sigma[\%] = \frac{|\partial^{\mathrm{cum}\_s} - \partial^{\mathrm{cum}\_S}|}{\partial^{\mathrm{cum}\_S}} \times 100 \tag{3.15}$$

where

- s : simulation run index.
- S : total number of simulation runs conducted during the convergence study.
- ∂^{cum_s} : partial ∂ cumulative mean at simulation run s.
- ∂^{cum_S} : total ∂ cumulative mean (from all simulations).

The cumulative mean of ∂ at simulation run s is given by:

$$\partial^{\mathrm{cum}\_s} = \frac{\sum_{i=1}^{s} \partial^{\mathrm{mean}\_i}}{s} \tag{3.16}$$

where:

- ∂^{mean_i} : mean of ∂ for simulation run i.

Additionally it is also useful to define the cumulative maximum of an evaluation metric:

$$\partial^{\mathrm{MAX}\_s} = \frac{\sum_{i=1}^{s} \partial^{\mathrm{MAX}\_i}}{s} \tag{3.17}$$

where:

- ∂^{MAX_s} : maximum value of ∂ obtained during simulation run s.

In Figure 3-9 the convergence of the ∂ function for the REF scenario is presented. It can be observed that the convergence exists for the four parameters involved, mainly due to the scenario employed where the higher demanding services presented a few part of the total (sum of all service data is equal to 50%, and the most exigent services like FTP or Streaming are just 24% of the total service penetration). Considering that the ∂ function has 7 variables, and considering that the variation of the ∂ parameters for REF scenario involves a total number of 20 simulations, one got the time consuming just for the simulator audit process. Total processing time used for all simulations assuming that a reasonable security margin is achieved with 5 simulation runs, raises the total number of simulations to 115 (since 20 were used for assessment and another 20 for tests), leading to 173 hours of simulation (7 days approximately).This simulation workload should be considered relevant enough to produce at least some reliable results. Since the present model results are supported by a discrete event simulator, the actual time that takes to complete a simulation is more than 70 minutes.

The total amount of time to run the total number of simulations just reinforces the need to establish an adequate number of simulations for each scenario computation, since it is a very time consuming task.

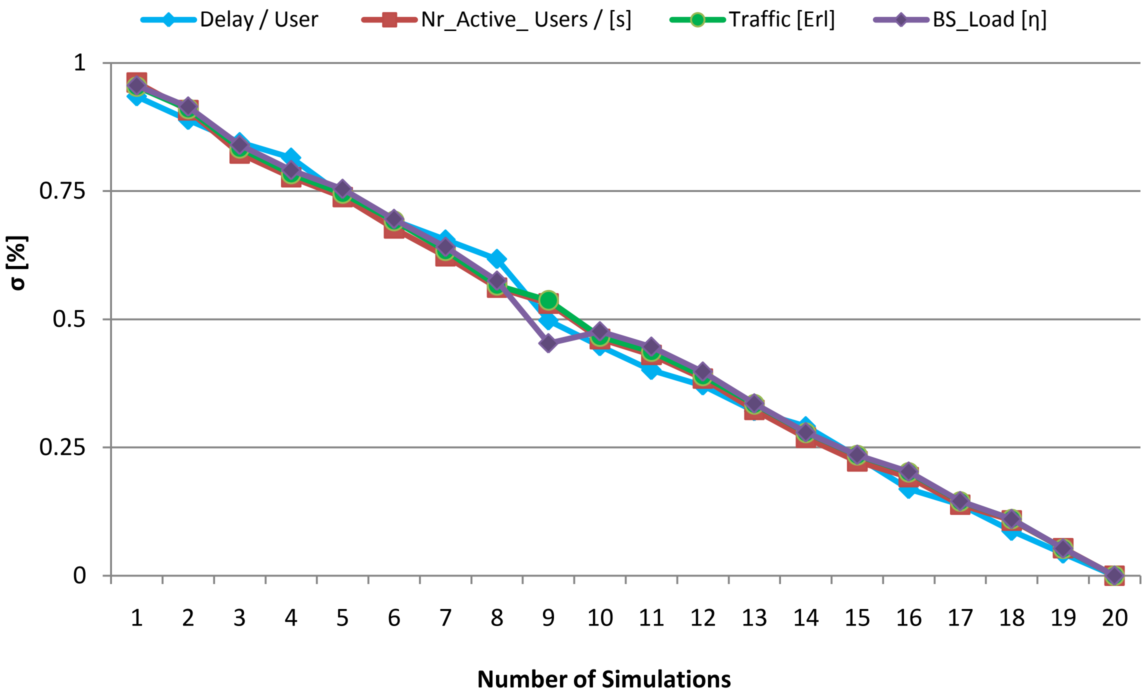

Figure 3-9 - Convergence of the ∂ function.

In Table 3-5, one shows the results for some studied parameters that characterised in part the simulator performance.

Table 3-5 - Average and standard deviation most relevant values after 20 simulations.

Cumulative simulations (total of 20)	N_u		η_{DL}		$\tau_{u\,[ms]}$		$T_{r\,[Erl]}$	
	Aver.	Std. Dev.	Aver.	Std. Dev.	Aver.	Std. Dev.	Aver.	Std. Dev.
5	81.0	4.28	0.16	0.017	99	0.015	36.7	3.601
10	81.3	1.12	0.17	0.014	94	0.012	38.1	3.037
15	81.5	0.55	0.16	0.013	91	0.012	37.8	2.635
20	83.0	0.10	0.17	0.013	91	0.011	38.2	2.513

As previously mentioned, voice users were just used to characterize in the best way present mobile networks where the increase in data services, makes the use of a second carrier, more common. In the analysed scenario during the assessment, voice and data have equal user densities. Once voice data has always precedence over data users, it is normal that the blocking rate is negligible, because to reach blocking voice users, with the employment of two carriers, it is needed higher values for “Busy Hour Call Attempt” (BHCA); this parameter is fixed to a value of 0.6, which means, a medium traffic network. In case of situations close to blocking, the impact is felt in data users - they are delay sensitive, each data service with different limits – that are delayed until voice transmission is finished and there are available resources to proceed with the setup process. In Table 3-6, one represents the simulation times for every simulation, defined in Figure 3-9. The sum of all values presented in Table 3-6 is 31h 43min 13s (approximately 1 day and a half), not including the time spent during initial studies (cumulative mean stability and “transitory” period). This simulation time justifies why the simulation time is considered a limiting factor in the process. From Table 3-6 it is also possible to observe a high convergence rate within all simulations; the standard deviation value is approximately 13% of the final mean value simulation.

Table 3-6 - Actual simulation times for simulator assessment.

Total of 20 Simulation	Average Simulation time	Standard Deviation Simulation Time	Total Simulation time
	1h 34 min 29 s	12 min 33 s	31 h 43 min 06 s

Chapter 4

Analysis of Results

In this chapter, the simulation results for different scenarios are presented and analysed in detail. In the first section, the reference scenario is characterised. In the following sections, the combination of services allows to evaluate network performance, which is described as a function of user-dependent and system-level parameters.

4.1 Scenarios Definition

In this section, the definition of the simulated scenarios is detailed. One starts to define the REF scenario, identifying the number of users, describing the applications and the usage profiles. Then, the simulated scenarios are described, identifying the changes made on the base scenario: variation of users (number, applications and usage profiles) and variation of system parameters (QoS parameters. like packet size).

The background is an urban environment, physically situated between Campo Pequeno and Saldanha, in Lisbon. Users are identified by two parameters: position and service (within each is the service priority). The position is defined according to a bi-dimensional uniform distribution in a 1000 x 1000 m^2 area. Users are assumed to be static, following the standard network behaviour (particularly for data, where the requested services are at the moment mostly performed without mobility). Each user has one service each time is going to transmit. In order to introduce service heterogeneity, five services are defined. Services penetration rates are associated to users according to Figure 4-1. By settling down a REF scenario, one seeks to reproduce a heterogeneous network integrating UMTS (R99 and R5) with 3 BSs (all 3 BSs performing voice and high speed data, according to different scenarios configuration), geographically distributed, as depicted in Figure 4-2.

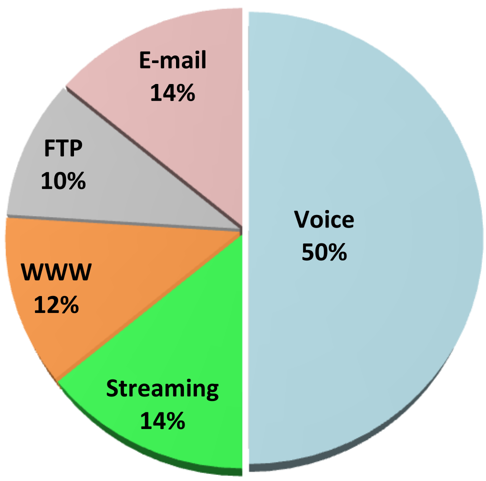

Figure 4-1 - Service Penetration for the REF scenario.

The REF scenario is generically based on the parameters defined in the AROMA project [AROM07]. From the 6 different applications employed in the set defined in AROMA, 5 were used: VoIP, Streaming, HTTP (WWW), FTP and E-mail; video was not studied, since until now is not a so used service. The number of users processing real-time applications (50% of voice users for the REF scenario), is the usual in current mobile networks.

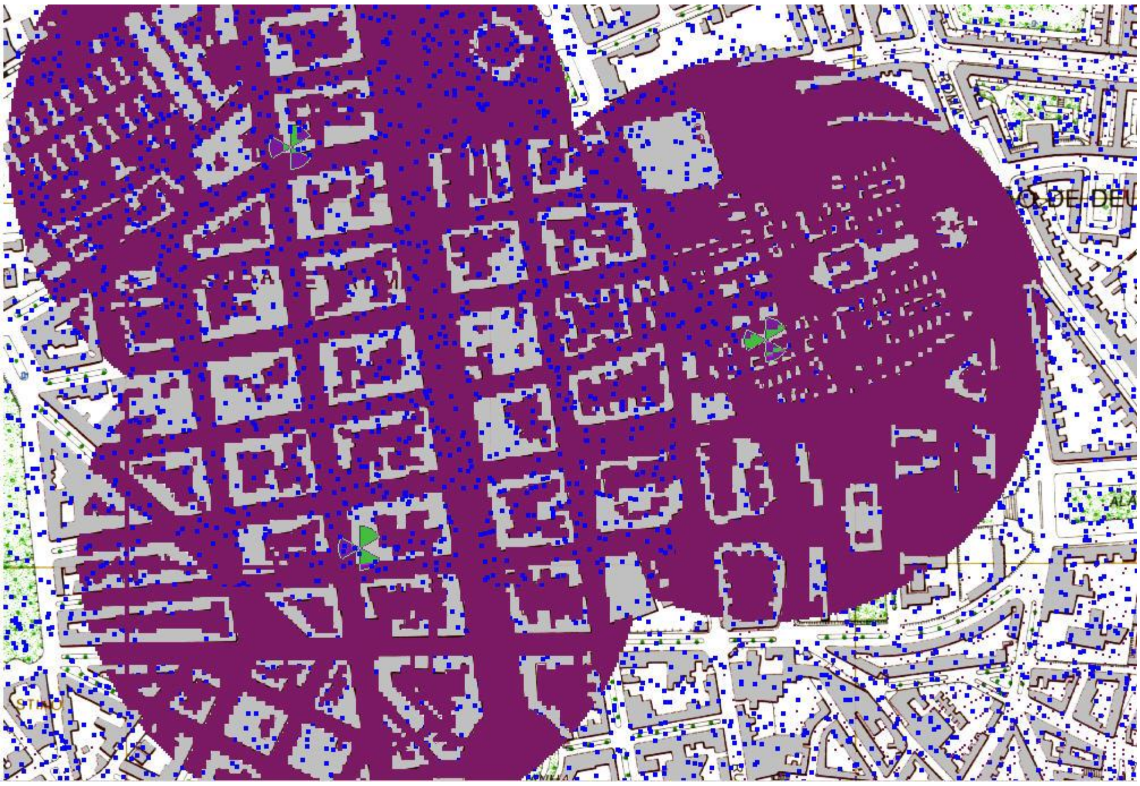

Figure 4-2 - BS distribution in the REF scenario and respective coverage areas.

In all scenarios, a correspondence between the bit rate and the number of codes employed was made, Table 3-2, which is inside a minimum and maximum threshold according to the radio resources availability at that moment, particularly load (power transmission) and number of available codes.

The goal of defining a service mix simulation set is to assess the impact of having different traffic patterns coming from the client side on the performance of applications and backbone network. The link budget inherent to a UE power transmission and its associated limitations is shown in Annex A. Traffic models description within each service can be consulted in Annex B; finally the propagation model is described in Annex C.

The behaviour of users in a network is somehow unpredictable, its density in a given area can be higher or lower, corresponding to a bigger or smaller number of users attached to the BSs on that area. From a service perspective, there can be variations both on the average duration of a voice call and on the packet volume of all data services employed. In Table 4-1, the most important parameter values used to characterise each service are presented. The service penetration is in this case an important issue, since different services require different level of resources.

In order to achieve a response of a typical network to this level of uncertainty affecting user dependent parameters, several simulation scenarios were created and processed; in each one, only one parameter is varied. This way, a more effective correspondence between input parameter variation and output results is more easily observed. The variations in the input parameters generate the following results for each scenario:

- Bit rate per service and in global network.
- Application codes attribution.

- Number of users, global and per service.
- Average delay per user.
- HSDPA/R99 usage service rate.
- Power output distribution between both releases R99 and R5.

Table 4-1 - Service parameters of the REF scenario.

Service / Parameter	**Voice call**	**Web browsing**	**Streaming**	**E-mail**	**FTP**
Penetration [%]	50	12	14	14	10
Mean Holding Time [s]	90	N/A			
Busy Hour Call Attempt (BHCA)	0.6	N/A			
Session/Service rate (Poisson)	N/A	3	3	3	3
Packet calls in a session	N/A	5	4	N/A	
Number of packets per call		25	200		
Reading/Viewing Time [s]		10	40		
Data Volume per packet (Pareto)		α=1.1; k=80			
Total data volume per session (Log Normal)		N/A		u_1=8.2; s_1^2=3.4	u_1=8.5; s_1^2=3.7
Inter-Arrival Time (Log Normal)				u_2=-4.4; s_2^2=4.5	u_2=-3.6; s_2^2=5.0

In some cases, a correspondence between different input parameters and a direct change on the correspondent output parameter exists, according to Table 4-2. In a more detailed analysis of those relations, one notes a separation between voice and data parameters, where in each case there is impact of changing a parameter associated to data services in the performance of main Key Performance Indicators (KPI) for voice, e.g., packet data volume does not cause any impact on Blocking Probability for voice users. Even if the direct impact of each parameter can be separated, globally both user types will compete for an access grant to the network. The REF scenario is the starting point, from which the other two scenarios are generated, both with higher density in packet services (less penetration rate in voice services). The two other scenarios, which derivate from the REF one are: Data Centric (DC) and Data Mostly (DM); in both cases, the mainly difference is the exclusively use of the second carrier for data services, in opposition to the REF scenario, Table 3-1, where both carriers are shared; the first carrier is always shared regardless of the scenario.

Table 4-2 - Inter dependency on inputs and outputs used in the simulator.

Inputs / Outputs	Users density	Packet data volume	Service penetration [%]	Scenario choice
Blocking Probability (P_b)	✓	✗	✓	✓
Average delay / user ($\overline{\tau_u}$(s))	✓	✓	✓	✓
Bit rate [Mbps]	✓	✓	✓	✓
BS Load	✓	✗	✓	✓
Number of active users / [s]	✓	✓	✓	✓
Codes / Power usage in HSDPA technology [%]	✓	✓	✓	✓
Traffic voice [Erl]	✓	✗	✓	✓

Data services gain in the two last scenarios an extra weight, by increasing their penetration rate: plus 10% in DC and plus 20% of data services in DM scenario, according to Figure 4-3 and Figure 4-4, respectively. In the three scenarios the input parameters under evaluation are:

- Number of users.
- Bit rate, global and per service.
- Codes attribution per service.
- Scenario choice (within service rate penetration).
- Packet data volumes.

One of the most relevant parameters in the planning of a mobile network is the maximum number of users it can support, while maintaining QoS inside previously defined thresholds. Three scenarios were defined, considering the following densities: 8 000, 10 000 and 12 000 users/km^2. All users are uniformly distributed over the area, where a given number of them falls out of coverage. Considering Blocking Probability of 2% as an acceptable level of GoS, the network would be under dimensioned for 12 000 users/km^2 and over dimensioned for 8 000 users/km^2. At the same time, the network has enough resources to deliver the demanded resources in all scenarios, keeping the average delay per user as low as possible; a delay bellow 100 ms can be considered a good performance in terms of QoS, once this can be accepted as a satisfactory metric for all applications, specially for most demanded data services like FTP or Streaming (maximum tolerable delay close to 250 ms).

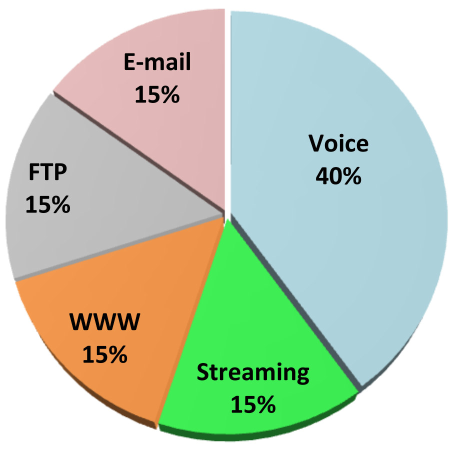

Figure 4-3 - DC scenario characteristics.

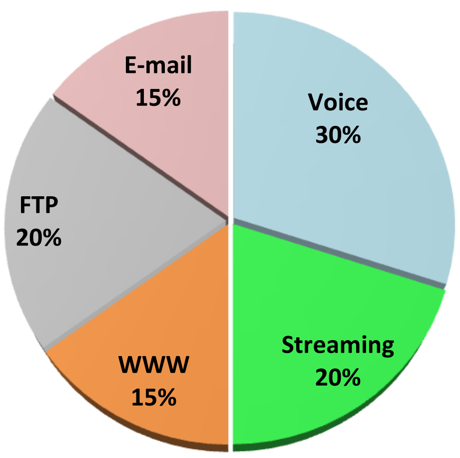

Figure 4-4 - DM scenario characteristics.

The number of codes per service was defined, making a direct correspondence: higher codes delivery (till a maximum of 15) to higher services demanding. Those scenarios will be distinguished by the following variables:

- Services Penetration (voice and data) per scenario.
- Packet size transmission.

The packet volume employed in data services (non-conversational) is generated according to a Pareto distribution with parameters α and k. In the REF scenario, the values for both parameters are respectively 1.1 and 80 bytes (Table 4-1); to create two additional scenarios for testing, the volume parameter, k was multiplied by 2 and 3, to obtain k =160 bytes and k =240 bytes. The k parameter value is common to Web browsing, Streaming, E-mail and FTP services. The increase of the packets volume causes additional consumption of the network resources and this effect reflects in network performance. In particular, direct impact is expected in $\overline{\tau_u}$.

The evaluation of the packet volume was performed for the three scenarios (REF, DC and DM), considering a user density of 10 000 users/km^2. Increasing the average size of the packet volumes to transmit means that on average each data user remains active during a longer time period. The relative weight of the data sessions in the overall network traffic is therefore increased. Data services includes FTP and Streaming, which are the most demanding ones in terms of radio resources to promote higher bit rates; for a network with a high average load (close to a maximum of 70% of its capacity) such situation could lead to a higher delay values per user.

4.2 Number of Users Variation

One of the input parameters to be analysed is the user density per km^2. The variation range (standard deviation) for each service is higher as expected for most demanding services, like Streaming and FTP, since their order of magnitude is higher and their minimum and maximum thresholds fluctuate between 384 kbps (minimum) to several Mbps – 5.6 in Streaming and 14.4 in FTP (theoretical values). In Figure 4-5, one shows the value for global bit rate that is on average close to 2 Mbps, which has to be considered a good value, since in the REF scenario there is a high penetration for voice (user density distribution of 50%) and, as previously defined, voice is processed at a constant bit rate of 15 kbps.

The highest average bit rate was obtained with FTP services – around 7 Mbps – followed by Streaming with values in order of 2 Mbps, and at last WWW and E-mail services, both with values around 300 kbps. If instead of average values, peak values are considered, the final results are quite different: FTP can reach values close to 14 Mbps and Streaming in the range of 5 Mbps; the remaining data services, like WWW and E-Mail registered values less than 384 kbps; such difference between average and peak values is not that large, since the variations (minimum and maximum values) inside each service, are of different magnitude; their variation is different (E-mail and web browsing have just some kbps differences between lower and upper limits, but on the other hand Streaming and FTP can present variations of several Mbps between each new request, Figure 4-6. A small variation within each scenario for each computed average bit rate per service is observed. This is an expected value, since this key performance indicator is to be kept as maximum as possible and inside defined thresholds; when it is not possible to guarantee such performance, data users are delayed and voice users are blocked, instead of applying any type of bit rate reduction to accomplish the purposed targets. This is a strategy employed in this algorithm being just a possible choice among several possible management strategies.

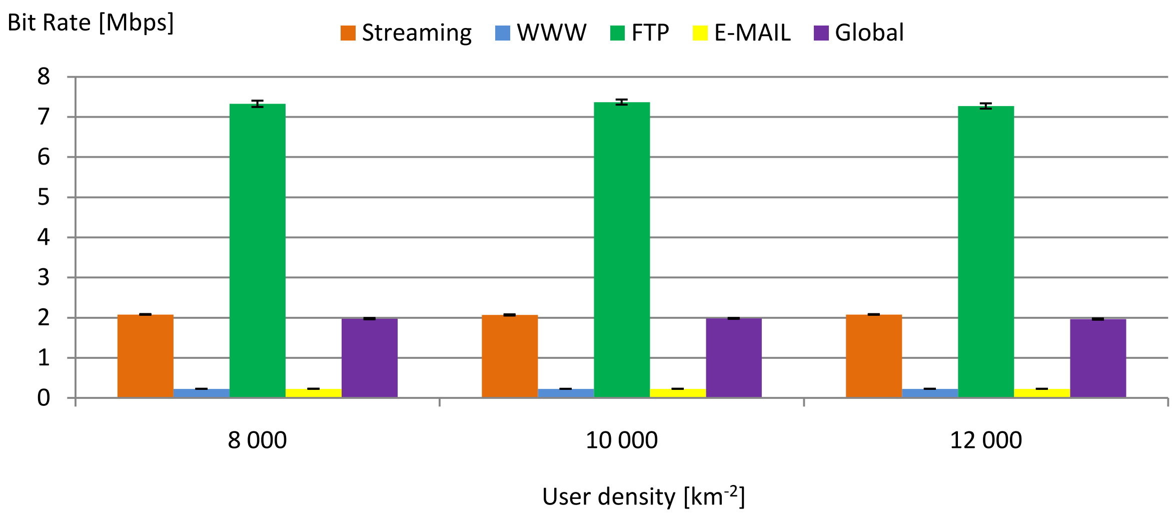

Figure 4-5 - Bit rate performance per service for the REF scenario.

As expected, between services like E-mail, that runs in background and just needs some hundreds of kbps to be realised, and on the other hand services like Streaming or FTP, which are performed online, there is a difference on radio resource needs, to finish each request successfully and as fast as possible.

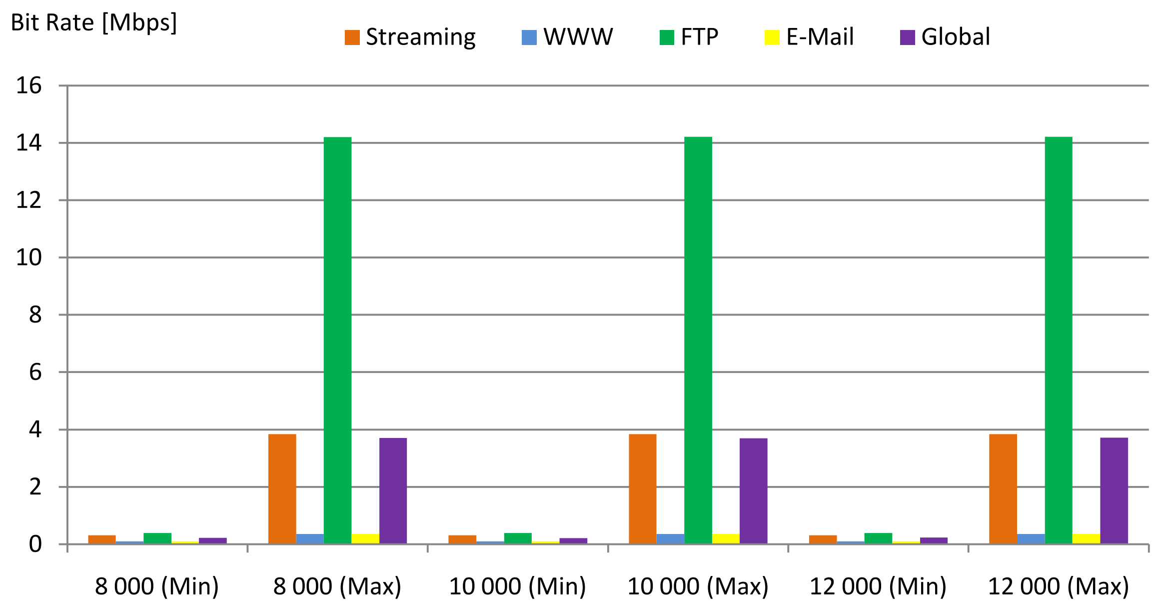

Figure 4-6 - Bit rate variation per service for the user density variation in the REF scenario.

The standard deviation between each different simulation is negligible for all data services, even if inside each simulation the difference between peak values (minimum and maximum) is significant for services like FTP and Streaming; among different simulations, the final results for bit rates presents a high convergence. In Table 4-3, one shows that on average, and regardless the user density value, the system provides satisfactory bit rates per service, every time there are enough resources (power and codes) available in the network. The main variation is in the average delay per data user, since to keep a high bit rate performance, when the number of simultaneous users in the network raises up, the system responds with higher average delays per user, resulting from an increase in competition for network resources.

Table 4-3 - Bit rate variation per service.

Bit Rate Global [Mbps] / service per user density						
User density [km^{-2}]	8 000		10 000		12 000	
Services	Min	Max	Min	Max	Min	Max
Streaming	0.307	3.839	0.307	3.839	0.307	3.840
WWW	0.096	0.360	0.096	0.360	0.096	0.360
FTP	0.384	14.207	0.384	14.208	0.384	14.208
E-Mail	0.096	0.359	0.096	0.360	0.096	0.360
Global	0.223	3.707	0.209	3.695	0.236	3.718

The codes allocation for each service in the REF scenario follows the mapping presented in Table 3-2, where, depending on the available resources (power and codes), each time a new user starts to transmit, it takes in consideration all of its neighbour mobiles, which are already performing a service.

In Figure 4-7, one can notes a small decrease in the codes allocation for higher bit rate services (FTP and Streaming) when there is an increase in the number of users in the network, this makes the CAC algorithm to be re-adapted to new circumstances: users requesting higher bit rates, will try to be supported, joining anytime is possible the maximum number of users in good radio performance. For low bit rate data services (WWW and E-Mail), such variation in user density has not so big impact, since in this cases any extra request motivated by a higher number of users does not cause, or is far from, a lack of resources. For most demanding services, one notes a higher variation in the number of allocated codes, for lower users densities. The resources allocation margin for data users, is higher for the REF scenario, since voice users always request a low number of codes; so when the number of voice users is significant and user density is not the highest, the network itself allocates more codes for all data services, as soon as the total number of users in the network goes to values higher than 10 000, such resource allocation variation is not so evident, Figure 4-7. The number of active users per second in the REF scenario has a significant presence of voice users; this user generation process is randomly built. The number of users per second and per service, as expected, grows up (particularly for voice, WWW and E-mail), following the higher density users tendency grow, from 8 000 to 12 000 users [km^{-2}].

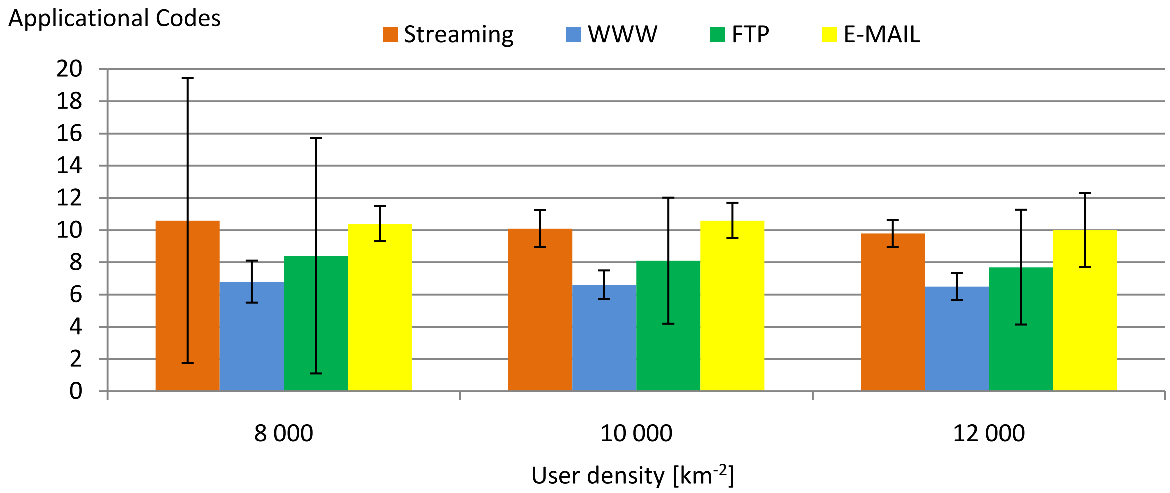

Figure 4-7 - Codes allocation for the REF scenario.

In Figure 4-8, one observes that the standard deviation in the REF scenario is higher for data services in higher user densities, which can be explained in some situations with an increase on the number of data users "delayed", as long as the user densities grow up. As a consequence, the RRM algorithm reacts by keeping the system under congestion limits, giving always priority to voice.

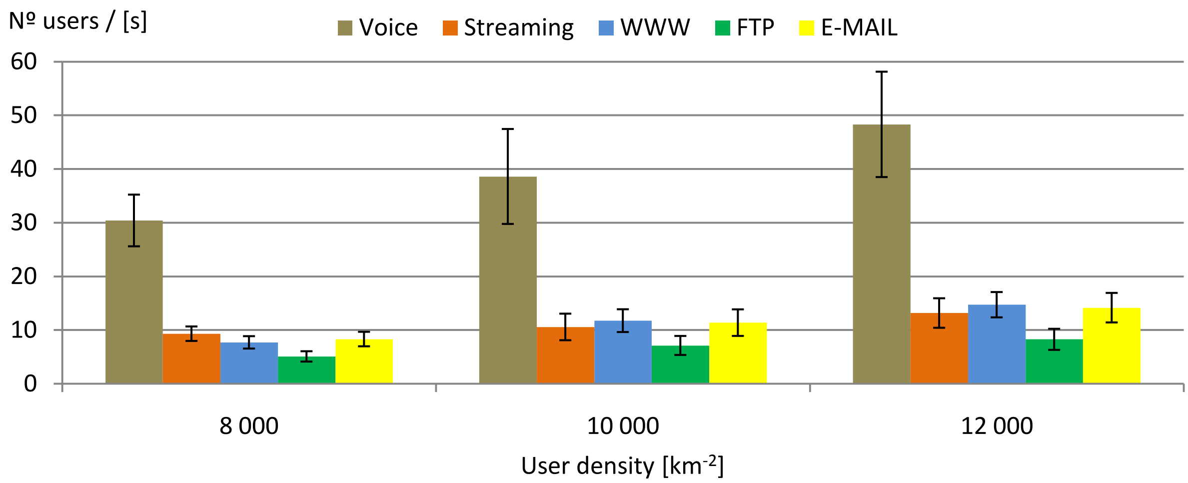

Figure 4-8 - Active users per service for the REF scenario

The average delay shows a coherent behaviour according to different user density evolution. In Figure 4-9, this parameter follows an increase tendency when the user density is incremented from 8 000 to 12 000 users [km^{-2}]; such slightly variance was computed by a ratio factor of 1.12, between each different user/density profile in the REF scenario. Despite of this variation, the absolute delay per user, is on average close to 100 ms, regardless of the user density. Despite of the ascending values for average delays, they are always maintained inside target values. The load factor, is on average kept under the upper limit (η_{DL}<0.7); in cases where this limit is overcome, users are delayed (PS services) or blocked (CS services). Another important conclusion is that when the voice services weight is close to 50%, the network capacity, with 2 UMTS carriers, is able to deal with the majority user requests, keeping the average delay per user close to 100 ms. The resulting average delay or block is often caused by lack of power distribution, instead of lack of codes.

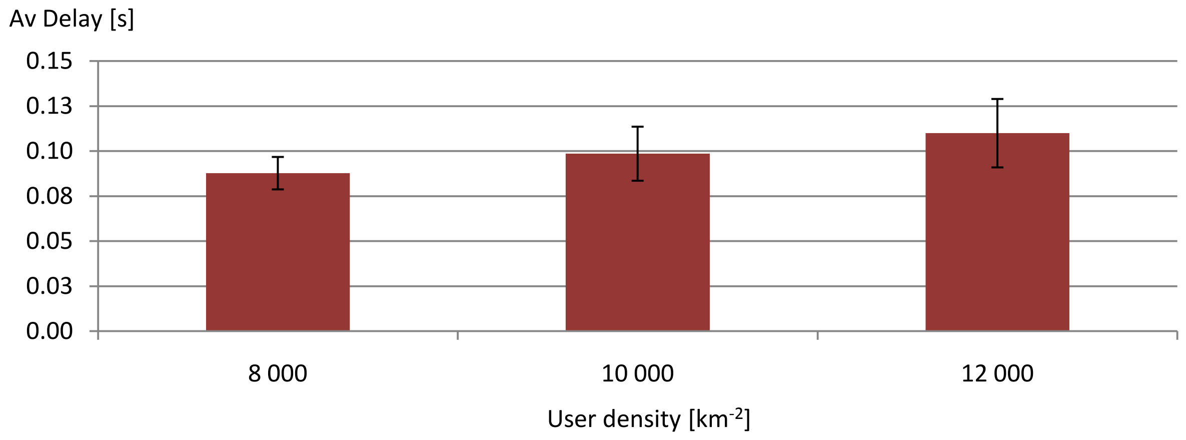

Figure 4-9 - Average delay for the REF scenario.

The DC scenario is characterised by an increase in data services, which means less 10% in voice user density per km^2. As previously pointed out, this increase in the penetration rate is accompanied by a concentration of voice service in the first carrier; data services being allocated in both carriers. In comparison with the REF scenario it is possible to see (on average) similar bit rates, with the same order of magnitude for the services under study, but with higher codes allocation for most demanding services (Streaming and FTP), which means that data users have more codes available to reach high bit rates without violating the pre-defined limits, Figure 4-10. Voice is a service that was planned at constant bit rate (15 kbps) and a number of codes constant for each user (always 1 code for each request), so there is no additional problem (congestion, causing voice blocking services) by concentrating voice in the first carrier. At the same time, data users gain extra resources (that in previous scenario are shared in the second carrier) to improve its code allocation that on average does not change too much the final bit rate for each data service. Since there is an increase in data users, normally the services with higher standard variation are those with greater range variation, as it can be confirmed in Figure 4-10.

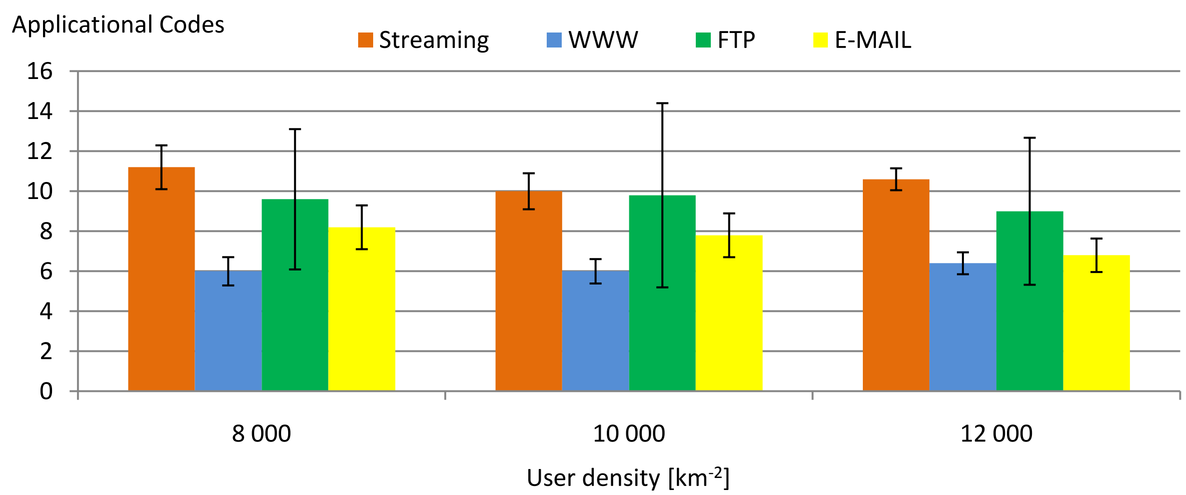

Figure 4-10 - Codes allocation for the DC scenario.

Data users in the present scenario represent more than 50% of the total population considered, which is why the number of data users per second increased in comparison with the previous scenario. In Figure 4-11, one shows that the number of voice users still increase as before, following the tendency of data user density increase; the same behavior happens in a homogeneous manner to all data services under study; voice users represent always less than 45% of the total amount of population is DC scenario. For data load, the difference with the previous scenario is similar, because the power transmission associated to each voice or data user is the same depending on power transmission any variation in the user profile affects on average the final load in the network.

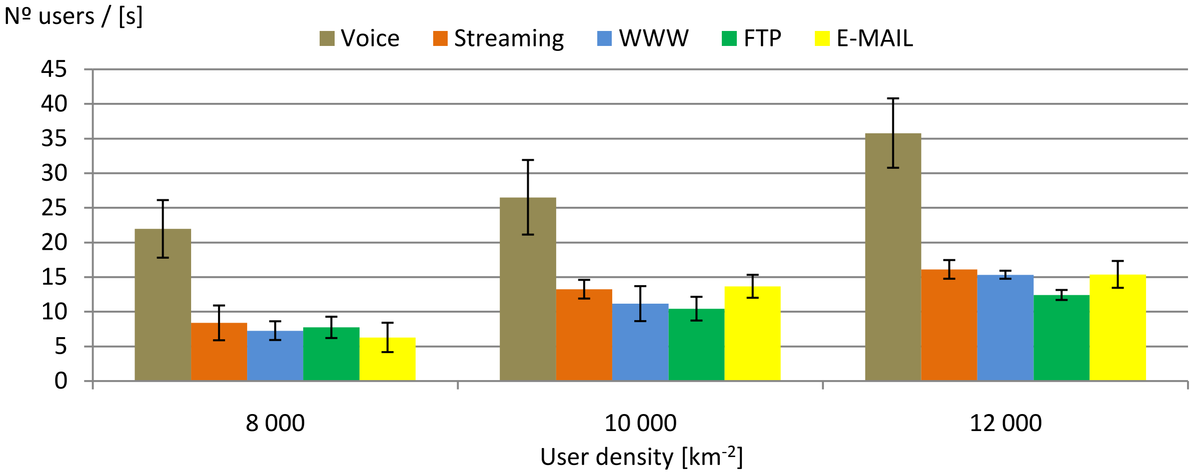

Figure 4-11 - Active users per service for the DC scenario.

On the other hand, the delay associated to each user presents on average a higher probability to happen, because with higher data services penetration, more radio resources (codes) are requested (particularly for FTP and Streaming), so the delay per user is probably higher whatever the user density is; one shows this situation in Figure 4-12. A difference has to be pointed to the REF scenario: for all user density values, the average delays per user are higher than 100 ms.

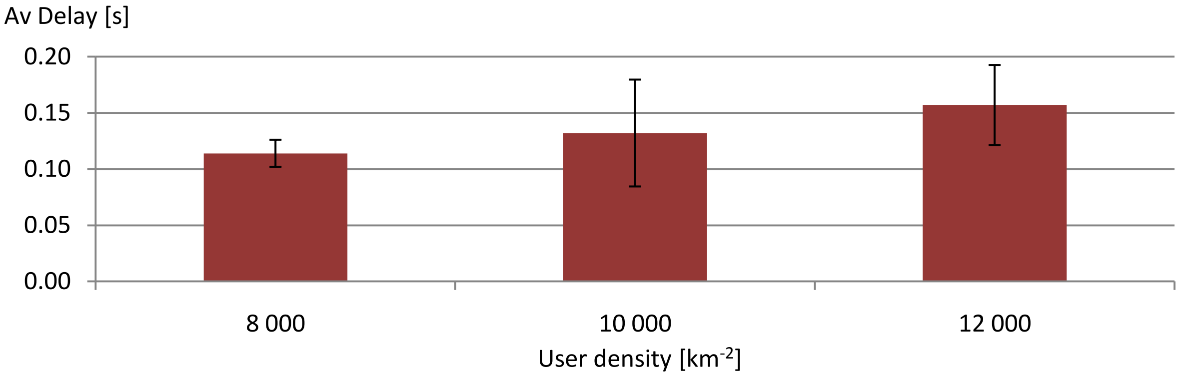

Figure 4-12 - Average delay for the DC scenario.

In the last scenario, a network with data user density equal to 70% is defined, keeping the main characteristics of DC scenario. As expected, the global bit rate and bit rate per service stay inside the same range values as for the REF and DC scenarios, which means that regardless of the scenario, the bit rate per service and the global network bit rate are not significantly affected by the input change in user density. One notes a higher number of codes employment in data services, from a higher demanding services: the network whenever is possible allocates a higher number of codes, shown in Figure 4-13, there are not many differences in bit rates per user, but for the maximum values per service, there are some samples where such higher codes allocation has a positive impact in the final bit rate. The standard deviation is lower when the network is more crowded (higher user density).

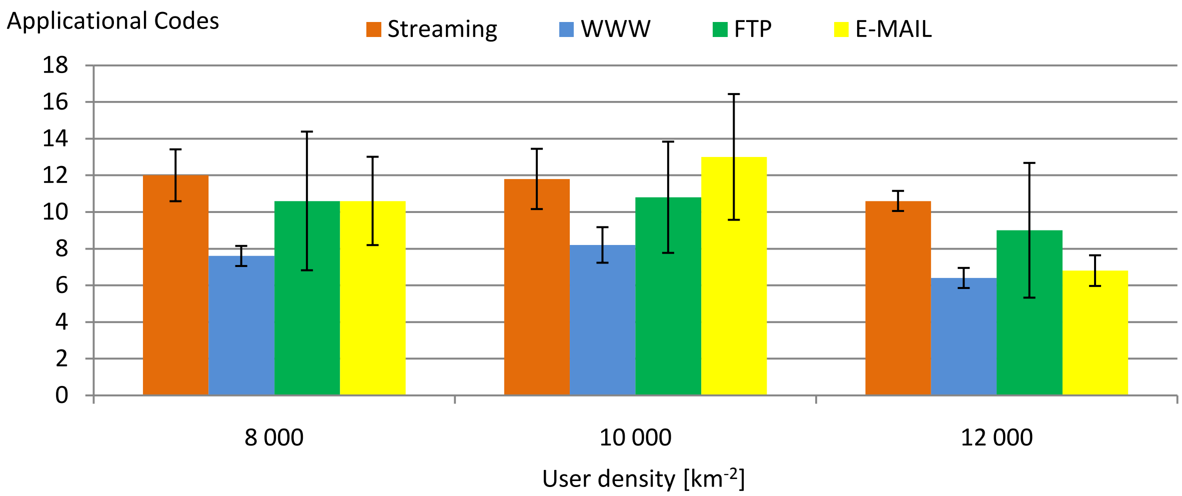

Figure 4-13 - Codes allocation for the DM scenario.

The number of data users increases in this scenario: on average, there are always more than 70% of data users active in the network for this scenario - Figure 4-14. The variation in the number of voice users still exists, and there is also a significant increase in all data services, with the exception of FTP. The services with higher variations are as usual FTP and Streaming, but in this scenario, also E-mail presents a higher standard deviation, mainly because the number of e-mail users has a particularly increase in the DM scenario. Average delay per user suffers a small increase in this scenario. With higher penetration rate for data services and greater rates for Streaming and FTP (both with 20%, WWW and E-Mail with 15% each) a higher number of data sessions occurs, which results in an increase in the total load network, and as a result and because the data users are dominant, in particular instants there is a slight increase in the delay per user, τ_u.

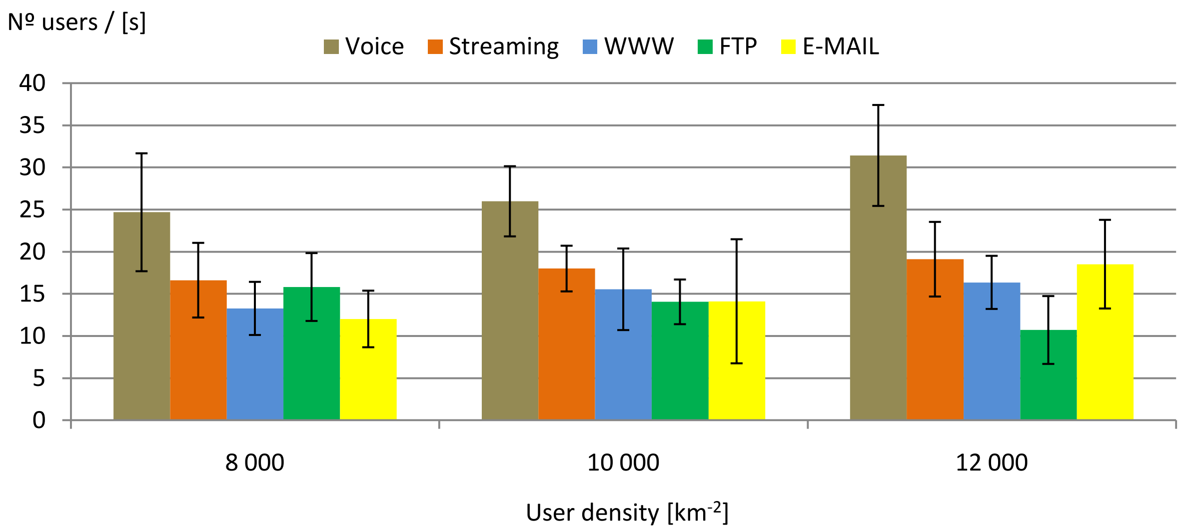

Figure 4-14 - Active users per service for the DM scenario.

With such increase, the final results for average delay per user, suffers a degradation - Figure 4-15. The optimal average value of 100 ms per user is just accomplished for a user density equal to 8 000; for the remaining user densities the final delay/user is always over 150 ms.

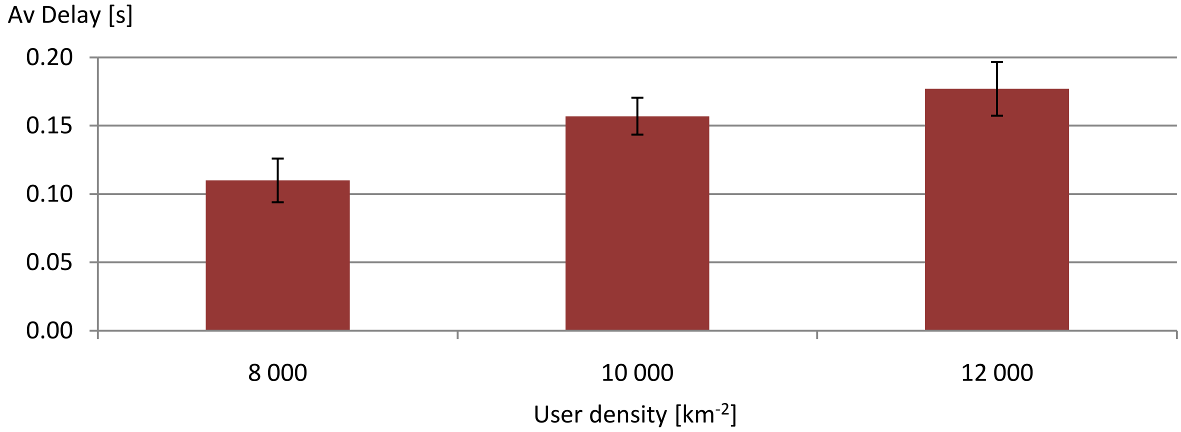

Figure 4-15 - Average delay for the DM scenario.

A comparative study is presented, just to retrieve the final conclusions, comparing the three main scenarios: REF, DC and DM. The parameters involved in the scenario comparison are: the number of codes employed per service, the resulting application bit rate per service, the evolution in the number of active users per service and the average delay: the later one results from the correlation of all previous results and indicates if each user stays a large amount of time in idle mode, with any possibility to transmit until the network has enough resources to allow a new transmission. In Figure 4-16, one shows that, despite of the scenario used, the number of codes does not present significant variations. When the scenarios have data predominance, there are more resources available to data services, and because the most priority voice service has lower user density, there is more freedom and probability to have data sessions on going, with higher performance whenever CAC algorithm allows it.

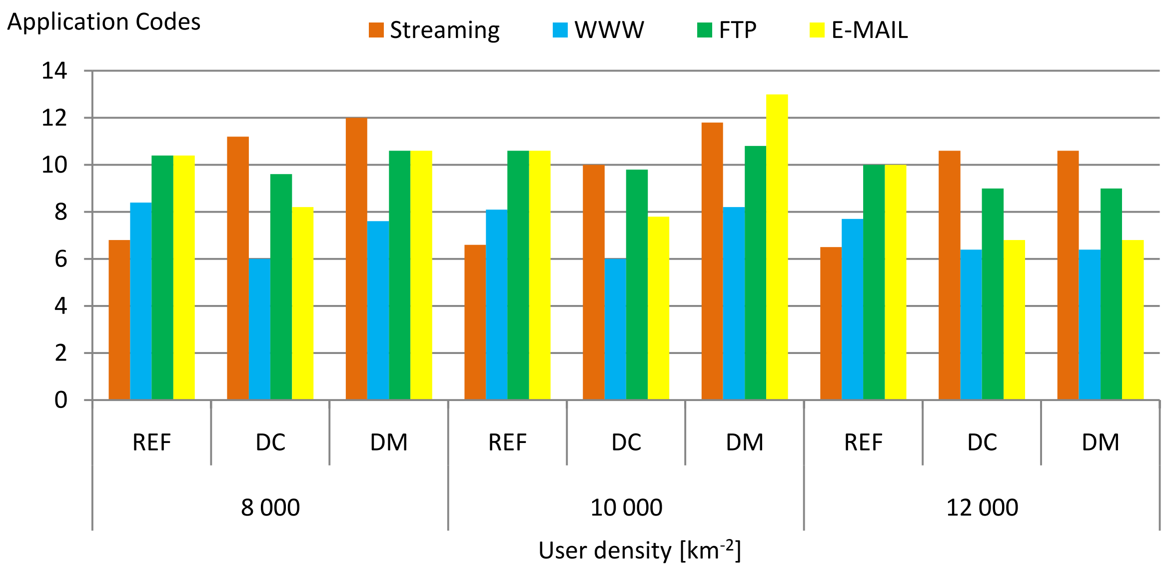

Figure 4-16 - Comparative study for the number of codes in the three scenarios.

The perspective pointed out for codes allocation, is applicable to bit rates and shown in Figure 4-17, since those parameters are correlated. As previously mentioned, the bit rate is not particularly influenced by the employed scenario.

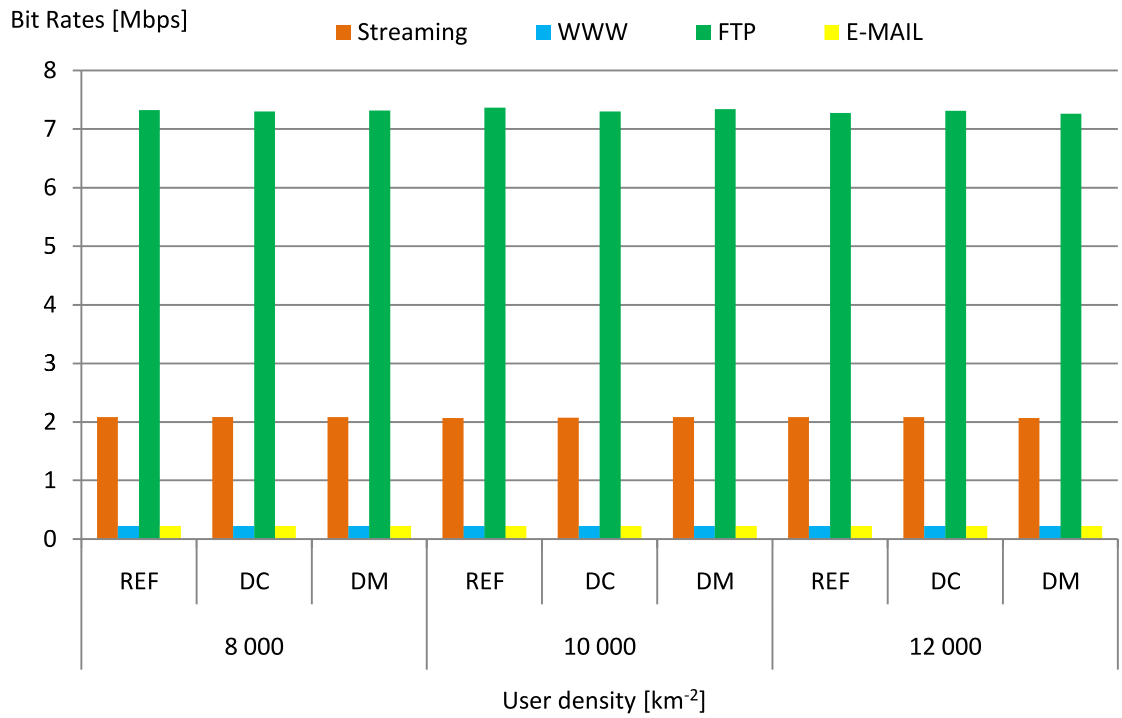

Figure 4-17 - Comparative study for the bit rate in the three scenarios.

In Figure 4-18, the number of active users follows the tendency of each user density scenario, since voice user population is just close to 50% of total active users in the REF scenario, whatever the user density involved.

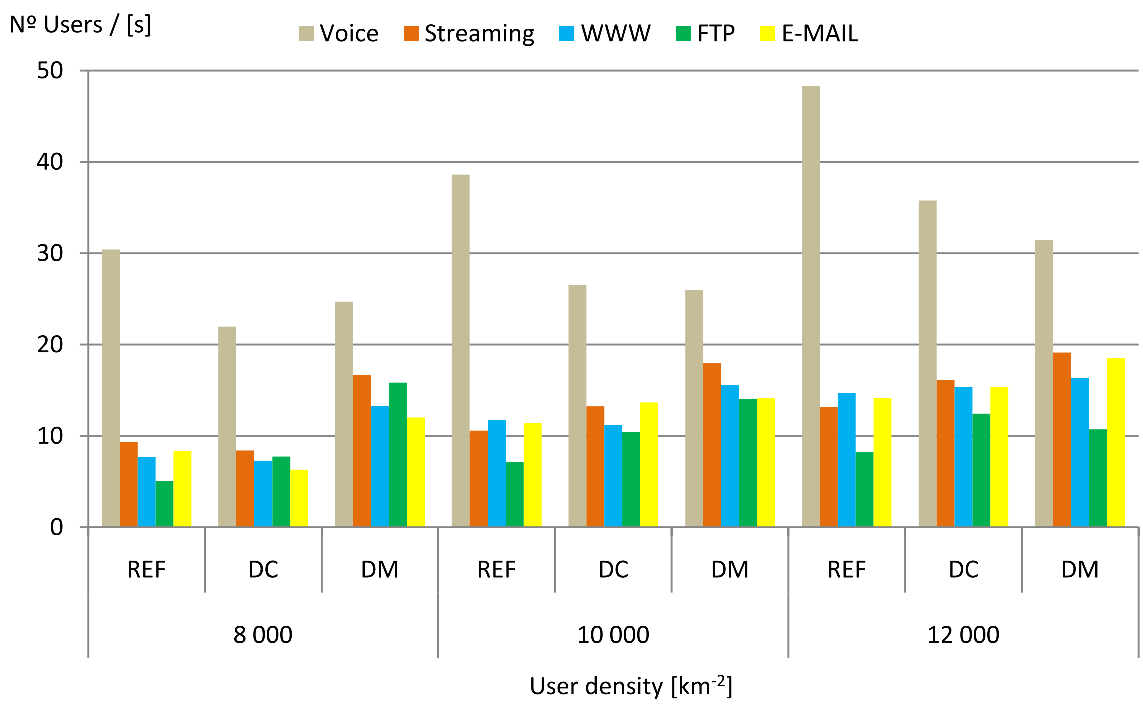

Figure 4-18 - Comparative study for the number of active users in the three scenarios.

In DC and DM, as expected, a higher unbalance between the number of voice and data users is seen. The average delay shown in Figure 4-19 shows that the best performance is obtained for the REF scenario, where regardless of the used scenario the average delay just exceeds 100 ms for a user density of 12 000 users [km^{-2}]. On the other hand, all data scenarios, with the exception of the one with a user density of 8 000 users [km^{-2}], have average delays close or higher to 150 ms.

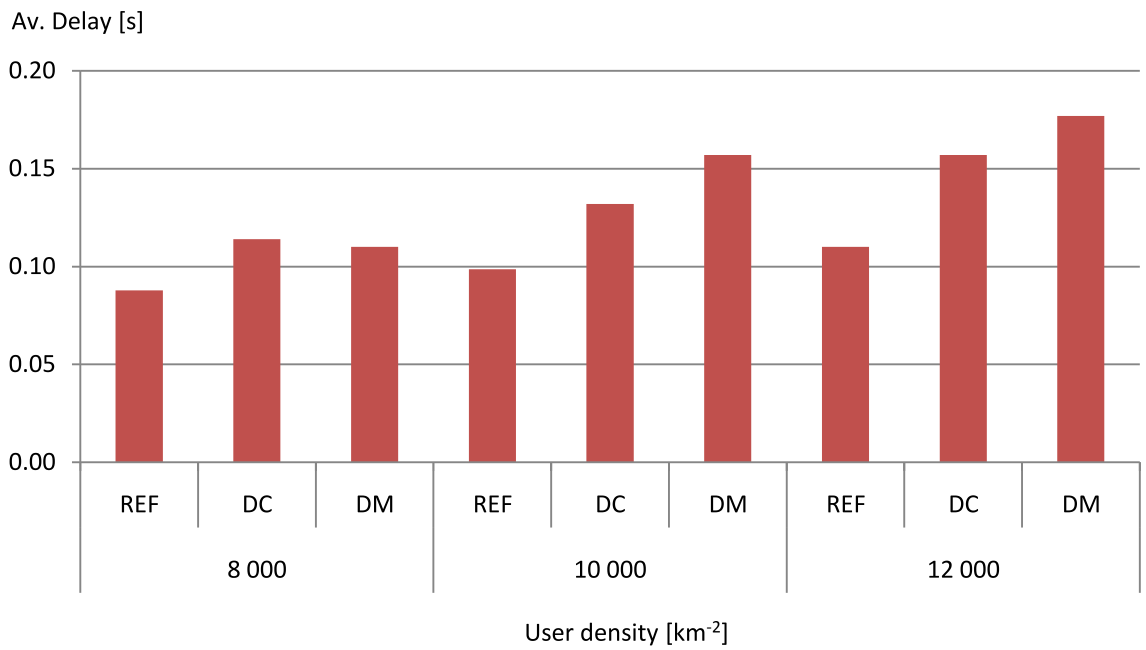

Figure 4-19 - Comparative study for the average delay in the three scenarios.

4.3 Services Penetration

In the REF scenario, users are assumed to be uniformly distributed in space. The characterisation of the ∂ function for the REF scenario is based on the result of 5 simulations. The error margin of the results is consequently lower, which is important, since they serve as the anchor for all the scenarios under analysis in this work. The final result for all scenarios is sumarised in Table 4-4. From each scenario the high level of convergence that globally affects the service indicators is shown. This tendency is observed in a relevant part of the results throughout this thesis, putting in evidence that even if the simulator has a random nature, the coherence of results is a reality. The observed dispersion can be understood just as a mirror of the reality itself. The blocking probability is negligible and in the majority of cases does not happen, due to the strategy employed, where voice has always precedence over data services. Regardless of the scenario, the global network bit rate is always close to 2 Mbps. The number of active users per second follows on average a tendency equal to the user density values, and the distribution of voice and data users rate in each scenario. In both scenarios where the user density for voice is lower (DC and DM), the number of data users increases as expected and the request for more radio resources (particularly application codes) also raises.

The most significant variations occur in the number of active users per service and the average delay associated with each scenario; the average delay increases as soon as the data users rate grows up.

Table 4-4 - ∂ function comparative analysis for all scenarios (10 000 users [km^{-2}]).

Scenarios		REF		DC		DM	
Main parameters		Mean value	St.dev	Mean value	St.dev	Mean value	St.dev
P_b[%]		0	0	0	0	0	0
$\overline{\tau_u}_{\text{[ms]}}$		99	0.015	120	0.05	145	0.014
$\overline{R_{b_s}}_{\text{[Mbps]}}$.	Voice call	0.015	0	0.015	0	0.015	0
	Web browsing	0.227	0.008	0.228	0.002	0.228	0.001
	Streaming	2.067	0.015	2.073	0.026	2.079	0.004
	E-Mail	0.228	0.008	0.227	0.001	0.228	0.001
	FTP	7.368	0.064	7.299	0.070	7.338	0.059
$\overline{R_{b_G}}_{\text{[Mbps]}}$.	Voice	0.015	0	0.015	0	0.015	0
	Data	1.981	0.011	1.968	0.018	1.978	0.011
$\overline{Nu_{srv}}$	Voice call	39	8.84	27	5.39	26	4.16
	Web browsing	12	2.12	11	2.53	16	4.84
	Streaming	11	2.48	13	1.35	18	2.71
	E-Mail	11	1.10	14	1.66	14	7.36
	FTP	7	3.91	10	1.71	14	2.65

One of the reasons for having negligible values of P_b is exactly explained by the algorithm structure: voice is the highest priority service, and even when there are delayed PS users trying to enter again in the network, they will first check if there are voice users to access the network at the same time and if there are, voice users will be firstly granted. The voice service is the lowest demanding one in terms of radio resources, which combined with input parameters parameterisation (BHCA=0.6 is a medium loaded network, not that much demanding and just chosen to follow a real network parameterisation) makes this P_b parameter not so relevant, mainly because the voice service is just employed to a perfect characterisation of present networks and also because R5 makes use of share resources with the voice service.

4.4 Packet Volume of Data Services

A change in the size of packets to be transmitted intends to analyse if there is a direct relationship between $\overline{\tau_u}$ and k with a Correlation Coefficient above 95%. When k is doubled from 80 to 160 bytes, $\overline{\tau_u}$ follows the tendency and is multiplied by 1.5 in each scenario. The following study was performed for all the three scenarios for a user density of 10 000 users [km^{-2}]; the same output parameters are being evaluated, for a packet size variation in two stages:

- k =160
- k =240

In the previous sections packet size variation k value is always equal to 80. In the REF scenario, the absolute average bit rates per service, keeps the same magnitude values, which means that if the packet size to be transmitted is higher and if the bit rate does not change too much, a longer time is needed to transmit all information, which will not result in significant changes in the average bit rate for each service. Packets with greater size result in higher channel occupation. In Figure 4-20, one shows the number of codes allocated to each service, which does not present great variations in comparison to all previous results. As before, the most demanding services employ a higher number of codes per user, but globally, because there is a higher number of users performing E-mail and WWW, on average the total amount of codes per second for each group of users performing the same service, does not present significant differences.

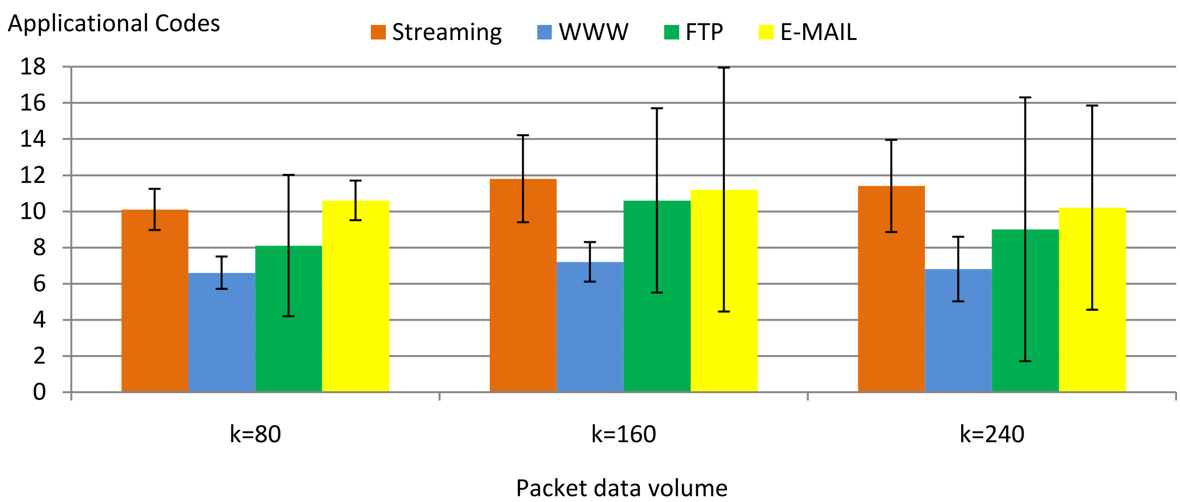

Figure 4-20 - Codes allocation for the REF scenario.

It is just important to note the higher standard deviation when k is changed from 80 to 160 and 240, even for the most demanding services, which normally in the REF scenario (with k=80) presents more stable results.The number of active users in the REF scenario has a new distribution when the value of k is changed. For voice users, as expected, its number stays similar to previous values, but the data services suffer a slight evolution with different weights when the value of k is multiplied by 2 and 3 – Figure 4-21. With k =160, the voice users just represent 33% of the total active users (with k =80, this value is 62%); the transition from *k*=160 to *k*=240, just causes a small adjustment in the number of data users: when transmitting with the highest packet size value for transmission (k =240), data users represent 64% of the total active population.

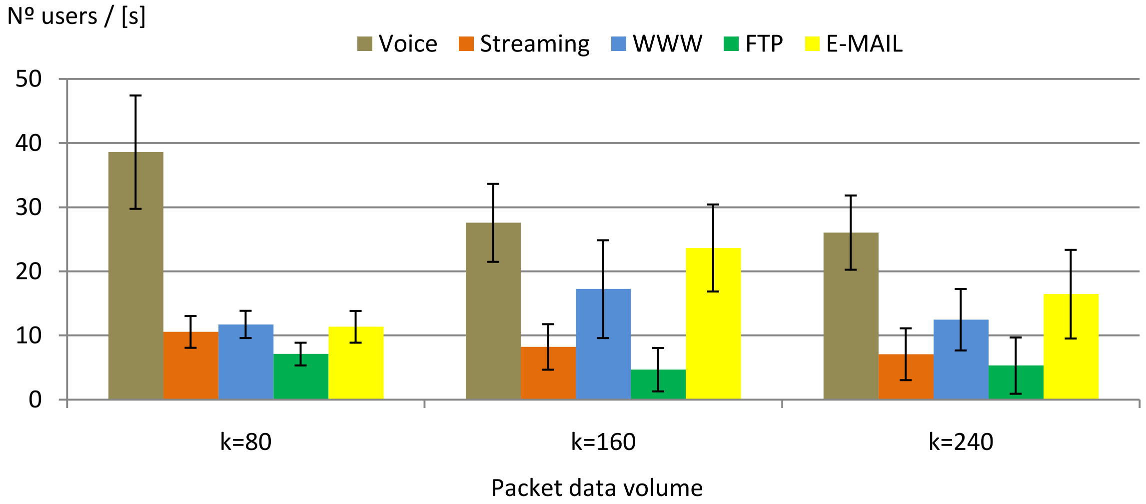

Figure 4-21 - Active users per service for the REF scenario.

One notes an increase in the average delay per user, when the packet size to transmit is the highest as possible: in this case the average delay increases almost 50 ms per user, when the value of k triplicates; with such results, it can be concluded that for the REF scenario, the best tradeoff between bit rate and delay is reached with k equal to 80 - Figure 4-22.

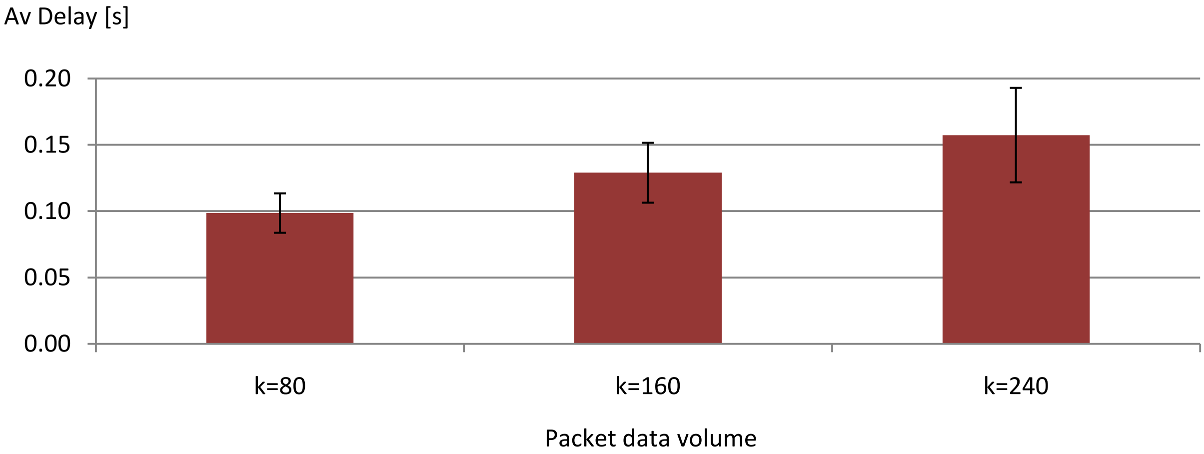

Figure 4-22 - Average delay for the REF scenario.

In scenarios where the penetration rate for data services is higher, it is interesting to see that the number of codes allocated on average had a slightly decrease, without a significant impact in the final average bit rate per service. This shows that the system algorithm, because the users stayed more time connected, did in some circumstances, adjustments particularly in Streaming and FTP, as one shows in Figure 4-23; this is the way to allow new accesses to the network: new users entering by the first time or re-entering again, coming from delayed state.

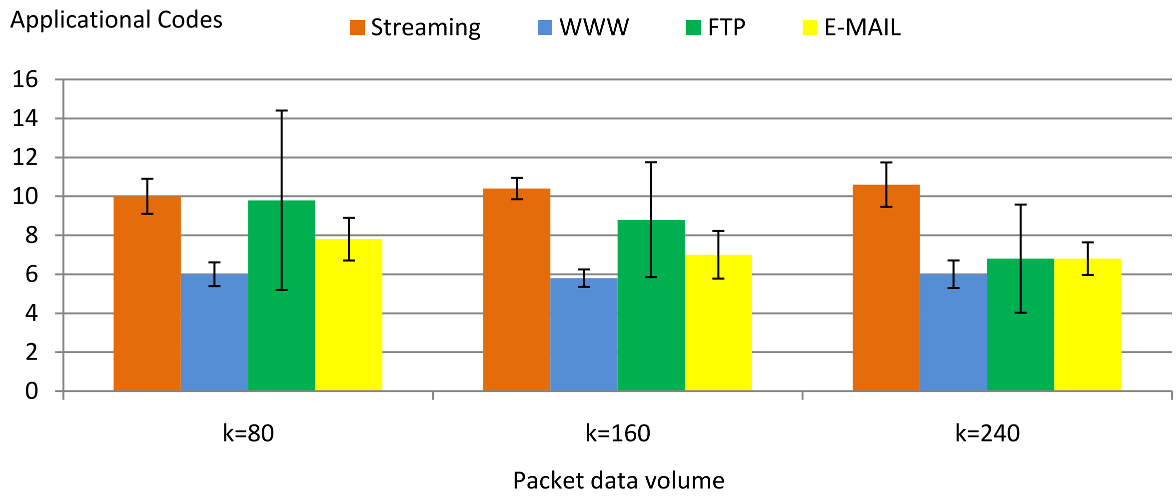

Figure 4-23 - Codes allocation for the DC scenario.

In Figure 4-24 one presents the number of active users per service, where there are significant differences when the value of k changes from 80 to 240. For higher packet sizes (k=160 or 240 bytes), there is an expected decrease in the number of simultaneous active users per service, since each active user is on average performing a higher usage of the channel transmission: in this case the most affected services are Streaming and FTP.

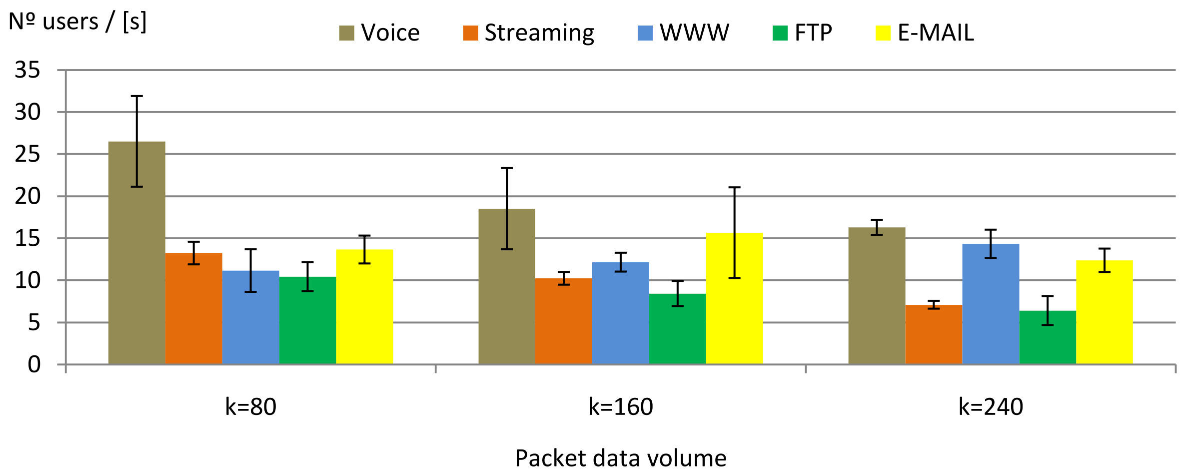

Figure 4-24 - Active users per service for the DC scenario.

The average load value does not present high variations: users when in delay mode are not taken in account to the final load network computation, since they still have a resource locked, in reality they are not transmitting, just suspended. Since an increase in the delay per user occurs, now with values that cannot be considered satisfactory, since they overcome the defined limit of 100 ms per user (for all k values), as seen in Figure 4-25. A target of 100 ms was considered the minimum acceptable in a

R5 network, to assure a satisfactory GoS. Once again, the best trade-off is achieved with k=80, which is undoubtedly the recommended value to be used in this type of network configuration. The delay associated to variations in k value, can be extrapolated to a R5 network, where in cases of over sized packets, the number of retransmissions will increase, resulting also in degradation in BLER, associated to each transmitted packet. The best effort is always to find the point where the optimal packet size is found to avoid an increase on BLER, but at the same time do not allow for too much blank times in the transmission intervals (TTI).

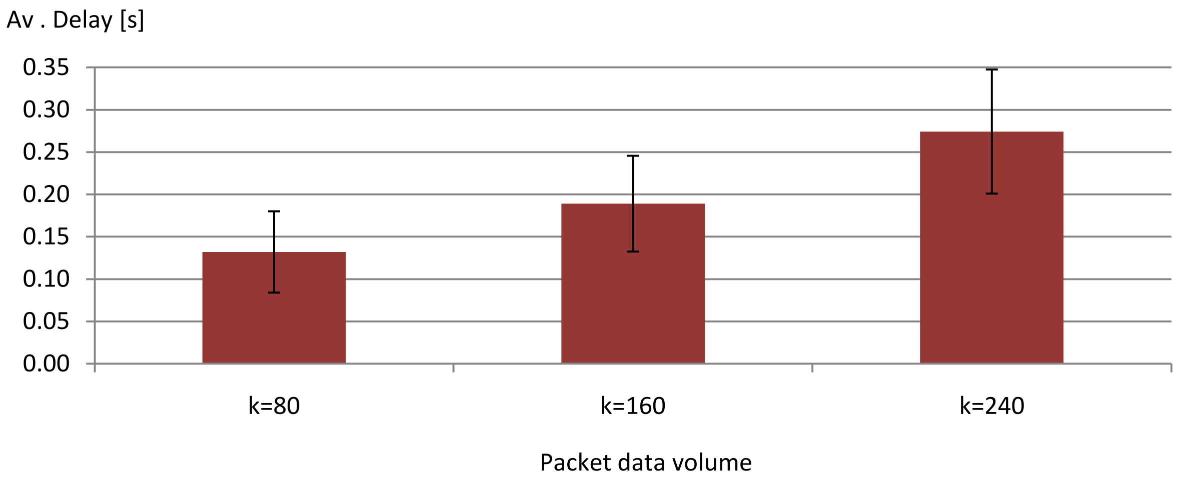

Figure 4-25 - Average delay for the DC scenario.

In Figure 4-26, one shows for the DM scenario that codes allocation is following all the previous studies where the bit rate and codes attribution present values that show a convergence of results, already expected.

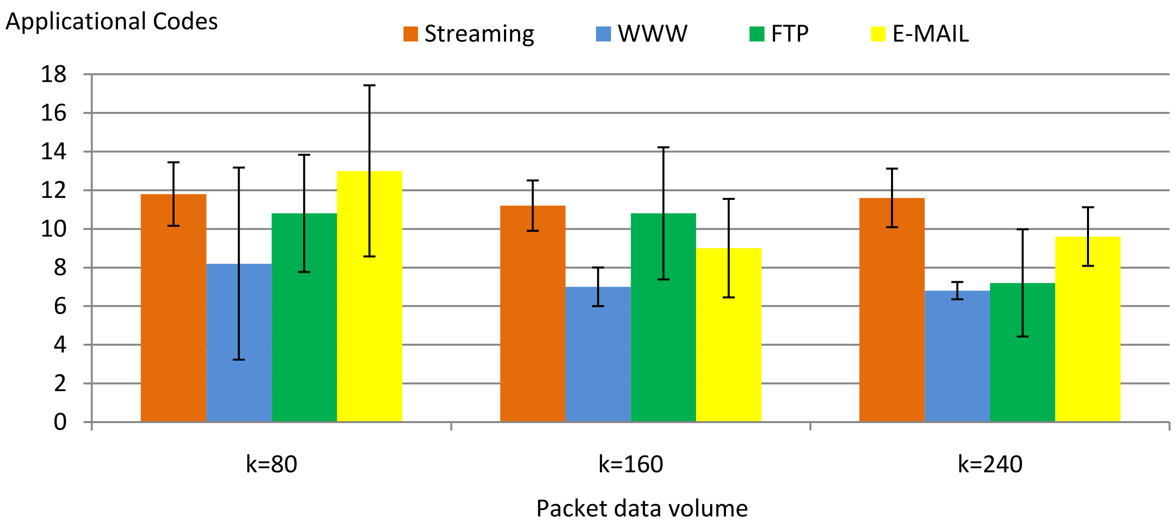

Figure 4-26 - Codes allocation for the DM scenario.

The number of active users/service per second shows a variation for voice service, more evident when k varies from 80 to 160, since the change on k value only affects voice indirectly (data users are using channel for transmission in a higher amount of time). For active data users there is a global decrease in all services on each scenario transition, particularly when k varies from 80 to 160, Figure 4-27; a stable evolution in the number of active users for WWW and E-mail is noted, with low standard deviations, in opposition to higher variations for Streaming and especially FTP services.

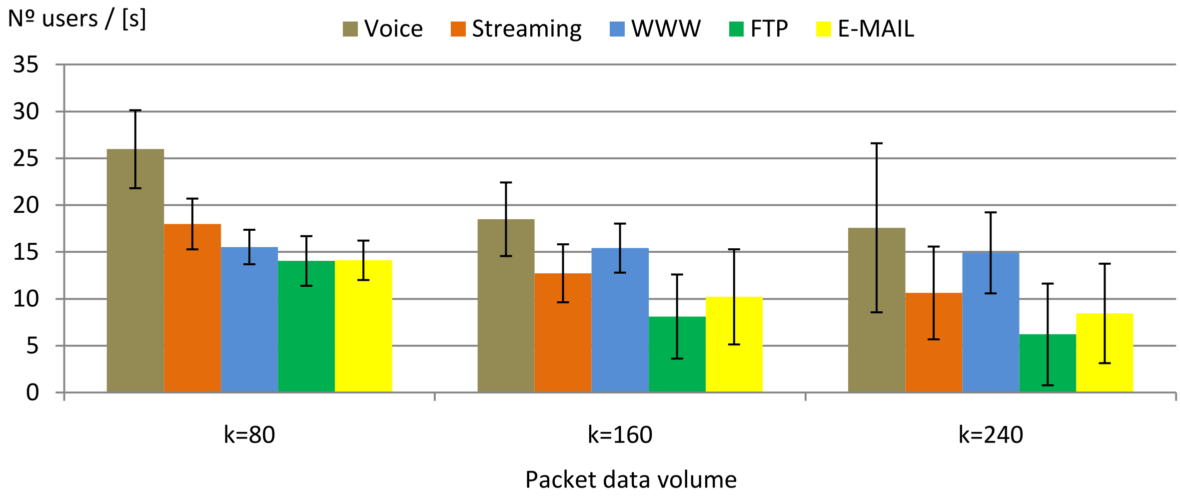

Figure 4-27 - Active users per service for the DM scenario.

In the last scenario, the optimal value of k for average delay is again 80, since the other 2 values used are in the limit of 200 ms per user, or even worst as with k =240, where the delay per user takes values far from the optimal limit, Figure 4-28.

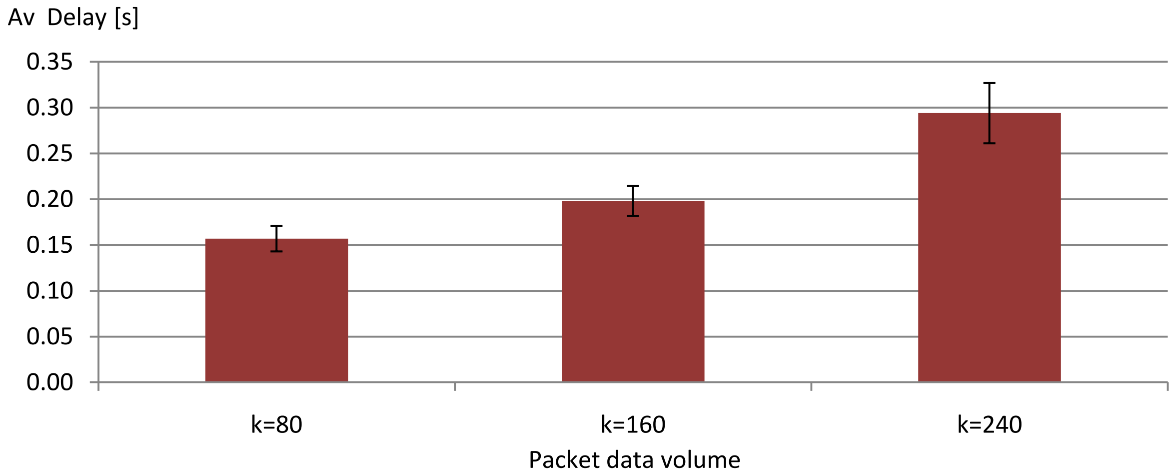

Figure 4-28 - Average delay for the DM scenario.

A comparative study is performed for the main KPIs indicators, for the following input parameters: number of allocated codes, bit rate, number of active users, all of them on a service perspective, and at the end the average delay. In Figure 4-29, one shows the variation in the number of application codes allocation, where a soft variation is seen, regardless of the used scenario, with lower codes delivery for the highest demanding service (FTP), for higher user densities scenario (12 000 users [km^{-2}]). In the remaining scenarios, the number of delivered codes is always up to 5, which is the minimum satisfactory value, when dealing with data services. The highest variation happens for E-mail and FTP services, even if in all data services, a stability is seen with values never under 6 until a maximum of 13, that can be considered a satisfactory result.

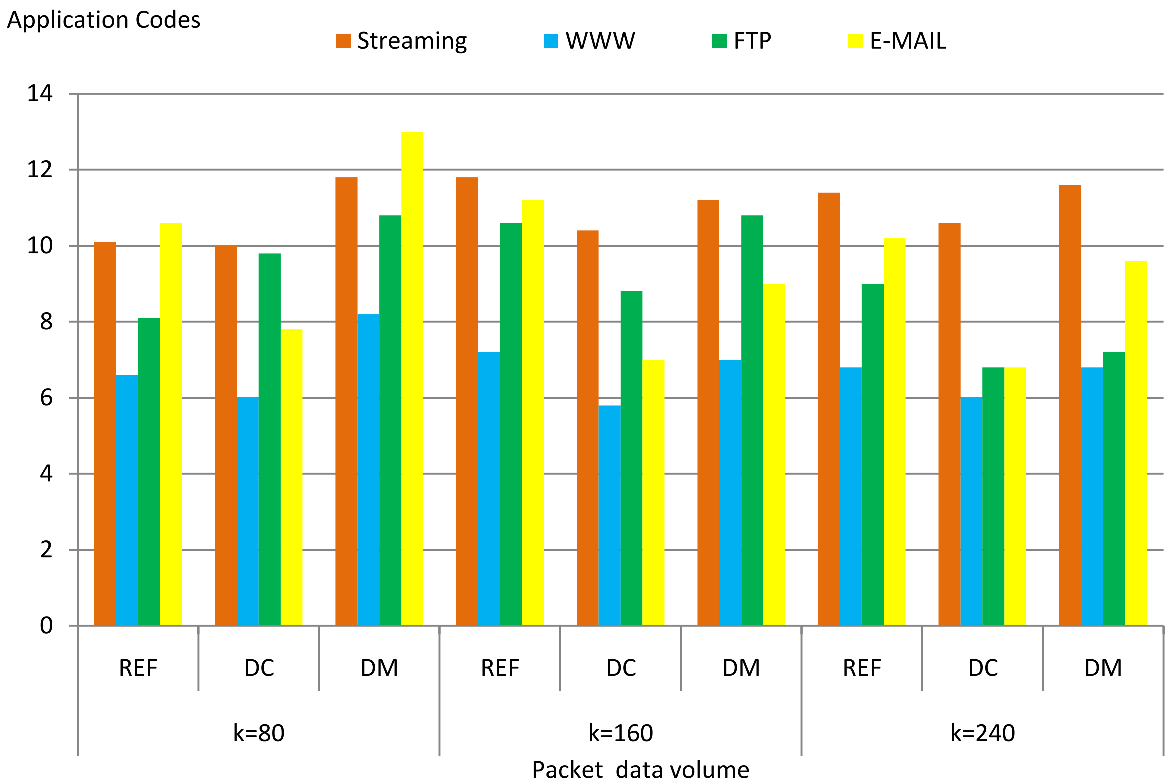

Figure 4-29 - Comparative study for codes allocation relatively to k variation.

In Figure 4-30 the bit rate per service is shown, which keeps the same profile already mentioned for a unique transmitted packet size (k=80): the average values remains the same for all the evaluated services

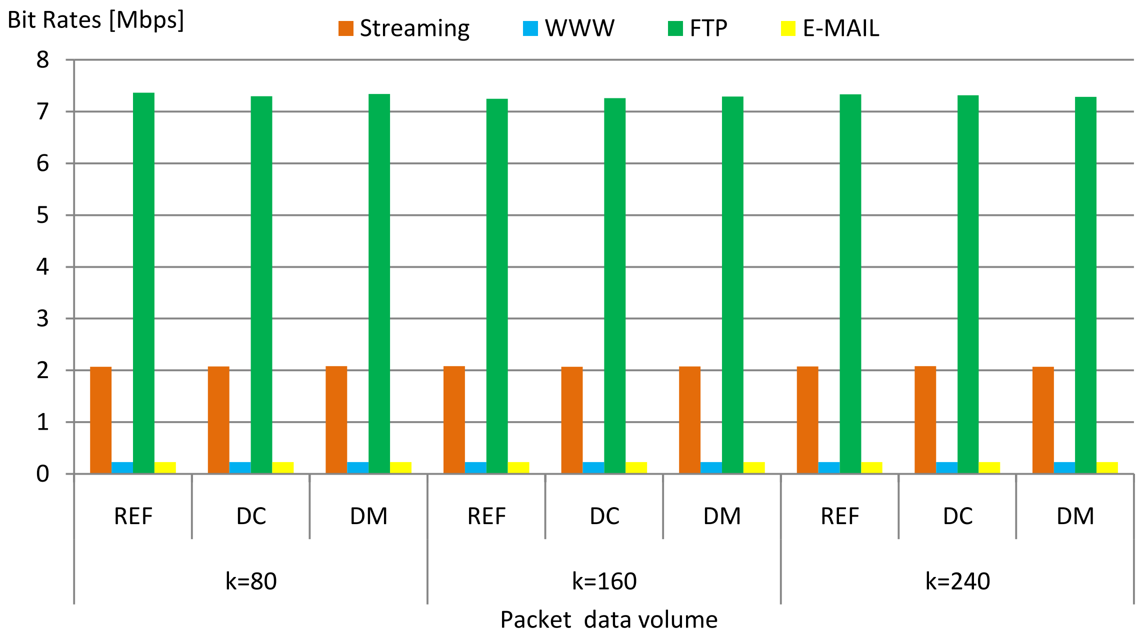

Figure 4-30 - Comparative study for bit rates relatively to k variation.

The number of active users variation is shown in Figure 4-31, where in the REF scenario voice users represent more than 50% of the total population. The number of data users that present a higher variation concerning the number of users simultaneously active are those with lower needs to start a transmission: WWW and E-mail. The remaining ones do not present so notorious oscillations, specially Streaming and FTP that for most high populated scenario (12 000 users [km^{-2}]), have their number of active users reduced to values close to 5 for each service.

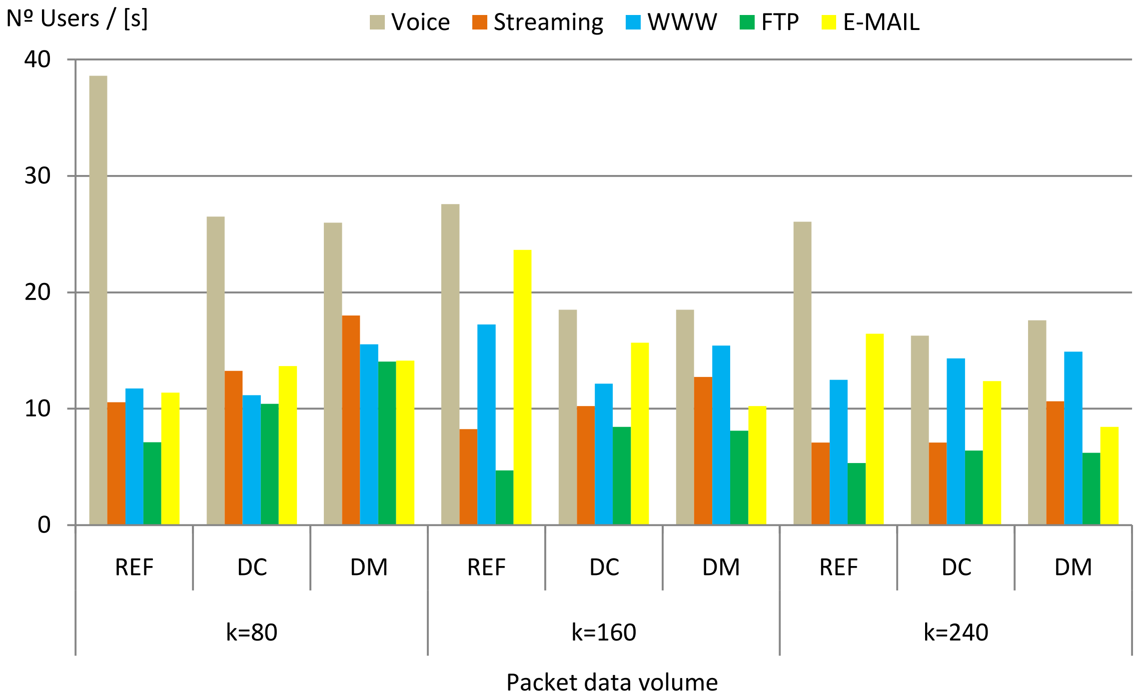

Figure 4-31 - Comparative study for number of active users relatively to k variation.

The last analysis is shown in Figure 4-32 and is relative to average delay variation for each particular scenario.

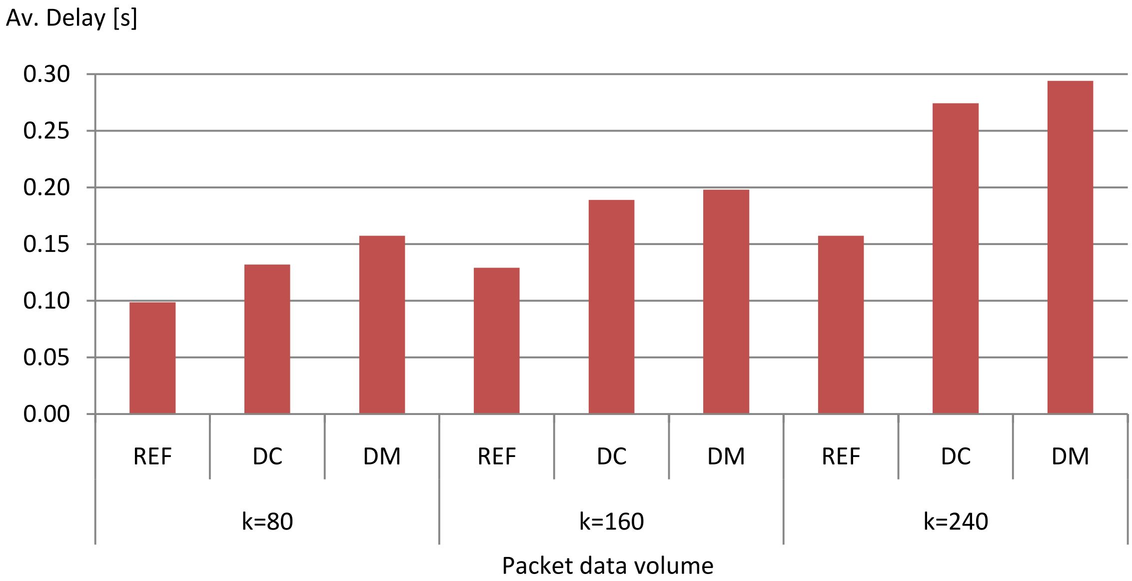

Figure 4-32 - Comparative study for average delay relatively to k variation.

As mentioned before, the REF scenario, regardless of the user density is the configuration scenario with the best average performance; DC and DM scenarios are just close or under 150 ms for user density of 8 000 user [km^{-2}], otherwise they show values over 200 ms, with the worst case for a user density of 12 000 user [km^{-2}].

4.5 Number of Users and Power Ratio

In UMTS, when voice and data users access the network, the power distribution has to be separated among those different user types in a dynamic way, according to the service needs requested by each user profile. Voice users have power control applied, which is why every time a user is becoming more distant from the BS (until a maximum distance), the normal behaviour from the mobile is to increase its power output, trying to stay connected to the BS and at the same time, to keep noise values as low as possible, to avoid excessive interference among the others active users. In HSDPA, there is no power control, but as soon as the mobile is in the periphery of a BS, there is no quality to establish or maintain a data session in good radio conditions. In this situation, a release command is sent, which disconnects the distant user from the BS where it was attached. Every time a new mobile enters in the network, the RRM entity (in HSDPA the Node B has an important function in this process), allocates part of its power to each UE, trying to keep balanced the amount of resources for future incoming users. When the user density is incremented, the probability to have more users accessing the network also raises. In HSDPA, normally the signal and quality levels to execute a service are more demanding than voice, which is why data users are normally not so far away to process a new service; because there is no power control associated to HSDPA, the power levels associated to each UE will not change too much, Figure 4-33.

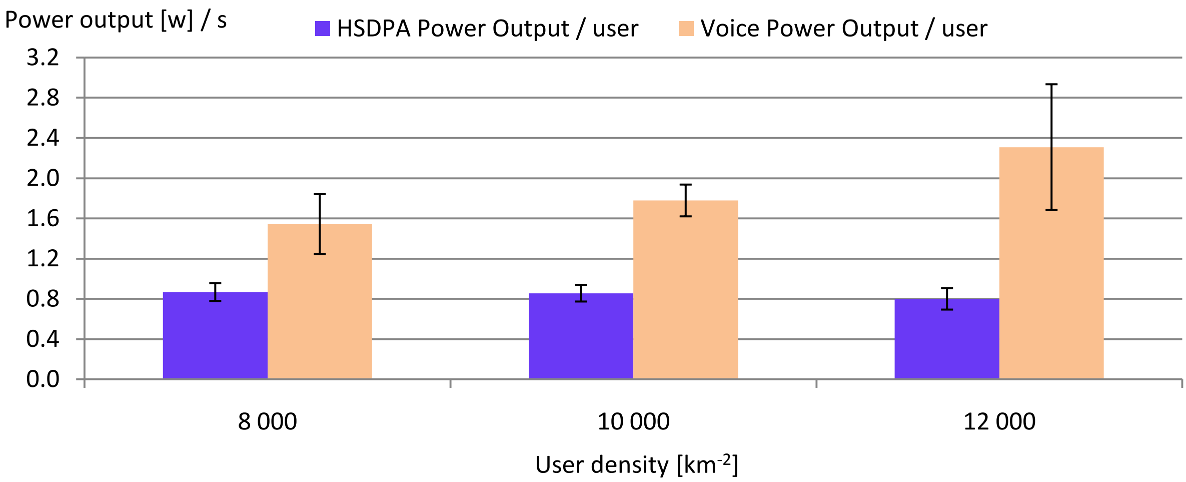

Figure 4-33 - Power distribution between voice and data for the REF scenario.

On the other hand, voice users, can access the network, even in wider ranges; the only compromise that should exist is that even in cases where each voice user has to increase its power output, to be listen by the BS to which it is attached, its transmission power should not contribute to excessive high noise levels that will make the associated communications to that channel impossible to be continued or initiated. With the increase in the number of users, it is expected that many of them are in the cell border areas; such situation implies an increase in the mobile power output, by a factor close to 1.2. Data user services does not experiment any power control, which can be explained by the fact that in areas far from the BS, it is not possible to establish an HSDPA connection in good radio conditions; with such impossibility, the mobile presents similar power output, regardless of the number of the users in the network, Figure 4-34.

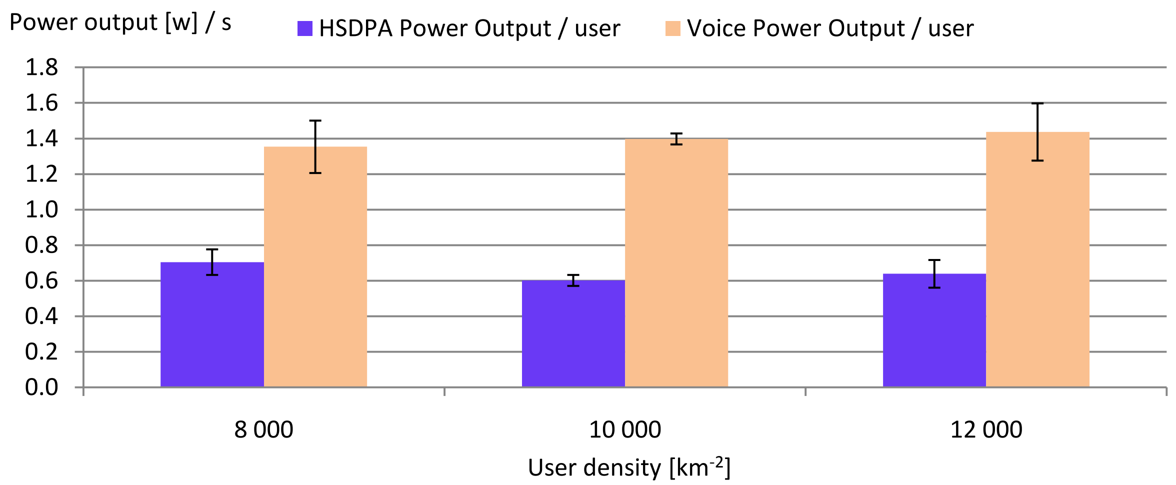

Figure 4-34 - Power distribution between voice and data for the DC scenario.

In comparison with the REF scenario, in the DC one voice users have less penetration rate, which is a reason for a not so aggressive power adjustment: voice users are probabilistically in less number, keeping higher priority to access the network, which is why their location is in this scenario (probably closer to the BS), does not need so high power output. In the DM scenario, the power control adjustment is even less sensitive, since voice users density decreased again (user density penetration for voice is equal to 30%), power control is applied to less users, and when applied, if users are not far away from the BS, the result is similar to DC scenario, where a predominance in data services, results in average lower needs in power output transmission. In Figure 4-35, one shows the power output evolution according to different users density profile and the result is quite similar to the DC scenario, for the same reasons pointed before.

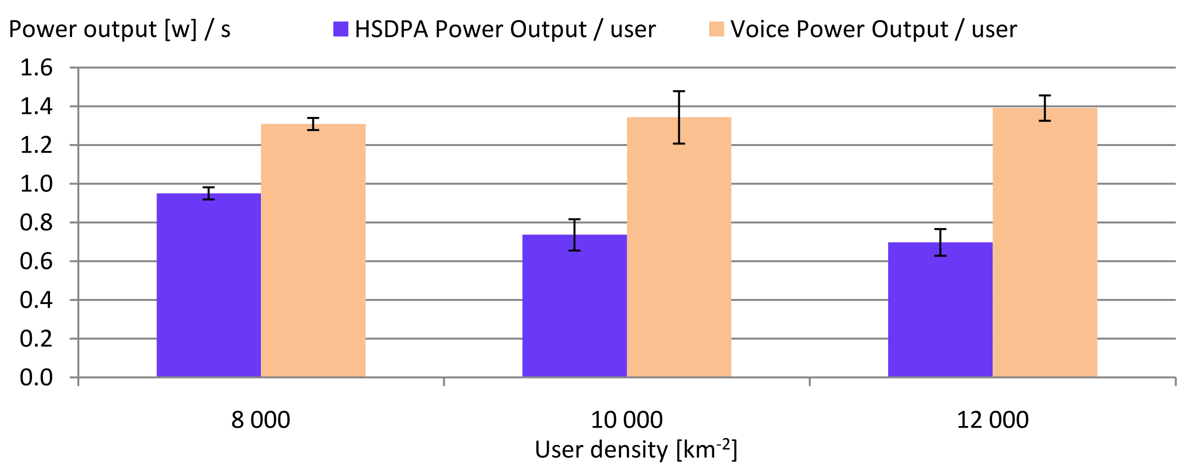

Figure 4-35 - Power distribution between voice and data for the DM scenario.

Particularly in comparison to the REF scenario, the output power for a higher user density (12 000 users [km^{-2}]) demands a tighter power control adjustment to keep the noise level inside acceptable levels, resulting in satisfactory GoS.

4.6 Number of Users and Releases Variation

The usage rate of both releases, R5 and R99, was defined simply by a separation of service types: voice is associated to R99 and all data services are associated to R5, since with 2 UMTS carriers it is possible to make such differentiation. For the REF scenario, the distribution of users for both technologies does not follow exactly the user penetration rate for each service, since voice has a penetration rate of 50%, but the service usage is close to 60%, having the remaining ones allocated to HSDPA, as it can be seen in Figure 4-36.

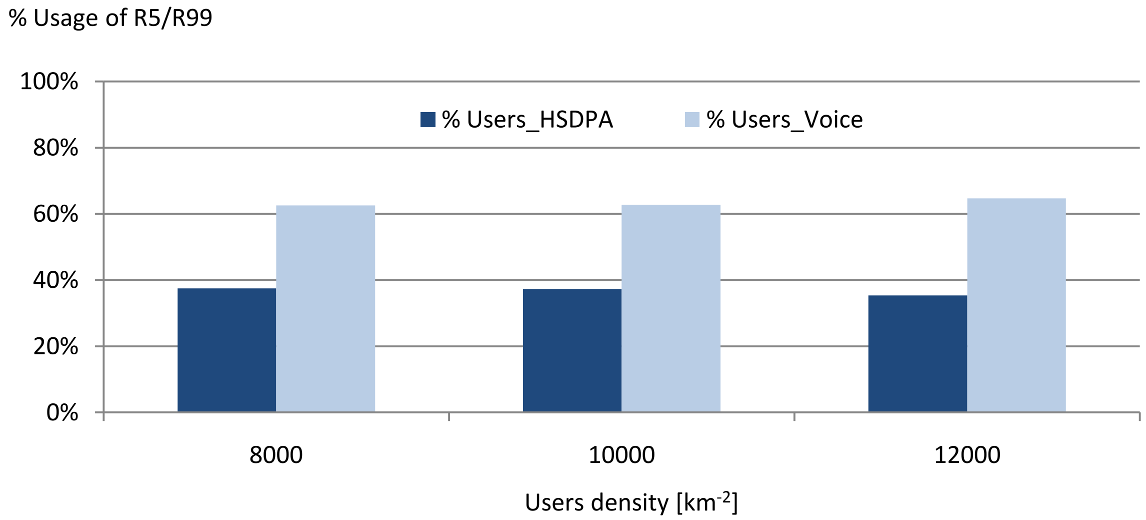

Figure 4-36 - Releases employment per user type in the REF scenario.

This result can be explained by the associated probabilistic model within each user generation. Even with different services penetration for each scenario, it could happen in some situations that small variations occur in the global characterisation of the dominant user profile, always inside reasonable limits. In the DC scenario, the distribution of data users is closer to the expected distribution, since the percentage of data user density in the network is increased in 10%, in comparison to the REF scenario, Figure 4-37.

Finally, the last scenario shows an increase in the data services to its maximum value of 70%, which is again close to the input definition of user density for data services in the DM scenario. In all scenarios, the user activation is made in a random way, which is why such results are representative of a group of simulations that follow the input values placed for voice and data user services.

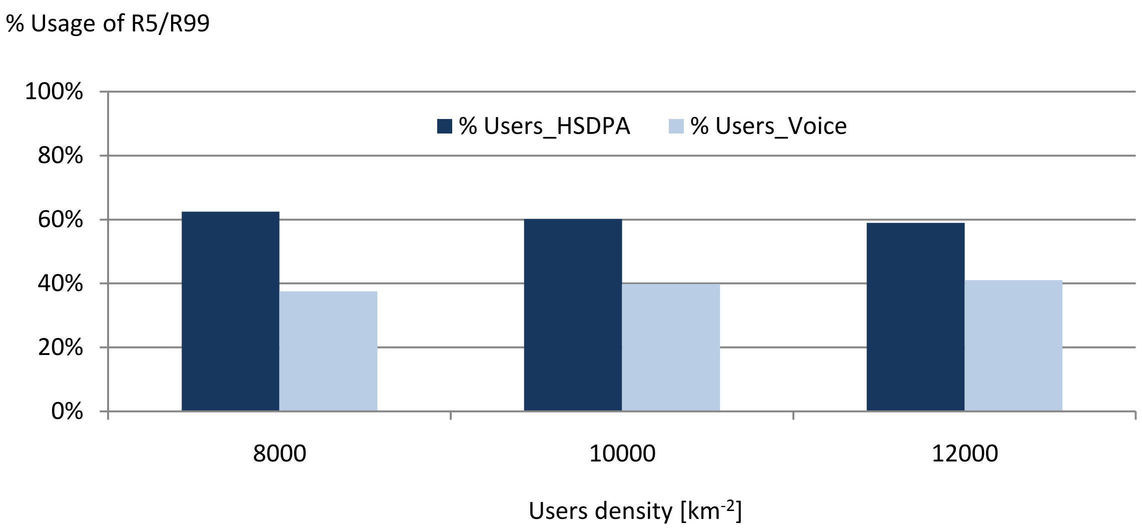

Figure 4-37 - Releases employment per user type in the DC scenario.

At the end, it can be concluded that whatever the input user density values are, the resultant release users distribution is not far from the planned inputs, perhaps with the exception of the REF scenario. In both data dominant scenarios, DC in Figure 4-37 and DM in Figure 4-38, a global tendency to follow the input values concerning the user density profile is shown.

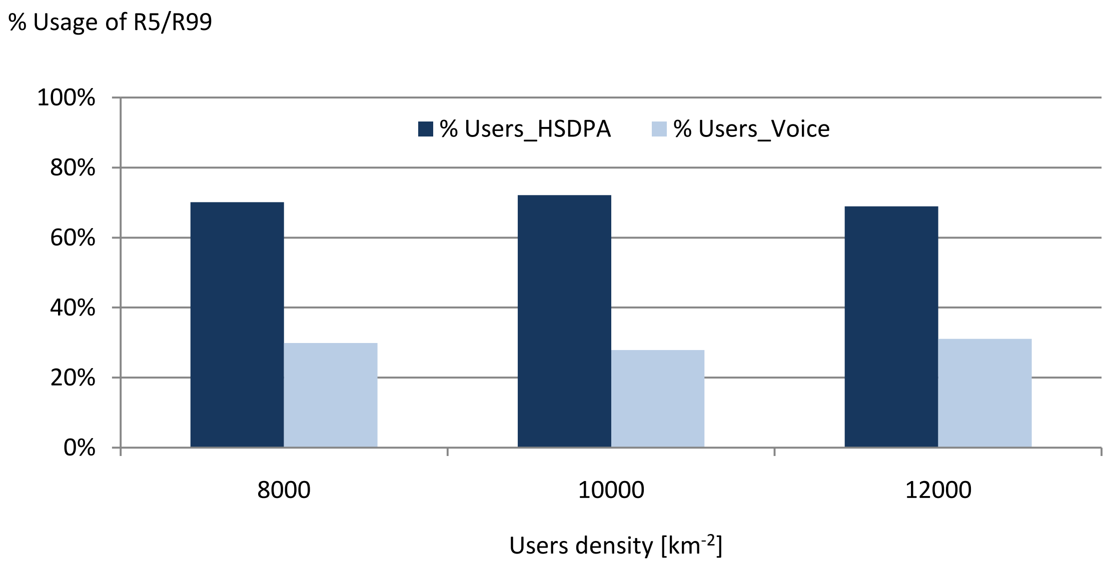

Figure 4-38 - Releases employment per user type in the DM scenario.

The position of each UE relative to each available BS is important for a successfully call or session establishment, since in the voice service, power control is applied and in data services the most distant users, if there is a lack of resources are delayed, result of cell breathing effect, which establishes a trade-off between coverage and capacity.

Chapter 5

Conclusions

This chapter finalises this work, major results are highlighted, summarising conclusions and pointing out aspects to be developed in future work.

In order to constantly enhance users experience and satisfaction, networks growth must guarantee higher bit rates keeping all QoS metric requirements as high as possible. From an operator point of view, Mobile Communications is one of the most dynamic and profitable market sectors in current economics, but is also one of the most demanding ones from the point of view of the required investment. The economic exploitation of the solutions directed towards the optimisation of heterogeneous networks performance is therefore a key issue. It can be simply understood that the final user becomes transparent to all optimisation efforts; he/she just requires to be served with the most radio access at each instant.

In Chapter 2, an introduction to UMTS, and HSDPA is done. Some basic aspects are explained for both releases, which are supported, by a common architecture. HSDPA new features, services and applications, capacity and coverage issues and radio resource management are described. At the end of the chapter, the state of the art is presented.
In Chapter 3, HSDPA RRM algorithms are described. The most important part of the algorithm (CAC) is detailed and explained: service separation between circuit and packet services, within priorities, schedule procedure and system limitations and reactions to overloaded situations (different responses for voice and data services). This chapter also addresses the characterisation of the functional blocks and input/output variables of the simulator; a separation between global and specific parameters is presented. The way that real systems features are adapted to a simulation environment is explained, together with the simplifications and approximations that are necessary in a work of this type.

The present work is dedicated to RRM procedures applied to a network processing voice and data services, and can be divided in two main stages. The first one was dedicated to build a software simulation tool. A system level, time based simulator has been developed over the Microsoft Visual Studio 2005 platform. The simulator is able to reproduce part of a complex RRM system, consisting of a module composed by UMTS technologies (R99 and HSDPA), implementing system functionalities like power control, load and access control and interference estimation. In the second phase, the simulator was used to produce simulations with different scenarios and extract results to evaluate the performance of a RRM algorithm based on a service/RAN priority table. A liquid total of 40 days of CPU processing time has been spent to produce the results. The different results are based on a REF Scenario, and in most other cases they differ from it in a single parameter, which can be user or system dependent. The performance of the network is defined according to a Medium Bit Rate Network performance function ∂, defined as a column matrix with entries corresponding to the output parameters of the simulator: number of active users, average delay, load and number of codes processed by service, bit rates, global and per service, rate of HSDPA usage and power output distribution between R99 and R5.

The REF Scenario is composed of 3 omni BSs, all of them with similar capabilities for both UMTS service types (voice and data). The background is an urban environment, physically located between Campo Pequeno and Saldanha, in Lisbon. Users are assumed to be static and uniformly distributed in a 1000 x 1000 m^2 area. A mono-service policy is adopted, in which only one type of service is allowed per user at each time. Five services are defined: Voice call, Streaming, Web browsing, FTP and E-mail. After the definition of the REF scenario, one defined all the remaining scenarios, DC and DM, deriving from the initial one. The first set of simulations had the objective of analysing the effect of using QoS mechanisms in a wide range of scenarios, varying in each time, the number of user density per km^2, service penetration rate in each scenario, always with 2 UMTS carriers (both shared or one shared and the second one exclusively for data services) and different packet volume transmission according to Pareto distribution. The REF scenario has 50% of real-time users and 50% of non-real time users density. Based on this scenario, several ones were defined by varying application profiles, while still maintaining the relative weight of each application, within each group. Three scenarios were built up ranging from RT one - 50% of both service types, CS and PS - REF scenario, to a maximum of NRT equal to 70% of users probability density – DM scenario). Between them, a scenario called DC, simulated with the purpose of creating an intermediate step between an equilibrate configuration, similar to nowadays mobile networks and an expected tendency that is the increase in data services penetration rate, by which present and future networks will be dimensioned. The number of users was also increased several times, in each scenario, to see the impact on network capacity. A maximum load was stipulated to be 0.7: whenever the amount of power distributed for all active users exceeds this threshold, PS users are delayed and voice ones are blocked.

All these definitions led to the definitions of 15 different scenarios, which amount to a total of 75 simulations (5 simulations per scenario, excluding tests and assessments). Some of the scenarios are computationally demanding, that is why the time reached more than 6 consecutive days, using a Intel (R) Core™ 2 Duo CPU P8600@ 2.40 GHz, 2 GB of RAM.

The simulator developed to serve as the basis for this thesis is a complex software system. Its main target is to reproduce as much as possible a real network environment in all of its particular aspects. Despite of the efforts to achieve that supreme objective, a software system of this type is only feasible while applying approximations and models that help on recreating reality without introducing significant distortion. The current simulator is not an exception and makes use of several simplifications, some of which are hereby presented and discussed.

One of the major simplifications of the simulator is the fact that it accounts for the downlink channel only. The thesis is based in HSDPA, but since nowadays there is a global tendency towards symmetry, it could be considered to have HSUPA; this is a particular point that could be improved in future.

The simulations performed for this thesis consider static users. Despite of the user interface of the simulator offering that possibility, introducing mobility in simulations represents a huge additional processing load that can raise processing times significantly. Optimisation work, perhaps based on more precise mathematical values is required in order to minimise that problem and produce more accurate results. Regarding the most important results, for the REF scenario, there is a satisfactory performance, as with an increase in the user density from 8 000 to 12 000 users [km^{-2}], the average delay per user is always in the margin of 100 ms; the average bit rate never has significant variations among different scenarios and is just a function of the service type (if it is high or low demanding and if it is a RT or NRT service): for highly demand services like Streaming and FTP, the bit rates reach several Mbps (maximum of 14.4 Mbps for FTP service). For E-Mail and web browsing, the bit rates never exceed 0.384 Mbps.

One of the conclusions is the small variation in the bit rate values, regardless of the chosen scenario, which means that the network, whatever the weight of voice and data services, always provides satisfactory results to the final user, considering the final bit rates to be generated. By increasing the user density probability in each scenario, the number of users for each service was also increased, but such growth has a main impact on the average delay per user, which varies from 99 ms for REF scenario (10 000 users) to close to 120 ms for a density of 12 000 users [km^{-2}]. In scenarios where data services are dominant and the second carrier is exclusively used for data services, an increase in the number of data users per service per second is noted. For the REF scenario, there is an increase of 20 data users/second, when the user density varies from 8 000 to 12 000 users [km^{-2}]; this variation for the DC scenario results on average in an addition of 30 users/second and finally for the DM scenario this value is on average equal to 6 users/second. This variation can be explained by the fact that, in case of data users, because the minimum quality for data services required to access the network is higher, the system allows the nearest users to access the network, to start transmitting; the remaining ones in external locations are delayed or simply considered out of coverage.

The second group of simulations had the objective of studying and analysing the impact of system parameters variations on network performance. This had the aim of improving network performance by tuning up the values of selected parameters. A relevant test was the variation of packet volume variation for each scenario, keeping the user density equal to 10 000 users [km^{-2}]. While duplicating packet volume (k), from 80 to 160, different results are obtained for each scenario, with effective impact on data users service, since voice users do not suffer any degradation related to higher probability blocking rates. For data services, it is clear that higher packet volumes to transmit, means that more users are simultaneously active in the network, since the bit rates per service are similar to previous scenarios. This situation results in an increase of the delay per user, particularly for the DM scenario: with a value of k =240, the worst results were obtained, regardless of the scenario. The delay for k =240, has a minimum value per user of 157 ms in the REF scenario, until a maximum value of 294 ms in the DM one. The best trade off happens for k =80 in the REF scenario where the delay per user, per second, does not exceed 100 ms.

The main conclusion from all simulations shows that the default system parameters guarantee good performance levels and balance for all studied profiles. On the other hand, all analysed scenarios allow one to observe that one might manage to improve network performance by tuning up system parameters, increasing performance for a given application or given under application profiles, but this being only possible in very specific situations. Regarding future work, could consider:

- Study the co-existence of different node B types in the final result.
- Scenarios study with users in mobility.
- Possibility of associating multi-service to each user.
- Traffic study per carrier, evaluation of each carrier usage particularly concerning data services.

- Rate of each modulation employed in each carrier (8 QPSK and 16 QAM).
- Create an intermediate stage to evaluate the bit rate and other parameters schedule and its difference with the effective final values attributed by the network.
- Impact on KPIs for having a number of carriers higher than 2.
- Consider other radio parameters to be included in the network analysis (like E_b/N_0 and CQI relations for example).

The subject of the present work, like Science in general, certainly does not end in a thesis. This Master Thesis represents the humble contribution of this author to go further towards the full knowledge on optimised management of heterogeneous networks of ours and future times.

Annex A

Link Budget

In this section the link budget is analysed.

The link budget used throughout this thesis is based on the R99 one, described in detail in [CoLa06] and [Sant05], adapted to HSDPA.

The path loss can be calculated by [Corr08]:

$$L_{P_{[dB]}} = P_{t_{[\mathrm{dBm}]}} + G_{t_{[\mathrm{dB}\,i]}} - P_{r_{[\mathrm{dBm}]}} + G_{r_{[\mathrm{dBi}]}} = EIRP_{[\mathrm{dBm}]} - P_{r_{[\mathrm{dBm}]}} + G_{r_{[\mathrm{dBi}]}} \quad \text{(A.1)}$$

where:

- P_t: transmitting power at antenna port.
- G_t: Antenna gain transmission.
- P_r: available receiving power at antenna port.
- G_r: receiving antenna gain.

If diversity is used (only diversity in UL is considered, since there is no space in the UE for spatial diversity, and polarisation diversity requires doubling the transmit equipment at the Node B [Sant05]), G_r in (A.1) is replaced by

$$G_{rdiv_{[\mathrm{dB}]}} = G_{r_{[\mathrm{dBi}]}} + G_{div_{[\mathrm{dBi}]}} \quad \text{(A.2)}$$

where:

- G_{div}: diversity gain.

The Equivalent Isotropic Radiated Power ($EIRP$) can be estimated for DL by (A.3):

$$EIRP_{[\mathrm{dBm}]} = P^{BS}_{TX_{[\mathrm{dBm}]}} - L_{c_{[\mathrm{dB}]}} + G_{t_{[\mathrm{dBi}]}} - P_{T_x sig\ ch_{[\mathrm{dBm}]}} \quad \text{(A.3)}$$

where:

- P^{BS}_{TX}: total Node B transmission power.
- L_c: cable losses between transmitter and antenna.
- $P_{T_x sig\ ch}$: signalling transmission power.

The received power can be calculated for DL by (A.4):

$$P_{RX_{[\mathrm{dBm}]}} = P_{r_{[\mathrm{dBm}]}} - L_{c_{[\mathrm{dB}]}} \quad \text{(A.4)}$$

where:

- P_{RX}: received power at receiver input.

The UMTS receiver sensitivity can be approximated by:

$$P_{RX_{min}\,[\mathrm{dBm}]} = P_{noise\ [\mathrm{dBm}]} - G_{P_{[\mathrm{dB}]}} + \rho_{[\mathrm{dB}]} \quad \text{(A.5)}$$

where:

- P_{noise} : Total noise power given by (A.6).
- G_P: processing gain.
- ρ: SINR.

The total noise power is given by:

$$P_{noise\ [\mathrm{dBm}]} = -174 + 10\,log\left(\Delta f_{BS\ [\mathrm{Hz}]}\right) + F_{[\mathrm{dB}]} + M_{I_{\eta_{DL}}\,[\mathrm{dB}]} \quad \text{(A.6)}$$

where:

- Δf_{BS} : signal bandwidth.
- F : receiver noise figure.
- $M_{I_{\eta_{DL}}}$: interference margin in DL.

The interference margin, not considered in the single user model, is calculated based on the total number of users of the Node B coverage area. It is most likely that the Node B with the higher number of users in its coverage area is the Node B with more served users. To calculate the number of served users, the interference margin is necessary to evaluate the bit rate due to the user distance and so, the latter approximation was considered. For the Node B with the higher number of users connected to, one assigns the maximum interference margin value, defined in (2.3) and for the other Node Bs, the interference margin for HSDPA at a given Node B is estimated by:

$$M_{I_{\eta\, DL}} = \frac{N_{uj}}{N_{u_{max}}^{NodeB}} \cdot \ \beta_{[\mathrm{dB}]} \tag{A.7}$$

where:

- β : maximum interference value considered.
- N_{uj} : number of users in the Node B j.
- $N_{u_{max}}^{NodeB}$: number of the users of the most populated Node B.

Some margins must be taken into account, to adjust additional losses due to radio propagation and others:

$$M_{[\mathrm{dB}]} = M_{SF\,[\mathrm{dB}]} + M_{FF\,[\mathrm{dB}]} + L_{int\ [\mathrm{dB}]} \tag{A.8}$$

where:

- M_{SF}: slow fading margin.
- M_{FF}: fast fading margin.
- L_{int} : indoor penetration losses.

The total path loss can then be calculated by:

$$L_{P_{total}\ [\mathrm{dB}]} = L_{P\,[\mathrm{dB}]} - M_{[\mathrm{dB}]} \tag{A.9}$$

The total path loss is used as input in the COST 231 Walfisch-Ikegami propagation model. The HSDPA frequency, Δf_{BS}, values used ([2110, 2170] MHz) exceed the frequency validation values and some of the calculated cell radius are below the distance validation values, namely for high data rates. Nevertheless, the model was used, since it is adjusted to urban non-line of site propagation.

The COST 231 Walfisch-Ikegami propagation model is valid for [DaCo99]:

- Δf_{BS} Є [800, 2000] MHz.
- Cell radius, R ϵ [0.02, 5 km]
- H_{BS} , is the Node B height between 4 and 50 m.
- H_{UE}, UE height between 1 and 3 m

Annex B

Traffic Source Models

In this annex, the traffic source models used throughout the present thesis are described.

For CS-based voice service, a 4-state model is used [VaRF99], which is based on measurement and includes not only the ON-OFF behaviour but also the effect of the voice encoder, compression device and air interface (AI).

The model described in [VaRF99] and [RaMe01] defines four-states: when the source is in state *l* it generates packets of size k each 10 ms, for a burst duration of α_k, k = 1,…,4. In the long time average, the probability that a packet is of size k is quantified as P_k. The burst duration τ_k is modelled as a random variable; with mean value m_k, with the Weibull Probability Density Function (PDF) as follows:

$$f(x) = \beta_k.\lambda_k.(x.\lambda_k)^{\beta_K-1}.e^{-(x.\lambda_k)^{\lambda_k}}. \quad \text{(B.1)}$$

where:

- $1/\lambda_k$ is the scale parameter.
- β_k is the shape parameter.

and both taking the values defined in Table B-1.

Table B-1 - Voice Source Model Parameters (partial extracted from [VaRF99]).

State l	Packet size k [Bytes]	Measured Probability P_k	Measured mean burst duration m_k [packet]	Weibull parameter λ_k	Weibull parameter β_k
1	2	0.5978	29.8	0.03	0.75
2	3	0.0723	2.5	0.45	0.80
3	10	0.0388	1.8	0.80	0.70
4	22	0.2911	38.8	0.05	0.90

After a *l* state, a new state is selected with probability Q_l, which is defined as [RaMe01]:

$$Q_l = \frac{\frac{P_l}{m_l}}{\sum_{j=1}^{4}\frac{P_j}{m_j}} \quad \text{(B.2)}$$

The voice calls generation process follows a Poisson process [Yaco93], and the duration of the calls is determined according to an exponential distribution.

VoIP services, in turn, present typically a symmetric or quasi-symmetric nature and require small end-to-end transmission delays. According to [Agui03], VoIP can be characterised through a traditional ON-OFF behaviour, in which sequences of speech-bursts are intercalated with silent bursts. Thus, a VoIP transmission can be modelled as a Markov model with two states of “silence” and “talk”: when in “silence”, no packets are generated, and when in “talk”, packets are generated at a constant rate. Particularly IP packets carrying the speech information are transmitted. Both activity and silent periods are generated by an exponential distributed random variable with mean values t_{ON} and t_{OFF}, respectively. The payload size of the IP packets carrying speech bursts depends on the considered speech codec and the packet rate. Typical VoIP codecs are G711, G732.1 and G729.A, all of these with their specific frame duration and frame sizes, Table B-2.

Table B-2 - Typical VoIP codecs (extracted from [Nune02]).

Codec	**Frame Duration [ms]**	**Frame Size [bytes]**	**Bit rate [kbps]**
G711	10	80	64.0
G723.1	30	24	6.4
G729.A	20	20	8.0

As VoIP uses UDP (User Data Protocol) and RTP (Real Time Protocol) at the transport layer, the size of a full IPv6 (Internet Protocol version 6) header together with a RTP/UDP header is 60 bytes, and 40 bytes if IPv4 (Internet Protocol version 4) is used instead. As the size of a typical voice packet is 20 bytes if G729.A is used, the RTP/UDP/IP overhead figures illustrate the typical problem of the header overhead in VoIP: in this case, instead of an 8 kbps bit rate, a final bit rate of 32 kbps case IPv6 was in use would be generated (24 kbps if IPv4 is used instead). When operating in a bandwidth limited system such as WiMAX, it is important to use the radio band as effectively as possible, and header overhead up to 60 bytes can seriously degrade the spectral efficiency of a VoIP service over such link. Without header compression, two-thirds of the transmission would be just headers. To handle this purpose, protocols such as “Robust Header Compression” (ROHC) have been developed to tackle this problem [IETF01]. According to [Agui03], one can assume that header bytes can be compressed to 8 bytes. Additionally, and following European Telecommunications Standards Institute (ETSI) recommendations [ETSI98], speech calls should be generated according to a Poisson process, with mean call duration of 120s. The resulting VoIP modelling is summarised in Table B-3.

Table B-3 - Modelling of VoIP Traffic.

Activity Factor [%]	50
Mean Active Phase, t_{ON} [s]	3
Mean Silent Phase, t_{OFF} [s]	3
Payload of IP Packets [bytes]	20
IP Overhead [bytes]	8
TTI [ms]	20
T_{vc} [s]	120

Non-real time applications typically present an asymmetrical nature, as they refer mostly to specific requests for information done by end users to remote machines. The most known applications are Web browsing, FTP and E-mail. Several models are studied and proposed to characterise web browsing but as the present work does not intend to focus specifically on traffic models, modelling of these applications is based on [ETSI98]. Figure B-1 illustrates a typical Internet surfing session, which consists of a sequence of packet calls. During a packet call several packets may be generated, which means that the packet call is composed by a bursty sequence of packets.

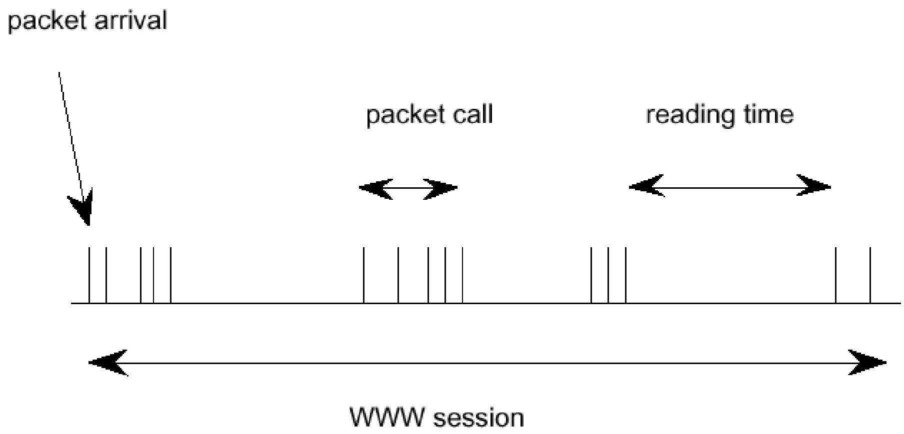

Figure B-1 - Typical WWW session (adapted from [ETSI98]).

A packet service session is modeled as a Poisson process. It typically contains one or several packet calls, depending on the application. In a WWW browsing session for instance, a packet call corresponds to the downloading of a web page, and after the document has entirely arrived to the terminal, the user takes some time for analysing the information, which is often called the reading time. On the other hand, in a FTP session, it is likely that the session contains only one packet call.

In order to fully characterise a packet session, the following parameters must be modeled [ETSI98]:

- The number of packet call requests per session, $\mathrm{N_{pc}}$; this is a geometrically distributed random variable with mean $\overline{\mathrm{N_{pc}}}$.
- The reading time between two consecutive packet call requests in a session, D_{pc}.; this is a geometrically distributed random variable with a mean $\overline{D_{pc}}$. Note that the reading time starts when the last packet of the packet call is completely received by the user. The reading time ends when the user makes a request for the next packet call.
- The number of packets within a packet call, N_d: although different statistical distributions can be used to generate the number of packets, it is assumed that N_d can be a geometrically distributed random variable, with mean $\overline{N_d}$.
- The time interval between two consecutive packets inside a packed call, D_d: This is a geometrically distributed random variable with a mean $\overline{D_d}$. Naturally, if there is only one packet in a packet call, this is not needed.
- The Packet size distribution model is based on Pareto distribution that suits best for the traffic case under study; Pareto distribution with cut-off is used.

Annex C

Propagation Models

The propagation model adopted and used in this thesis is described in this annex. This model is used for cellular and wireless local networks propagation estimation.

For a good estimation of the received average power, one may use the well know Walfisch-Ikegami propagation model adapted by COST 231 for microcell environment [DaCo99]. This model has the following input parameters:

- H_{BS}: BS height.
- H_B: Building height.
- H_{UE}: UE height.
- S_w: Street width.
- f_{BS}: BS frequency.
- *d*: Distance between BS and UE.
- *b*: Building separation.
- Ψ: Street orientation angle.

The following default values are recommended:

- *b*: [20, 50] m.
- S_w: *b*/2.
- H_B: 3 m × [number of floors]+roof.
- Ψ: 90 º.

The path loss when in LoS is calculated by:

$$L_{p_{[\mathrm{dB}]}} = 42.6 + 26\log d_{[\mathrm{km}]} + 20\log f_{BS\,[\mathrm{MHz}]} \tag{C.1}$$

The path loss in the case of NLoS is given by (all path loss values are expressed in dB):

$$L_p = \begin{cases} L_0 + L_{rts} + L_{msd}, & L_{rts} + L_{msd} > 0 \\ L_0, & L_{rts} + L_{msd} < 0 \end{cases} \tag{C.2}$$

where:

$$L_{0_{[\mathrm{dB}]}} = 32.4 + 20\log d_{[\mathrm{km}]} + 20\log f_{BS\,[\mathrm{MHz}]} \tag{C.3}$$

$$L_{rts_{[\mathrm{dB}]}} = -16.9 - 10\log S_{w_{[\mathrm{m}]}} + 10\log f_{BS_{[\mathrm{MHz}]}} + 20\log \Delta h_{Mobile\,[\mathrm{m}]} + L_{ori_{[\mathrm{dB}]}} \tag{C.4}$$

$$\Delta h_{Mobile} = H_{UE_{[\mathrm{m}]}} \text{-} H_{B_{[\mathrm{m}]}} \tag{C.5}$$

$$L_{ori_{[\mathrm{dB}]}} = \begin{cases} -10 + 0.34\,.\,\Psi_{[°]} & ,0° \le \Psi \le 35° \\ 2.5 + 0.075\,.\,(\Psi_{[°]} - 35) & .35° \le \Psi \le 55° \\ 4.0 + 0.114\,.\,(\Psi_{[°]} - 55) & .55° \le \Psi \le 90° \end{cases} \tag{C.6}$$

$$L_{msd_{[dB]}} = L_{bsb_{[\mathrm{dB}]}} + K_{a_{[\mathrm{dB}]}} + K_d\log d_{[\mathrm{km}]} + K_f\log f_{BS_{[\mathrm{MHz}]}} - 9\log b_{[\mathrm{m}]} \tag{C.7}$$

where:

$$L_{bsb_{[\mathrm{dB}]}} = \begin{cases} -18\log(1+\Delta H_{BS}), & H_{BS} > H_B \\ 0 & , \quad H_{BS} \leq H_B \end{cases} \tag{C.8}$$

$$\Delta H_{BS_{[\mathrm{m}]}} = H_{BS_{[\mathrm{m}]}} - H_{B_{[\mathrm{m}]}} \tag{C.9}$$

$$K_{a_{[\mathrm{dB}]}} = \begin{cases} 54 & , \quad H_{BS} > H_B \\ 54 - 0.8 \times \Delta H_{BS} & , \quad d \geq 0.5 \; and \; H_{BS} < H_B \\ 54 - 0.8 \times \Delta H_B \times \left(\frac{d}{0.5}\right) & , \quad d < 0.5 \; and \; H_{BS} \leq H_B \end{cases} \tag{C.10}$$

$$K_d = \begin{cases} 18, & H_{BS} > H_B \\ 18 - 15 \, . \frac{\Delta H_{BS} \, .}{H_B} & , \quad H_{BS} \leq H_B \end{cases} \tag{C.11}$$

$$K_f = \begin{cases} -4 + 0.7\left(\frac{f_{BS}}{925} - 1\right), for\ medium\ size\ cities\ and\ suburban\ centres\ with\ moderate\ tree\ density \\ -4 + 1.5\left(\frac{f_{BS}}{925} - 1\right) \quad for\ metropolitan\ centres \end{cases} \tag{C.12}$$

where:

- L_0: free space attenuation.
- L_{rts}: "roof-to-street diffraction and scatter loss".
- L_{ori}: attenuation caused by main street orientation with respect to the direct radio path
- L_{msd}: "multi-screen diffraction loss".
- L_P is the output parameter of the model in dB.

Some parameters have a validity range, Table C-1:

Table C-1 - Valid parameters range.

Frequency, f_{BS}	[800, 2100] MHz
Distance NLoS, *d*	[0.02, 5 km]
Distance LoS, *d*	[0.02, 0.2] km
BS antenna height, H_{BS}	[4, 50] m
UE antenna height, H_{UE}	[1, 3] m

References

[3GPP01] 3GPP, *Technical Specification Group Services and System Aspects, Service aspects; Services and Service Capabilities (Release 99),* Report TS 22.105, V3.10.0, Oct. 2001 (http://www.3gpp.org/ftp/Specs/html-info/22105.htm).

[3GPP02a] 3GPP,*Technical Specification Group Radio Access Network, UTRAN Overall Description (Release 99)*, Report TS 25.401, V3.10.0 June 2002 (www.3gpp.org/).

[3GPP03a] 3GPP*, Technical Specification Group Services and System Aspects, Service aspects; Services and Service Capabilities (Release 6),* Report TS 22.105, V6.2.0, June 2003 (www.3gpp.org/).

[3GPP03b] 3GPP,*Technical Specification 25.201 v6.2.0. Physical layer* – General Description, Apr. 2003 (www.3gpp.org/).

[3GPP03c] 3GPP,*Technical Specification 25.308 "High Speed DL Packet Access (HSDPA);* Overall Description" Release 5 Mar. 2003 (www.3gpp.org/).

[3GPP03d] 3GPP,*Technical Specification 21.902 v6.0.0. Evolution of 3GPP System*, Sep. 2003 (www.3gpp.org/).

[3GPP03e] 3GPP,*Technical Specification Group Services and System Aspects, Service aspects; Services and Service Capabilities* (Release 6), Report TS 22.105, V6.2.0, June 2003 (www.3gpp.org/ftp/Specs/html-info/22105.htm).

[3GPP05a] 3GPP,*Technical Specification 25.401 v6.7.0.* UTRAN Overall Description, Mar. 2005 (www.3gpp.org/).

[3GPP05b] 3GPP,*Technical Specification Group Radio Access Networks, BS Radio* transmission and *Reception (FDD) (Release 99)*, Report TS 25.104, V 13.3.0, Sep. 2005 (www.3gpp.org/ftp/Specs/html-info/25104.htm).

[3GPP05c] 3GPP, *Services and service capabilities*, Report TS 22.105, V6.4.0¸ Sep. 2005 (www.3gpp.org).

[3GPP05d] 3GPP, *Quality of Service (QoS) concept and architecture*, Report TS 23.107, V6.3.0, June 2005 (www.3gpp.org).

[3GPP06] 3GPP,*Quality of Service (QoS) concept and architecture*, Report TS 23.107, v6.4.0, Mar. 2006 (www.3gpp.org).

[3GPP07] 3GPP, (www.3gpp.org), 2007.

[Agil07] White Paper, *Agilent Technologies Concepts of High Speed DL Packet Access: Bringing Increased Throughput and Efficiency to W-CDMA*, Application Note, USA, Jan. 2007.

[Agui03] Aguiar,J., *Traffic Analysis at the Radio Interface in Converging Mobile and Wireless Communication Systems*, M.Sc. Thesis, IST-TUL, Lisbon, Portugal, Dec. 2003.

[AKRS01] Andrews, Kumaran,M.,Ramanan,K.,Stolyar,A., "Providing QoS over a Shared Wireless Link", *IEEE Communications Magazine*, Vol. 39, Feb. 2001, pp. 150-154.

[AROM07] http://www.aroma-ist.upc.edu/, Nov. 2007.

[ChAb07] Chimeh,J.,Abdoli,M., New Radio Resource Management Algorithm for Mixed Traffic Transmission in Mobile System - *IEEE 18th International Symposium, Athens*, Greece Sep. 2007.

[CoLa06] Costa,P.,Ladeira,D., "*Planning of UMTS Cellular Networks for Data Services Based on HSDPA (in Portuguese)"*, Graduation Project, IST-TUL, Lisbon, Portugal, June 2006.

[Corr08] Correia,L.M., "*Mobile Communication Systems – Course Notes*", IST-TUL, Lisbon, Portugal, Mar. 2008.

[DaCo99] Damasso,E.,Correia,L.M., "*Digital Mobile Radio Towards Future Generation",* COST231 Final Report, 1999 (http:/www.lx.it.pt/cost231/).

[DJMW06] Derksen,J.,Jansen,R.,Maijala,M.,Westerberg,E.,"HSDPA Performance and Evolution*"*, *Ericsson Review* nº3 2006, pp. 117-120.

[DPSB07] Dahlman,E., Parkvall,S., Sköld,J. and Beming,P., "*3G Evolution: HSPA and LTE for Mobile Broadband,* Academic Press", Jordan Hill, Oxford, UK, 2007.

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

[ETSI98] ETSI, *Universal Mobile Telecommunications System (UMTS); selection procedures for the choice of radio transmission technologies of the UMTS*, ETSI Technical Report No*.* TR 101 112 Ver. 3.2.0, France, Apr. 1998.

[FLBW07] Folke,M.,Landstrom,S.,Bodin,U. and Wanstedt,S., "Scheduling Support for Mixed VoIP and Web Traffic over HSDPA", *65th IEEE Vehicular Technology Conference*, Sweden, Apr. 2007, pp. 814-818.

[GoPa05] Gougousis,A. and Paterakis, M., "*Scheduling with QoS support for multi rate wireless systems with variable channel conditions*", Technical University of Crete, Greece, 2005 (http://www.telecom.tuc.gr/~pateraki/Papers/Extended%20Paper.pdf).

[Guti03] Gutiérrez,P.,*Packet Scheduling and Quality of Service in HSDPA*,M. Sc. Thesis, Institute of Electronic Systems , Aalborg, Denmark, Oct. 2003.

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

[HoTo06] Holma,H. and Toskala,A., *HSDPA/HSUPA for UMTS*, John Wiley & Sons, Chichester, UK, 2006.

[HoTo07] Holma,H. and Toskala,A., *WCDMA for UMTS – HSPA Evolution and LTE* (4rd Edition), John Wiley & Sons, Chichester, UK, 2007.

[HuNH07] Hussain S. Nasser N and Hassanein H. "*Efficient Delay-Based Scheduling Scheme for Supporting Real-Time Traffic in HSDPA Networks*" , *32th IEEE Vehicular Technology Conference*, Dublin, UK, Oct. 2007, pp.685-691.

[HuSA07] Huibin,l.,Shuping,C. and Asimakis,K., "Scheduling and Performance of Multi carrier TD-SCDMA HSDPA", in *Communications and Networking in China, 2th International Conference* on Shangai, Aug. 2007, pp. 1042-1046.

[IETF01] IETF, *RObust Header Compression (ROHC): Framework and four profiles*, IETF RFC 3095, Center Oak Plaza, VA, USA, July 2001 (www.ietf.org/rfc.html).

[LCCS06] Ladeira,D.,Costa,P.,Correia,L.M. and Santo,L., "*Planning of UMTS Cellular Networks for Data Services Based on HSDPA*", IST/IT - TUL, Portugal, 2006.

[MaHa07] Al-Manthari,B. and Hassanein,H., "*Packet Scheduling in 3.5G High-Speed Downlink Packet Access Networks: Breadth and Depth*", University of Guelph, Canada, Jan./Feb. 2007, pp. 41-46.

[Nune02] Nunes,M.S., *Integrated Services Networks – 4th section – Voice over IP* (*in Portuguese*), AEIST, Lisbon, Portugal 2002.

[PeMi06] Pedersen,K. and Michaelsen,P., "*Algorithms and Performance Results for Dynamic HSDPA Resource Allocation*", *64th IEEE Vehicular Technology Conference*, Montreal, Canada, Sep. 2006, pp. 1-5.

[QUAL04] White Paper, QUALCOMM CDMA technologies*, "HSDPA for improved downlink data transfer"*, (http://www.qualcomm.com/), Oct. 2004.

[QUAL08] (http://www.qualcomm.com/products_services/mobile_content_services/index.html), Oct. 2008.

[RaMe01] Rábanos,J. and Mesquida, C., "*Mobile Communications of Third Generation* (in Spanish)", Telefónica Móviles España. Madrid, Spain, 2001.

[RSAG05] Romero,J.,Sallent,O.,Agusti,R. and Guerra,M., "*Radio Resource Management Strategies in UMTS*", John Wiley & Sons, Chichester, UK, 2005.

[Sant05] Santo,L.,"*UMTS Performance in Multi-Service Non-Uniform Traffic Networks",* M.Sc. Thesis, IST-TUL, Lisbon, Portugal, Mar 2005.

[Serr02] Serrador,A., "*Optimisation of cell radius in UMTS-FDD Networks"*, M.Sc. Thesis, IST-TUL, Lisbon, Portugal, Dec. 2002.

[Shar06] Sharma,D., "*Introduction to High Speed DL Packet Access (HSDPA)"*, May 2006. (http://www.3g4g.co.uk/Tutorial/DS/ds_hsdpa.html).

[SoBa04] Solana,A., Bardaji,A., "Scheduling and Call Admission Control Schemes in Soft Handoff for Packet Switched Transmission in WCDMA Networks", *in Proc. Of VTC 04 Fall – 60th IEEE Vehicular Technology Conference*, Los Angeles, CA, USA, Sep. 2004, pp. 3486-3490.

[VaRF99] Valkó,A.G.,Rácz,A. and Fodor,G., "Voice QoS in Third-Generation Mobile Systems", *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 1, Jan. 1999, pp. 109-123.

[VeCo02] Velez,F.,Correia,L.M., "Mobile Broadband Services: Classification, Characterisation, and Deployment Scenarios*"*, *IEEE Communications Magazine*, Vol. 40, No. 4, Apr. 2002, pp. 142-150.

[Wann06] Wännström,A.,"High Speed DL Packet Access (HSDPA)*"*, *Ericsson TEMS ^{TM} news*, No 1 Jan. 2006 (www.ericsson.com/tems).

[WPKM05] B. Wang, K. I. Pedersen, T. E. Kolding, and P. E. Mogensen, "Performance of VoIP over HSDPA," *in IEEE VTC-Spring*, Vol. 4, May 2005, pp.2335-2339.

[Yaco93] Yacoub,M.D.,"*Foundations of Mobile Radio Engineering"*, CRC Press, Florida, Boca Raton, Florida, USA, 1993.