TECHNICAL UNIVERSITY OF LISBON
INSTITUTO SUPERIOR TÉCNICO

WiMAX 802.16e Network Performance Analysis for Different Radio Resource Management Algorithms

David Sacadura Cardoso Antunes
(*Licenciado*)

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

Supervisor: Doctor Luís Manuel de Jesus Sousa Correia

Jury

President: Doctor Luís Manuel de Jesus Sousa Correia
Members: Doctor Francisco António Bucho Cercas
Doctor António José Castelo Branco Rodrigues

April 2008

To my family

Acknowledgements

First of all, I need to thank Prof. Luís M. Correia for supervising this work. Prof. Correia has provided me with all the means for the development of the present thesis: discipline, permanent availability, weekly follow-up meetings and critical guidelines. His supervision was certainly a key factor for the completion of my Masters Thesis.

I also need to thank Martijn Kuipers, who co-supervisioned the present work. Our discussions during our follow-up meetings were very useful concerning the consideration of additional aspects which, without Martijn's perspective, would have been left out of discussion.

My appreciation also to the GROW members, especially António Serrador, for helping me by providing useful data from some of his previous and present work.

To my good friends from the University world: Mangas, Serras, Barriguinha, Pinho, Eduardo.

A special thank you to my good old friends'puritanos': without any special order, 'Troca', Bruno, Casquiço, Hugo, Miguel, 'Nico', César, 'Safio', and respective wifes and girlfriends. I hope that we stay together as we have been all these years, and that we can see our children to be the future Masters of this country.

To my parents and my sister, for their unconditional love, and for providing me all the means that I needed throughout all my scholar life, including the present journey.

To my little niece, Matilde, for bringing us all extra joy and happiness.

To my beloved wife Ana: Thank you for your company, love, friendship, patience, encouragement and support. And sorry for the many long nights spent in front of the computer.

To my little daughter, Leonor: You were born during the middle of my thesis and many nights were 'wasted' since then. But the joy of looking at you growing and smiling everyday certainly gave me the extra strength and motivation to proceed with my work.

To my grandfather, who left us recently. Thank you for your teachings, advices, wisdom, education and love. I will never forget you…

Abstract

This work analyses the overall performance of 802.16e radio interface. A simulator was implemented, including the development of different scheduling approaches: Round-Robin, Maximum Throughput and Proportional Fair. Six different applications are considered: VoIP, Video Call, Streaming, WWW, Email and FTP.

When QoS is to be guaranteed for Real Time services (Proportional Fair approach), Best Effort throughput is degraded but accessiblity for the former is maximised. The opposite trend is seen when Real Time services are not prioritised over Best Effort ones, which is case of Round-Robin or Maximum Throughput approaches, with the last one maximising system capacity by prioritising users with the best instantaneous radio conditions.

The Round-Robin approach is the one that is more subjected to the influence of varying traffic mix: if, from a reference traffic mix, VoIP weight increases 15%, throughput can increase up to 6% and VoIP blocking can be improved by up to 50%. The opposite trend is seen when VoIP weight decreases.

Within the Proportional Fair approach, Admission and Congestion Control provide additional QoS mechanisms, allowing for blocking/droppping to occur under acceptable levels, leveling accessibiliy/retainability of Real Time services and throughput/delay of Best Effort services.

Keywords

WiMAX, RRM, Scheduling, Delay, Throughput, Network Load.

Resumo

O presente trabalho analisa o desempenho da interface rádio baseada em 802.16e. Foi desenvolvido um simulador, incluindo algoritmos de agendamento: *Round-Robin*, *Maximum Throughput* e *Proportional Fair*. Seis diferentes aplicações foram consideradas: VoIP Vídeo-Chamada, *Streaming*, *WWW, Email e FTP*.

Quando é necessário garantir Qualidade de Serviço para serviços em tempo real (aproximação *Proportional Fair*), o ritmo binário para *Best Effort* é degradado mas a acessibilidade para serviços em tempo real é maximizada. Observa-se a tendência oposta quando não há prioritização de serviços em tempo real, caso das aproximações *Round-Robin* ou *Maximum Throughput*, com a última a maximizar a capacidade do sistema através da prioritização dos utilizadores que sofrem as melhores condições rádio.

A aproximação *Round-Robin* é mais sujeita à influência de diferentes perfis de tráfego: se a partir de um cenário de referência o peso de VoIP aumentar 15%, o débito binário pode aumentar até 6%, podendo reduzir-se o bloqueio de VoIP em 50%. Verifica-se a tendência contrária quando o peso de VoIP se reduz.

Dentro da abordagem *Proportional Fair*, os algoritmos de controlo de admissão/congestão adicionam critérios adicionais de Qualidade de Serviço, assegurando níveis aceitáveis de acessibilidade/fiabilidade, nivelando o desempenho dos serviços em tempo real face a *Best Effort*.

Palavras-chave

WiMAX, Gestão de Recursos Rádio, Agendamento, Atraso, Ritmo Binário, Carga de Rede.

Table of Contents

List of Figures

List of Tables

List of Acronyms

16-QAM	16-Quadrature Amplitude Modulation
3G	3^{rd} Generation
3GPP	3rd Generation Partnership Product
3GPP2	3rd Generation Partnership Product 2
64-QAM	64-Quadrature Amplitude Modulation
AAA	Authentication, Authorisation and Accounting
AAS	Adaptive Antenna Systems
AMC	Adaptive Modulation and Coding
AROMA	Advanced Resource management solutions for future all IP heterOgeneous Mobile rAdio environments
ARQ	Auto Repeat Request
AS	Active Set
ASN	Access Service Network
ATM	Asynchronous Transfer Mode
AWGN	Additive White Gaussian Noise
BE	Best Effort
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
BTC	Block Turbo Coding
BS	Base Station
CBR	Constant Bit Rate
CC	Convolutional Coding
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CDMA2000	Code Division Multiple Access 2000
CID	Connection Identifiers
CINR	Carrier-to-Interference-plus-Noise Ratio
CP	Cyclic Prefix
CPS	Common Part Sub-Layer

CS	Convergence Sub-Layer
CSN	Connectivity Service Network
CTC	Convolutional Turbo Coding
DAC	Data Centric Scenario
DIUC	Downlink Interval Usage Code
DL	Downlink
DSCP	Differentiated Services Codepoint
DSL	Digital Subscriber Line
EIRP	Equivalent Isotropic Radiated Power
ETSI	European Telecommunications Standards Institute
EVDO	Evolution-Data Optimised
FBSS	Fast Base Station Switching
FCH	Frame Control Length
FDD	Frequency Division Duplex
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FTP	File Transfer Protocol
FUSC	Fully Used Subchannelisation
FWA	Fixed Wireless Access
GBAR	Gamma Beta Auto-Regressive
GOP	Group of Pictures
GSM	Global System for Mobile communications
HARQ	Hybrid Automatic Repeat Request
HD	Half-Duplex
HHO	Hard Handover
H-NSP	Home Network Service Provider
HSDPA	High-Speed Downlink Packet Access
HSxPA	High-Speed Downlink/Uplink Packet Access
HSUPA	High-Speed Uplink Packet Access
HTTP	Hyper Text Transfer Protocol
HUMAN	High-Speed Unlicensed Metropolitan Area Networks
IE	Information Element
IEEE	Institute of Electrical and Electronics Engineering
IFFT	Inverse Fast Fourier Transform
IMS	IP Multimedia Subsystem

IP	Internet Protocol
IPR	Intellectual Property Rights
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
ISDN	Integrated Services Digital Network
ISI	Inter-Symbol Interference
ISP	Internet Service Provider
ITU-R	International Telecommunications Union – Radio Sector
KPI	Key Performance Indicator
LAN	Local Area Network
LLC	Logical Link Control
LOS	Line of Sight
LTE	Long-Term Evolution
MAC	Medium Access Control
MAN	Metropolitan Area Network
MAP	Media Access Protocol
MBS	Multicast and Broadcast Services
MCM	Multicarrier Modulation
MDHO	Macro Diversity Handover
MIMO	Multiple Input Multiple Output
MLT	Maximum Latency Tolerance
MPEG	Moving Picture Expert Group
MPLS	Multilayer Protocol Switching
MRTR	Minimum Reserved Rate
MSTR	Maximum Sustained Rate
NRTPS	Non-Real-Time Polling Services
NSP	Network Service Provider
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OFUSC	Optional Downlink Fully Used Subchannelisation
OPUSC	Optional Uplink Partially Used Subchannelisation
OSI	Open Systems Interconnection
PDF	Probability Density Function
PER	Packet Error Rate
PF	Proportional Fair

PHY	Physical Layer
PPP	Point-to-Point Protocol
PUSC	Partially Used Subchannelisation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RCS	Radio Channel Simulator
REF	Reference Scenario
RNG	Random Number Generator
ROHC	Robust Header Compression
RRM	Radio Resource Management
RS	Reed-Solomon
RSSI	Radio Signal Strength Indicator
RTG	Receive/Transmit Transition Gap
RTP	Real Time Protocol
RTPS	Real-Time Polling Services
SC	Single Carrier
SCM	Single Carrier Modulation
SDU	Servide Data Unit
SIMO	Single Input Multiple Output
SIR	Signal-to-Interference Ratio
SM	Spatial Multiplexing
SNR	Signal-to-Noise Ratio
SOFDMA	Scalable Orthogonal Frequency Division Multiple Access
SOHO	Small Office Home Office
SPC	Speech Centric Scenario
SSCS	Service Specific Convergence Sub-Layer
ST	Subscriber Terminal
STC	Space Time Coding
TCP	Transmission Control Protocol
TDD	Time Division Duplex
TDM	Time Division Multiplex
ToS	Type of Service
TTG	Transmit/Receive Transition Gap
UDP	User Data Protocol
UGS	Unsolicited Grant Service

UL	Uplink
UMTS	Universal Mobile Telecommunications System
UTG	User Traffic Generator
VBR	Variable Bit Rate
VLAN	Virtual Local Area Network
VO	Video Object
VoIP	Voice Over IP
VOL	Video Object Layer
VOP	Video Object Plane
WiMAX	Wordwide Microwave Interoperability Forum
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area network
WWW	World Wide Web

List of Symbols

α	Parameter for GBAR model
β	Shape parameter of *Ga*
η	MAP efficiency
λ	Scale parameter of *Ga*
λ_i	Parameter of the GOP GBAR model
μ_{Dd}	Mean value of D_d
μ_{Dpc}	Mean value of D_{pc}
μ_{LN}	Natural log of the mean of the Lognormal PDF
μ_{Nd}	Mean value of N_d
μ_{Npc}	Mean value of N_{pc}
ρ	Lag 1 autocorrelation coefficient of the sample data
ϱ_i	Parameter of the GOP GBAR model
φ	Walfisch-Ikegami's angle measured between street direction and incident wave direction
$a(h_{ST})$	Okumura-Hata's correction factor due to ST antenna height variation
a_I	Parameter of the GOP GBAR model
A_{3s}	Area of a tri-sectorised site
A_n	Random variable associated to a stochastic process $\{X_n\}$ with marginal $Ga(b,l)$ distribution
A_r	Attenuation between receiving antenna and receiving radio unit
A_t	Attenuation between transmitting antenna and transmitting radio unit
a_p	Parameter of Pareto distribution
B	Bidirectionally Predicted Video Object Plane
b	Walfisch-Ikegami's inter-building distance
b_s	Uncoded block size per symbol
B_W	Channel Bandwidth
B_E	Beta distributed random variable
B_n	Random variable associated to a stochastic process $\{X_n\}$ with

	marginal Ga(b,l) distribution
C_R	Cell radius
d	Distance between BS and ST
D_d	Interarrival time between packets within a packet call
D_{pc}	Reading time between packet calls
F_N	Noise Figure
f	Frequency
f_s	Sampling Frequency
G_{ar}	Gain of receiving antenna
G_{at}	Gain of transmitting antenna
G_r	Additional gains at reception
G_t	Additional gains at transmission
$G(x)$	Gamma function
Ga	Gamma distributed random variable
h_{BS}	BS antenna height
h_{ST}	ST antenna height
h_{roof}	Walfisch-Ikegami's average building height
I	Intracoded Video Object Plane
K	Minimum packet size
L_{CWI}	Walfisch-Ikegami propagation attenuation
L_{CH}	Okumura-Hata propagation attenuation
L_0	Free Space Propagation Loss
L_{msd}	Walfisch-Ikegami's multi-screen diffraction loss
L_{pen}	Penetration loss
L_{rts}	Walfisch-Ikegami's rooftop-to-street diffraction and scatter loss
M	Spacing between successive anchor Video Object Planes
M_{FF}	Fast-fading margin
M_{int}	Interference margin
M_{SF}	Slow-fading margin
M_{ps}	Maximum allowed packet size
N	Spacing between successive Intracoded Video Object Planes
N_{BC}	Number of Calls Blocked
N_{CA}	Number of Call Attempts
N_d	Number of packets within a packet call
$N_{DataSlots}$	Number of slots busy for user data

N_{DBEP}	Number of Best-Effort Delayed Packets
$N_{DL\text{-}MAP}$	Number of slots occupied by DL-MAP, per frame
N_{DC}	Number of Dropped Calls
N_{FFT}	Fast Fourier Transform Size
N_{LRTPSP}	Number of Real-Time Packet Services Packets Lost
$N_{MaxSlots}$	Maximum number of slots available for downlink, per frame
$N_{OHSlots}$	Number of slots busy due to overhead
N_{pc}	Number of packet calls per session
N_{slots}	Number of available slots for downlink before simplifications, per frame
$N_{subcarriers}$	Number of subcarriers
$N_{subchannels}$	Number of subchannels
$N_{symbols}$	Number of symbols in a frame
N_{TRTPSP}	Total Number of Real-Time Packet Services Packets
N_u	Number of users
N_{tot}	Total noise power
$N_{UL\text{-}MAP}$	Number of slots occupied by UL-MAP, per frame
$N_{usedsubcarriers}$	Number of used subcarriers
n_f	Over-sampling factor
n_{fl}	Number of Floors
P	Forward Predicted Video Object Plane
P_D	Percentage of Best-Effort Packets Delayed
P_L	Percentage of Real-Time Packet Services Lost
P_m	Median of Received Power
P_o	Slow Fading Received Power Threshold
$p_B(x)$	Beta PDF
$p_G(x)$	Gamma PDF
$p_{Ge}[n]$	Probability of generating n packet call requests during a session
$p_P(x)$	Pareto PDF
P_{PC}	Probability of generating one packet call
P_r	Power received at the input of the receiving antenna
P_t	Power at the exit of the transmitting unit
$P_{tx,max}$	Maximum Transmit Power
RB_{BE}	Throughput of Best-Effort Session
RB_{BS}	BS Instantaneous Throughput

R_{DC}	Drop Call Rate
R_{RC}	Refused Call Rate
S_d	Size of a packet
s_{LN}	Natural log of the standard deviation of the Lognormal PDF
S_L	Network Load
T_g	Guard Time
t_{OFF}	VoIP Mean Silent Phase
t_{ON}	VoIP Mean Active Phase
T_s	Total Symbol Period
T_u	Useful Symbol Period
W	Walfisch-Ikegami's street width
X_n	Size of the n^{th} frame for video coding
Z_{ik}	Sample function of a GBAR(a_i,r_i) process

Chapter 1

Introduction

This chapter gives a brief overview of the work. Work targets are established, the scope and motivations are brought up. At the end of the chapter, the work structure is provided.

1.1 Overview

Use of data services has recently been subjected to a significant growth. Democratisation of Internet access boosted not only the usage of services/applications, like Internet surfing itself, but also new applications like www.youtube.com, skype, messenger, file sharing, etc.. This data services increase was initially lead by the increased deployment of traditional fixed access, like cable, but more recently cellular operators have already boosted it even further, through new offers supported mostly by their 3G (3rd generation) networks.

Provisioning of this new traffic mix trend requires efficient technlogies in order for new services to be delivered within certain requirements, either from the operator point of view, who wants to address as many customers as possible, keeping them happy and expend as less as possible, or the user point of view, who wants to have quick Internet access at reduced costs. In this context, many different options exist, e.g., UMTS/HSxPA (Universal Mobile Telecommunications System/High-Speed Downlink/Uplink Packet Access, which combines both HSDPA - High-Speed Downlink Packet Access - and HSUPA - High-Speed Uplink Packet Access), [WiMA06b]. However, wireless broadband service availability is still limited today, as few subscribers have a true broadband experience with throughputs above 1 Mbps, and the cost of service is still too high. Nevertheless, the increasing penetration of notebooks and other portable, data-centric devices is creating a strong propensity for the adoption of mobile Internet services. Mobile data revenues are growing quickly, due to the introduction of these technologies, whereas voice revenues are stagnant or falling, due to strong competition between operators and to market penetration reaching 100% or more.

Original UMTS based cellular infrastructures are optimised to carry circuit switched voice traffic, but are not designed to cope with the growing amount of traffic generated by high-speed and real-time applications. HSDPA introduction minimised this problem, by allowing bitrates to reach up to 14.4Mbps, although only very recently the first terminal equipments capable of reaching 7.2Mbps were made available. Nevertheless, costs for the operators for delivery of such bitrates are very high, due to intrisic matters related to UMTS. Among these, Code Division Multiple Access (CDMA) related Intellectual Property Rights (IPR) is one key issue, as royalties paid by manufacturers on Wideband CDMA (WCDMA) phones are around 10% to 15% of the average selling price of a handset, compared to a telecommunication industry norm of 2% to 5%, [WiMA06e]. Another issue is the network related cost, e.g., high backhaul based on ineficient Time Division Multiplex (TDM) technology, well suited for voice but not efficient for data, which is only now being migrated towards more efficient and cheap Internet Protocol (IP) based technology. For UMTS to be truly data-centric and more cost-efficient, it is necessary to wait until the rollout of Long Term Evolution (LTE), which is expected only by 2010. To meet the demand for wireless broadband, mobile operators and other service providers have to explore new technologies when planning for next-generation networks.

The Institute of Electrical and Electronics Engineering (IEEE) has developed a new standard, 802.16, [IEEE02], with the objective to make (fixed) broadband wireless access more widely and cheaply available. Later on, the original standard was enhanced so that improved radio features and support for mobility was addressed as well. The latest revision is the 802.16e standard, [IEEE06].

The IEEE 802 committee, which has set, among others, networking standards such as Ethernet 802.3 and Wireless Local Area Network (WLAN) 802.11 has published a set of standards that are applied to the concept of Wireless Metropolitan Area network (WMAN) within the 802.16 group. In the 802.16 scope, the Wordwide Microwave Interoperability (WiMAX) was created. WiMAX is a non-profit corporation formed in 2001 by leading communications, component and equipment companies, intending to promote and certify compatibility and interoperability of broadband wireless access equipment conforming to IEEE 802.16 standard. The WiMAX Forum intends to remove an important barrier to adoption of the standard by assuring demonstrable interoperability between system components developed by the different manufacturers. Roughly, the role of WiMAX for 802.16 is the same role that the Wi-Fi Alliance has taken for 802.11.

The IEEE 802.16 project started in 1998 with the objective of making fixed broadband wireless access more widely and cheaply available through a standard for WMANs. These standards specify the radio interface for Broadband Wireless Access systems, characterised by being wireless, fixed or mobile, by providing multiple access to multiple simultaneous users, and by allowing multiple services to be delivered to end users.

The 1st standard version for 802.16 was published in April 2002, basically addressing fixed line of sight (LOS) wireless connections for the *last-mile* link, operating in the 10 to 66 GHz bands, [IEEE02]. The first revision, 802.16a, was later released in April 2003, extending the operation to the 2-11 GHz bands, [IEEE03]. The 802.16a standard specified protocols that provide connectivity without requiring a direct LOS between subscriber terminals (STs) and the base stations (BSs), allowing a single BS to support hundreds or even thousands of STs. Later, by October 2004, a new revision, 802.16d [IEEE04] was issued, addressing interoperability, by providing detailed system profiles and specifying combinations of options for compliance and interoperability tests. Finally, the latest revision, 802.16e [IEEE06] was released in February 2006, bringing mobility and additional features to the standard. Targets of the 802.16e standard also include increased peak data rates, improved Quality of Service (QoS), and enhanced spectral efficiency for both uplink and downlink (UL and DL) traffics. 802.16 genealogy is illustrated in Figure 1.1. The present report only concentrates on the 802.16e version, which basically acts as a WMAN system offering mobility and high bitrates.

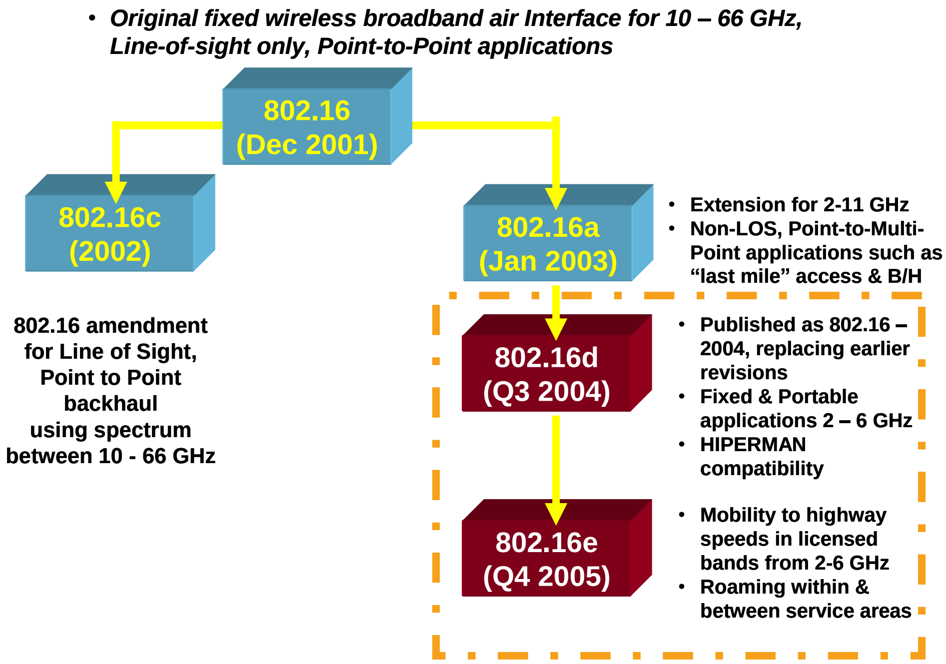


Figure 1.1 – IEEE 802.16 genealogy (extracted from [Kirk06]).

As illustrated in Figure 1.2, existing cellular standards were originally focused on providing lower bitrates in wide areas, with full mobility support, but are now moving towards the support of higher bitrates, through, e.g., HSDPA, while still providing the same coverage of mobility support. On the contrary, IEEE 802 based standards addressed originally high bitrates within low coverage areas, with no mobility support at all, but are now moving towards wide area coverage with mobility support, at the cost of bitrate.

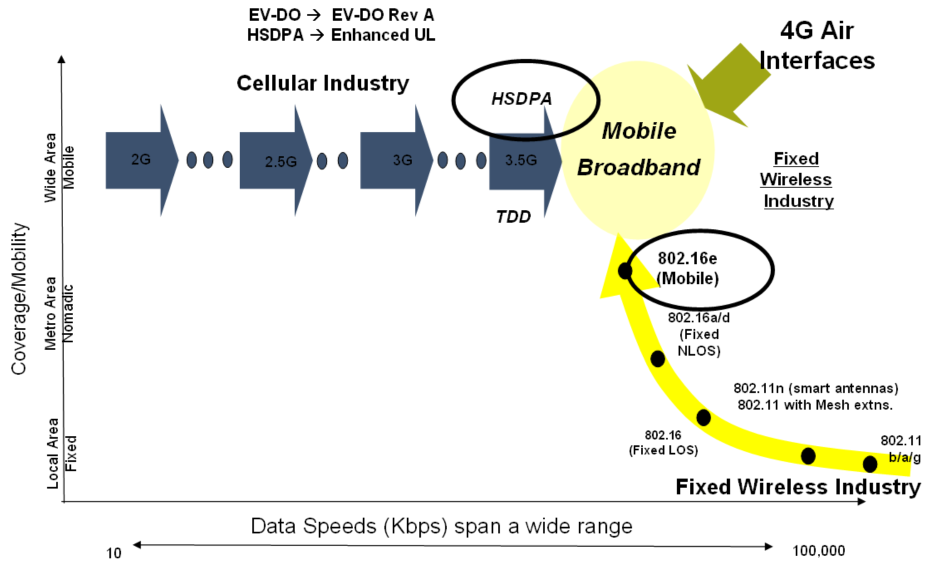


Figure 1.2 – Relation between existing wireless technologies (extracted from [Kirk06]).

802.16e is the technology that theoretically meets the demand for broadband services. It is based on a next-generation all-IP core network, that offers low latency, advanced security, QoS, and worldwide roaming capabilities. Service providers also benefit from the low costs of a technology based on open standards, and favorable IPR. The theoretical advanced performance of 802.16e is largely tied to its use of Orthogonal Frequency Division Multiple Access (OFDMA), a multiplexing technique well suited to multipath environments that gives network operators higher throughput and capacity, great flexibility in managing spectrum resources, and improved indoor coverage. OFDMA has clearly emerged as the technology of choice for next-generation mobile networks. The Third Generation Partnership Project (3GPP) has incorporated OFDMA in its LTE specification, and the Third Generation Partnership Project Two (3GPP2) is moving in the same direction, but WiMAX may have a two-to-three year time advantage over LTE, which is still in the early stages of development. In addition, the use of OFDMA in 802.16e networks makes it substantially less complex and more cost-effective to implement technologies like Multiple Input Multiple Output (MIMO) and beamforming, compared to CDMA-based ones.

WiMAX will coexist and interwork with existing and emerging technologies, both wired and wireless. Even though it can support Voice over IP (VoIP), WiMAX is not expected to replace or compete with 3G technologies for voice services. Cellular networks provide the extensive coverage that circuit-switched voice services require, and that the WiMAX infrastructure is not designed to support. 3G networks cover many urban and suburban areas, but they may not offer sufficient capacity or throughput for data applications. Similarly, WiMAX and Wi-Fi are complementary, being expected to be incorporated in dual-mode chipsets in mobile devices, as

WiMAX provides wider coverage, while Wi-Fi is better suited for high-throughput, indoor Local Area Network (LAN) applications. WiMAX also addresses the requirements of those subscribers that want to be able to use their broadband connection regardless of location, a functionality that Digital Subscriber Line (DSL) and cable modem services do not support. The WiMAX Forum has taken a proactive role in ensuring that WiMAX will be capable of interworking with these technologies and in supporting emerging architectures like IP Multimedia Subsystem (IMS) that enable operators to make the same applications and services available across multiple wired and wireless interfaces.

1.2 Motivation and Contents

The introduction of a new technology such as 802.16e puts the question about its performance capabilities: is 802.16e a radio access technology that can facilitate the evolution of the wireless communication market by enabling the support of high data rate services? In order to be able to facilitate the evolution of the wireless communication market, 802.16e based networks should be able to increase system capacity over other competing technologies. This technology should also be able to meet the QoS demands for different mobile services. The concept of QoS is a key factor in WiMAX evaluation, because there is a fundamental trade-off between the QoS provided to the end user and the way how the different traffic sources are managed by the system.

The main objective of this M.Sc. thesis is to investigate the system level performance of 802.16e, as well as the existing trade-off between the QoS provided to the end user and the type of scheduling approach, for different traffic types. This evaluation will enable to set the boundaries of the performance capabilities of 802.16e technology. The traffic classes targeted in the investigation described above include Real-Time Packet Services (RTPS), as voice, and best-effort (BE) services, e.g., file transfer.

In the early stages of this M.Sc. thesis, the focus concentrated on building understanding on the main theoretical aspects that concern 802.16e, including details on the Physical (PHY) layer, frame structure, spectrum, Radio Resources Management (RRM), QoS support, etc. The 802.16e standard appears as an umbrella of features to improve both user and system performance, including Adaptive Modulation and Coding (AMC) schemes, fast PHY layer Hybrid Auto Repeat Request (HARQ) mechanism, fast channel quality feedback, allowing for the implementation of advanced scheduling algorithms, and Scalable Orthogonal Frequency Division Multiple Access (SOFDMA) based radio interface.

Focus was also concentrated on understanding the services that could potentially be conveyed by 802.16e: VoIP, Video Telephony, World Wide Web (WWW), Streaming, File Transfer Protocol (FTP) and Email services. Relevant characteristics such as the QoS demands of the different traffic classes and the statistical properties of the traffic are investigated.

Prior to the 802.16e evaluation, it is of major importance to identify the traffic characteristics of the traffic classes under investigation, as well as the QoS requirements that these services impose on the communication, and especially on the radio access network. The investigation concentrates on the scheduler, because it is the central entity of the 802.16e design, as this functionality governs the distribution of the radio resources available in the cell among the users. Due to its function, the scheduler has a direct impact on system performance. Similarly, it also determines end user performance, and more specifically the relative performance among users in the cell. Hence, it is of paramount interest to find suitable scheduling algorithms that can optimise the aforementioned trade-off between system capacity and end user performance for the different traffic classes. Due to the influence of this functionality on the relative performance among users, special emphasis is put on the concept of fairness in the distribution of the radio resources among users in the cell.

The assessment of 802.16e technology at the network level requires a method that can provide absolute cell capacity figures under realistic conditions. There are several options to evaluate the performance of a cellular network, including, e.g., analytical analysis, static simulations or dynamic simulations. Considering the need of the M.Sc. objectives, and the early stage of the 802.16e system design (which implies that operational networks are not available yet), dynamic system level simulations were selected as the assessment methodology in the present M.Sc. thesis. Dynamic simulations, by including the time dimension, are very appropriate for investigating time dependent mechanisms or dynamic algorithms. RRM functionalities, such as the scheduler, can be properly analysed with this assessment methodology.

As 802.16e is a very recent standard, with very few products and commercial networks based on it available on the market, there are still a lot of areas to study and investigate on this field. This thesis innovates by the fact that different 802.16e based scheduling policies suited for multi-traffic mix scenarios were studied, implemented within a simulator and assessed, with the major advantages and disadvantages of each scheduling approach being highlighted. This thesis has also contributed for the Advanced Resource management solutions for future all IP heterOgeneous Mobile rAdio environments (AROMA) project, [Ljun06].

This thesis is organised as follows:

- Chapter 1 gives a short introduction and outlines the objectives.
- Chapter 2 presents a description of 802.16e main concepts. Additional concepts concerning implementation of traffic sources and propagation models used in the developed simulator are introduced as well.
- Chapter 3 describes the main mechanisms and algorithms studied and implemented within the developed simulator.
- Chapter 4 summarises the main results for the different scenarios considered for simulations, scenarios being listed as well. The chapter discusses different scheduling policies for the different traffic types and evaluates the network level performance.
- Chapter 5 draws the main conclusions and discusses future research topics.

Chapter 2

802.16e System Description

This chapter provides a description of 802.16e, namely architecture, radio interface, supported services and applications, as well as a theoretical performance comparison with other wireless broadband systems, such as UMTS/HSxPA and Code Division Multiple Access 2000 Evolution-Data Optimised (CDMA2000 EVDO).

2.1 System Architecture

The IEEE 802.16 standard, ([IEEE04], [IEEE06]), is part of the whole set of IEEE 802 standards for both Location Area Networks LANs and Metropolitan Area Networks (MANs), in which both physical and logical levels of the OSI (Open Systems Interconnection) model are defined. The relationship between the 802.16 standards and the remaining IEEE 802 standards is illustrated in Figure 2.1.

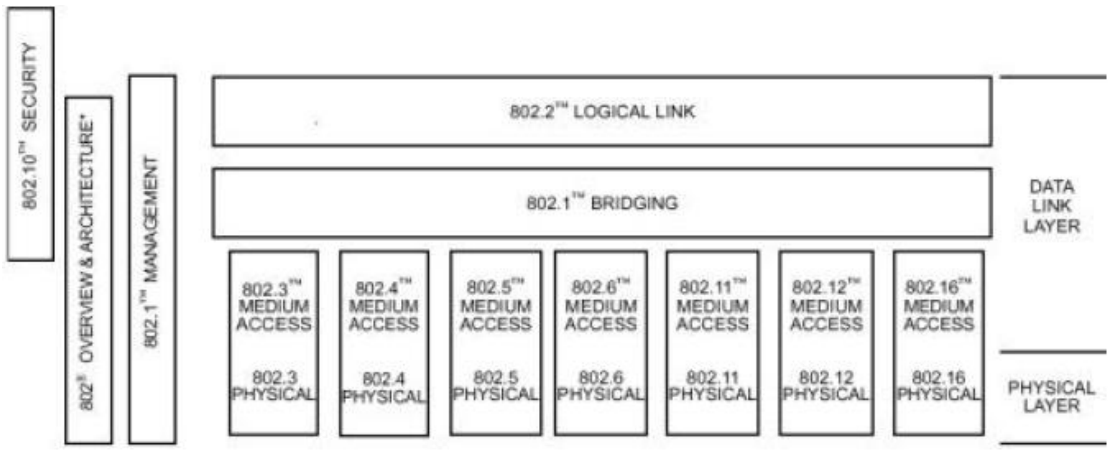


Figure 2.1 - IEEE 802.group of standards, including 802.16 (extracted from [Nune03]).

The lowest layer of the IEEE 802 reference model corresponds to the PHY layer of the OSI model, including functions such as encoding and decoding of signals, preamble generation and removal, or bit transmission and reception. The functions associated to providing service to LAN/MAN users are above the PHY layer, including, [Stal00]:

- On transmission, assembly of data into frames with address and error-detection fields.
- On reception, disassembly of frames, address recognition and error detection.
- Government of access to the transmission medium.
- Provision of interfaces to higher layers and perform flow and error control.

These functions are typically associated to OSI layer 2. Above the Medium Access Control (MAC) layer, the bridging and Logical Link Control (LLC) protocols exist within the IEEE 802 model, and several MAC options may be provided for the same bridging/LLC. Thus, the specific 802.16 standards for MAN specify only PHY and MAC layers, while the LLC protocol specified through the IEEE 802.2 standard is kept. The specific IEEE 802.16 reference model is illustrated

Figure 2.2.

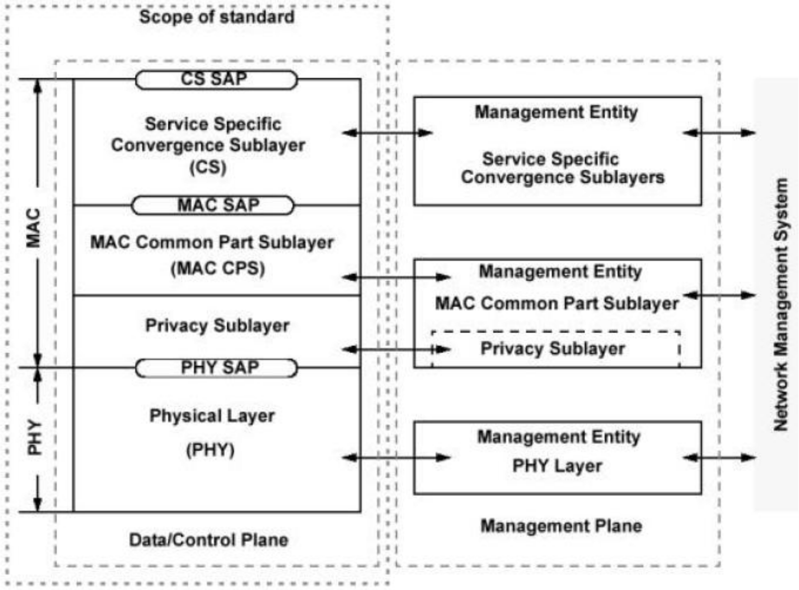


Figure 2.2 – IEEE 802.16 protocol stack (extracted from [Nune03]).

As illustrated, the MAC layer includes three sub-layers:

1. The Service Specific Convergence Sub-Layer (SSCS), which provides translation or mapping of external network data received from upper layers onto MAC Service Data Units (SDUs), including classifying these SDUs into the proper service flows and Connection Identifiers (CIDs).
2. The Common Part Sub-Layer (CPS), providing the core MAC functionalities of system access, bandwidth allocation, connection establishment and connection maintenance.
3. The Privacy Sub-Layer, providing authentication, secure key exchange and data encryption.

Within the Convergence Sub-Layer (CS), two specifications are defined, [IEEE06]: the Asynchronous Transfer Mode (ATM) CS and the packet CS, which specify how either ATM packets or IP packets (including not only IP but Point-to-Point Protocol (PPP) or 802.3 based packets as well), respectively, should be classified, processed and delivered to CPS. This is a major role within the whole MAC protocol, because it affects packet prioritisation and system efficiency, critical for providing the necessary conditions for fulfilling QoS requirements.

WiMAX components can be grouped into two major blocks: the Access Service Network (ASN) and the Connectivity Service Network (CSN) subsystems. The ASN is the usual access network, comprising elements such as BSs, STs and gateways, and may be shared by more than one CSN.

ASN is defined as a complete set of network functions needed to provide radio access to a WiMAX subscriber, providing some mandatory functions such as Layer-2 connectivity with STs, transfer of Authentication, Authorisation and Accounting (AAA) messages to subscriber's Home Network Service Provider (H-NSP), network discovery and selection of the subscriber's preferred Network Service Provider (NSP), relay functionality for establishing Layer-3 connectivity with a ST (i.e., IP address allocation), RRM or ASN-CSN tunnelling. ASN and CSN anchor mobility is a function of the ASN as well, [WiMA06c].

The CSN comprises elements such as routers, AAA proxy/servers, user databases or inter-working gateway devices. The CSN is defined as a set of network functions that provide IP connectivity services to subscribers, enabling functions such as ST IP address and endpoint parameter allocation for user sessions, Internet access, AAA proxy or server, policy and admission control based on user subscription profiles, ASN-CSN tunnelling support, subscriber billing and inter-operator settlement, inter-CSN roaming and inter-ASN mobility [WiMA06c].

WiMAX can be integrated either as stand-alone system, or together with other access networks on existing 3GPP based systems such as UMTS. 3GPP has defined in its specifications (and more generally in the .234 family, [3GPP06], [GPP07a], [GPP07b], an architecture describing inter-working between 3GPP systems and WLANs. In order to keep consistency in inter-working between 3GPP systems and Wireless Access Networks, inter-working with WiMAX will be based on the same model. There are ongoing discussions aiming at extending the scope of [3GPP02a] to a wider range of IP-based Access Networks, including the IEEE 802.16 family [WiMA06d]. The ultimate goal, based on an IMS core network, is to inter-connect multiple access technologies to the same common core network, as illustrated in Figure 2.3.

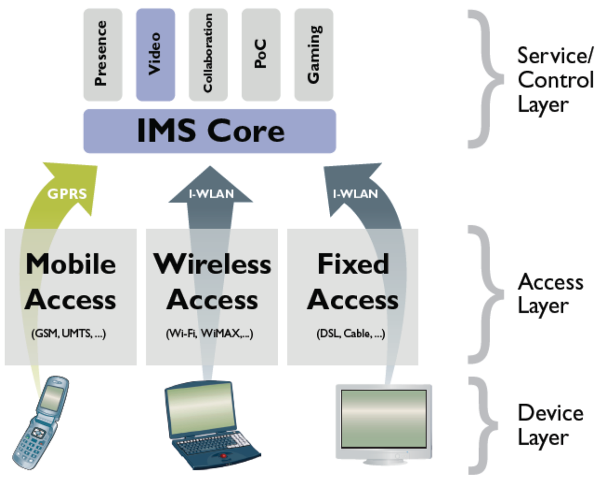


Figure 2.3 – WiMAX inter-connection with other access networks under a common core network (extracted from [Kine07]).

2.2 Radio Interface

This section summarises the main topics that characterise the 802.16e radio interface. The first sub-section describes briefly the different specified air interface procedures. After that, a brief introduction to the differences between OFDMA and SOFDMA is done. Aspects related to the PHY layer are introduced later, as well as specific mobility and frame structure topics. The different possibilities in terms of subcarrier allocation modes are summarised in another sub-section, the section ends with an introduction to 802.16e specific link budget calculations.

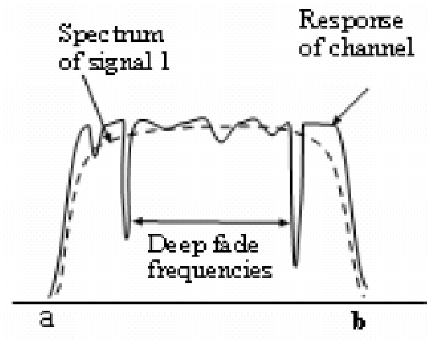
2.2.1 OFDMA and SOFDMA

The 802.16e standard specifies mainly 4 different air interface procedures, [IEEE04]:

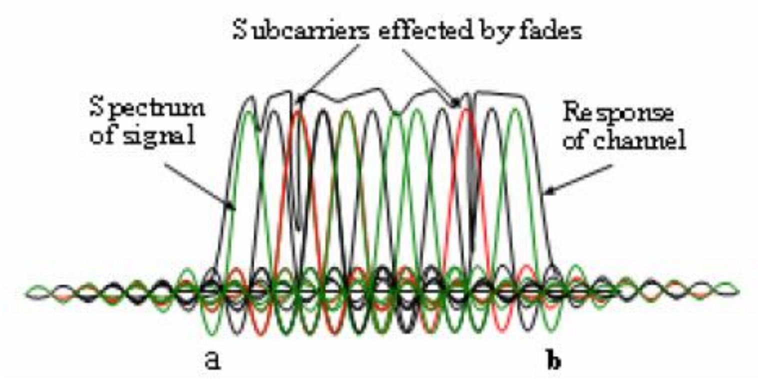
1. WMAN Single Carrier (SC), operating in [10, 66] GHz in either Frequency Division Duplex (FDD) or TDD.
2. WMAN Orthogonal Frequency Division Multiplexing (OFDM), operating below 11 GHz, also in either FDD or TDD.
3. WMAN OFDMA, operating below 11GHz, in either FDD or TDD.
4. Wireless High-Speed Unlicensed Metropolitan Area Network (HUMAN), operating below 11 GHz in license-exempt bands in TDD only.

OFDM is a special form of multi-carrier modulation (MCM), where a single data stream is transmitted over a number of lower rate subcarriers. OFDM brings an obvious advantage compared to a traditional single carrier modulation (SCM), which is the increased robustness against frequency selective fading and narrowband interference: in a SCM system, a single fade or interferer can cause the entire link to fail, while in an MCM, only a small percentage of subcarriers is affected, [Hara03], Figure 2.4.

OFDM differs from OFDMA in the fact that with OFDMA packets can be scheduled for different users on both frequency (subchannels) and time (symbols) domains. One of the majors setbacks to the OFDM static multiple access scheme is the fact that different users see the wireless channel differently. When OFDMA is used instead, multiple users are allowed to transmit simultaneously on the different subcarriers per time symbol, and since the probability that all users experience a deep fade in a particular subcarrier is very low, it can be assured that subcarriers are assigned to users that see good channel gains when OFDMA is used, [WZEA04]. Thus, OFDMA provides higher granularity in resource allocation, more degrees of freedom in scheduling, and improved fairness and QoS.



a) SC signal



b) MCM signal

Figure 2.4 – Comparison between multi-path frequency selective fading influence on SC and MCM signals.

The present thesis focuses only on OFDMA as the access and multiplexing mode for the PHY layer, since the 802.16e amendment was developed to cover mobile applications, which is boosted by OFDMA adoption instead of OFDM. OFDMA is capable of providing the flexiblity to deal with varied scenarios and challenges associated with rapidly moving mobile users in a NLOS environment, [Alva06]

Furthermore, the concept of SOFDMA is introduced, in order to support scalable channel bandwidths from 1.25 to 20 MHz. 802.16e is designed to be able to work in different channelisations complying with varied worldwide requirements, as efforts proceed to achieve spectrum harmonisation in the longer term, [WiMA06a]. An introduction to OFDM is given in [Pras98], this modulation and corresponding multiple access technique being detailed in Sub-section 2.2.2.

The fact that the resulting subcarriers have reduced rate, thus, increased symbol duration, improves the robustness of the system against delay spread. Additionally, a cyclic prefix (CP) is introduced in each symbol, which can completely eliminate inter-symbol interference (ISI) as long as the CP duration is longer than the channel delay spread. If a guard period is used instead of the CP, the same ISI mitigation effect can also be achieved, but that would result in a sudden change of the waveform, which introduces higher spectral components, thus, resulting in intersubcarrier interference [Hara03]. However, the use of CP has a drawback, since it introduces overhead, reducing bandwidth efficiency. Nevertheless, this is moderated, because OFDM has a very sharp spectrum, meaning that a large fraction of the allocated channel bandwidth can be used for data transmission, [WiMA06a], Figure 2.5.

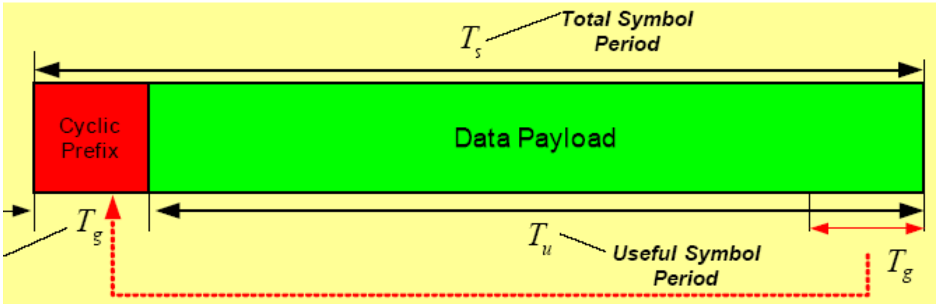


Figure 2.5 – Cyclic Prefix Insertion in OFDMA Symbol (extracted from [WiMA06a]).

OFDM can be realised via Inverse Fast Fourier Transform (IFFT), which enables a large number of subcarriers with relatively low complexity. However, while in OFDM there is no user multiplexing across multiple subcarriers within the same time slot, in OFDMA resources are available in both time and frequency domains, by means of OFDM symbols and subcarriers, respectively. These time and frequency domain resources can be organised into subchannels for allocation to individual users.

OFDMA scalability is supported by adjusting the Fast Fourier Transform (FFT) size, N_{FFT}, while keeping a constant subcarrier frequency spacing, which minimises the impact to higher layers. Table 2.1 summarises the main parameters that characterise SOFDMA, for some of the supported channel bandwidths.

Table 2.1 – Main SOFDMA Parameters (taken from [WiMA06a]).

Parameters	**Values**			
System Bandwidth [MHz]	1.25	5	10	20
Sampling Frequency (f_s) [MHz]	1.40	5.60	11.20	22.40
Sampling Time [ns]	714.29	178.57	89.29	44.64
FFT Size (N_{FFT})	128	512	1024	2048
Subcarrier Frequency Spacing [kHz]	10.94			
Useful Symbol Time (T_u) [μs]	91.40			
Guard Time (T_g) [μs]	11.43			
OFDMA Symbol Time (T_s) [μs]	102.83			
Number of OFDMA Symbols (5ms frame)	48			

Some remarks regarding Table 2.1 are in order, [WiMA06a], [Yagh04]:

1. As mentioned previously, subcarrier spacing is independent of the bandwidth.
2. N_{FFT} scales to the bandwidth.
3. The sampling frequency is obtained by multiplying the system bandwidth by the over-sampling factor (n_f), given by [IEEE06]:
   a. $8/7$ for channel bandwidths multiple of 1.75 MHz;
   b. $28/25$ for channel bandwidths multiple of 1.25, 1.5, 2 or 2.75 MHz;
   c. $8/7$ for channel bandwidths not specified.
4. The guard time can varies between $1/4$, $1/8$, $1/16$ or $1/32$. In Table 2.1, $1/8$ is used as an example, meaning that a maximum of roughly 11.4 ms delay spread can be tolerated, with an overhead of 12.5%.

According to the International Telecommunications Union – Radio Sector (ITU-R) Vehicular Channel coherence bandwidth Model B [ITUR97] for mobile environments, for delay spreads up to 20 μs, a coherence bandwidth of 10 kHz is achieved. The subcarrier spacing design requires a flat fading characteristic for worst-case delay spread values of 20 μs with a guard time overhead of no more than 10% for a target delay of 10 μs. By combining these requirements, it can be seen that:

1. The defined subcarrier frequency spacing of 10.94 kHz copes with the coherence bandwidth for the worst case delay spread.
2. The possibility of adopting a maximum guard time of $1/4$, if the total symbol allows tolerating delay spreads of up to 22.8 μs, again coping with the worst case delay spread. The WiMAX operator can thus set a trade-off between the CP overhead and the delay spread tolerance.

The advantage of allowing different channel bandwidths is to make the standard compliant with the varied worldwide requirements in terms of spectrum. By standardising scalable OFDMA, spectrum harmonisation can be more easily achieved. Other benefits include the fact that the operator can use different channel bandwidths according to the specific needs in different areas: for instance, the operator can choose to adopt narrower channel bandwidths in rural areas, where demand for capacity is lower and coverage is the main requisite, while in urban areas larger bandwidths are likely to be needed for capacity, at the cost of reduced coverage. Details on trade-off between capacity and coverage depending on channel bandwidth are given in Section 2.3.

Although the standard refers to all channel bandwidths listed in Table 2.1, the system bandwidths for the initial planned profiles being developed by the WiMAX Forum Technical Working Group for Release-1 are mainly 5 and 10 MHz [WiMA06a].

2.2.2 Physical Layer

The standard supports TDD, FDD and Half-Duplex (HD) FDD operation modes, but it is expected that the profiles from the initial release of 802.16e will only include TDD [WiMA06], due to which the present work focuses mainly on the TDD approach. As already known from other technologies, TDD requires system-wide synchronisation to cope with interference issues, but it presents several advantages against FDD, namely, [WiMA06a]:

- TDD requires only an unpaired channel for both UL and DL, while FDD requires two paired channels, which naturally means that TDD is more flexible in order to get the required spectrum.
- Balancing of traffic between UL and DL is further more efficient when TDD is used, with the ratio between UL and DL allocations being done dynamically according to the actual needs. On the contrary, when FDD is used, fixed channel bandwidths are permanently allocated for both UL and DL, regardless of the actual channel activity.
- Contrary to FDD, TDD ensures channel reciprocity, providing better support of closed loop advanced antennas technologies.
- TDD transceivers are cheaper, due to the lower required complexity.

The specifications define a combined variable-rate Reed-Solomon (RS) / Convolution Coding (CC) scheme as mandatory, supporting code rates of $1/2$, $2/3$, $3/4$ and $5/6$. Variable rate Block Turbo Code (BTC) and Convolution Turbo Code (CTC) are specified as being optional [NCMP04]. The standard supports multiple modulation levels as well, including support for Binary Phase Shift Keying (BPSK), Quadrature Phase Shift Keying (QPSK), 16-Quadrature Amplitude Modulation (16-QAM) and 64-QAM, [NCMP04]. When the radio signal is strong and less interfered, a less robust combination of coding and modulation may be applied and higher bit-rates can be achieved, which is the case of, e.g., 64-QAM with $5/6$ CTC. On the other hand, if the radio channel is highly attenuated and interfered, the received signal is weak and likely to have a higher bit error rate, due to which a more robust combination of coding and modulation is required, which is the case of, e.g., BPSK with $\frac{1}{2}$ CC. Normally, higher modulation and coding rates will be used by STs located close to the BS, while lower ones will be used by STs away from the BS, Figure 2.6.

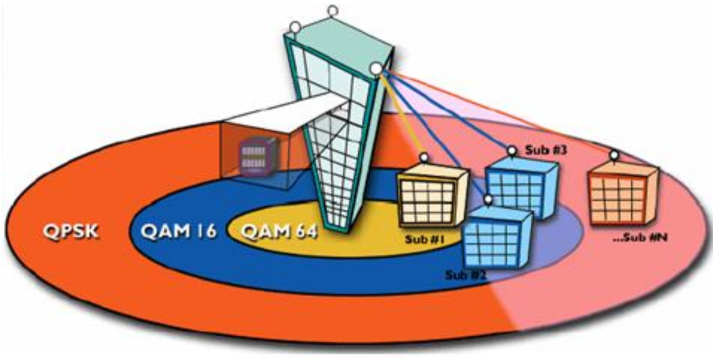


Figure 2.6 – Illustration of Modulation and Coding Rate allocation across cell radius (extracted from [Mark03]).

The specification also includes the optional support for smart antenna technologies [WiMA06a]:

- Beamforming, in which multiple antennas systems, i.e., Adaptive Antenna Systems (AAS) are used to transmit weighted signals in order to improve both coverage and system capacity.
- Space Time Coding (STC), in which DL transmit diversity is achieved, reducing fading margins, thus, improving coverage by reduction of path loss.
- Spatial Multiplexing (SM) techniques, like MIMO, in which multiple streams are transmitted over multiple antennas, leading to higher peak rates. If the receiver also has multiple antennas, it can separate the different streams and thus achieve higher throughputs. On the other hand, in the UL, each user has only one transmit antenna, but two users can transmit collaboratively in the same slot as if two streams are spatial multiplexed from two antennas of the same user.

Adaptive switching between the different advanced antenna options can be performed, allowing the system to better adapt the benefit of smart antennas to the different channel conditions. It is expected, for instance, that while SM improves the peak throughput, it can degrade the Packet Error Rate (PER) if the radio channel suffers from bad quality. On the other hand, STC provides larger coverage areas even under bad channel conditions, but does not improve peak throughputs, [WiMA06a].

OFDMA supports the use of subchannelisation in both UL and DL, with the standard supporting five different schemes (Sub-section 2.2.4). This is particularly important for UL, since without subchannelisation, the regulatory restrictions as well as the need for cost-effective STs typically would cause the link budget to be asymmetrical and, particularly, UL limited. The use of subchannels enables the balancing of the link budget by concentrating the ST transmit power into fewer OFDM subcarriers, thus, overcoming, e.g., indoor penetration losses, at the expense of UL capacity, since fewer carriers can be used. This mechanism is illustrated in Figure 2.7.

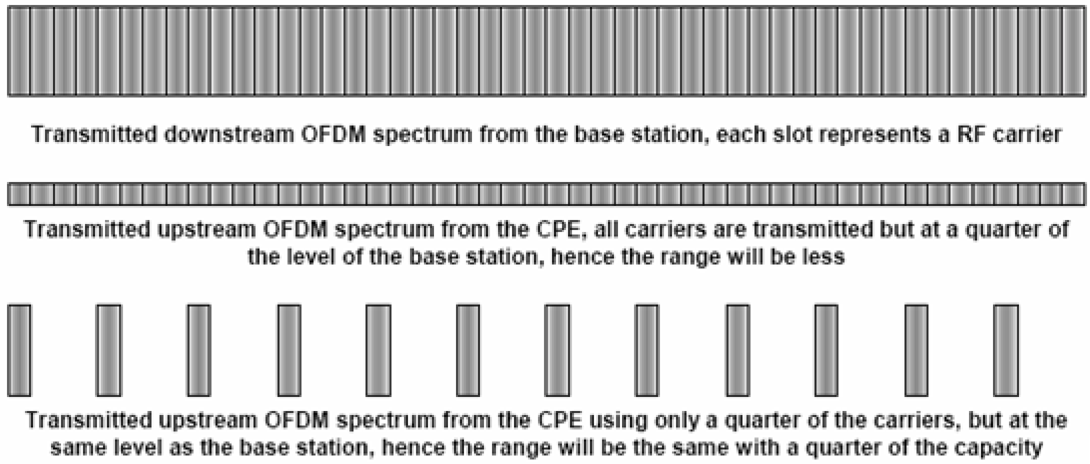


Figure 2.7 – Illustration of subchannelisation effect when used on the UL (extracted from [WiMA04b]).

Furthermore, thanks to subchannelisation, the frame is divided into zones, each using a different subchannelisation scheme, with the MAC layer being responsible for dividing the frame into these zones and by communicating this structure to STs in the DL and UL Media Access Protocol (MAP) messages. Further details on this subject are given in Sub-section 2.2.3.

HARQ is supported by 802.16e was well. HARQ is similar to traditional Auto Repeat Request (ARQ) stop-and-wait procedures, except that it operates over simultaneous and parallel channels, providing faster responses to packet errors at the PHY layer. The support for chase combining is mandatory for HARQ, and incremental redundancy is an optional feature for 802.16e. Chase combining means that every retransmission matches the coded word employed for the first transmission, [Guti03], thus, the decoder at the receiver combines these multiple copies of the transmitted packet weighted by the received Carrier-to-Interference-plus-Noise Ratio (CINR) prior to decoding. This type of combining provides time diversity and soft combining gain at a low complexity cost and imposes the least demanding ST memory requirements of all HARQ strategies. On the other hand, when incremental redundancy is used, retransmissions include additional redundant information that is incrementally transmitted if the decoding fails on the first attempt, which causes the effective coding rate to increase with the number of retransmissions. This implies more demanding requirements on the ST memory against, e.g., chase combining.

IEEE standards have not identified a specific band for WiMAX to be deployed. Instead, and particularly since the 802.16a standard was issued, WiMAX standards evolved so that the system can work within the [2, 11] GHz band. Within this band, it is up to the different manufacturers, regulators and operators to work jointly in order to define the best solution. Under this scope, Figure 2.8 roughly illustrates the worldwide spectrum allocation in the [2, 6] GHz band:

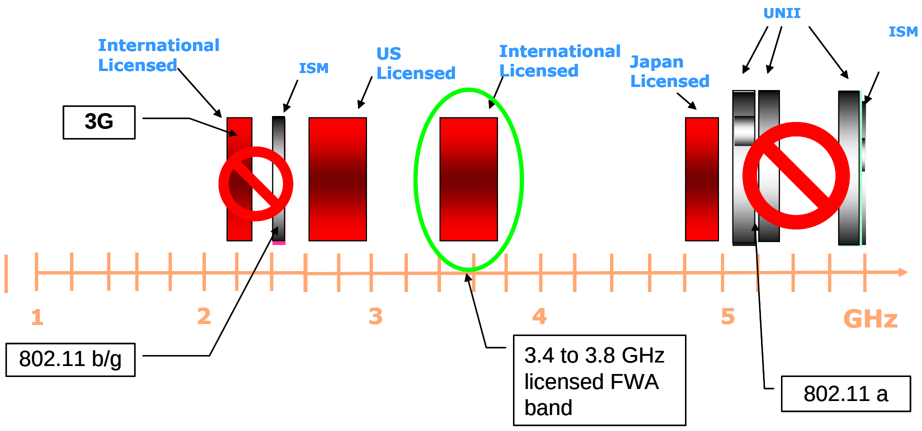


Figure 2.8 – WiMAX spectrum analysis (extracted from [Orr04]).

Within the different possibilities, some remarks are in order, [Eric06]:

- The band [2300, 2400] MHz is already allocated to WiMAX in some countries, e.g., South Korea, but needs to be made available in others. This band mainly employs TDD systems, but in some countries FDD ones can be deployed.
- The band [2500, 2690] MHz (in United States of America, the band [2496, 2690] MHz) is already allocated to International Mobile Telecommunications (IMT) 2000 compliant access systems in many countries. However, the band might be suitable for WiMAX in, e.g., Western European countries, especially since WiMAX was also designated as an IMT 2000 standard.
- The band [3400, 3600] MHz is already allocated for broadband wireless access systems, including WiMAX, but this band needs clearly to be made available for WiMAX in its full range in more countries, as this is the main spectrum profile for WiMAX. The limitation of RF power for mobile devices should be increased to 5 W Equivalent Isotropic Radiated Power (EIRP), allowing for extended and reliable communications.

- The band [5725, 5825] MHz (in some countries [5725, 5875] MHz) should be made available for WiMAX in more countries and the regulatory conditions clarified. The limitation of RF power should be increased to 4 W EIRP allowing for extended and reliable communication.

In general, it is very difficult to find spectrum that is globally available. The openness to use either FDD or TDD is another flexibility that may limit product availability and volumes. The spectrum situation is probably one of the major challenges for WiMAX to become a mainstream technology benefiting from scale economy.

2.2.3 Frame Structure

When TDD is adopted, the 802.16e frame is divided between DL and UL sub-frames through time gaps named Transmit/Receive Transition Gap (TTG) and Receive/Transmit Transition Gap (RTG), respectively. These gaps are required to avoid collisions between transmissions on the two links. The frame structure is illustrated in Figure 2.9.

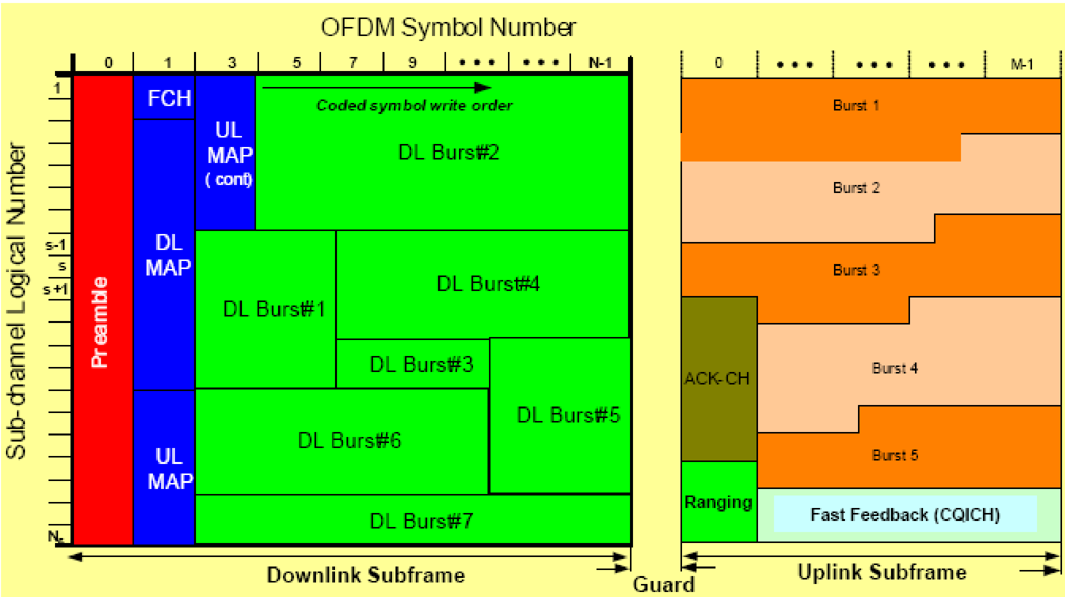


Figure 2.9 – TDD Frame Structure (extracted from [WiMA06a]).

- The **Preamble** is used for synchronisation, being the first symbol of the whole frame.
- The **Frame Control Head (FCH)** provides information related to the frame configuration, including length of MAP messages, allocated subchannels and respective

coding schemes.

- The **DL-MAP** provides additional information on DL subchannel allocation, as well as additional control information for the DL sub-frame.
- The **UL-MAP** has similar goals to DL-MAP, but related specifically to UL.
- The **UL Ranging** subchannel is used by the ST to perform several adjustments, including closed-loop time, frequency and power adjustment, being also used for bandwidth requests.
- The **UL CQICH** channel is used by the ST for feedback about the radio channel quality and reception, which is is used by the BS to evaluate the specific conditions that each ST is experimenting, and based on that, scheduling options, proper modulation and coding schemes can be decided by the BS. The items available to report radio link quality include Physical CINR and effective CINR.
- The **UL ACK** is used by STs to provide feedback on DL HARQ processes, particularly acknowledgement or not of the received data blocks.

Both DL-MAP and UL-MAP messages are critical for scheduling of multipe users within the same radio frame: every ST addressed in the same TDD frame needs to be fully aware of which part of the DL subframe it must listen. The DL-MAP message contains, among others, special Information Elements (IEs) for that purpose ([IEEE04]), including used modulations, coding schemes, identification of target users, symbol and subchannel offsets within the frame, etc.

As for each addressed ST in a single frame a specific DL-MAP IE is needed, if more users are to be simultaneously addressed in the same frame, more IEs are included in the frame, thus, more overhead exists, which needs to be accounted in overall system performance. Additionally, the DL-MAP may be repeated in order to ensure higher decoding probability. The standard specifies a repetition factor of 1, 2, 4 or 6. The tradeoff is simple: a higher repetion factor increases correct decoding probability at the expense of increased overhead, i.e., wasting resources; a lower repetion factor minimises MAP overhead but decoding errors may happen. The present work considers a factor of 2, i.e. DL-MAP overhead is actually duplicated, which represents a fair tradeoff between MAP overhead and reliability of correct decoding of MAP IEs content, [WiMA06a]. It is also assumed that DL-MAP is QPSK ½ modulate and coded, in order to increase its reliability.

The specified mechanism allows the existence of simultaneous DL allocations, which can be not only unicast but also multicast or broadcast. It can also include allocations for other BSs rather than the serving one. Multiple STs can also be given allocation for transmission of data on the UL or request for bandwidth.

The MAP message size is variable, since it depends on the number of allocated users in a frame. If the nature of the traffic is a kind of FTP or Hyper Text Transfer Protocol (HTTP) one, it is expected that the number of users being scheduled per frame should be small (less than 10) [WiMA06a], while if the dominant traffic is, e.g., VoIP, more simultaneous users are expected. As result, additional overhead exists in the case of VoIP due to larger MAP messages, meaning that a lower end-user throughput will be available.

2.2.4 OFDMA Subcarriers and Allocation Modes

There are three types of OFDMA subcarriers [Yagh04]: data subcarriers for data transmission, pilot subcarriers for channel estimation and synchronisation purposes, and null subcarriers, for guard bands. Among these subcarriers, the active ones (data and pilot) are grouped into subsets called subchannels. The minimum frequency-time resource unit of subchannelisation is one slot, which is equal to 48 data tones.

The subcarriers forming one subchannel may be adjacent, although this is not absolutely necessary. Additionally, pilot allocation may be performed differently, depending on the subcarrier allocation mode used [Yagh04]:

- In **DL Fully Used Subchannelisation (FUSC)**, pilot subcarriers are allocated first, the remaining subcarriers being divided into subchannels. In this case, the pilot subcarriers are used from a common set. Subcarriers can be scattered throughout the frequency channel range.
- In **DL PUSC**, the set of used subcarriers, data and pilot, is first divided into subchannels, with pilot subcarriers being allocated within these. DL PUSC uses a cluster-based structure that spams over two OFDM symbols (in time domain) of fourteen subcarriers, each one with a total of four pilot subcarriers per cluster Several scattered clusters of subcarriers can thus be used to form a subchannel.
- **UL PUSC** is similar to DL PUSC, but a tile structure is used instead, which spams over three symbols (in time) of four subcarriers, each one with a total of four pilot subcarriers.
- **Optional DL FUSC (OFUSC)** and **Optional UL PUSC (OPUSC)** can also be

employed, being similar to 'standard' DL FUSC and UL PUSC, but making use of a larger number of subcarriers thanks to different pilot allocation mappings.

- In **DL and UL AMC**, adjacent subcarriers are used to form subchannels.

With PUSC and FUSC, the allocation of subcarriers to subchannels is done in a pseudo-random fashion, such that the subcarriers for a given subchannel in a certain cell are different than the subcarriers for that same subchannel in another cell. This pseudo-random permutation provides an interference averaging effect similar to Global System for Mobile communications (GSM) frequency hopping, further reducing the adverse effects of co-channel interference between neighbor cells. Generally, PUSC and FUSC are more suited for mobile applications while AMC is a better alternative for stationary or low mobility applications, [Alva06].

The IEEE802.16 standard specifies that for the DL only PUSC is mandatory, due to which the present thesis only considers the support of PUSC. Table 2.2 summarises how the available subcarriers are divided among data, pilot and guard subcarriers for N_{FFT} of 512 and 1024 in case of DL PUSC.

Table 2.2 – DL PUSC Subcarrier allocations.

N_{FFT}	1024	512
Number of DC subcarriers	1	
Number of Guard subcarriers	183	101
Number of used subcarriers	841	421
Number of pilot subcarriers	121	61
Number of data subcarriers	720	360
Number of subchannels ($N_{subchannels}$)	30	15
Number of data subcarriers per subchannel ($N_{subcarriers}$)	24	

In DL PUSC, the minimum allocation unit is the slot, which spams over 2 symbols in the time domain and over 1 subchannel in the frequency one, i.e., over 24 data subcarriers. This is the minimum DL allocation unit that can be assigned to a user within a radio frame and every resource allocation within a frame is made in multiples of this unit.

2.3 Link Budget

The Mobile WiMAX link budget differs slightly from traditional single carrier link budgets, like GSM ones, because the transmit power needs to be divided among the multiple subcarriers.

Additionally, since multiple modulations and coding rates are supported, different received sensitivities are required to support each one, which results in different coverage areas for the different combinations of modulation and coding, Figure 2.6.

The specific power class profiles for both BS and ST in OFDMA are summarised in Table 2.3, which refers to the maximum transmit power, but both BS and ST should be capable of reducing their transmission power by 30 dB and 10 dB, respectively, with adjustment steps below 1 dB for both. Typically, BSs are class 4, while STs are class 3.

Table 2.3 – Power Classes (extracted from [IEEE04]).

Class Identifier	Transmit Power [dBm]
1	$17 \leq P_{tx,max} < 20$
2	$20 \leq P_{tx,max} < 23$
3	$23 \leq P_{tx,max} < 30$
4	$30 \leq P_{tx,max}$

Cell coverage in WiMAX depends naturally on the used modulation and coding rate. The less robust modulation and coding rates demand a higher CINR, because the radio link needs to be less interfered in order to support those less robust combinations, which results in smaller cell radius, the trade-off is higher throughputs. On the contrary, more robust modulation and coding rate combinations require a lower CINR, because extra coding is applied for Forward Error Correction (FEC), meaning that they are supported over longer distances from the BS, resulting in larger cell radius, at the cost of lower throughputs. For an Additive White Gaussin Noise (AWGN) channel, the standard mandates that the Bit Error Rate (BER) shall not exceed 10^{-6} for the Signal-to-Noise (SNR) levels summarised in Table 2.4.

Channel coding and modulation assigned for a given user is decided by the BS, based on the feedback received from each ST, which is based on the quality of the DL received signal. The ST measures the signal quality and compares it with the allowed range of operation for each modulation and coding scheme: if the received signal is too good for the used modulation and coding scheme, the ST can request the BS to use a less robust physical mode; on the contrary, if the signal is weak and/or interfered, the ST can inform the BS that a more robust physical mode should be used instead. This is illustrated in Figure 2.10.

Table 2.4 – Receiver SNR assumptions (extracted from [IEEE06]).

Modulation	Coding Rate	Receiver SNR [dB]
QPSK	1/2	5
	3/4	8
16-QAM	1/2	10.5
	3/4	14
64-QAM	1/2	16
	2/3	18
	3/4	20

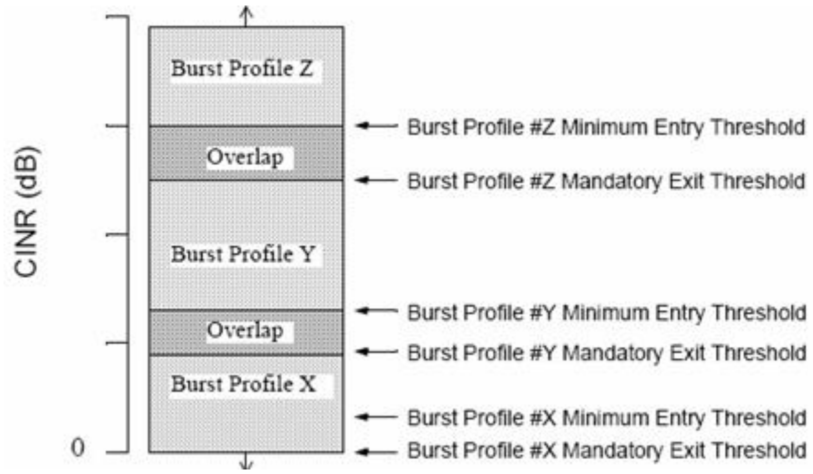


Figure 2.10 – Burst profile threshold usage (extracted from [IEEE06]).

According to the TDD frame structure (Sub-section 2.2.3), the first blocks of the DL subframe contain control information, such as the Preamble, the FCH and the MAP messages that need to be listenable over the whole cell radius, so they require more robust combinations of modulation and coding rate [WiMA06a].

Since diversity techniques such as Transmit and Receive diversity are supported, additional gains are possible on both DL and UL. If, e.g., two antennas are used on the DL for transmission and two antennas are used on the ST for reception as well, a total of up to 6 dB gain can be achieved due to the combination of both [WiMA06a].

Since the system is based on OFDM, the total thermal noise power is lower than the corresponding one for a single carrier technique, because the calculation is performed over the bandwidth of a single subcarrier and not over the whole channel bandwidth. For instance, if a 5 MHz channel is assumed, the calculation of the total thermal noise power is not performed over the whole 5 MHz, but rather over the 10.9 kHz bandwidth of each independent subcarrier (see Table 2.1), which represents an improvement of roughly 26.6 dB.

The margins needed to account for fast fading differ as well in case calculations are performed for, e.g., MAP messages or traffic. When MAP messages are transmitted, higher fast fading margins should be considered, while if effective user traffic data is transmitted, the margin can be lower, because fast frequency selective techniques can be used in order to adapt the used modulation and coding rate to the quasi-instantaneous radio channel conditions. According to [WiMA06a], 6dB should be considered for MAP DL, 4 dB for ranging on UL, while 2 dB should be enough for user traffic on both ways.

There are some additional exceptions due to the use of features such as subchannelisation (see Section 2.2), which allow additional gains on UL as the ST can concentrate its available transmit power on a subset of subcarriers, at the cost of a lower aggregate throughput.

Link budget calculations are presented in Annex A, allowing the estimation of the maximum cell path loss, which, based on a proper propagation model, results in a cell radius.

Cell radius for the different environments results from combination of data summarised in Table 2.6 applied to the respective propagation models, associated to the considerations done in Annex A concerning specific link budget details. For the calculated cell radius, the site area for a tri-sector configuration is calculated through (2.1), [Corr03], where C_R stands for the cell radius. Results are listed in Table 2.5.

$$A_{3s[m^2]} = \frac{1.5 \times C_{R[m]}^{2} \times \sqrt{3}}{1000^2} \tag{2.1}$$

Table 2.5 – Cell radius and site area per radio enviroment considered for simulations.

Environment	**Cell Radius [m]**	**Tri-Sectorised Site Area [km^2]**
Dense Urban	206	0.110
Urban	359	0.335
Suburban	738	1.415
Rural	2842	20.985

2.4 Propagation Models

Propagation models are used for estimation of the received signal strength in a given point in space. There are not well known propagation models to work in the 2.5 or 3.5GHz bands, precisely the bands targeted for 802.16e. The 3.5 GHz band, for example, is mainly used for Fixed Wireless Access (FWA) systems, in which typically LOS exists, thus, standard free-space formulas are used with additional losses due to obstruction of the Fresnel ellipsoids. But there are no specific models for propagation in typical cellular networks.

One the most used propagation models for signal strength estimation is the COST 231 - Hata model, [DaCo99], which is an empirical model with urban scenarios as a standard environment, but that can be used as well on suburban or rural ones. This model is valid for the [1500, 2000] MHz band and for distances within [1, 20] km. Model calculations are detailed in Annex B.

Another commonly used propagation model is the well known COST 231 – Walfisch-Ikegami model [DaCo9], which combines the results from two other models: the Ikegami model [IkYU84], used to estimate signal strength in streets, and the Walfisch-Bertoni one [WaBe88], used for propagation in urban areas, which considers propagation over buildings. This model is valid for the [800, 2000] MHz band and for distances within [0.02, 5] km. Model calculations are detailed in Annex B.

Although these two models are not intended to be applied for frequencies above 2 GHz, both are used under the scope of the present work. The main focus of the work is on RRM and scheduling algorithms, and not on the specific propagation subjects, so these models are used, knowing that the estimated propagation losses can suffer from some lack of accuracy due to the fact that the models are used on bands above the 2 GHz upper threshold.

These propagation models allow the estimation of the average value for path loss, and, consequently, the average value for the received signal strength. The existence of fading results in variations of the received signal strength around its average value, thus, an additional transmission power is required so that this negative effect caused by fading is mitigated. Typically, fading can be divided between slow and fast. Slow fading typically follows a Log-normal Distribution, being mainly a function of the distance between the BS and the ST as a consequence of the terrain configuration, while fast fading is associated to the movement of the

ST, being caused by the multi-path originated by multiple reflections that the transmitted signal suffers on buildings or other structures, being described by Rayleigh or Rice Distributions.

Under these assumptions, propagation models for the present work are used according to the parameterisations shown in Table 2.6, which reflect typical Portuguese environments and cellular networks.

Table 2.6 - Parameterisations used for propagation models.

Environment	Dense Urban	Urban	Suburban	Rural
Propagation Model	COST231 Walfisch-Ikegami			COST231 Hata
BS Antenna Height (h_{bs}) [m]	30			35
ST Height (h_{st}) [m]	1.8m for indoor and pedestrian, 1.2m for vehicular			
Distance between Buildings (b) [m]	60	70	100	
Buildings Height (h_{roof}) [m]	23		15	
Streets Width (w) [m]	30	35	50	
Morphology [m]				Forest / Semi-Open

Specifically for indoor users, the propagation model is computed for 1.8 m height, but a random floor is assigned to each user (3 m height considered per floor). Based on the floor, a penetration loss is considered for each user, as describe in Annex A.

2.5 802.16e Services and Applications

With the growth of UMTS/HSDPA networks, which allow the combination of voice, video and data services, a lot of studies have been conducted, and the resulting literature is widely available regarding the subject of modelling these services and the applications that can be carried on these networks. However, as far as WiMAX is concerned, the same type of studies is not yet available in large scale. Regarding the present work, studies on traffic modelling is not under the scope of the final goals, thus, existing studies from other systems are used.

WiMAX is flexible enough so that different QoS profiles can be set almost on a per application basis, as long as the higher layers, such as TCP (Transmission Control Protocol) /IP or ATM, can somehow differentiate the different packets generated by the different applications. As explained in Section 2.1, there are two specifications defined within the MAC CS, [IEEE06]: the

ATM and the packet ones. Within the present work, only the packet CS is addressed, due to the growing world wide adoption of IP and the gradual reduction of ATM adoption.

The packet CS is designed to cope with any protocol using packets for transporting data [EMPS06]. Currently, the service flow signalling used for setting up the packet CS supports only Ethernet and IP. However, there are no fundamental limitations preventing additional support of other protocols, such as multiprotocol label switching (MPLS), either as part of the standard or as a vendor extension. Ethernet packet classification, based on, e.g., IEEE 802.1Q Virtual LAN (VLAN) or Type of Service (TOS) (based on Differentiated Services Codepoint (DSCP) marking of IP packets [EMPS06]), can be used to favour certain users against others, or to differentiate the type of application being used, allowing for more demanding applications to get most of the system resources.

802.16e can address a wide range of applications, as summarised in Table 2.7.

Table 2.7 – WiMAX Service Classes (extracted from [WiMA05]).

Class Description	**Application Type**	**Real Time**	**Data Rate**
Interactive Gaming	Interactive Gaming	Yes	50-85 kbps
VoIP, Video Conference	VoIP	Yes	4-64 kbps
	Video Phone		32- 384 kbps
Streaming Media	Music / Speech	Yes	5-128 kbps
	Video Clips		20 - 384 kbps
	Movies Streaming		> 2 Mbps
Information Technology	Instant Messaging	No	< 250 bytes / message
	Web Browsing		> 500 kbps
	E-mail (with attachments)		> 500 kbps
Media Content Download (Store and Forward)	Bulk Data, Movie Download	No	> 1 Mbps
	Peer-to-Peer		> 500 kbps

Due to the fast radio interface capacity to handle balancing between UL and DL, as well as fast changing radio conditions, the 802.16e MAC layer is capable of providing flexible mechanisms that can meet the QoS required for a wide range of services and applications. The 802.16 standard defines the following types of services [NCMP04]:

- Unsolicited Grant Services (UGS), designed to support Constant Bit Rate (CBR) services, characterised by fixed size data packets on a periodic basis, such as CBR voice without silence suppression or E1 emulation. The BS schedules regularly, in a pre-emptive manner, grants of the size defined at connection setup, without an explicit request from the ST, which eliminates the overhead and latency of bandwidth requests in order to meet the delay and jitter requirements of the underlying service.

- RTPS, designed to support real-time services, characterised by periodical variable size data packets that require a guaranteed rate and delay, like video streaming or VoIP with silence suppression. These services are dynamic in nature, but offer periodic dedicated requests opportunities to meet real-time requirements. Since the ST issues explicit requests, the protocol overhead and latency is increased, but capacity is granted only according to the real needs of the connection. Service parameters include the Minimum Reserved Transfer Rate (MRTR) and the Maximum Latency Tolerance (MLT), among others.
- Non-Real-Time Polling Services (NRTPS), targeted to support non-real-time services that require variable size data packets on a regular basis. It is very similar to the real-time polling service, except that connections may use random access transmit opportunities for sending bandwidth requests. These NRTPS services, such as Internet access with a minimum guaranteed rate, are characterised by requiring a guaranteed rate, but can tolerate longer delays and are rather insensitive to jitter. Service parameters include MRTR and Maximum Sustained Transfer Rate (MSTR).
- BE services, typically working in background and normally associated to applications like web browsing, i.e., tailored for services where neither throughput nor delay guarantees are provided. The ST sends requests for bandwidth in either random access slots or dedicated transmission opportunities. The occurrence of dedicated opportunities is subjected to network load and the ST cannot rely on their presence. Service parameters include MSTR.

A service class is assigned to each connection between a ST and a BS. When packets are classified in the CS (Section 2.1), the connection into which they are placed is chosen according to the requirements needed for the specific application that belong to and the QoS guaranteed by that connection. This is achieved through the use of service flows and CID, which are unidirectional flows of packets provided with a specific set of QoS parameters. The QoS parameters associated to each service flow define how scheduling should be performed, as well as link adaptation decisions, and the ultimate goal is to provide proper QoS treatment to the traffic carried across the radio interface, at the same time optimising resources usage. To achieve this, scheduling generally operates at two levels, [NCMP04]:

- Depending on the existing radio interface conditions for each ST, the scheduler must determine the appropriate burst profile for each ST, either increasing or decreasing the coding and modulation through CINR monitorisation.

- Depending on the bandwidth requirements of the individual STs and based on the service classes of the connections and instantaneous status of the traffic queues at both BS and ST, divide the DL and UL sub-frames into zones, accommodating the different STs according to the needs.

Figure 2.11 illustrates these procedures.

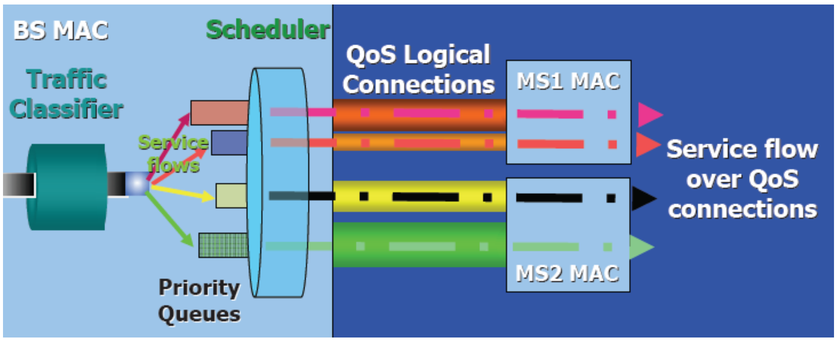


Figure 2.11 – Classification of upper layer packets and CID mapping (extracted from [Alam06]).

The available 802.16e QoS categories and respective specifications are listed in Table 2.8, [WiMA06a].

Table 2.8 - 802.16e QoS Categories and Specifications.

QoS Category	Typical Applications	Qos Specifications
UGS	Constant Bit Rate Speech Constant Bit Rate Video Call	1) Maximum Sustained Rate 2) Maximum Latency Tolerance 3) Jitter Tolerance
RTPS	VoIP Variable Bit Rate Video Call Video Streaming Audio Streaming	1) Minimum Reserved Rate 2) Maximum Sustained Rate 3) Maximum Latency Tolerance 4) Traffic Priority
NRTPS	Web Browsing	1) Minimum Reserved Rate 2) Maximum Sustained Rate 3) Traffic Priority
BE	FTP, E-Mail, Peer-to-Peer	1) Maximum Sustained Rate 2) Traffic Priority

Multicast and Broadcast Services (MBS) are supported by 802.16e as well, which can be done by constructing a specific zone within the DL sub-frame along with the unicast services. In the limit, the whole DL sub-frame can be allocated for stand-alone broadcast services. The DL MAP is used to inform the STs about the location of the MBS zones. A frame containing both unicast and MBS services is illustrated in Figure 2.12.

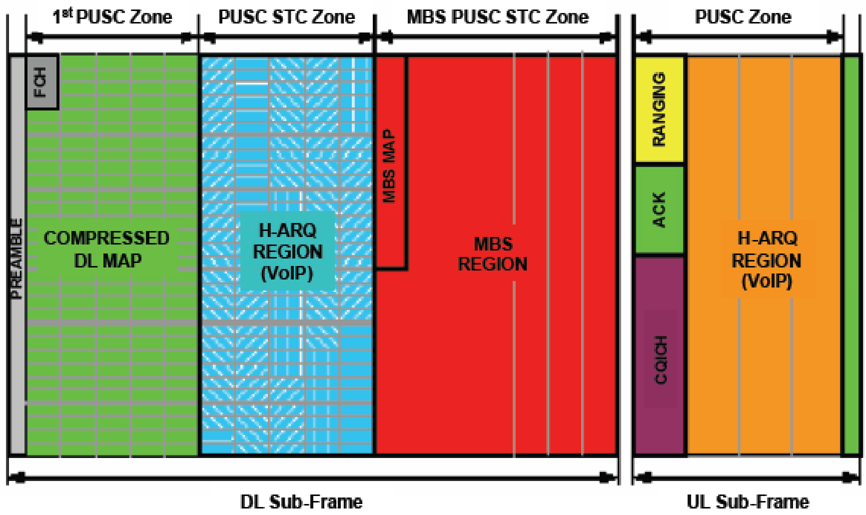


Figure 2.12 – Illustration of TDD frame with embedded MBS support (extracted from [WiMA06a]).

The present thesis considers the following services: VoIP, Video Telephony, Streaming, WWW, FTP and Email, as, beyond being typical services over cellular networks, these are also key applications considered by the AROMA project, [Ljun06]. The related traffic source models are detailed in Annex C.

2.6 Mobility

The 802.16e standard specifies three distinct handover mechanisms [WiMA06a]: Hard Handover (HHO), which is a mandatory feature, Fast BS Switching (FBSS) and Macro Diversity Handover (MDHO), these last two being optional features.

HHO is the most basic handover mechanism, used, e.g., in GSM, in which the link between the BS and the ST can be changed under certain conditions. In this case, the existing radio link is completely replaced by a brand new radio connection, and there may be a cut in the transmission flow, because the switch procedure is not instantaneous.

When FBSS is supported, ST and BS maintain a list of the BSs that are involved in the procedure with the ST, which is called an Active Set (AS), a designation also adopted in UMTS. Among the BSs in the AS, an Anchor BS is defined, which is the only BS with which the ST can communicate for both UL and DL messages, either management or traffic related. The ST continuously monitors the BS in the AS, scanning neighbour BSs and selecting those that are suitable to be included in the AS. The ST reports the selected BS, and the AS update procedure is performed by the BS and ST. The ST continuously monitors the signal strength of the BSs that are in the AS, and selects one BS from the set to be the Anchor BS. The ST reports the selected Anchor BS on CQICH (see Sub-section 2.2.3) or ST initiated handover request message. An important requirement of FBSS is that the data is simultaneously transmitted to all members of an AS of BSs that are able to serve the ST.

For STs and BSs that support MDHO, the ST and the BS maintain an AS of BSs that are involved in MDHO with the ST. Among the BSs in the AS, an Anchor BS is defined. The regular mode of operation refers to a particular case of MDHO with the AS consisting of a single BS. When operating in MDHO, the ST communicates with all BSs in the AS of UL and DL unicast messages and traffic. A MDHO begins when a ST decides to transmit or receive unicast messages and traffic from multiple BSs in the same time interval. For SL MDHO, two or more BSs provide synchronised transmission of ST DL data, such that diversity combining is performed at the ST. For UL MDHO, the transmission from a ST is received by multiple BSs where selection diversity of the information received is performed.

2.7 802.16e vs UMTS/HSxPA and CDMA200/EVDO

It is important to compare 802.16e against other wireless broadband technologies, particularly with CDMA2000/1xEVDO and the UMTS/ HSxPA. Table 2.9 summarises the main differences regarding supported features and specific system characteristics. Naturally, the differences between these systems are far beyond the small list presented but the goal is simply to highlight some of the main features with relevant impact on the radio interface characteristics.

Table 2.9 – System differences between 802.16e, UMTS/HSxPA and CDMA2000/EVDO (extracted from [WiMA06b]).

Technology		CDMA2000/1xEVDO	UMTS/HSxPA	802.16e
Duplex Method		FDD	FDD	TDD
Multiple Access	**DL**	TDM	CDMA-TDM	OFDMA
	UL	CDMA	CDMA	
Channel Bandwidth [MHz]		1.25	5	Scalable (5, 7, ...)
Frame Size [ms]	**DL**	1.67	2	5
	UL	6.67	2, 10	
HARQ		Fast 4-Channel Synchronous IR	Fast 6-Channel Asynchronous CC	Multi-Channel Asynchronous CC
Scheduling		Fast Scheduling in DL		Fast Scheduling in DL and UL
Handover		Virtual Soft Handover	Network Initiated Hard Handover	Network Optimised Hard Handover
Tx Diversity and MIMO		Simple Open Loop Diversity	Simple Open & Closed Loop Diversity	STBC, SM
Beamforming		No	Yes (dedicated pilots)	Yes
Coding	**DL**	Turbo Coding, Rates: 1/3, 1/5	Turbo Coding, Convolutional Coding, Rates: 1/4, 1/2, 3/4, 1	Turbo Coding, Convolutional Coding, Rates: 1/12, 1/8, 1/4, 1/2, 1/3, 2/3, 3/4, 5/6
	UL	Turbo Coding Rates: 1/2, ¼	Turbo Coding, Convolutional Coding, Rates: 2/3, 3/4, 1	Turbo Coding, Convolutional Coding, Rates: 1/12, 1/8, 1/4, 1/2, 1/3, 2/3, 3/4, 5/6 (optional)
Modulation	**DL**	64-QAM (Rev B only), 16-QAM, 8-PSK, QPSK	16-QAM, QPSK	64-QAM, 16-QAM, QPSK
	UL	BPSK (fixed in Rev 0), QPSK, 8-PSK	QPSK, BPSK	64-QAM (optional) 16-QAM, QPSK

Before comparing the theoretical performance of these systems, some comments are in order:

- Both CDMA2000/EVDO and UMTS/HSxPA operate in the 2.0 GHz band, whereas 2.5 GHz was assumed for 802.16e, which brings disadvantages to 802.16e from the propagation in the air interface point of view.

- UMTS/HSxPA operates in two 5 MHz channels due to its FDD nature, while CDMA2000/1xEVDO-Rev A operates in two 1.25 MHz channels. CDMA2000/EVDO-Rev B is also considered assuming an operation with 3 carriers of 5 MHz each, in order to make it comparable with both UMTS/HSxPA and 802.16e from the spectrum allocation point of view. Additionally, like 802.16e, CDMA2000/EVDO Rev B supports 64-QAM. On the contrary, 802.16e operates in a single 10 MHz assuming a TDD implementation.
- A Single Input Multiple Output (SIMO) 1 by 2 implementation is assumed for both UMTS/HSxPA and CDMA2000/EVDO. For 802.16e, a MIMO system 2 by 2 is assumed, although results are also presented for SIMO (expected in the initial deployments of 802.16e [WiMA06b]).
- A 1 by 3 frequency re-use pattern is assumed for all systems.
- A scenario of 19 sites of 3 sectors each is assumed, with 2.8 km distance between sites.

Under these assumptions, the results presented in Table 2.10 are achieved from simulations performed by theWiMAX Forum, [WiMA06b].

Table 2.10 – Performance comparison between 802.16e, CDMA2000/EVDO and UMTS/HSxPA (taken from [WiMA06b]).

Technology		CDMA2000/ 1xEVDO Rev A	CDMA2000/ 3xEVDO Rev B	UMTS/ HSxPA	802.16e SIMO	802.16e MIMO
Spectrum [MHz]	**UL**	1.25	5.00	5.00	DL/UL=3 (TDD)	DL/UL=3 (TDD)
	DL	1.25	5.00	5.00		
	Total	2.50	10.00	10.00	10.00	10.00
Spectrum Efficiency [bps/Hz]	**UL**	0.36	0.28	0.30	0.61	0.84
	DL	0.85	0.93	0.78	1.23	1.91
Net Throughput by Sector [Mbps]	**UL**	0.45	1.39	1.50	1.60	2.20
	DL	1.06	4.65	3.91	9.10	14.10

These results predict that even the first 802.16e deployments based on SIMO will offer performance far above the ones achieved with either CDMA2000/EVDO or UMTS/HSxPA. Using the same bandwidth, 802.16e with SIMO will provide throughputs up to 96% better than the ones given by CDMA2000/3xEVDO Rev B and up to 133% better than UMTS/HSxPA. If MIMO is considered for 802.16e, these improvements are even better. Nevertheless, and since these results are taken from WiMAX Forum sources, some natural restrains shall be taken into consideration as they come from an organisation that backs up the WiMAX technology, thus, they may be over-estimated.

Regarding spectral efficiency, 802.16e advantages are clear as well, with values around 1.91 bps/Hz in DL, being reached in case MIMO is used, and 0.84bps/Hz in UL, against only 0.78 e 0.30 bps/Hz for DL and UL, respectively, regarding UMTS/HSxPA.

These results basically show that the principles on which OFDMA is based on are ahead of the ones in which other systems such as UMTS/HSxPA or CDMA2000/EVDO are based on, i.e., CDMA. In fact, both 3GPP and 3GPP2 are considering to adopt the same technology for its long term evolution, since it has clear advantages for broadband wireless access.

Chapter 3

Simulator Description

This chapter presents a description of the simulator developed for the current thesis, with the objective of allowing the evaluation of 802.16e capability for delivery of different services considering the main concepts within its scheduling functions.

3.1 Overview

The main objective of the present thesis is to study WiMAX 802.16e performance considering the adoption of different scheduling policies for delivery of the usual services provided by cellular wireless networks, particularly voice, video and data services. Within this main idea, a simulator was designed and implemented, the main goal being the development of a platform enabling the following main tasks:

- To implement traffic source models to be used for the main foreseen applications.
- To simulate changing radio channels, by implementing simple mobility models and by allowing the different users to experiment different radio conditions (clutter type, penetration losses, etc.).
- To implement an 802.16e simulator with some of the main concepts behind WiMAX, including FFT size, TDD DL to UL split influence, MAP efficiency, scheduling, etc., and the consequent system performance analysis when varying traffic conditions.

The selected programming language was Visual C++, using object-oriented programming (e.g., each user is an object). The overall program has approximately 8000 code lines, and constitutes a flexible platform for the simulation of different scenarios, while being easily upgradeable for future work.

The simulator is constituted by the following main functional blocks, as illustrated in Figure 3.1:

- Setting of simulator inputs, which include network specific information, characterisation of users, configuration of the general simulation parameters and characterisation of the different services. Due to the object oriente nature of Visual C++, the setting of these inputs is very straight forward.
- User Traffic Generator (UTG) module, which produces the generated traffic based on statistical distributions that take users' and services' characterisation and a time related random seed as inputs. This module produces arrays of users, each with its specific and unique characteristics (e.g., position, distance to BS, etc.). Each array of users includes further arrays of services, again with each service being characterised by specific and unique properties (e.g., VoIP call holding time, WWW session data volume, etc.).
- Radio Channel Simulator (RCS) module, which computes the signal strength and the

interference felt by the different users according to its position, distance to BS, type (indoor, incar or outdoor pedestrian), etc.

- The RRM Module, which is the main module within the simulator, including Admission Control, Congestion Control, Resource Allocation and Traffic Shaping sub-modules. This module takes as input the information produced by the UTG and by the RCS, and, based on these, determines how network resources should be distributed among active users.
- The Performance Analysis module, which basically computes the final figures that summarise how the network behaved according to the outcome of the RRM Module. The outcome of the simulation is dumped to comma separated files, so that easy post-processing can be achieved through standard applications like, e.g., Microsoft Excel.

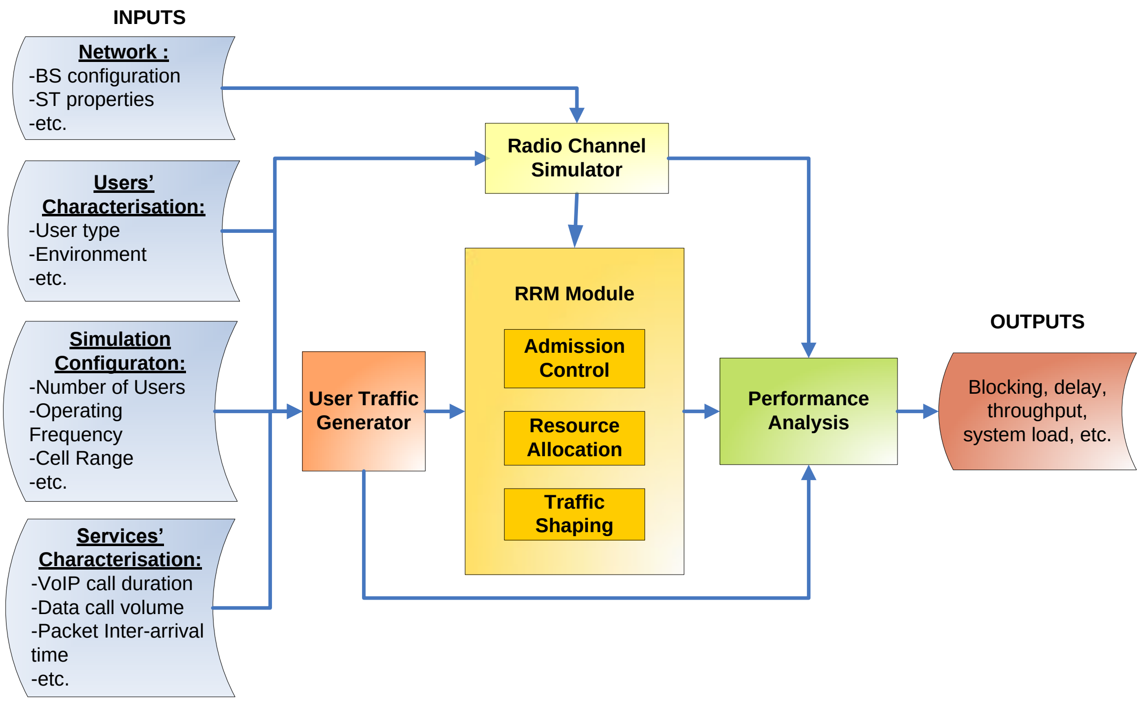


Figure 3.1 – Generic Simulator Structure.

3.2 Simulator Modules

In this section, more detailed descriptions of the simulator models are presented. The UTG module is presented in the first sub-section, followed by a description of the RCS module. Finally, the RRM main module is described.

3.2.1 User Traffic Generator Module

The main function of the UTG module is to generate the unique characteristics of the services associated to each user, based on statistical distributions and on inputs used to characterise these distributions. Six different applications are supported by the simulator, corresponding to those presented and described in Annex C, which can be enabled or disabled for analysis of the impact of different traffic mixtures, by setting the respective penetration rate to zero. Ultimately, the UTG module will associate to each user a given service, properly and uniquely characterised.

The total number of users is defined within the simulation configuration, being an input to the UTG module. Each user is limited to a single service, and to a single call within the simulation. Different arrays of users are defined, one array per each considered service, meaning that there is one array of VoIP users, one for WWW, one for FTP, etc.. Thus, a 'user' is an object that aggregates the main common characteristics to all users. Additionally, 'WWW users', 'VoIP users' and others are defined as being objects that inheretitate from the 'user' object, but that implement specific attributes related to the type of service associated to each user.

Each user is assigned to a user type classification: Business, Small Office Home Office (SOHO) and Mass Market (other user types can easily be added, given the objected-oriented nature of the program). A user type corresponds to a particular user profile, with specific traffic usage characteristics, particularly the type of services used (for instance, a Mass Market user may be more focused on VoIP, while a Business one may use data services more often). The allocation of a specific user type to each user is fixed, and depends on the penetration of each user type, which is a simulator input. In the same way, the assignment of a specific service to a specific user is done through straight simulation inputs.

Each service is characterised by the following main properties:

- For VoIP, Video Telephony and Streaming – the beginning (in seconds) and duration of each session.
- For WWW, FTP and Email applications – the beginning (in seconds) of each session, the number of packet sessions within each session (1 in FTP case), the beginning (in ms) of each packet session, the size (in bytes) of each packet within a packet session and the time between consecutive packets.

Figure 3.2 illustrates the main ideas summarised in the previous paragraphs regarding UTG module behaviour.

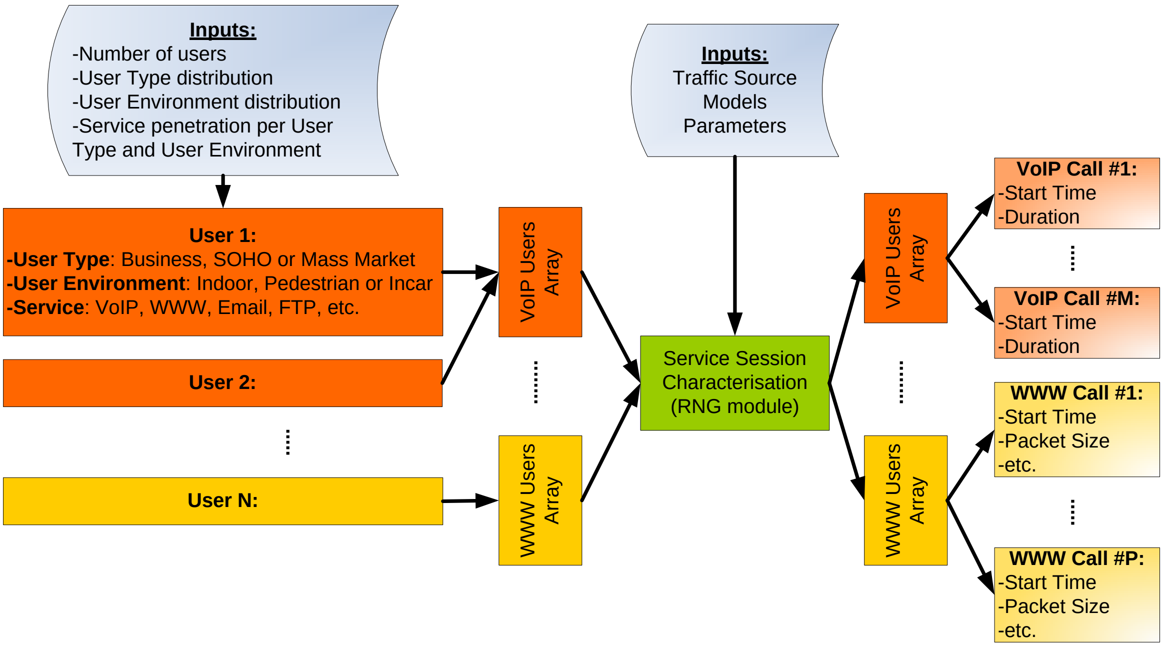


Figure 3.2 – Generic User Traffic Generator Module Structure.

3.2.2 Radio Channel Simulator Module

The function of the RCS module is to simulate a radio channel in DL, in order for each user to experiment specific Radio Signal Strength Indicator (RSSI) and SNR levels as a function of its location and environment. RSSI estimation depends on:

- Configuration of the propagation model. Propagation models employed in the simulator are described in Annex B and they need to be fully characterised in terms of their parameters (BS height, ST height, etc.). For the simulations run under the present thesis, Table 2.6 summarises the configurations adopted for the propagation models, but these can be fully customisable for further simulations.
- The choice between one of the two models depends on the clutter type under analysis, which is an attribute of the BS, and that can be either dense urban, urban, suburban or rural.
- The operating frequency is a simulation parameter, fully customisable, but for the present thesis it is set to 3.5 GHz, as this is the main target band for Western European countries.
- User location is related to the BS, which includes the distance as well as the orientantion.

Distance is used as direct input to propagation model calculations, while the orientantion is used to get the antenna gain applied in link budget calculation. For the latter, an 360 lines by 2 columns long tab separated text file is used as input, with the first column containing the azimuth and the second column the gain related to the maximum antenna gain at the main lobe.

- Penetration losses are considered as being fixed for pedestrian and for incar users, and can both be customisable in the simulation configuration attributes. For the present thesis, they are set to 0 and 5 dB, respectively. For indoor users, penetration losses depend on the floor where they are located. The main assumptions for this calculation are summarised in Annex A, and for the present thesis 20 dB is considered for the ground floor, although this value can be set to any other value.
- Additional gains and losses are considered, as summarised in Annex A for link budget calculations. This includes cable losses, fading, diversity gains, etc..

In the developed simulator, mobility is only considered for the variation of the path loss, due to increasing or decreasing distance between the BS and ST, and for variaton of slow and fast fading. A simple mobility model is implemented, [Corr03], in which user's speed is dictated by a triangular distribution, updated every second.

At the beginning of the simulation, a random initial and final position is generated for every user, inside the cell radius used for the simulation. For every second within the whole simulation time, speed is calculated and each non-indoor user is moved linearly from their position to the final position calculated in the beginning of the simulation. If during the simulation time a user reaches its final position, a new final position is randomly calculated and the procedure goes on in the same way.

In the simulator, variation of fast fading in space depends on the users's speed. It is assumed that fast fading changes every quarter of wavelength, thus, depending on the user type and considering that the granularity of simulation cycles is at frame level (every 5 ms), it takes less or more frames for fast fading to change for each user type. Although indoor users are static, fast fading is always present for such users, due to non-existence of line-of-sight, and because the movements around the ST cause fast fading variations. According to Table 3.1, a new fast fading calculation on a per user basis is performed every 5 frames for pedestrian and indoor users, while for vehicular ones a new calculation is performed every frame.

Table 3.1 – Number of frames needed for fast fading changes per user type.

User Type	Speed [km/h]	Speed [m/s]	Speed [(λ/4) / s]	Speed^{-1} [(λ /4) / ms]$^{-1}$	Speed^{-1} frames / (λ /4)
Pedestrian	3	0.833	38.9	25.7	5.1
Vehicular	50	13.889	648.1	1.5	0.3

Concerning slow fading, a new calculation is performed on a per user basis whenener a user moves 100 m. This is valid for pedestrian and vehicular users, but not for indoor ones, for which it is considered that the shadowing effects are always accounted through the penetration loss. For the pedestrian and vehicular cases, and according to the users' speed, it takes a variable number of frames for each user to move 100 m, Table 3.2.

Table 3.2 - Number of frames needed for slow fading changes per user type.

User Type	Speed [km/h]	Speed [m/s]	Speed^{-1} [(100m) / ms]$^{-1}$	Speed^{-1} frames / (100m)
Pedestrian	3	0.8	24000	24000
Vehicular	50	13.9	7200	1440

Ultimately, a DL RSSI figure is generated for each user, for each of the BS's sector. The user is assigned to the strongest sector, which basically depends on its orientation concerning the BS, since the antenna gain changes accordingly. This impacts on the RRM module, as the user takes resources from the specific sector to which it is anchored to.

For computation of DL CINR, the effect of interference needs to be considered as well, since the adjustment of modulation and coding is based on the quality of the received signal, which depends not only on the own signal path loss, but also on the existing interference. The present thesis considers a simple model for co-channel inteference, in which only the first tier of neighbour sites is considered as being significant for interference calculation. Additionally, a 1 by 3 frequency reuse model is assumed. The effect of the radiation pattern of the antennas of neighbour stations is considered as well: according to the user position within its serving BS coverage area, the received interfering signals are affected by the specific gain of the interfering antennas for the calculated azimuths. Additionally, and since BSs are not transmitting 100% of the time nor all subcarriers are used every instant, a 'load' effect is considered as well: a uniform distribution between 0 an 1 is applied for each of the 6 interfering cells and the effect of each interfering cell is only consider if the outcome of that distribution is between 0.5 and 1. On top of this, noise is added to the interference calculation. The noise factor of the ST is an attribute of the object ST, and can be fully customisable (for the present work a value of 7.0 dB is considered

– Annex A), changing SIR into CINR. At the end, depending on the resulting CINR, an AMC is selected for each user, according to Table 2.4.

The whole DL RSSI and DL CINR calculation process is illustrated in Figure 3.3.

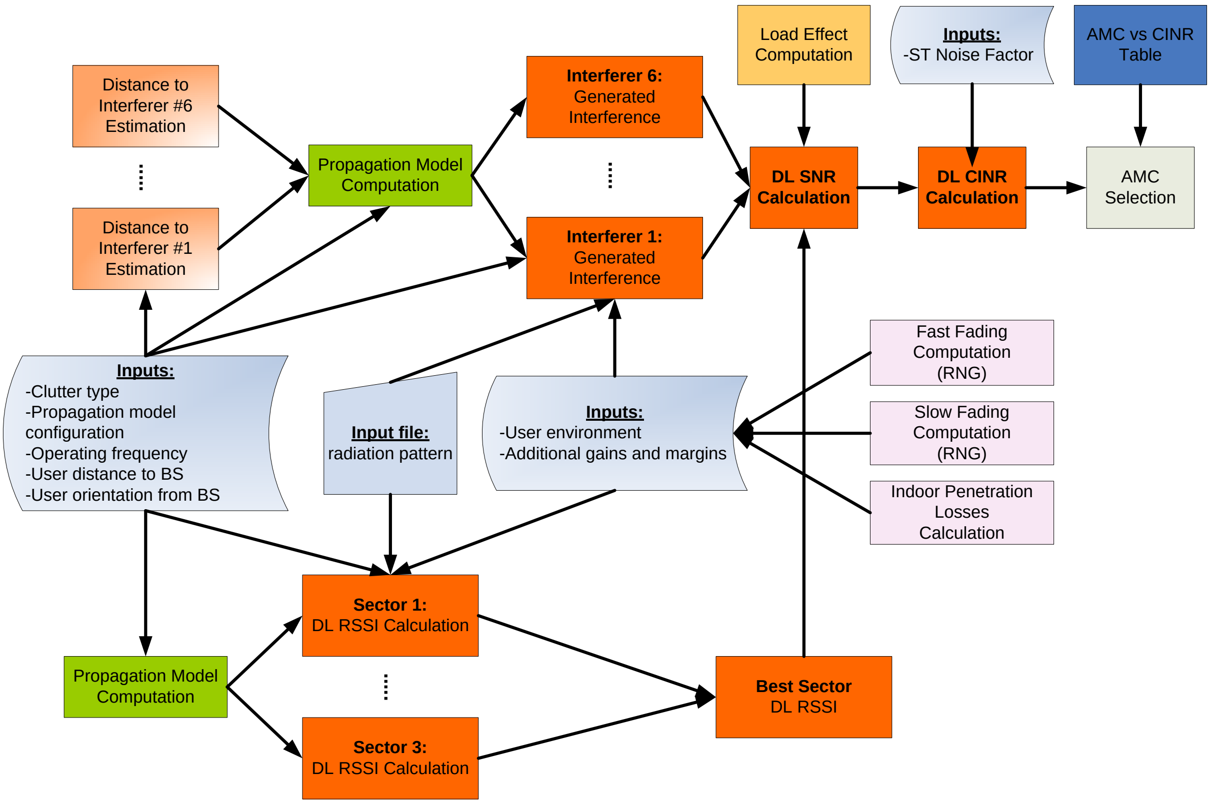


Figure 3.3 - Generic Radio Channel Simulator Module Structure.

3.2.3 Radio Resource Management Module

RRM is the main module of the simulator, as it comprises the main WiMAX concepts. It includes the following functions required at the network layer:

1. Admission Control: considering that any network has a certain capacity in terms of the maximum amount of traffic and services it can handle simultaneously at a given instant, it is important to regulate which users can access the network and under which conditions.
2. Congestion Control: similar functioning as for Admission Control, but applied to on-going connections, discarding packets in case of overload.
3. Traffic Classification: for the network to be able to provide differential treatment based on QoS requirements, it must be able to classify the incoming packets.

4. Shaping and Policy: while the admission control manages how the network controls the admission of new users, the already on-going data flows from active users need to be controlled as well, so that these data flows meet the agreed service specifications. Shaping may, e.g., queue non-compliant packets and delay the release of these into the network. Policing can merely discard or re-classify as low priority the non-conforming packets.
5. Traffic Scheduling: this function is responsible for determining the transmission order of incoming packets, according to the defined QoS requirements, especially when fluctuations in the incoming rates results in traffic queueing.

The general block structure of the RRM module is illustrated in Figure 3.4.

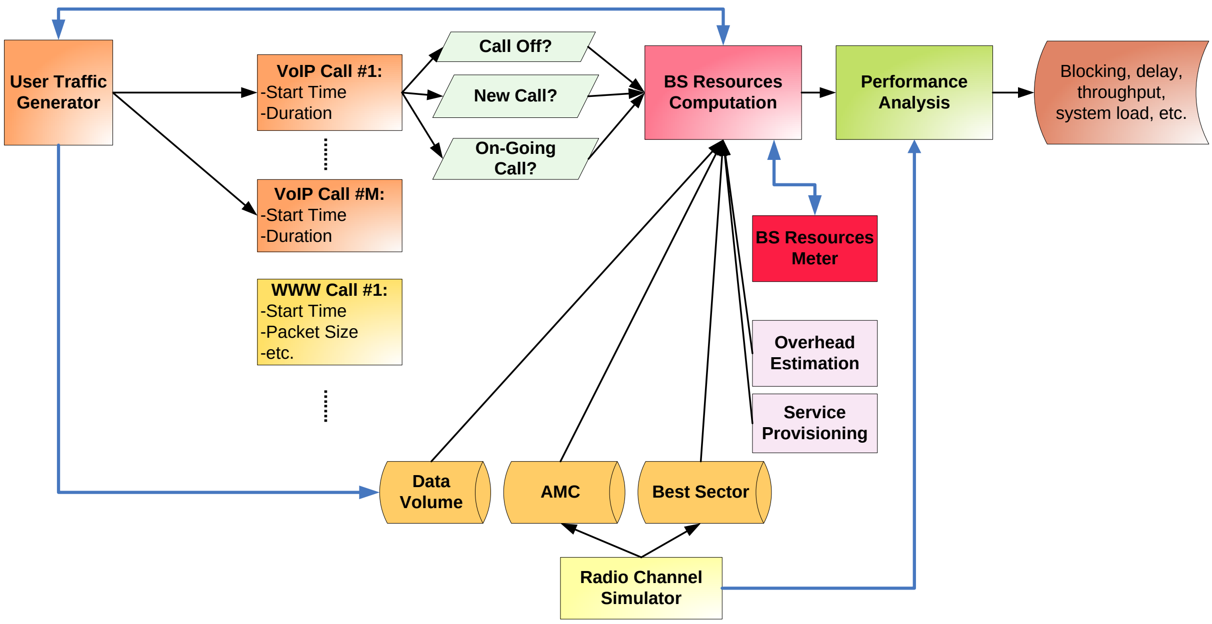


Figure 3.4 – Generic Radio Resource Module structure.

Service Provisioning handles the management of resources as a function of the type of application to be carried by the system. Combinining the services modelled in Annex C with the figures summarised in Table 2.7 and Table 2.8, services characterisation summarised in Table 3.3 are considered under the present thesis.

Table 3.3 - Services characterisation adopted for simulations.

QoS Class	Application	Traffic Priority
RTPS	VoIP	1
RTPS	Video Call	2
BE	Streaming	1
BE	Web Browsing	2
BE	FTP	3
BE	Email	4

Traffic Priority decides upon which application within the same QoS class is address first in case of coexistence of multiple simultaneous applications. There are other QoS parameters in the standard, like maximum latency, tolerated jitter, request and retransmission policy, fixed vs. variable-length SDU indicator, maximum throughput, etc.. However, for simplicity of the present thesis, the simulator only considers the QoS class and traffic priority summarised in Table 3.3, as well as maximum throughput per user, which is limited to 2 Mbps.

When QoS is implemented under the Proportional Fair (PF) scheduling, described later in this section, RTPS services have priority over the BE ones. This means that, when allocating resources to the different services, RTPSs are served first, and only after all RTPSs have the respective instantaneous resources allocated, will BE services be served. This also means that RTPS services will only be blocked at call setup if, within Admission Control limits, system resources are exhausted by other RTPS services. Prioritisation between services belonging to the same QoS Class is handled through the Traffic Priority field, as summarised in Table 3.3.

In every cycle or frame, the available resources are divided among the calls that are active and have packets to be transmitted in that particular cycle. This division is performed depending on the Service Provisioning sub-module, which looks at the QoS classes of the different services when QoS is implemented. As previously described, if a QoS mechanism is implemented, the first services that are looked for are the RTPS ones, and within these, the ones with higher priority. After these services are handled, BE services follow.

The consequence of lack of BS resources to serve a given call has different consequences, depending on whether the call is for an RTPS service or for a BE one. An incoming new RTPS call is blocked in case there are no resources available for its first packet to be transmitted when the call is to be started. This is managed by Admission Control, which works according to Figure 3.5. A threshold is set for VoIP, and another is set for Video Call. A new call is only accepted if the system load at the time instant in which the call is to be set up, plus the additional load introduced by the incoming call, is below the admission control threshold; otherwise the call is blocked.

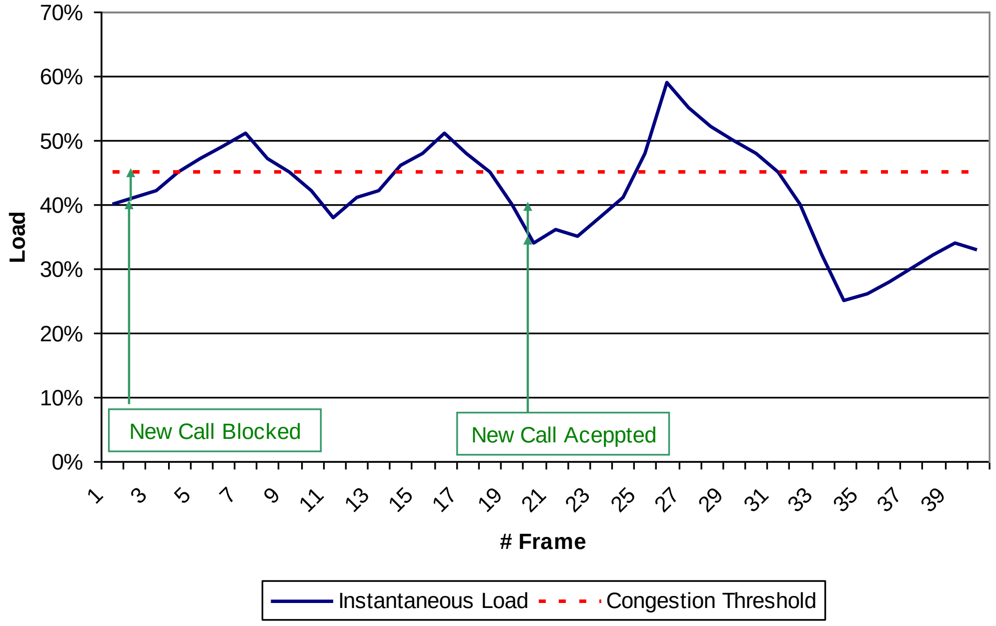


Figure 3.5 - Admission Control function for RTPS calls.

For on-going calls that where already accepted by the network, Congestion Control is used to control availability of resources, Figure 3.6. There is a threshold for VoIP and for Video Call. If the sum of the instantaneous system load plus the additional load needed to transmit a given packet is above the threshold, the packet is discarted and lost. To cope with situations in which too many packets are lost for RTPS services, a criteria is set in order for theese calls to be released in case too many consecutive frames are lost.

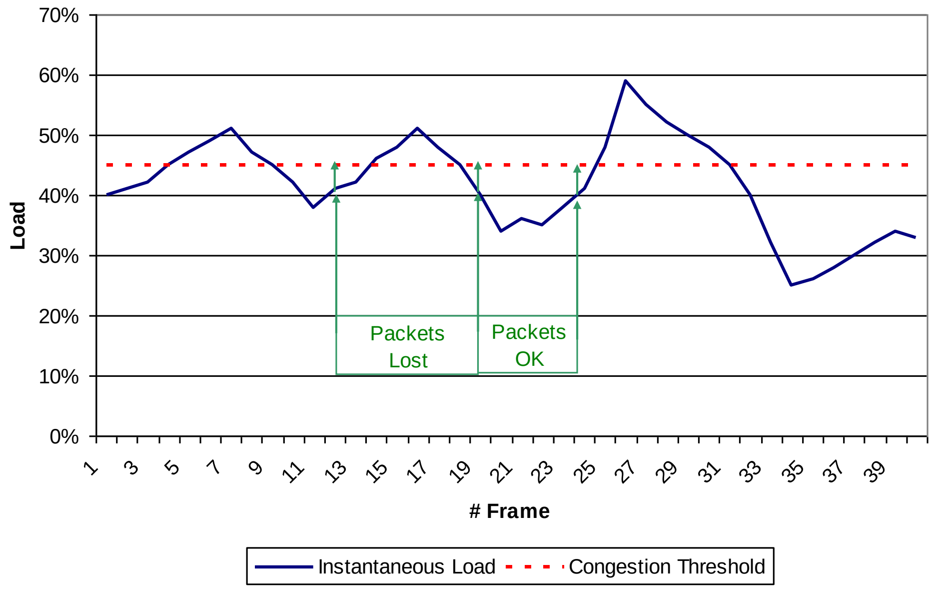


Figure 3.6 – Congestion Control function for RTPS calls.

Concerning BE calls, management is completely different. For, e.g., FTP calls, if at a given moment a given FTP packet cannot be transmitted due to lack of resources, either at session start or when the session is already on-going, the packet is delayed, buffered in the BS and its transmission will be attempted in the next cycle. The delay is measured in these situations but the call is neither blocked nor dropped. When a QoS mechanism is implemented, BE calls can use the remaining free system resources left free by RTPS ones.

As illustrated in Figure 3.1 and in Figure 3.4, one of the main inputs for the RRM module comes from the UTG one. The output from this module is a set of arrays of users, with a specific service being associated to each one. During a simulation, in every cycle, these arrays are analysed in search for new calls to be established, already established calls, and calls that are to be finished. A cycle corresponds to a frame, i.e., 5 ms.

If, in a given cycle, a given call is active and it has a specific packet to be transmitted on that cycle (which depends on the packet interarrival time, another output of the UTG module), the BS Resources Computation algorithm computes the necessary BS resources needed to transmit that packet. This depends on further inputs, particularly the packet volume (given by the UTG module), the AMC and the best sector used by the ST (which are outputs of the RCS module. Additionally, it interacts with two further algorithms: the Service Provisioning, which basically defines the minimum and the maximum number of resources allocated for each service, regardless of the packet size (Table 3.3); and the Overhead Estimation algorithm, which estimated the resources that are needed for MAP overhead purposes, and thus cannot be assigned for user data transmission. The BS Resources Computation algorithm interacts with the BS Resources Meter, which keeps track on the BS's resources. The result of the BS Resources Computaton algorithm is further analysed and the consequent output is stored in the output files described, in Sub-section 3.3.

For the BS Resources Computation algorithm to work properly, one of the critical inputs is the useful capacity of a single subcarrier combined with a single symbol, which depends on aspects such as the modulation, overall coding rate, symbol duration and guard period. It is given by (3.1), [IEEE06].

$$Rb_{BS} = b_s / \left[\frac{N_{FFT}}{B_W \times n_f} \times \left(1 + T_g\right) \right] \tag{3.1}$$

Where:

- B_W stands for the channel bandwidth.

- b_s stands for the uncoded block size per symbol, as given by Table 3.4.

Within the present thesis, N_{FFT} is restricted to 512 or 1024 (B_W of 5 and 10 MHz, respectively), as these are the main values considered in the initial WiMAX certification. The assumed T_g is 1/8, which represents a tradeoff between the highest value of ¼, which covers the expected worst case delay spread, and a lower value, which reduces guard time overhead. Regarding coding, the present work only considers CC, since this is the only mandatory coding in the standard, [IEEE06]. Table 3.4 summarises the uncoded and coded block size per symbol/subcarrier. As it can be seen, higher order modulations like 64-QAM allow the delivery for more useful bits in the same frequency/time resources:

Table 3.4 – Mandatory Channel Coding per Modulation

Modulation	**Uncoded Block Size per Symbol/Subcarrier [bits]**	**Coded Block Size per Symbol/Subcarrier [bits]**	**Overall Coding Rate**
QPSK	1	2	1/2
QPSK	1.5	2	3/4
16-QAM	2	4	1/2
16-QAM	3	4	3/4
64-QAM	4	6	2/3
64-QAM	4.5	6	3/4

With these assumptions, the combination of Table 3.4 with equation (3.1) results in the per slot capacity summarised in Table 3.5 (assuming PUSC), noticing that a slot spams over 2 symbols in time domain and over 24 subcarriers in the frequency domain:

Table 3.5 – Per slot capacity (PUSC).

Modulation	**Overall Coding Rate**	**Slot Capacity [bits]**
QPSK	½	48
QPSK	¾	72
16-QAM	½	96
16-QAM	¾	144
64-QAM	⅔	192
64-QAM	¾	216

As clearly expressed in Table 3.5, the capacity assigned to each ST per frame depends on the used channel coding and modulation, and on the number of slots allocated to the user. As an example, an active FTP call for which a 1500 bytes (12000 bits) long packet needs to be transmitted in a given moment requires 84 slots in case 16-QAM ¾ AMC is used by the respective ST. If the AMC is worse, e.g., QPSK ½, more slots are needed for the same packet volume, quickly exhausting system resources.

Another critical input for the BS Resources Computation algorithm is the number of free slots within the BS. The number of available slots for allocation per frame per user depends on the channel bandwidth (Table 2.2), on frame overhead, and on the actual needs of a given user within the radio frame specific time instant. Thus, DL slot allocation for different users in every DL subframe must be done with consideration for several aspects:

- The existing overhead limits, on a frame by frame basis, the number of available symbols in the time domain.
- The minimum frequency domain unit in DL PUSC is one slot, which corresponds to one subchannel by two symbols.
- The DL slots must fit within the DL subframe so that capacity is maximised. These cannot overlap with the preamble, the FCH and the MAP fields.
- QoS parameters must be fulfilled, i.e., a user cannot have a slot allocation that surpasses the provisioned QoS parameters.

The useful capacity for DL of a single radio frame depends on the DL/UL TDD split, and on control fields per frame. Due to the TDD nature of the system, a subset of the available symbols is reserved for the UL subframe, and it cannot be used for DL bursts. Under the present work, a 2:1 ratio is considered for this spit (typical current assymetry factor for cellular operators), resulting in 32 symbols available for DL and 16 for UL. Other options would be, e.g., to consider a 3:1 split, assigning less weight to UL in favour of DL, more suited for networks dominated by WWW like traffic, or, on the contrary, a 1:1 split, more suited for symmetric applications, like VoIP or Peer-to-Peer.

Concerning allocation of slots by multiple users, WiMAX presents a strong challenge of resolving a complex 2-dimension problem of addressing users in both frequency and time domains. It gets even more complex by considering the fact that, within the same radio frame, the same user can experience different radio channel conditions on the different subcarriers, meaning that non-adjacent slots may provide the best allocation for a given user, at the cost of increased DL-MAP overhead. The challenge here is to provide that the BS is aware of the different channel conditions, experienced throughout the whole range of subcarriers for the different users, and that complex calculations can be performed by the BS, in order to find the best commitment between overhead increase and non-adjacent slot allocation for the different users.

For simplicity, the present work only considers a 1-dimension problem. This is achieved by replicating the time domain on the single frequency domain, translating two-dimension allocation arrays of [$N_{subchannels}$, $N_{symbols}$] into single-dimension allocation arrays of $N_{subchannels}$ x $N_{symbols}$, where $N_{symbols}$ stands for the number of symbols per frame. This is illustrated in Figure 3.7. With this approach, some additional assumptions are considered as well:

- It is assumed that the user experiments the same radio channel conditions for all the subcarriers of the radio frame, and throughout the whole frame duration. This aprroach does not reflect the frequency selective nature of the radio channel, thus, it does not allow to take the full advantage of an eventual frequency selective scheduling, resulting in a system performance below the ideal one.
- As this approach makes it easier to fit more users in the same radio frame, the final results can be too optimistic. To cope with this, it is assumed that 1 subchannel is always wasted, as well as 2 DL symbols.

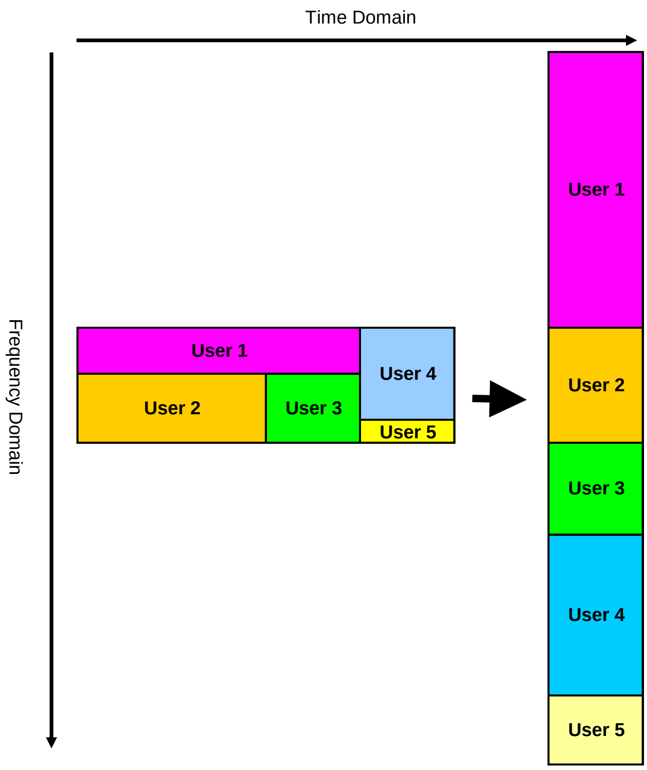


Figure 3.7 – Translation of 2-Dimensional Time and Frequency Domains to Frequency Domain Only.

With this simplification, the present simulator considers:

- From the available 32 symbols for DL, only 30 can be used, meaning that up to 15 users can be time multiplexed on the same subchannel.

- From the available 15 (30) $N_{subchannels}$ for N_{FFT} of 512 (1024), only 14 (29) can be used.

The combination of these into a 1-dimension problem results in the maximum number of available slots for DL transmission per radio frame, $N_{MaxSlots}$, summarised in Table 3.6.

Table 3.6 – Available DL slots per radio frame.

Channel Bandwidth [MHz]	5	10
$\boldsymbol{N_{MaxSlots}}$	210	435

As previously stated, an Overhead Estimation algorithm exists as well, in order to compute how resources are affected by the MAP overhead. The per frame capacity summarised in Table 3.6 still includes overhead, which must then be subtracted. Control fields like the preamble, FCH, DL and UL MAP, ACK-CH, Ranging and CQICH, as illustrated in Figure 2.9, are responsible for this overhead. Some of these fields have fixed size, while others have variable size, like the case of the DL-MAP message. The simulator adopts a simplified approach for overhead considerations:

- All the subchannels within the first symbol of the DL subframe are always assigned for the preamble, meaning that within the 30 available symbols per 5 ms frame, 1 is always used for the preamble (which includes all subcarriers), thus, resulting in a reduction of 15 or 30 slots, depending on if N_{FFT} is 512 or 1024.
- The FCH always takes the first four subchannels of the second and third DL subframe symbols, occupying 8 slots.
- Each DL-MAP message contains fixed and single fields, occupying fixed 104 bits (208 bits with repetition factor of 2) plus additional 60 bits for each addressed ST (120 bits with repetion factor of 2). As DL-MAP is QPSK ½ coded, according to Table 3.5, the number of slots needed for DL-MAP is given by:

$$N_{DL-MAP} = \left\lceil \frac{208 + N_u \times 120}{48} \right\rceil \tag{3.2}$$

  Where:

  - $N_{DL\text{-}MAP}$ is the number of allocated slots for DL-MAP.
  - N_u the number of users addressed in the radio frame.

- Although the present thesis is only concerned with DL performance, the UL-MAP message also needs to be considered for overhead calculation in the DL subframe. Each

UL-MAP takes a single 64 bits block (128 with repetion factor of 2, plus a variable block size per address ST, depending on the goal of the message. For the present work, 52 bits (104 with repetion factor of two) is considered for each assigned ST, as this is the value that covers most of the situations. The full details for UL-MAP message construction are available in [IEEE06].

$$N_{UL-MAP} = \left\lceil \frac{128 + N_u \times 104}{48} \right\rceil \tag{3.3}$$

- 1 symbol is also 'wasted', because it is needed for the TTG, but this is taken from UL allocation and thus does not affect available slots for DL.

Summarising, the number of slots consumed by overhead is given by (3.4).

$$N_{OHSlots} = 23 + \left\lceil \frac{336 + N_u \times 224}{48} \right\rceil \tag{3.4}$$

As an example, if only 1 user is to be addressed in a given frame, and if N_{FFT} = 512, 12 symbols are occupied by MAP overhead, meaning that 35 slots are consumed by overhead, thus, leaving only 175 slots free for user data.

Finally the Scheduling algorithm determines which users are allocated to the available resources, considering all the inputs described previously. Different approaches may be implemented for this purpose. Within the present work, three different algorithms are implemented in order to treat differently the co-existence of a service mix and the different radio conditions felt by each user:

- Round-Robin (RR).
- Maximum Throughput (MAX)
- PF.

RR is a simple and easy to implement algorithm that handles users alternatively, without concern for the particular radio conditions felt by each user. Thus, most likely, e.g., a VoIP user that suffered a packet lost in a frame will successfully be allocated system resources in the next frame. In the same way, a WWW user whose last packet was delayed will most likely have a successful allocation on the next frame. This algorithm is simple, but it does not maximise system capacity, as it does not give preference to users feeling the best radio conditions at each frame. Additionally, as in each frame the different users are prioritised in a sorted order, there is no QoS implemented, e.g., VoIP services are not prioritised over FTP ones. Consequently, neither Admission Control nor Congestion Control are applied for load control due to RTPS services. Figure 3.8 illustrates how RR is implemented (the illustrated services are only an example, and only represent a given cycle: in another cycle, other users would be served in first hand, following the round-robin mechanism).

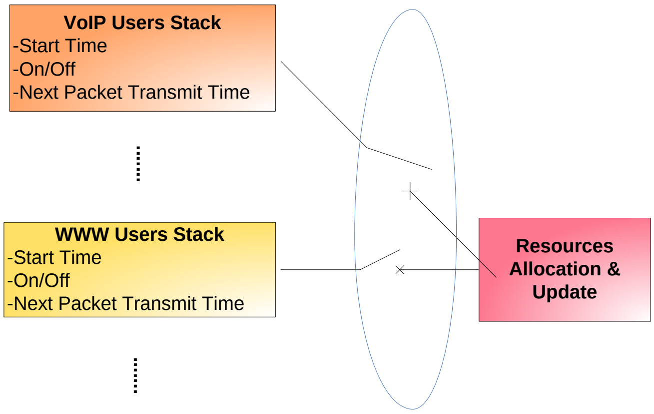


Figure 3.8 – RR scheduling mechanism illustration.

MAX takes advantage of prioritising users that experiment a better AMC at that particular moment. The algorithm searches for those users and attempts to serve them in first hand, without concern for the different QoS priorities among services (for instance, a VoIP user suffering a bad AMC is served later than any WWW user). This algorithm maximises system capacity, but it can lead to extreme situations in which users that are constantly affected by bad radio conditions may have to wait a long time before successfully being allocated resources, resulting in unfair situations. Additionally, there are no QoS guarantees for any user, either RTPS or BE based, meaning that again neither Admission Control nor Congestion Control is applied for RTPS calls. Figure 3.9 illustrates how the MAX scheduling is implemented (again, the indication of the users's application is simply an example and could be different).

PF, like MAX, looks for users experiencing the best radio conditions, but in order to avoid that users that experiment long-term worse radio conditions remain long periods without any assignment, the algorithm keeps track of the recent delays or packet losses felt by each user, and overrides the search for the better AMC in such situations, providing some fairness to the system. The simulator has an input that specifically determines from which cumulated delay these users are served in first hand, and this input is different for RTPS and BE services. Within the BE class, Streaming may be differentiated from FTP, WWW and Email as well. This algorithm also implements QoS mechanisms, i.e., RTPS services, such as VoIP or Video Call, are served before BE ones, regardless of their AMC, and controlled by both Admission Control and

Congestion Control. Figure 3.10 illustrates how PF scheduling is implemented.

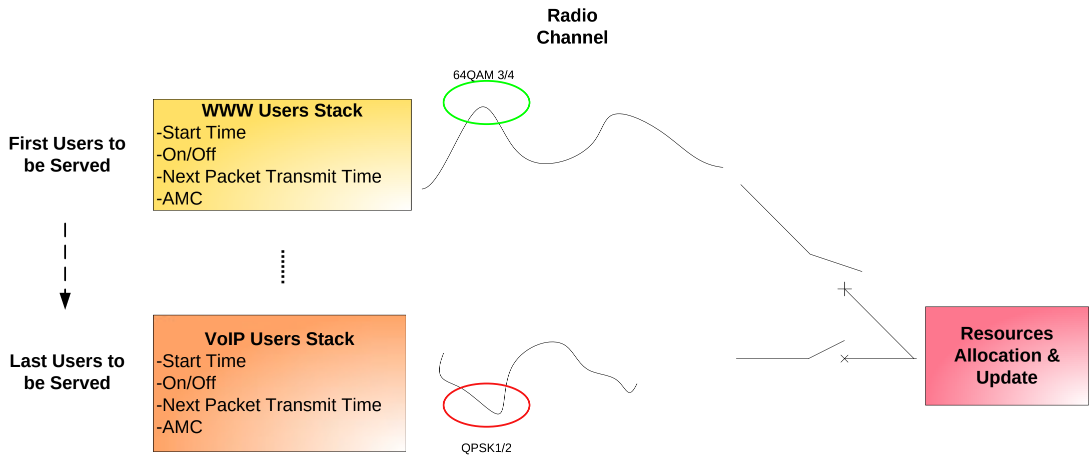


Figure 3.9 - MAX scheduling mechanism illustration.

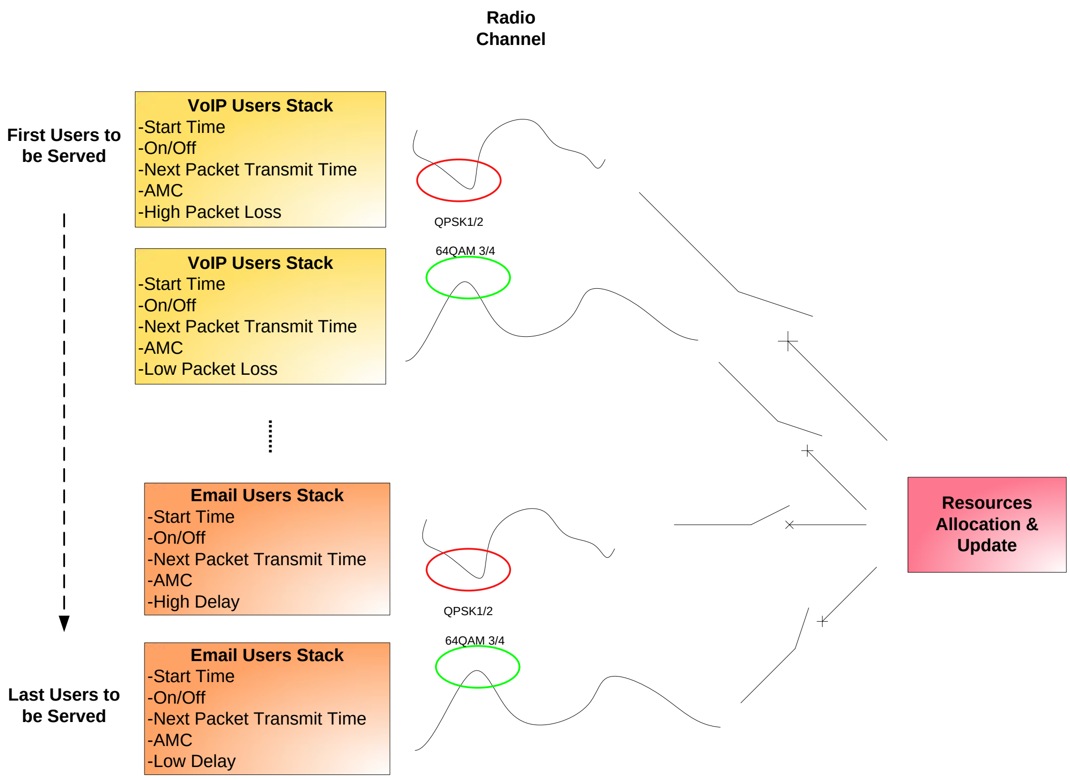


Figure 3.10 - PF scheduling mechanism illustration.

3.3 Input and Output Parameters

In this section, the main simulator's inputs and outputs are introduced.

3.3.1 Simulator Inputs

The input files of the simulator are comma separated type ones, thus, easily treated in common applications, such as, e.g., Microsoft's Excel. Log files format and examples are shown in Annex E.

A generic System Configuration file is used as input to configure general simulator attributes. These include information concerning number of users to consider, clutter type for simulation, cell radius, operating frequency, channel bandwidth, link budget generic information (including BS transmit power, cable losses, additional gains provided by, e.g., power boosters, etc.), and the CINR thresholds needed for a given AMC to be used by the BS for a specific ST.

One of the files concerns the different services. This file summarises the main statistics that are used as inputs for generation of the different services' traffic profiles, as detailed in Annex C. Generation of traffic profiles is based on Random Number Generators (RNGs), whose validation is illustrated in Annex D.

Another input file concerns users' characterisation. It includes generic information on users' mobility per environment, penetration losses affecting the different user types (indoor, incar, etc.), and the way how users are distributed within the different segments, different environments and the usage of the different services.

Propagation models are characterised through another input file, which basically summarises the parameters to be used by the two propagation models implemented in the simulator, described in Annex B.

Different antenna beam widths can be considered for coverage performance. Although the maximum gain value is part of the System Configuration input file, there is a specific Antenna Radiation Pattern input file that summarises, for each azimuth (0º to 359º), the attenuation from the maximum gain value, defined as a positive value.

3.3.2 Simulator Outputs

The output files of the simulator are generated in order to store useful data needed to assess system performance. The generated files are also comma separated type files. Log files format and examples are shown in Annex F.

Logs are generated on a per application basis, in order to store all the traffic-generated patterns, application performance and impact of each session on system resources. The objective of these files is to easily post-process the main indicators associated to each service type in 3rd party tools, particularly to assess system performance and system response to different applications in different conditions. According to each service own characteristics, each per application generated file is compliantly unique: for instance, the VoIP session summary does not concern with application throughput but rather with call blocking, while, e.g., FTP session summary directly addresses application throughput and delay.

Two files are generated concerning BS specific indicators: one that saves the BS information periodically (every 1 second) and a second that includes the simulation aggregation counters. The first is a log file that tracks that registers the number of served users per application, the resources used per application, the slots needed for overhead, the system load and the instantaneous throughput due to each application. The BS counters file is a summary file that stores statistical information about some of the quantities that are detailed in the log file, particularly the ones regarding system resources used. Additionally, the BS counters file summarises statistical information on the radio channel, particularly RSSI and CINR.

A Key Performance Indicators (KPI) summary file is produced as well, summarising the main statistics on the performance of each application, including blocking of VoIP calls, throughput of BE kind applications (e.g., FTP), the traffic volume carried (discriminated by service), the number of sessions attempted for each service, and information concerning delay of BE kind applications.

A final file is produced summarising only the start and end time of the simulation.

3.3.3 Main Performance Indicators

Some key concepts concerning the overall performance are issued by the simulator. Network load stands for the use of system resources, being defined in (3.5).

$$S_L = \frac{N_{OHSlots} + N_{DataSlots}}{N_{MaxSlots}} \tag{3.5}$$

Where:

- $N_{DataSlots}$ stands for the number of slots busy for user data transmission.
- $N_{OHSlots}$ is introduced in (3.4).
- $N_{MaxSlots}$ is introduced in Table 3.6.

MAP efficiency (η) stands for the percentage of resources available for user data transmission only, being given by (3.6).

$$\eta = 1 - \frac{N_{OHSlots}}{N_{MaxSlots}} \tag{3.6}$$

RTPS Refused Call Rate and Drop Call Rate are defined in (3.7) and (3.8), respectively.

$$R_{RC} = \frac{N_{BC}}{N_{CA}} \tag{3.7}$$

$$R_{DC} = \frac{N_{DC}}{N_{CA} - N_{BC}} \tag{3.8}$$

Where:

- N_{BC} stands for the number of calls blocked due to no resources available for transmission of the first packet.
- N_{CA} stands for the number of call attempts.
- N_{DC} stands for the number of calls dropped due to reaching a consecutive number of lost packets, as described in Sub-section 3.2.3.

RTPS percentage of lost packets due to lack of resources is given by (3.9).

$$P_L = \frac{N_{LRTPSP}}{N_{TRTPSP}} \tag{3.9}$$

Where:

- N_{LRTPSP} stands for the number of RTPS packets that are lost.
- N_{TRTPSP} stands for the total number of RTPS packets.

Percentage of BE delayed packets is defined in (3.10). The average delay of delayed packets is defined as being the number of frames for which the packet is not successfully transmitted, due to lack of resources.

$$P_D = \frac{N_{DBEP}}{N_d} \tag{3.10}$$

Where N_{DBEP} stands for the number of packets that were not transmisted at the first attempt, due to lack of resources.

The overall throughput of a single BE session is given by (3.11)

$$Rb_{BE\,[kbps]} = \frac{N_d \times S_{d\,[bits]}}{5_{[ms]}} \tag{3.11}$$

Alternatively to (3.1), the instantaneous throughput of the BS can be obtained from (3.12), where the first sum concerns the 6 aplications considered for the simulator and the second sum concerns the number of users served at a given radio frame for each applications.

$$Rb_{BS\,[kbps]} = \frac{\sum_{i=1}^{6}\sum_{n=1}^{N_u} S_{d\,[bits]}}{5_{[ms]}} \tag{3.12}$$

3.4 Simulator Assessment

The developed simulator was previously assessed through dedicated debugging tests. The following modules were properly validated:

- Use of propagation models for link budget computation, including distance between BS and ST, influence of antenna radiating pattern, indoor penetration losses, fading, etc..
- Mobility module.
- Traffic sources, including validation of RNGs (Annex D), influence of traffic mix, number of users, etc..
- Scheduling algorithms, including MAP overhead calculations, correct assignment of resources to users as a function of the type of traffic, scheduling approach, instantaneous AMC, etc.
- Alignment of generated output files with simulation results.

As a basis for the implementation of each service, the following different RNGs were used: Uniform, Beta, Gamma, Exponential, Geometric, Normal, LogNormal, Pareto and Poisson. A validation of each of these RNGs is presented in Annex D. In this annex, it can be confirmed that the correlation between the theoretical and the generated Cumulative Distribution Functions

(CDF) for the different statistical distributions used in the simulator is pratically perfect, as the correlation figures given by Microsoft's Excel CORREL function, are very close to 100%.

Another required test concerns the assessment of the time interval to consider in order to simulate the typical 60 minutes period, known as the 'busy hour' period. Within this scope, simulations of 70 minutes where conducted, as the initial 10 minutes period must be neglected. This happens because initially, when the simulator is started, it is not stabilised yet. It is necessary to check how many of the initial minutes must be discarded. To perform this assessment, the Reference scenario (REF) summarised in Table 4.1 is assumed, for a suburban environment (Table 4.2 and Table 4.3), assuming 2000 users served during the simulation interval. As illustrated in Figure 3.11, roughly the first 300 s of simulations are characterised by a continuous growth of the system load towards its average simulation value. However, after this initial period, system load changes around its average value, meaning that these 300 s may be enough as guard period.

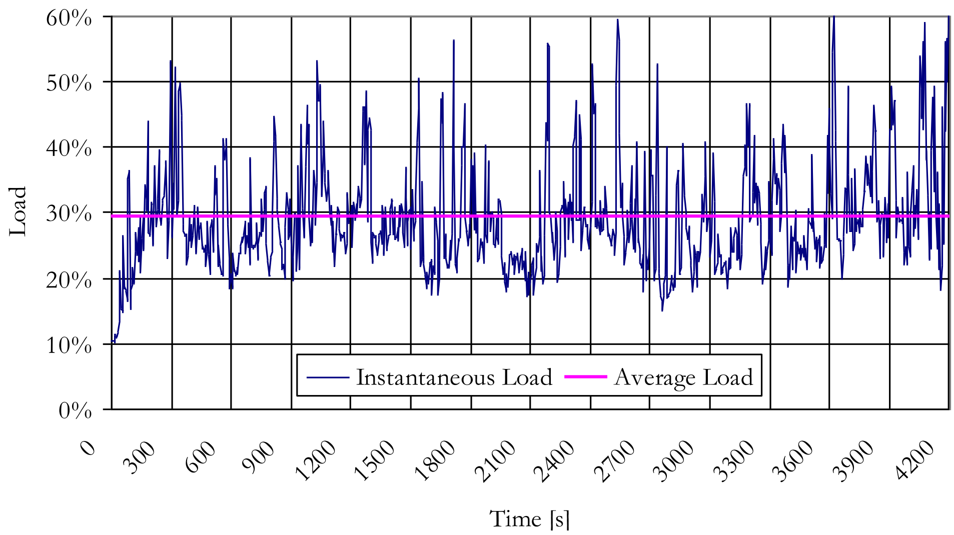


Figure 3.11 – System Load Variation in Time.

For this particular example, the average system load when considering only the 'useful' period of simulation, for a 'useful' simulation time of 60 minutes and with varying 'guard periods' of 0, 5 and 10 minutes, one should note that the final average system load only changes marginally when considering either a 5 or a 10 minutes guard time, as illustrated in Figure 3.12. Since a guard time of 10 minutes would imply the need for longer simulation periods, which would be time costly, it was decided to use a 5 minutes guard time only, resulting in a total of 65 minutes per simulation.

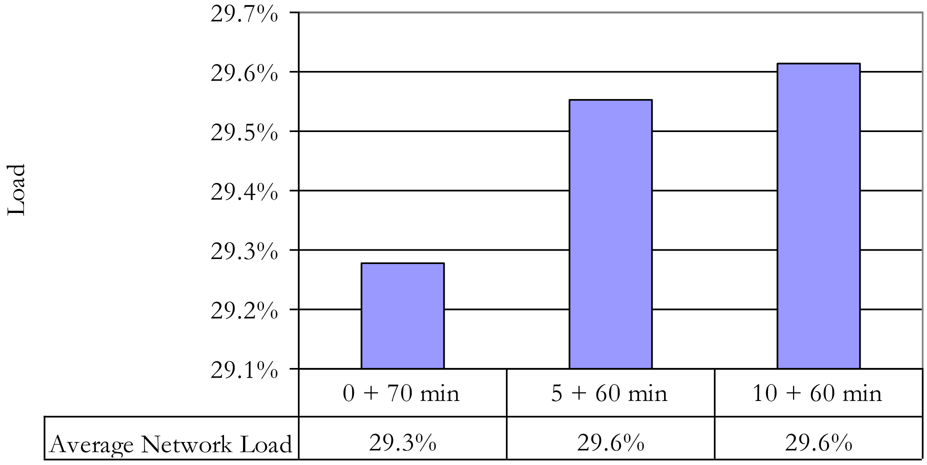


Figure 3.12 – Impact of guard time in total simulation time on average system load.

It is also necessary to assess the number of simulations required to have enough convergence of results. This is achieved when the standard deviation is minimum, and when a minimum average value fluctuation is achieved. This test was performed assuming the same scenario as before.

Figure 3.13 illustrates both average load and average delay of delayed BE packets (other quantities could have been studied, e.g., system throughput, but all these are related and results can be extrapolated). It is easily seen that although the values fluctuate around an average value, there are important variations that must be accounted for, thus, the final figures must be computed over a set of simulations.

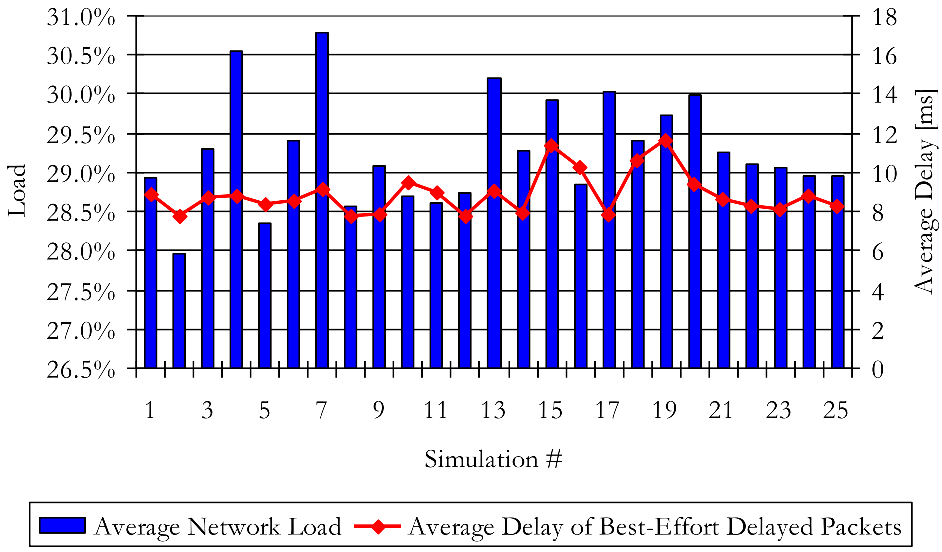
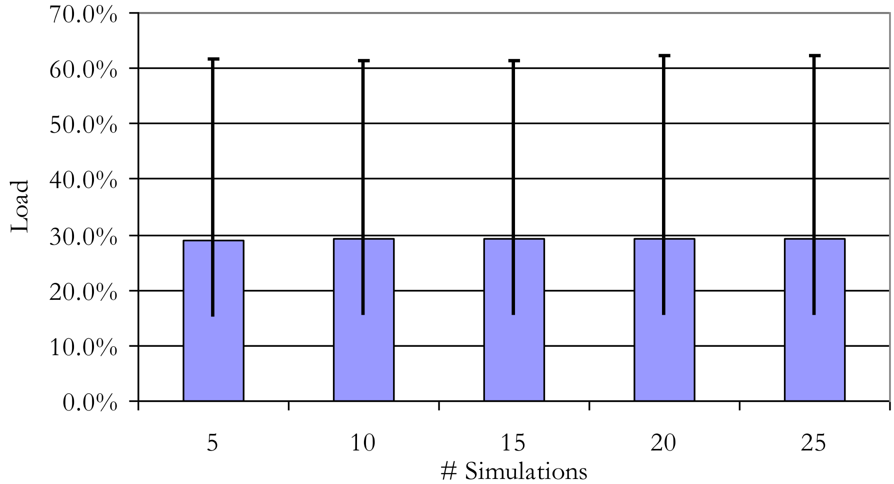
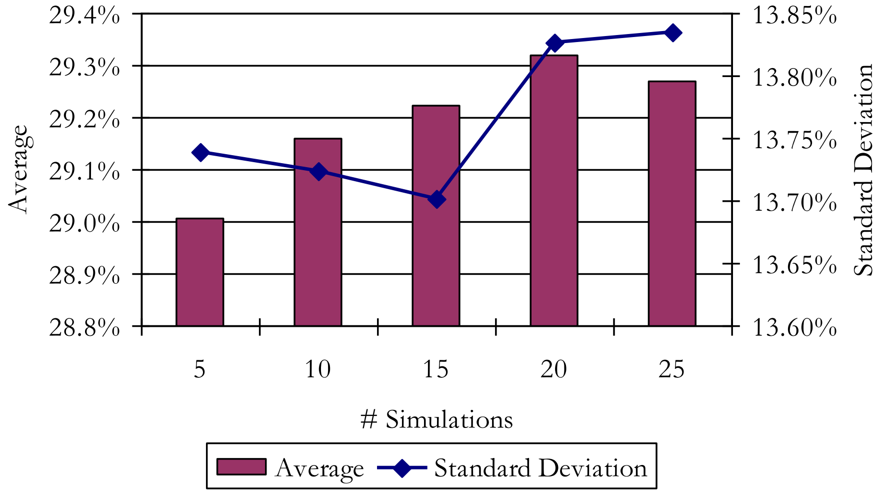


Figure 3.13 – Load and Average Delay for BE Services – 25 simulations.

Considering only network load, after aggregation of 5, 10, 15, 20 and 25 simulations, the average standard deviation as well as the 90% confidence interval are compared, in order to assess convergence results. From Figure 3.14, it is clear that the fluctuation for both average and standard deviation are marginal, even when comparing 5 to 25 simulations. When considering the confidence interval, no visible variations are also seen for the different number of simulations considered. Without penalty for the convergence of the results, 10 simulations were considered per scenario.



a) Average and 90% confidence interval



b) Average and standard deviation

Figure 3.14 – Convergence of system load values.

Chapter 4

Analysis of Results

This chapter summarises the main results of the simulations performed. An analysis of results for the different scenarios is done. An introduction to simulation scenarios is presented initially, followed by results obtained for the pre-defined reference scenario. A sensitivity analysis on results is addressed when the scenario conditions are changed.

4.1 Simulation Scenarios

In order to analyse the different system responses when different traffic mixes are offered to the network, several scenarios were conceived in order to fully understand the main system capabilities and issues according to the different services characteristics.

WiMAX is clearly designed to support data services with performances that are typically associated to 'fixed' technologies, such as DSL or cable, likely outperforming 3GPP standards for this purpose. However, as in any other technology, data services are typically more demanding than regular speech ones, due to which different traffic mixes must be analysed, particularly when the service distribution changes from a speech dominant scenario to a data one. Inputs from the AROMA project [Ljun06] are taken for this purpose.

Two classes of users are considered for each service: consumer (associated to simulator's Mass Market class) and business classes (associated to simulator's Business class), each indicating different traffic and service patterns. Although the simulator provides also capability for a 3rd class (SOHO), this was not considered for the presented simulations and respective results. The total traffic and users are depending on the type of environment (dense urban, urban, suburban and hotspot), as well as on the balance between user classes. The detailed service mix per user class is summarised in Table 4.1, where reference figures for the reference (REF) scenario are taken from [Ljun06], while for the Speech Centric scenario (SPC), an increase of 15% in speech service weight is assumed, the remaining services being reduced evenly. For the Data Centric scenario (DAC), a reduction of 15% in speech service weight from the REF scenario is assumed, the remaining services being increased evenly.

The distribution between Mass Market and Business Users per radio environment is summarised in Table 4.2. These inputs are also taken from [Ljun06], except for the Rural environment, as the AROMA project does not consider this environment. Due to the lack of figures for such environment, the trend concerning the reduction of the Business class (and consequent increase of Mass Market class) when moving from higher density environments, like Dense Urban, to lower density ones, like Suburban, is assumed.

Table 4.1 – Service distribution per user class for the different scenarios.

Service	Service Distribution					
	Mass Market			Business		
	REF	SPC	DAC	REF	SPC	DAC
VoIP	60.0%	75.0%	45.0%	50.0%	65.0%	35.0%
Video Telephony	10.0%	6.3%	13.8%	10.0%	7.0%	13.0%
Streaming	4.0%	2.5%	5.5%	4.0%	2.8%	5.2%
FTP	3.0%	1.9%	4.1%	10.0%	7.0%	13.0%
WWW	10.0%	6.3%	13.8%	12.0%	8.4%	15.6%
Email	13.0%	8.1%	17.9%	14.0%	9.8%	18.2%

Table 4.2 – User class distribution per radio environment.

Environment	User Class Distribution	
	Mass Market	Business
Dense Urban	50%	50%
Urban	60%	40%
Suburban	70%	30%
Rural	80%	20%

The assumed distribution of users between indoor, pedestrian and vehicular is summarised in Table 4.3. These are typical values that reflect the higher existence of indoor users in strongly densified areas, and the higher existence of outdoor/vehicular users in less densified ones.

Table 4.3 – User densities per radio environment.

Environment	User Type		
	Indoor [%]	Pedestrian [%]	Vehicular [%]
Dense Urban	80	10	10
Urban	60	20	20
Suburban	40	30	30
Rural	20	40	40

Concerning mobility aspects, the average values considered for user speed were 0 km/h for indoor, 3 km/h for pedestrian and 50 km/h for vehicular users.

Regarding simulation times, the time needed for a single simulation depends mostly on the number of users addressed by the BS and on the scheduling algorithm implemented. Table 4.4 summarises the simulation times according to those conditions, when an Intel Pentium 1.86GHz processor with 1Gbyte of RAM is used.

Table 4.4 – Simulation times according to number of users and scheduling algorithm

# Users	Scheduling Algorithm	Simulation Time [min]	Scheduling Algorithm	Simulation Time [min]	Scheduling Algorithm	Simulation Time [min]
3500	RR	7	MAX	10	PF	11
4000		8		11		12
4500		9		12		14
5000		10		14		16
5500		11		16		18
6000		13		18		20
6500		15		20		22
7000		18		23		25
7500		22		27		29
8000		26		31		33
17000						180

Roughly, the approximated total time spent in simulations, considering that 10 simulations were run for each scenario was:

- 14680 minutes for simulations concerning results summarised in Section 4.2 and Annex G, corresponding to 9 full days of simulations.
- 6340 minutes for simulations concerning results summarised in Section 4.3 and Annex H, corresponding to around 4.5 full days of simulations.
- 6490 minutes for simulations concerning results summarised in Section 4.4, Annex I and Annex J, corresponding to around 4.5 full days of simulations.

In total the equivalent to 18 full days were spent in simulations.

4.2 Variable User Density Scenarios

In this section, one analyses the system response to variable user density figures. Simulations were carried out for the 4 considered radio environents: Dense Urban, Urban, Suburban and Rural. However, as expected, since the trends are more or less the same for these environments, only the figures for Dense Urban environment are presented in the current section, while the figures for Urban, Suburban and Rural environments are summarised in Annex G. For each of these scenarios, results are obtained when using individually each of the three implemented scheduling algorithms. For each set of simulations, the number of users covered by the BS is increased so that the effect of increasing load can be seen for each scenario. The traffic mix is kept unchanged, with the REF scenario being used. For PF scheduling, Admission Control and Congestion Control thresholds are set to 100%, meaning that both VoIP and Video Call services can fully use all the available system resources.

4.2.1 Network Load

Figure 4.1 illustrates the number of average served users per each of the BS's sector, for an increasing number of total users covered by the BS. As expected, the number of average served users increases when the total number of users covered by the BS increases as well. It is also clear that one gets a higher number of average served VoIP users, which is a consequence not only of the higher weight of this application for the REF scenario (Table 4.1), but also from the fact that a VoIP call lasts longer than, e.g. a WWW call served with high throughput. For the same total number of users covered by the BS, when considering the different scheduling algorithms, it can be seen that there is a slight increase of VoIP users when the PF algorithm is in use, around 2% more. This is also an expected trend, since the PF algorithm is the one implementing QoS mechanisms that favour VoIP against the remaining services, meaning that such users are served in first hand, whereas in the other approches, RR or MAX, such users compete with the remaining for the system resources.

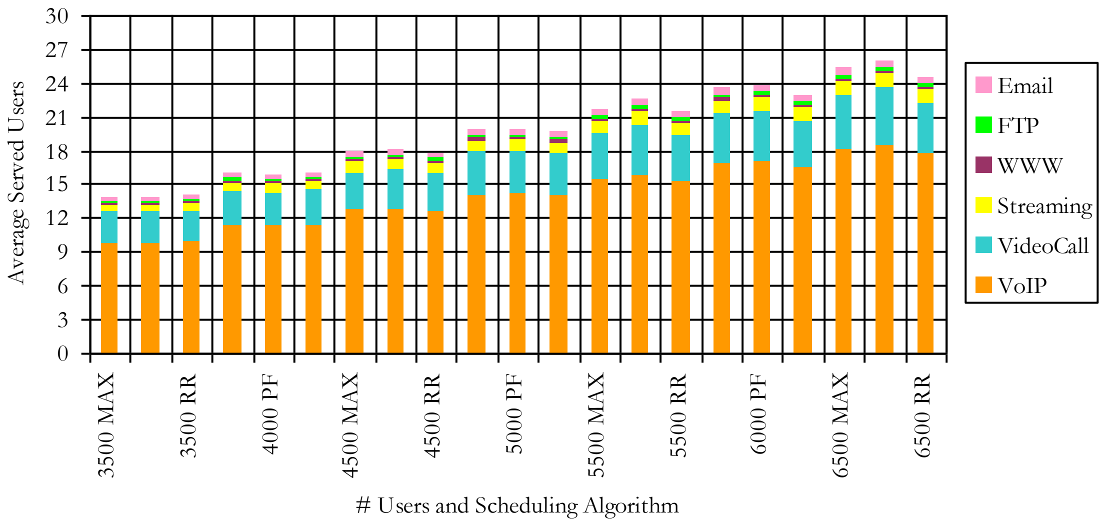


Figure 4.1 – Average number of users served per application for variable number of users covered by the BS – Dense Urban.

Whereas Figure 4.1 stands for the average number of served users, Figure 4.2 illustrates the average network (or system) load, as defined in Sub-section 3.3.3. The differences are obvious: although the average number of VoIP users prevails over the remaining applications, this is not directly translated into a higher network load due to VoIP. Applications such as streaming or, mainly, FTP are the ones that demand more from the system. Also important to notice is the high weight of overhead in system load. As explained in Sub-section 3.2.3, the MAP message size varies with the number of allocated users in a frame. According to [WiMA06a], when the network is dominated by bursty data traffic, such as FTP and HTTP, the number of users being scheduled per-frame is typically small, meaning that, in this case, the resource allocation is done most efficiently and the MAP message mainly contains the fixed MAP overhead.

On the contrary, when the network is dominated by VoIP traffic, the number of users scheduled per-frame may be large. [WiMA06a] states that the MAP overhead increases linearly as the number of scheduled users increases. This is compliant with the results provided by Figure 4.2. To fight this issue, in particular to better control the MAP overhead, the 802.16e standard introduces multicast sub-MAPs, which allow multiple sub-MAP messages to be sent at different data rates to users with different SINR. Therefore, while broadcast messages are sent at the highest reliability needed to meet the cell edge coverage, the common control ones, e.g., traffic allocations, can be delivered more efficiently according to the user SINR condition. Remember that, as detailed in Sub-section 3.2.3, in the simulator, a fixed QPSK ½ coding with repetion factor of ½ was assumed for MAP overhead; with this sub-MAPs mechanism, the weight of overhead in network load would be decreased and more system capacity would be available for ‘real’ traffic, and the need for such implementation is proved by the results issued in Figure 4.2.

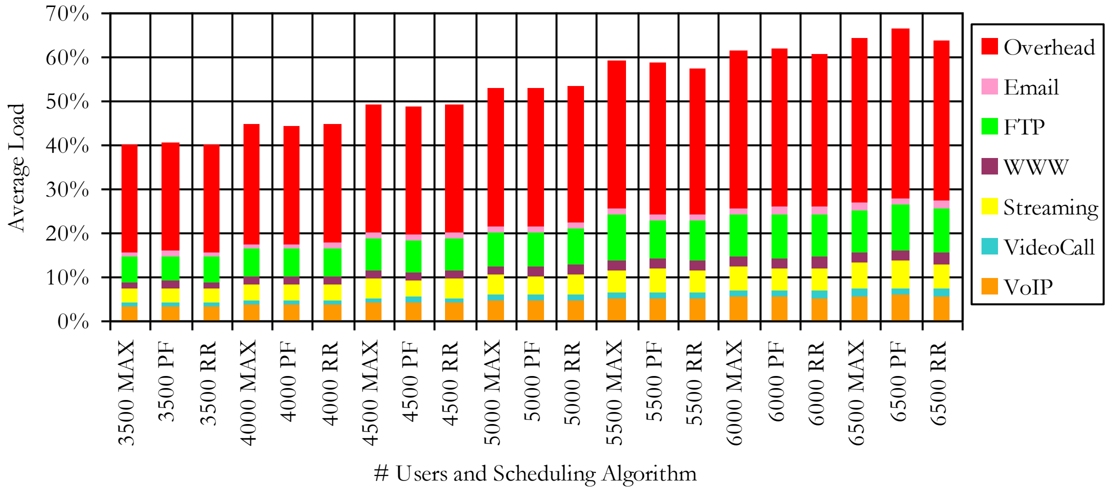


Figure 4.2 - Average load for variable number of users covered by the BS – Dense Urban.

The overhead is translated into a reduction of system efficiency, defined in Sub-section 3.3.3, as illustrated by Figure 4.3. An extrapolation shows that each addressed user per frame introduces on average a decrease of 1.13% in MAP efficiency, with a near-perfect linear trend. These results also show that this issue is not mostly related to the type of scheduling adopted, as the same trend is seen for the three considered implementations, and only with slight changes exist between them. However, results could be significantly more different for distinct scheduling approaches in case the concept of sub-MAPs would have been implemented, as with this concept the overhead due to a single user is less in case this user is allowed to use a higher AMC. In this case, scheduling approaches like PF or, especially, MAX, would likely result in a better MAP efficiency.

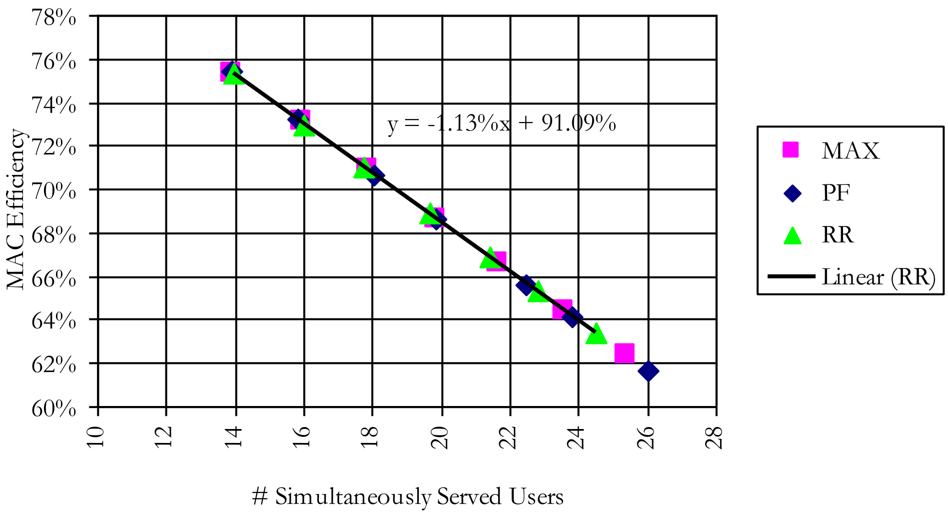


Figure 4.3 – MAP efficiency for increasing number of simultaneously served users.

Figure 4.4 summarises the total traffic volume carried by each sector of the BS, divided by service. As expected, although there are more VoIP and Video Call users, the reigning applications are the BE ones, like FTP. This is an expected trend, following the trend illustrated in Figure 4.2 as well. Naturally, more traffic is carried when more users are served, which is an obvious result. Typically, more traffic is carried by the system when the MAX approach is considered, around 2% more than with PF or RR.

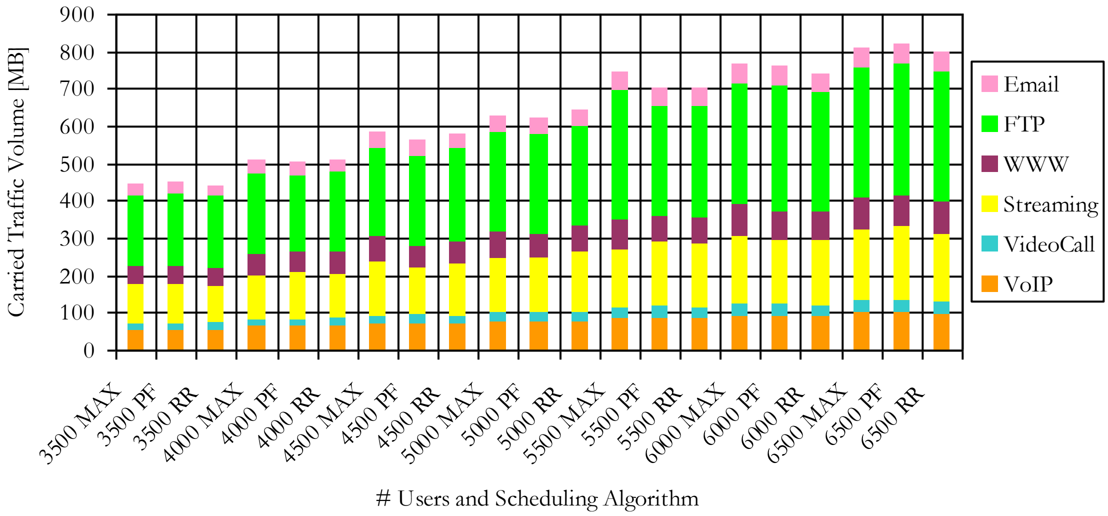


Figure 4.4 – Traffic volume carried by the system – Dense Urban.

As an indicative result, Figure 4.5 illustrates the weight of each AMC. It is interesting to notice that the worst and the best AMC, QPSK1/2 and 64QAM3/4, respectively, are the dominant ones. However, these figures are highly dependent on factors such as the distribution of users between indoor or outdoor environments or the channel model, and their fine analysis is outside the scope of the present thesis.

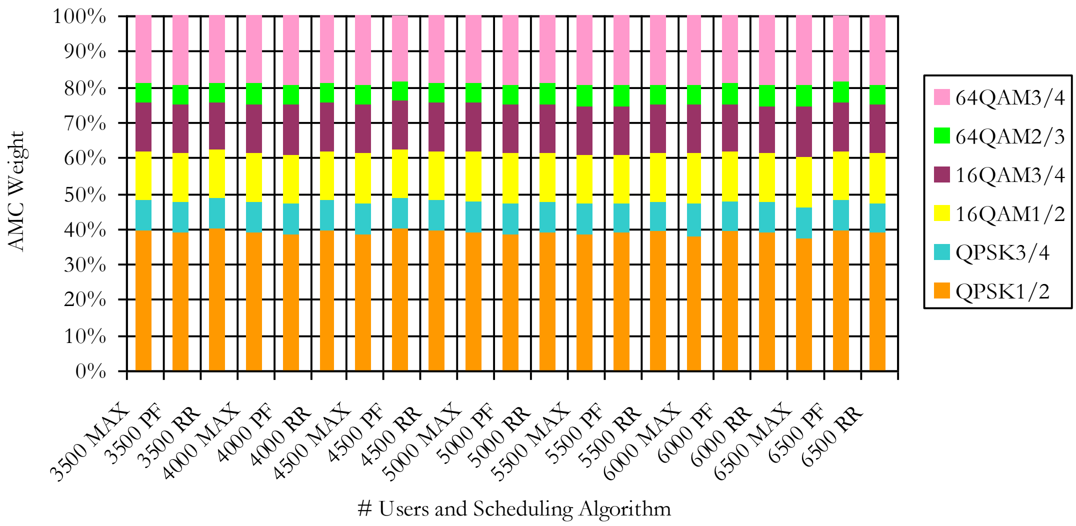


Figure 4.5 – AMC distribution – Dense Urban.

In conclusion, the expected trends concerning the different scheduling approaches are seen in the main results. The MAX approach is the one maximising system capacity, as it always seeks for users suffering best AMC conditions, regardless of their application. Thus, more traffic can be carried by the system if AMC is implemented. The PF approach, although also looking first for users suffering best AMC, prioritises RTPS in first hand, meaning that when PF is implemented, a clear increase of simultaneous served VoIP users occurs. The MAP efficiency issue is also addressed, being shown that there is roughly a 1.1% inefficiency introduced by each user addressed per radio frame. Additional features targeting minimisation of MAP overhead, such as sub-MAPs, are part of the standard.

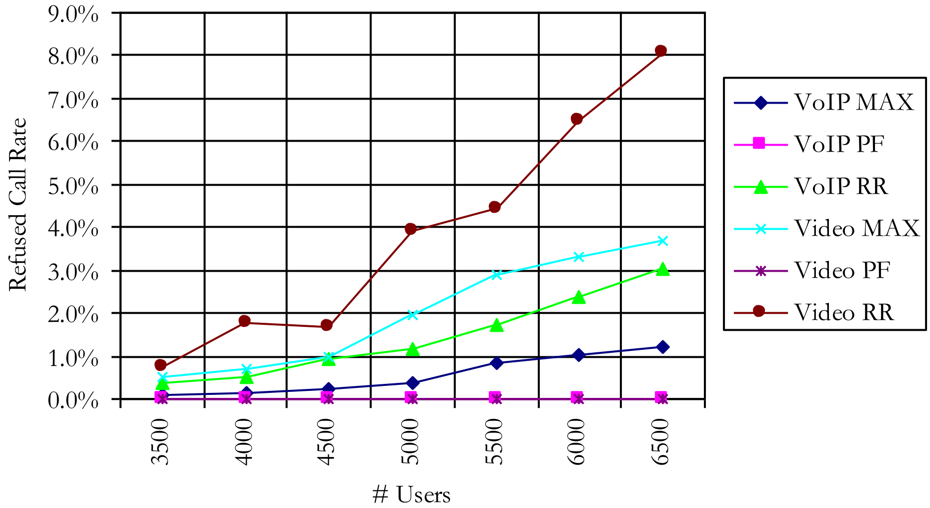
4.2.2 Real-Time Applications Performance

In this section, one analyses in detail the performance results of real-time applications, particularly VoIP and Video Call, as defined in Sub-section 3.3.3. Figure 4.6 summarises Refused Call and Drop Call Rate results. Concerning Refused Call Rate, the differences between the three algorithms are clear. The basic RR approach is the one providing worst Refused Call Rate figures for VoIP, whereas PF provides the best ones, with MAX in the middle. Notice that for the same number of users, e.g., 6500, VoIP Refused Call Rate with RR is already at 3.0%, MAX is around 1.2% and PF is still at 0%. These results are easily explained:

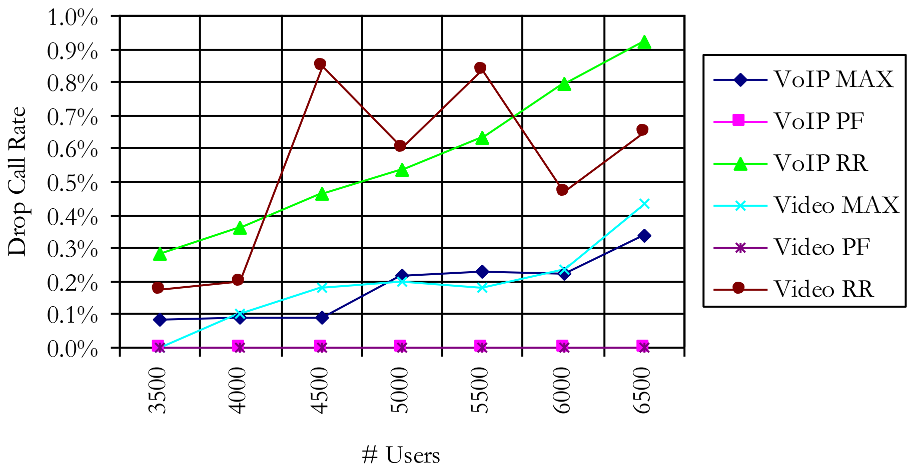
- With the basic RR approach, no QoS mechanism is implemented and users are served, regardless of their application, without considering radio channel conditions to which each user is subjected. Thus, VoIP users are not differentiated from the remaining ones and, additionally, it may happen that the system is flooded with users that, due to low AMC assignations, may quickly exhaust system resources.
- With the MAX approach, still no QoS mechanism is implemented, and the same round-robin approach is taken. But within the MAX approach, the system already looks at the radio conditions felt by each user, prioritising, firstly users that are subjected to higher AMCs. This results in an increased system capacity, and the immediate consequence is a decrease of Refused Call Rate felt by VoIP users.
- The PF approach is the one implementing QoS mechanims, prioritsing firstly VoIP and Video Call applications, instead of BE ones. Since Admission Control and Congestion Control are set to 100%, i.e. RTPS calls are only blocked in case all the network resources

are exhausted by these two applications combined, the difference in the results is remarkable, as the Refused Call Rate for VoIP decreases to pratically 0% for the same network load.

Concerning Drop Call Rate, the trends are the same as for Refused Call Rate, with PF providing the best results, whereas the RR approach provides the worse. When comparing Video Call with VoIP, it is clear that Video Call suffers from congestion problems before VoIP, which is due to the fact that Video Call packets are larger than VoIP ones, as detailed in Annex C, respectively.



a) Refused Call Rate



b) Drop Call Rate

Figure 4.6 – RTPS performance indicators for increasing number of served users by the BS – Dense Urban.

These Refused Call Rate and Drop Call Rate trends are supported by the VoIP and Video Call packet loss figures, Figure 4.7. Again, and due to the same reasons already given for the explanation of the Refused Call Rate and Drop Call Rate trends, lower packet loss is achieved when more inteligence is implemented on the scheduling algorithms, i.e., when either MAX or, mostly, the PF approaches are in place. It is important also to highlight that the PF approach configured for this section's results foresees a full prioritisation of real time applications, due to Admission Control and Congestion Control.

As illustrated in Figure 4.7, this implementation often results in 0% packet loss for either of the applications. However, both VoIP and Video Call services quality can still be assured even where a low packet loss rate exists. Since these applications are supported by protocols that do not foresee the retransmission of packets, at any protocol layer at all (contrary to, e.g., data services, which are controlled at least at the transport layer), the acceptable packet loss is low, typically up to 1% for VoIP, and even lower for Video Call. In GSM, for instance, the negative impact of missed speech frames is mitigated through the repetition of the previous frames, within certain limits, of course. This approach allows for the system resources to be better distributed within the different applications or within the different subscribers using the same type of applications, as the BS can, e.g., put less power on a given radio link (thus, decreasing interference) and assign less resources to the same user all the time (thus, increasing room for further users), without jeopardising the overall QoS parameters. The impact of using this intelligence is analysed in Sub-section 4.4.2.

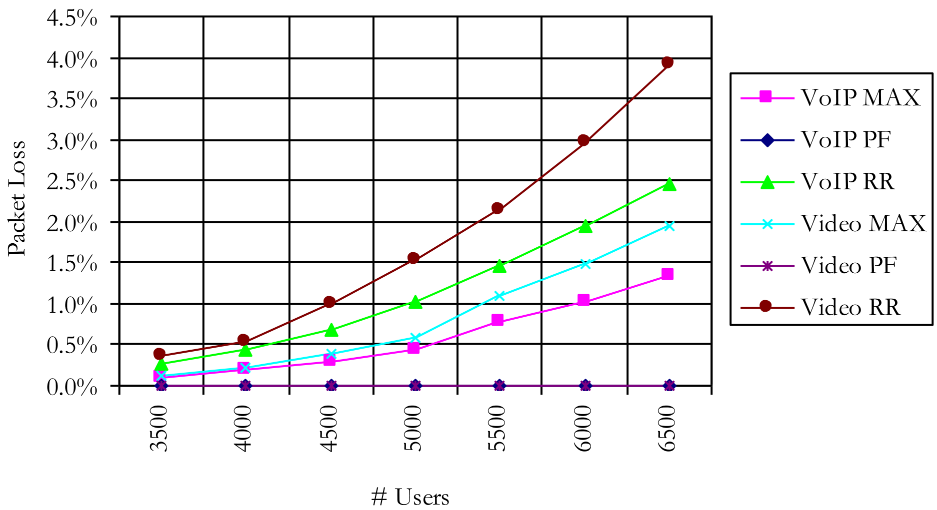


Figure 4.7 – RTPS packet loss for increasing number of served users by the BS – Dense Urban.

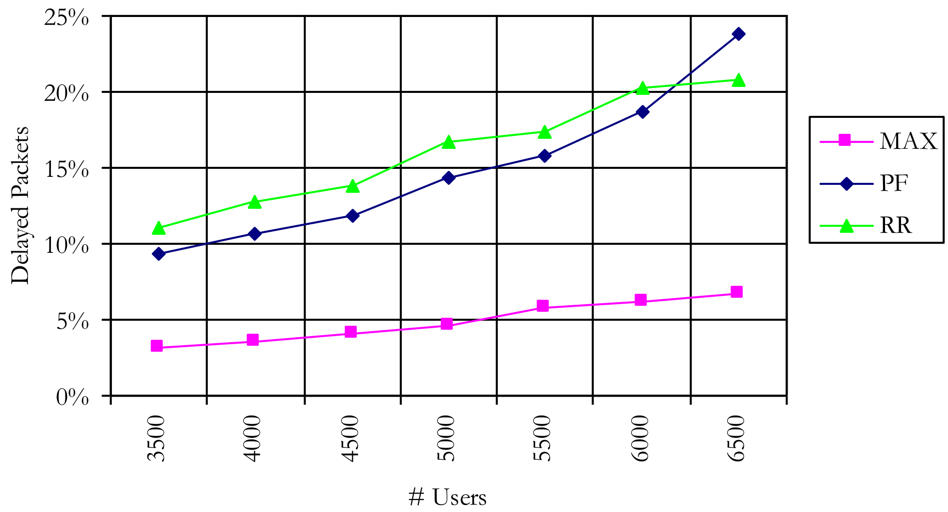
In order for VoIP blocking to occur with the PF configuration in use, it is necessary to go up to 17000 users covered by the BS. At this point, VoIP blocking starts to occur already, but by then the system is completely loaded with VoIP traffic, leaving very few resources for the remaining services. For instance, by then Video Call blocking is already around 25.7% and the average FTP throughput is around only 39 kbps, as near 100% of BE packets are delayed, with average delays around 1.250 s.

4.2.3 Best-Effort Applications Performance

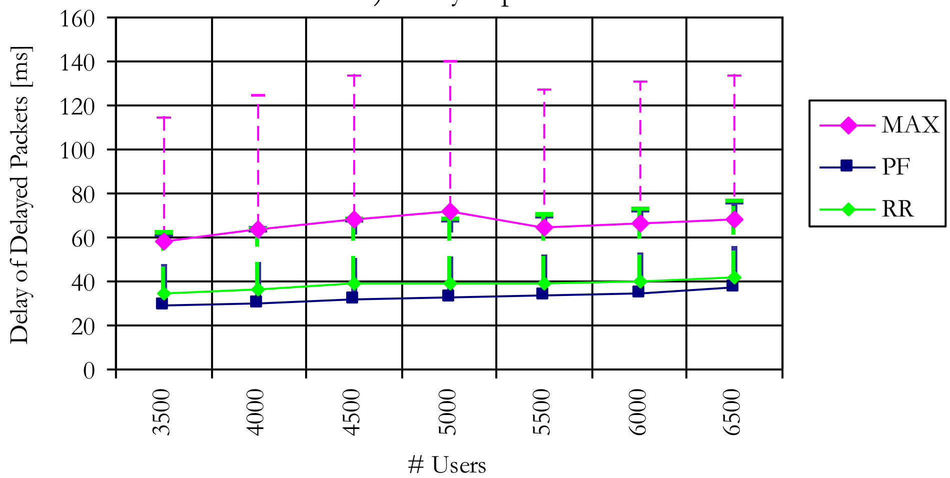
In this section, a detailed analysis of the performance results of BE applications, which include streaming, WWW, FTP and Email is done. These being BE services, instead of packet loss, packets are delayed when there are no system resources available to carry them, which affects the overall user perception of the service.

Figure 4.8 illustrates the percentage of delayed packets for BE applications, as well as the average delay and respective standard deviation of the delayed packets. It is clear that the percentage of BE delayed packets increases when the number of users increases as well, as expected. If the MAX scheduling is implemented, the percentage of delayed packets decreases significantly, varying between 4 to 6%, which is expected, since, by serving in first hand the users having the best channel conditions, the overall system capacity increases. The tradeoff concerns the average delay of delayed packets: since users suffering worse channel conditions are put in the end of the queue, and since there is no additional criterion concerning historical delay, these users are likely to wait long time periods, between 60 to 70ms, before being assigned any resources. This commitment between avoiding too long delays and increased system capacity is achieved through the PF scheduling: it provides a lower percentage of delayed packets compared to RR, while it provides reduced delays for delayed packets compared to the MAX approach. This is due to the fact that PF, like MAX, seeks for users feeling better channel conditions, but, concerning BE users, the first priority is to serve users that already felt long delays, minising average delay for delayed packets compared to MAX.

In order to better understand the impact of delay, a throughput analysis is needed. Throughput results are better interpreted when considering only streaming or, especially, FTP services. WWW throughputs, for instance, are strongly variable and depend on the size, type and number of objects of the web pages, whereas FTP throughput usually is limited by the 'last mile' part of the network, in this case the WiMAX radio interface.



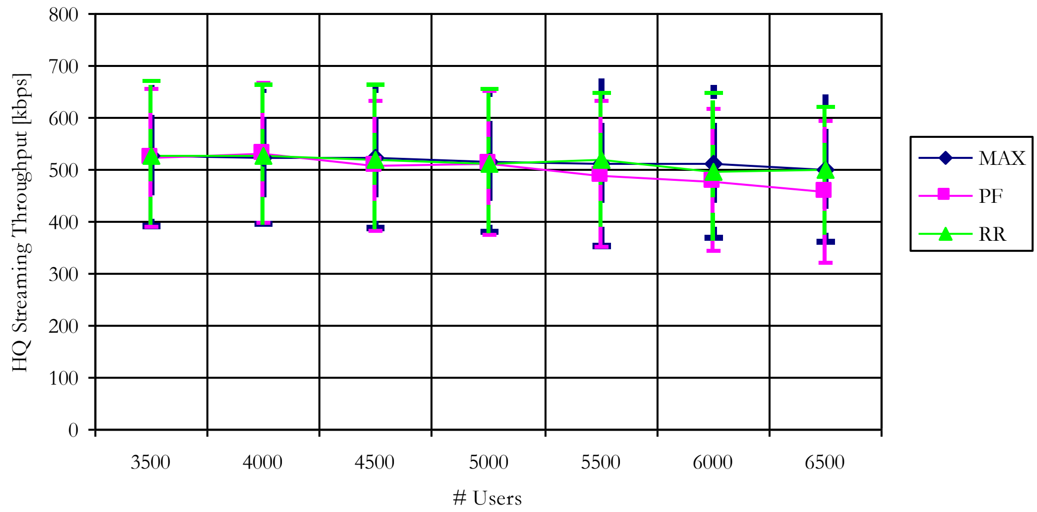
a) Delayed packets



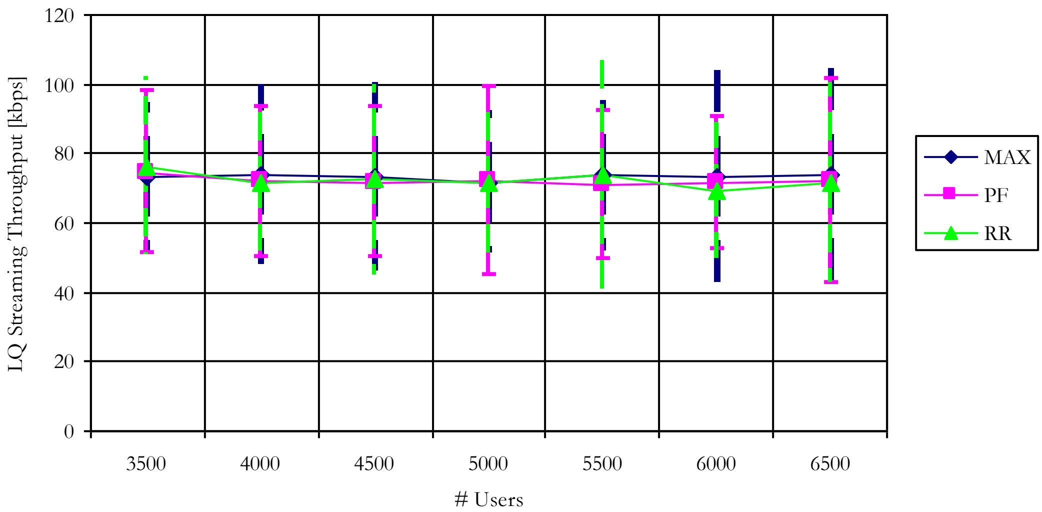
b) Delay of delayed packets

Figure 4.8 – Percentage of BE delayed packets and average delay of delayed BE packets for increasing number of served users by the BS – Dense Urban.

Figure 4.9 illustrates the average and standard deviation throughput for high and low quality streaming video. The low quality video throughput is not significantly affected by the increase of system load, as the video throughput is around 70kbps, a value that is easily provided by a WiMAX network, even in high load scenarios. On the contrary, high quality video has clearly a slight degradation when the total load increases. In any case, no clear differences are seen when considering either of the different scheduling approaches.



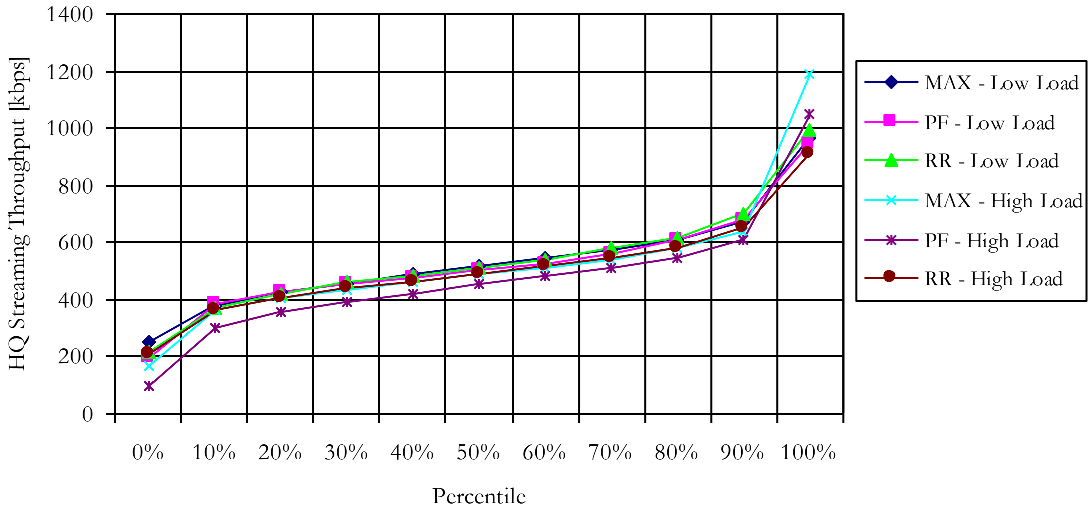
a) High quality video



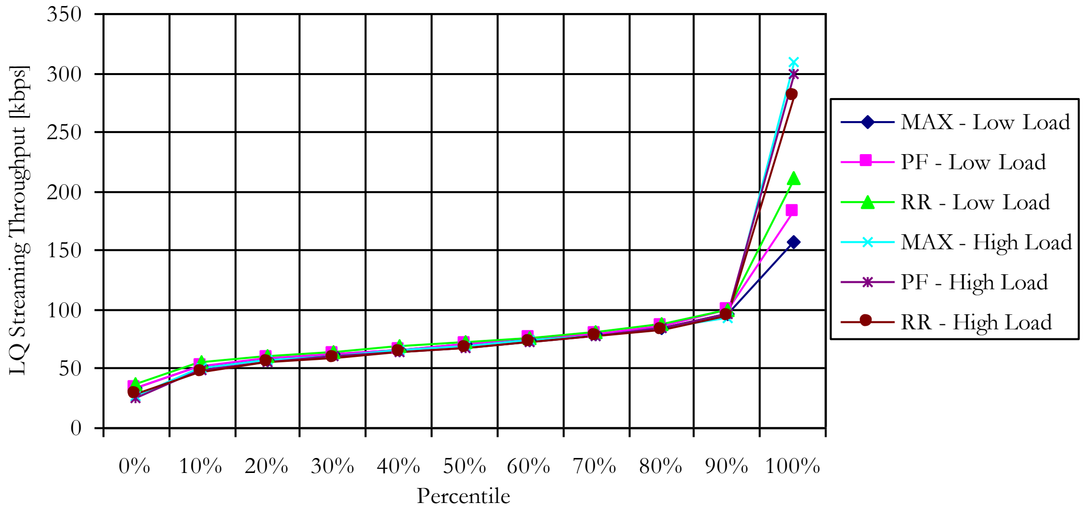
b) Low quality video

Figure 4.9 – Average streaming throughput and standard deviation for increasing number of users served by BS.

Figure 4.10 illustrates the cumulative distribution of streaming throughput, for both a low load scenario (3500 users) and a high load scenario (6500 users). While for a low load scenario no perceptible differences are seen within the different scheduling algorithms, for a high load scenario and high quality video it is clear that the PF approach provides worse throughput figures than RR or MAX. This is explained by the fact that when PF is considered, both VoIP and Video Call services have the priority in system resources allocation, leaving less room available for BE applications.



a) High quality video



b) Low quality video

Figure 4.10 - Streaming throughput cumulative distribution for low and high load scenarios.

As explained previously, the best application for throughput evaluation is FTP and Figure 4.11 illustrates the average (and respective standard deviation) FTP throughput. In this case, it is already clear that with PF set with 100% thresholds for both Admission Control and Congestion Control of RTPS calls, worse FTP performance is achieved on average: MAX and RR are, roughly, 6.2 and 7.7% faster, respectively. Again, this is explained by the fact that with PF, real-time applications are prioritised when assigning system resources, being allowed to fully use them. Figure 4.11 also shows that the basic RR mechanism provides an average throughput higher than MAX. This means that although the MAX approach provides a lower percentage of delayed packets than RR, the fact that the MAX delayed packets are subjected to a higher average delay results in an average lower throughput (Figure 4.8).

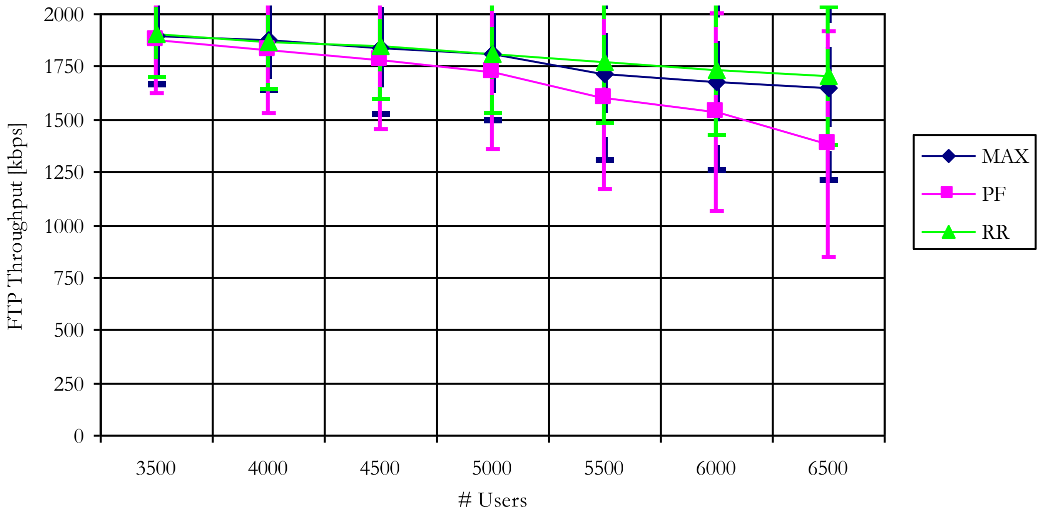


Figure 4.11 – Average per user FTP throughput and standard deviation for increasing number of users served by BS.

When looking at the cumulative distribution of FTP throughput, while for a low load scenario no clear differences are seen within the three scheduling implementations, for a high load scenario this is not the case, Figure 4.12:

- MAX scheduling is the one providing the lowest throughputs, which is due to the fact that users permanently suffering from very bad radio conditions may have to wait long periods before being assigned any resources at all. Up to user's percentile 6%, MAX provides the worse performance.
- For intermediate throughput figures, PF clearly performs worse, due to the already given reasons related to real-time services prioritisation. In this interval, RR is the best approach. Degradation in throughput when PF is considered is the 'price to pay' for having QoS mechanisms that consider RTPS services in first hand.
- When one approaches the maximum possible throughput, the MAX scheme takes the advantage, i.e., there are more users having very good throughputs when MAX is adopted. Above user's percentile 50%, MAX provides the best performance.

Thus, MAX appears as an 'all or nothing' approach, in which either a user capable of handling a very high AMC will have very good throughput, because he will be prioritised first (the reason why the worse throughputs are also achieved when MAX is used), or a user capable of only handling a very low AMC will have a very bad performance, because he will be prioritised in last.

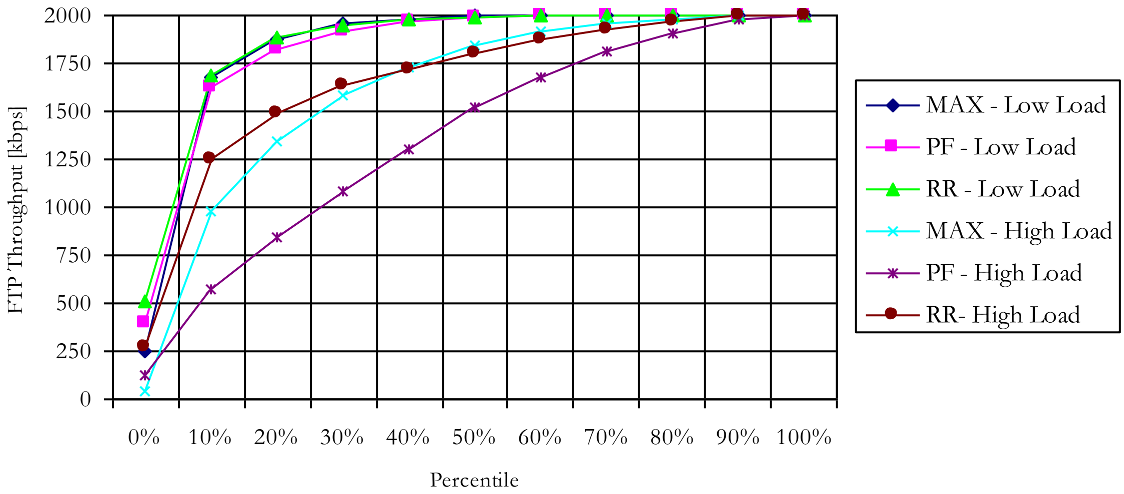


Figure 4.12 – FTP throughput cumulative distribution for low and high load scenarios.

In conclusion, regarding applications performance, there is a clear tradeoff between the throughputs achievable by BE applications and the accessibility and retainability of real-time applications, with these being mutually exclusive. PF is the policy providing best performance for real-time applications, due to its QoS mechanism combined with 100% thresholds for Admission Control and Congestion Control of RTPS services, providing pratically 0% packet loss, Refused Call Rate or Drop Call Rate for VoIP, while RR or MAX policies provides the worse performance by reaching Refused Call Rate, Drop Call Rate and packet loss figures well above 1% for the same conditions.

However, the opposite occurst when considering BE services: throughput achievable with PF can be up to 6 and 7% worse than the one achievable through RR or MAX, respectively. It is clear that the PF implementation, although useful to guarantee real-time applications accessibility, shall be smart enough in order to discard part of the real-time traffic, as both a reduced packet loss and reduced blocking probability are tolerable. Such PF configuration can, perhaps, still provide good accessibility for VoIP and Video Call, while releasing resources to cope with the demand of the remaining BE services, without jeopardising the overall QoS requirements of the different services. This is analysed in Sub-section 4.4.2. The MAX approach maximises the number of users feeling higher throughputs but it can severely penalise users suffering from worse AMC. Comparing MAX with RR, the percentage of packets delayed with MAX is, roughly, 1/3 of the one with RR, but the average delay of these delayed packets with MAX is more or the less the double than with RR.

These basic scheduling implementations clearly show that the system does not have infinite resources, and, in order to favour a given kind of applications, others will be penalised. A

WiMAX system supplier or operator must then decide how the system shall be designed or configured, according to the defined targets in terms of overall QoS, knowing that, regardless of the approach assumed, there are always advanges and disadvanges associated to each option. Additionally, the decision between, e.g., a RR approach, MAX or PF is not merely technical. From the previous results one can conclude that the RR is not advisable at all, because it does not guarantee any requirerments from a real-time services point of view. But the operator can, e.g., decide to implement such algorithm, because it is cheaper than buying a more intelligent algorithm, such as MAX or PF, and decide to deploy BSs closer to each other, limiting the number of users covered by each BS, thus, mitigating the accessibility problems for real-time applications, while still taking advantage of the RR capabilities to serve BE applications.

On the contrary, another operator may wish to be guaranteed in terms of real-time services QoS, and decide to use a PF scheduling algorithm, which certainly has an additional cost for the operator, since network vendors usually charge for the use of such algorithms. But then the operator has a margin to deploy less BSs, reducing the need for investment, since more users can be served by one BS if PF implemented, without jeopardizing real-time services accessibility and retainability. All these remarks are useful to remember that one cannot simply look at the results provided by a given algorithm: in every circunstance, an operator must consider both the technical and the economical aspects associated to each available possibility in terms of network deployment and configuration.

4.3 Variable Traffic Mix Scenarios

In this section, one analyses the system response to variable traffic mix figures. Simulations are again carried out for the 4 considered radio environents, but again only the results for Dense Urban environment are presented in the main sections: Urban, Suburban and Rural results are summarised in Annex H, with the main remarks concluded for Dense Urban being roughly the same for the remaining environments. For each of these scenarios, results are obtained when using individually each of the three implemented scheduling algorithms. For each set of simulations, the number of users covered by the BS is kept unchanged, based on the results of Section 4.2: 5250 users are considered, which is the number that roughly leads to a VoIP Refused Call Rate around 1% for both RR and MAX scheduling (Figure 4.6). The variable analysed in the present section is the traffic mix, according to the data summarised in Table 4.1. Admission

Control and Congestion Control settings for the PF approach are unchanged, i.e., set to 100% for both VoIP and Video Call. The effect of increasing data services against a decrease of VoIP, or vice-versa is assessed, as well as the system response for the different scheduling algorithms used. This type of analysis is particularly interesting, considering that, within a cellular network, the traffic mix varies considerably along the 24 hours of one day: typically there are more VoIP users during the morning and afternoon periods, whereas data services are mostly used during the night ones. Thus, operators need to dimension their networks in order to cope with the different demands along these periods, otherwise a system may, e.g., behave very well to cope with the voice traffic during voice peak hours, and very bad to cope with the data traffic during data traffic peak hours.

4.3.1 Network Load

Figure 4.13 illustrates the average number of served users when the traffic mix varies (Table 4.1). The trend is the expected one when comparing SPC or DAC traffic mixes with the REF scenario, analysed in Section 4.2:

- Compared to the REF traffic mix, the SPC one results in a higher number of served VoIP users, whereas the DAC traffic mix results in an increased number of average served data users.
- When the DAC traffic mix is considered, although the total number of users covered by the BS is the same, the average number of users simultaneously served is lower, meaning that the data services are held during shorter time periods.

Figure 4.14 illustrates network load variation following traffic mix changes. Clearly, when data services have a higher weight, the network load is higher, despites the fact that the average number of simultaneously served users is lower, as given by Figure 4.13. Figure 4.14 also shows that the overhead is higher when speech service prevails, which is explained by the fact that MAP overhead is proportional to the number of users addressed per radio frame. This clearly highlights the MAP efficiency problem, when applications like VoIP prevail in the system, as more resources are wasted due to overhead. These results are aligned with the remarks already summarised in Sub-section 4.2.1.

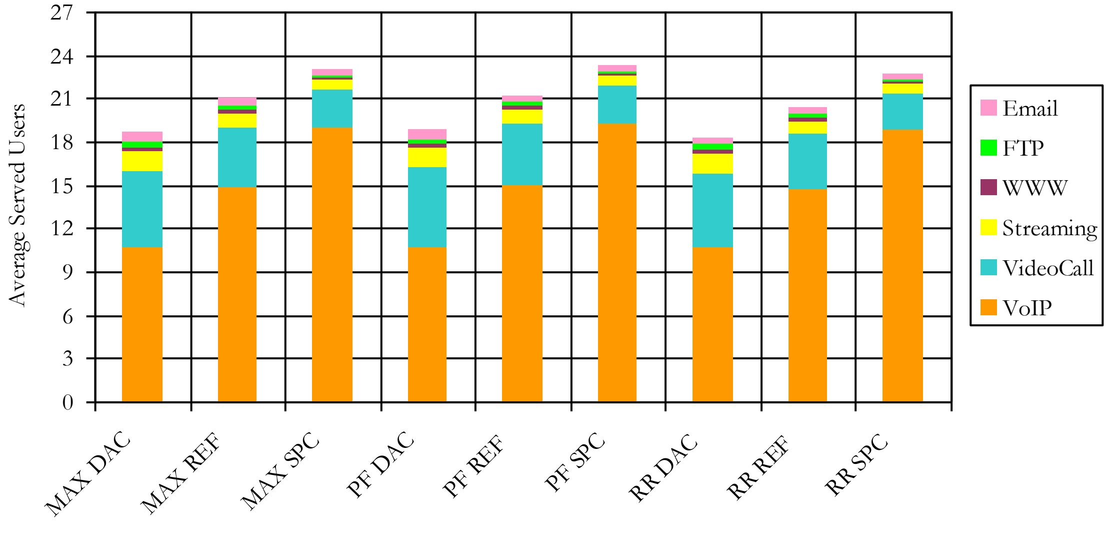


Figure 4.13 - Average number of users served per application for variable per user traffic mix – Dense Urban.

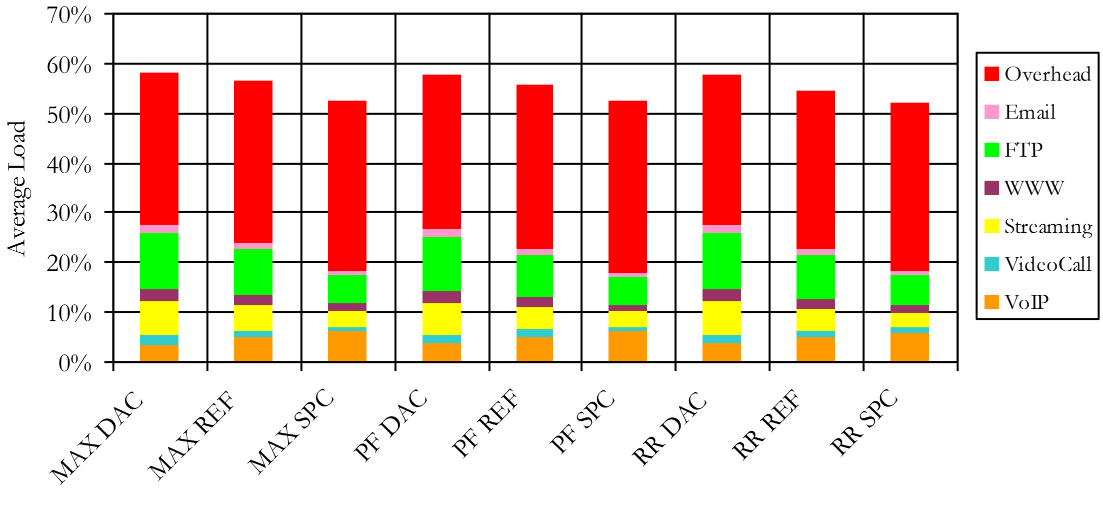


Figure 4.14 - Average load for variable per user traffic mix – Dense Urban.

The relation between the dominant services and MAP efficiency is illustrated in Figure 4.15. Clearly, a lower system capacity is available when speech applications prevail, which is due to the higher number of simultaneously served users per frame. The opposite occurs when data services have higher weight, showing that WiMAX is more suited to address data based applications. Thus, an operator with a WiMAX network must consider that the system capacity will be different according to the period of the day. It is likely that one network may be capable of handling a given traffic volume at a given hour, while in a different time period the same network may suffer from congestion problems, only due to a different behaviour in terms of services usage by the same base of clients.

Following the previous remarks, Figure 4.15 also shows that the PF approach is the one leading to higher MAP inneficiency. This is explained by the fact that with PF, VoIP users are served in first hand, which increases the number of simultaneously served users. The MAX approach, although not implementing QoS mechanisms that would serve VoIP in first hand, as PF does, is still not as efficient as the RR algorithm. This is explained by the fact that, by looking first to users having higher AMC, the MAX approach allows the BS to address more simultaneous users, regardless of these being VoIP, Video Call or data users. RR appears as being clearly the policy that leads to better MAP efficiency, since it does not look for better AMCs nor for particular services, thus, it cannot serve as many users simultaneously as the other scheduling policies do.

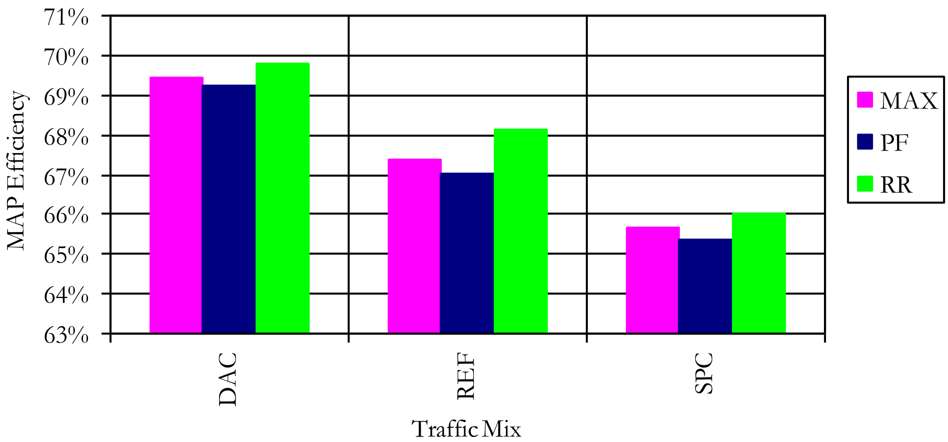


Figure 4.15 – MAP efficiency for different traffic mix, for different scheduling algorithms.

The total carried traffic volume for the different traffic mixes is illustrated in Figure 4.16. Naturally, it is much higher for the DAC scenario than for the REF one or, especially, the SPC scenario: more than 60% more traffic can be carried with DAC than with SPC. Again, this is due to fact that data applications origin a higher traffic volume than applications like VoIP do, even though there are more VoIP users addressed simultaneously.

Figure 4.17 shows how AMC is distributed. It can be seen that the distributon does not change significantly from one traffic mix to another, which is natural, as it mostly depends on link budget related issues, rather than from the traffic mix itself.

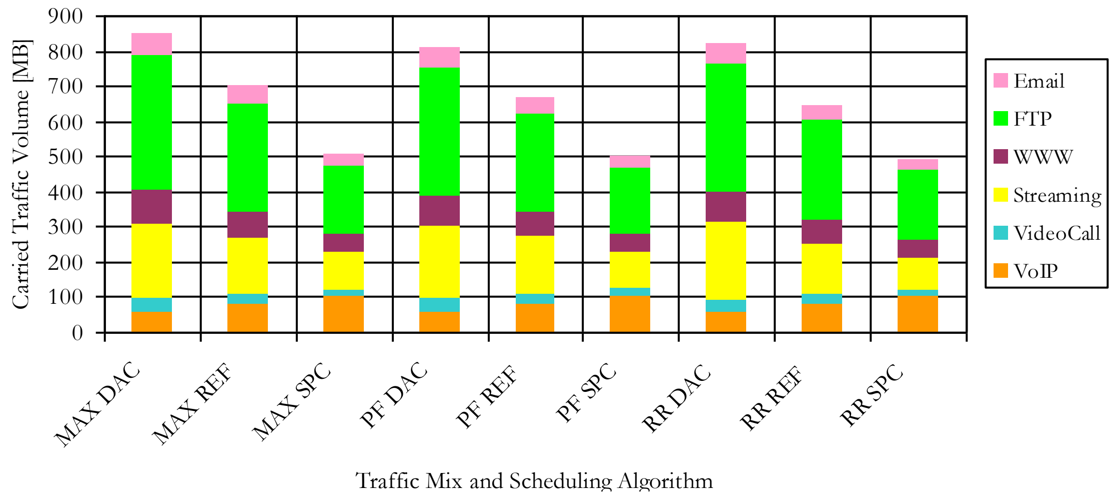


Figure 4.16 – Traffic volume carried by the system for variable per user traffic mix – Dense Urban.

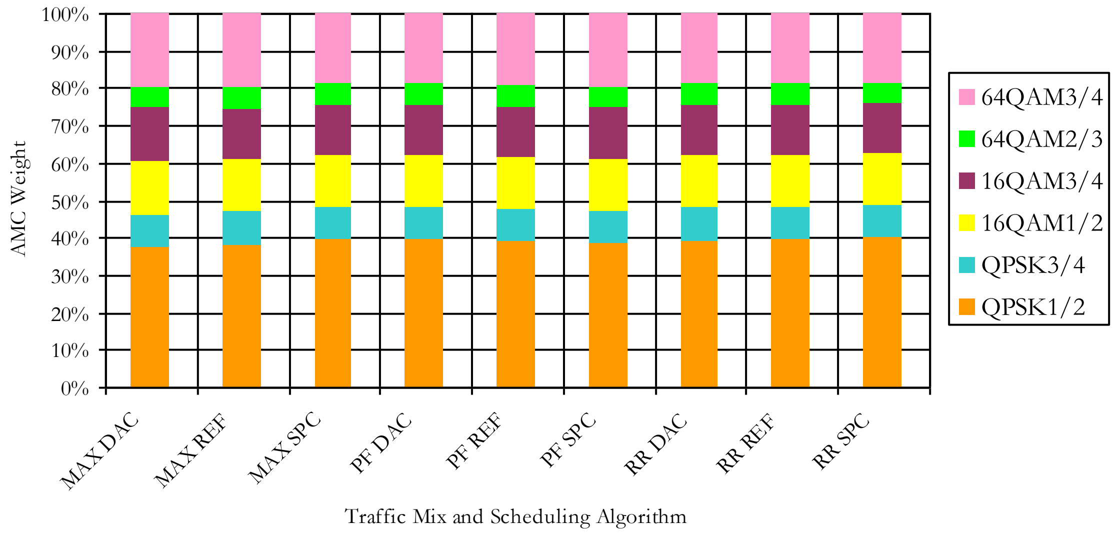


Figure 4.17 – AMC distribution for variable per user traffic mix – Dense Urban.

The average and standard deviation of instantaneous throughput is illustrated in Figure 4.18.

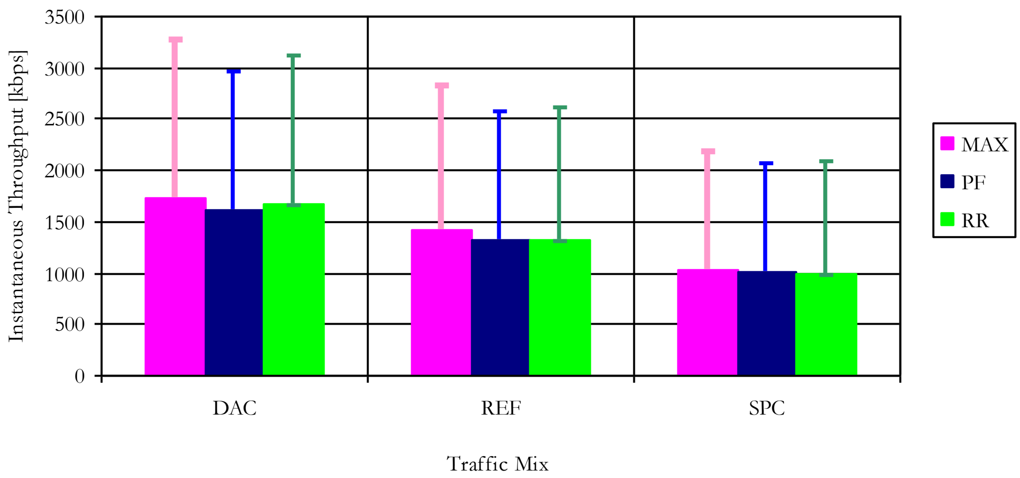


Figure 4.18 - Average instantaneous throughput and standard deviation for variable per user traffic mix – Dense Urban.

Naturally, the instantaneous throughput is higher when a DAC traffic mix is in place, being lower when the SPC traffic mix is considered. On average, the instantaneous throughput is around 1.6 Mbps for DAC and 1 Mbps for SPC. The highest throughputs are achieved when the MAX scheduling is adopted, due to the highest capacity provided by the best AMCs.

The cumulative distribution for the instantaneous throughput per sector is illustrated in Figure 4.19. One can observe that, when either RR or PF scheduling is used, the maximum instantaneous throughput does not go beyond 10 Mbps, regardless of the traffic mix. On the contrary, MAX scheduling can provide instantaneous throughputs up to 14 Mbps. The reason why the average throughputs summarised in Figure 4.18 are below the instantaneous figures illustrated in Figure 4.19 is simply due to the services' characterisation consirered for the present work: if services were characterised as being more 'heavy' (e.g., larger FTP volume sizes, heavier WWW pages, etc.), then the cumulative distributions would be shifted towards higher instantaneous throughput figures, which is compliant with the higher throughputs seen for DAC. Taking the RR approach, for instance, one can see that for the SPC traffic mix, only 30% of the time the instantaneous throughput is above 1 Mbps, whereas when the DAC traffic mix is considered, more than 2 Mbps is achieved during 40% of the time, i.e., higher throughput during larger periods.

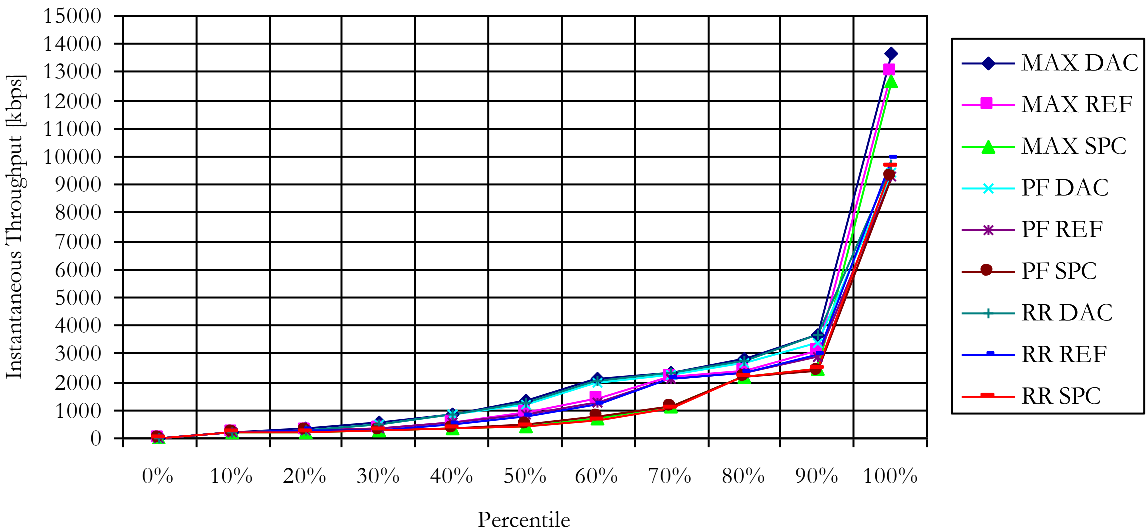
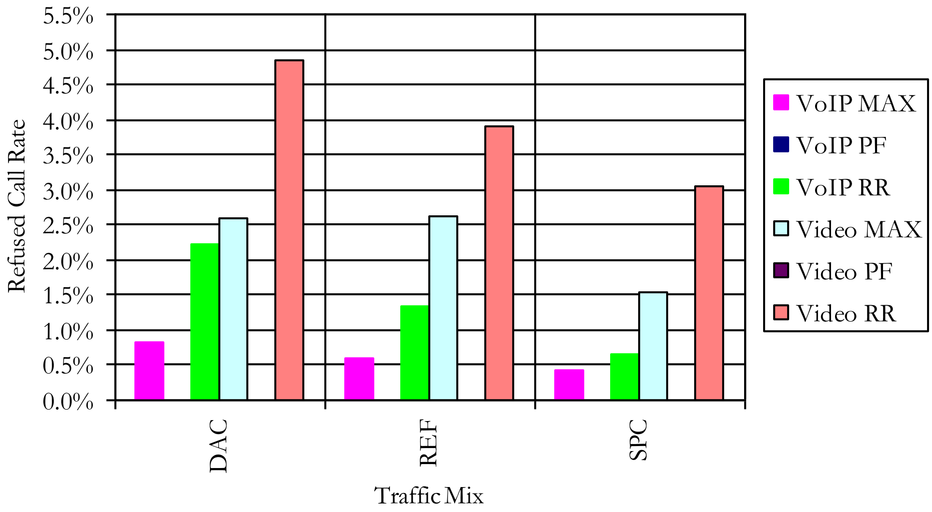


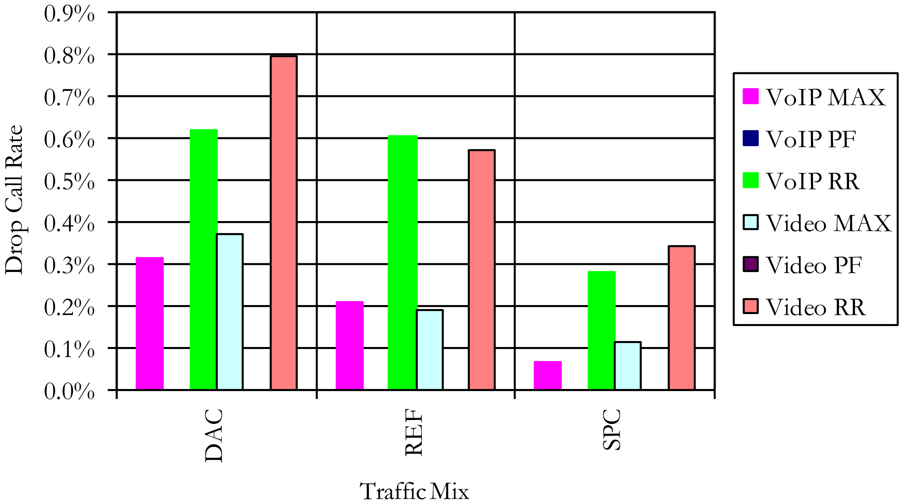
Figure 4.19 - Instantaneous throughput cumulative distribution for variable per user traffic mix – Dense Urban.

4.3.2 Real-Time Applications Performance

Figure 4.20 illustrates how RTPS Refused Call Rate and Drop Call Rate vary according to the traffic mix. The differences are clear: VoIP performance is much worse when the traffic mix is more balanced towards data services (same trend for Video Call), which is natural, since it results in a higher network load, leaving fewer resources available. It is interesting to notice that, e.g., when a RR mechanism is in use, the Refused Call Rate for VoIP can easily double just upon an increase of the data services usage. Thus, if the RR approach is not already the most suited to cope with real-time traffic, it may become quite worse if the usage of data traffic increases. It is also clear that the impact of traffic mix when the MAX approach is implemented is less severe. If the PF scheduling is used, still no blocking is seen with any of the traffic mixes considered.



a) Refused Call Rate



b) Drop Call Rate

Figure 4.20 – RTPS performance indicators for variable per user traffic mix – Dense Urban.

Packet loss figures, illustrated in Figure 4.21, follow the same trend for both VoIP and Video Call. The RR approach is very volatile, according to the traffic mix, more than the MAX one, which is more stable. The PF approach remains around 0% packet loss for any of the traffic mixes considered, being more suited for this type of traffic than any of the other scheduling approaches.

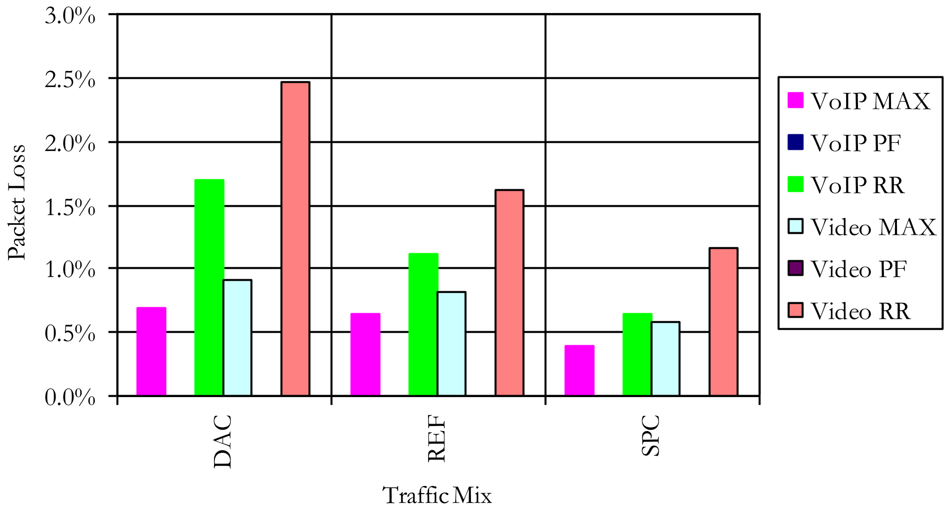
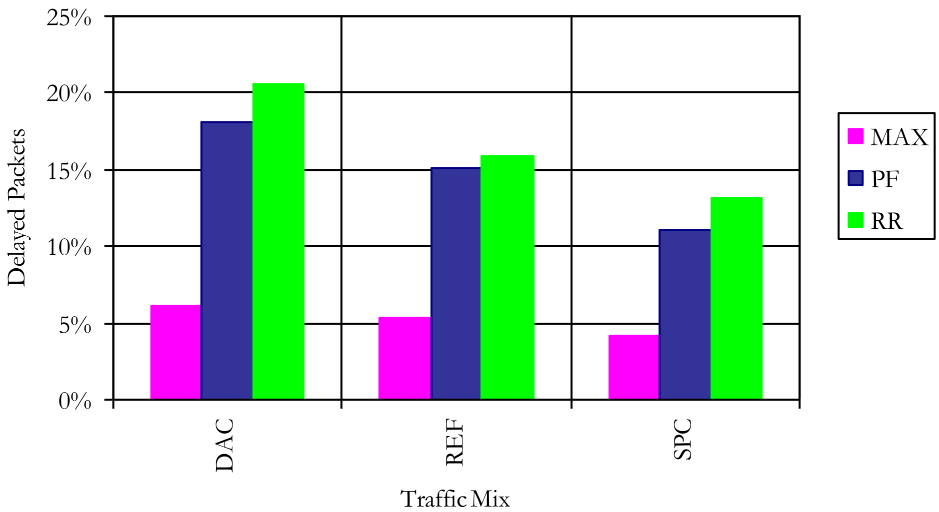


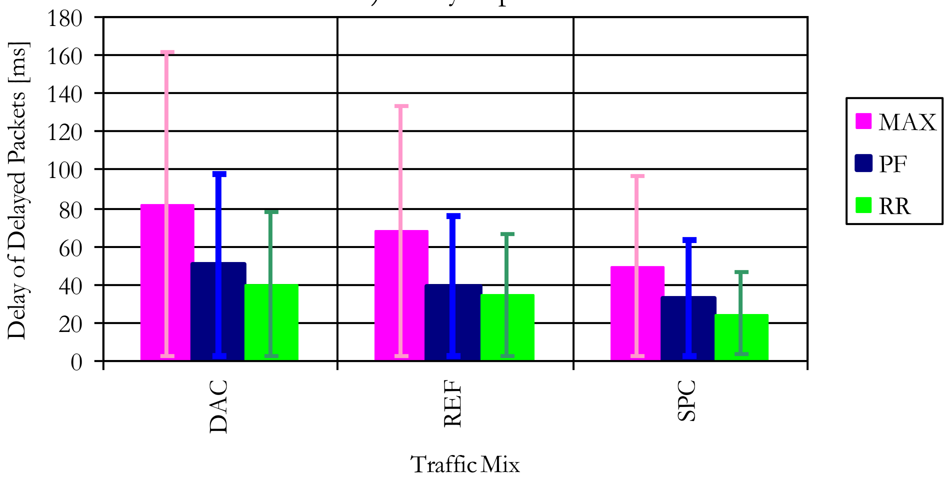
Figure 4.21 – RTPS packet loss for variable per user traffic mix – Dense Urban.

4.3.3 Best-Effort Applications Performance

BE applications performance also depends on the traffic mix. As illustrated by Figure 4.22, the percentage of BE delayed packets changes significantly when one moves from a SPC traffic mix to a REF one or to a DAC traffic mix. This is mostly true for either RR or the PF approach, whereas when MAX is adopted, this impact is not so severe. The MAX approach is less volatile, because this scheduling algorithm is the one that maximises system capacity, thus, making it more capable of addressing the more demanding data based applications. However, when looking at the average delay of the delayed packets, Figure 4.22 shows that for the MAX approach, the average delay can be significantly degraded when the weight of data services increase: if this approach is better from the percentage of delayed packets point of view, it can be very unfair to users that are located farther from the BS or that are more interfered, and this unfairness may become even worse if the system load increases. It is clear that the gap between the average delay for MAX, compared to either PF or RR, tends to increase when the weight of data services increases as well.



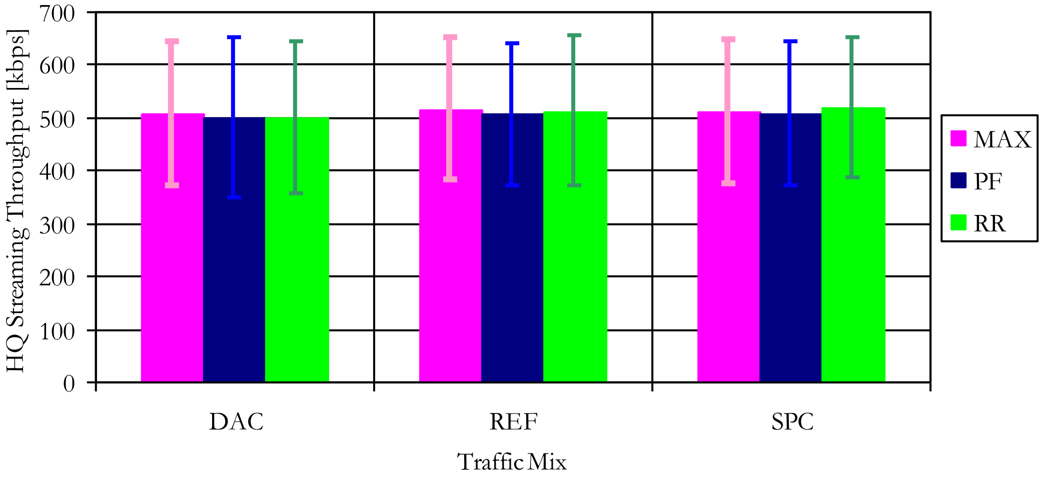
a) Delayed packets



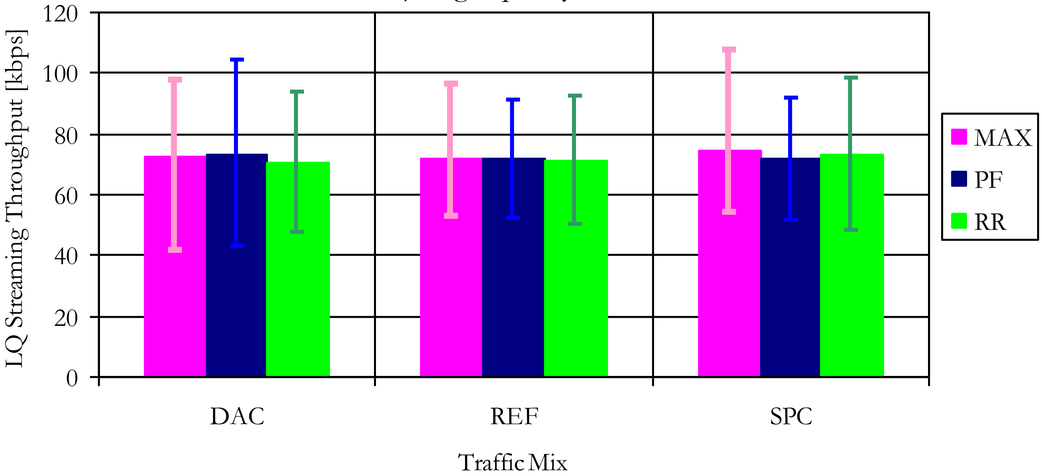
b) Delay of delayed packets

Figure 4.22 – Percentage of BE delayed packets and average delay of delayed BE packets for variable per user traffic mix – Dense Urban.

Throughput figures for streaming video are illustrated in Figure 4.23. Following the same trend already described in Sub-section 4.2.3, no clear differences are seen for streaming throughput when the traffic mix changes. This is supported also by the data provided by Figure 4.24 and Figure 4.25, meaning that, again, impact on user throughput must be analysed through FTP performance.



a) High quality video



b) Low quality video

Figure 4.23 – Average streaming throughput and standard deviation for variable per user traffic mix – Dense Urban.

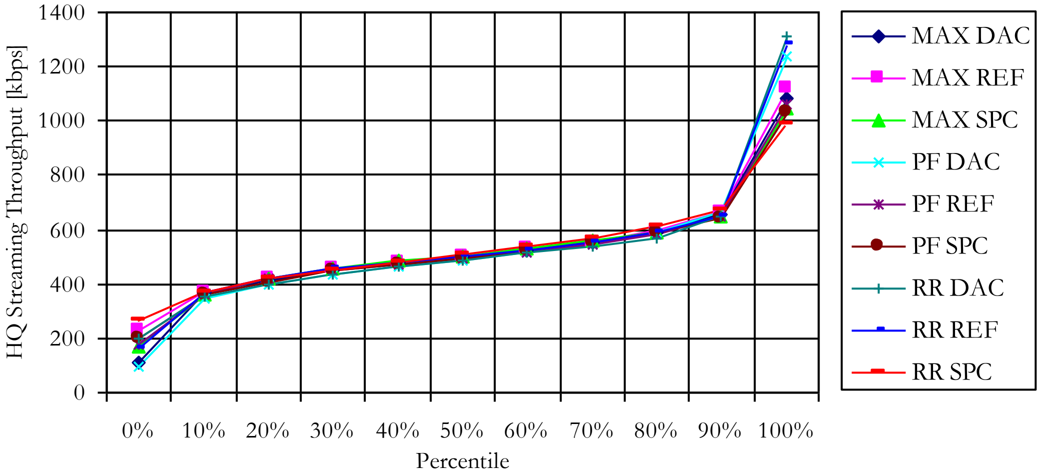


Figure 4.24 – High quality streaming throughput cumulative distribution for variable per user traffic mix – Dense Urban.

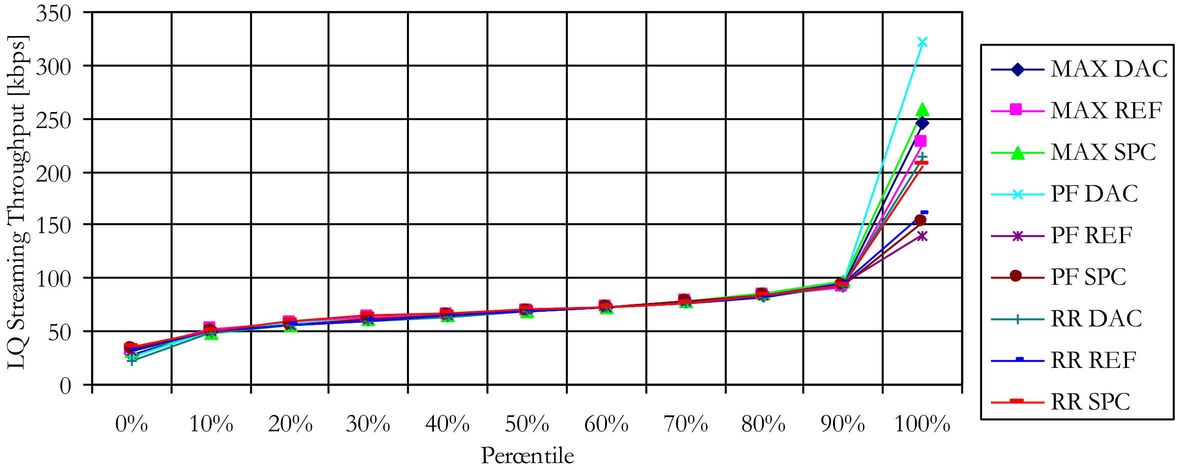


Figure 4.25 – Low quality streaming throughput cumulative distribution for variable per user traffic mix – Dense Urban.

FTP performance is degraded when the weight of data services increases, regardless of the scheduling algorithm considered. While for the REF scenario, the RR approach was the one providing the best average throughput, when the DAC traffic mix is considered, this advantage vanishes, and RR has a similar average throughput as the MAX algorithm. PF remains as the approach providing worse FTP average performance for any of the traffic mixes considered.

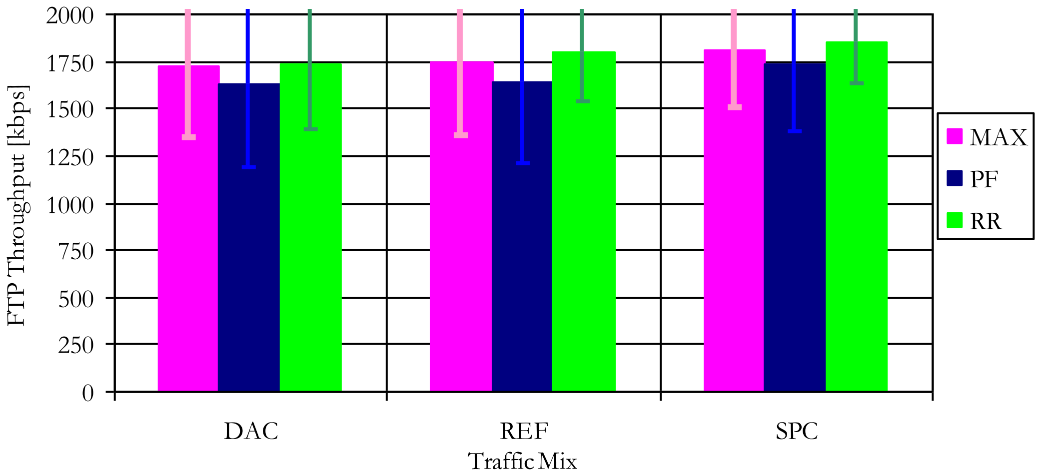


Figure 4.26 – Average per user FTP throughput and standard deviation for variable per user traffic mix – Dense Urban.

The cumulative distributions for FTP throughput for the different traffic mixes depend on the scheduling algorithm implemented, as shown by Figure 4.27.

- When the RR scheduling is in place, there is a clear dependence between throughput and the weight of data services: the throughput has a decrease tendency when one moves from a SPC traffic mix towards a DAC one.

* When either MAX or PF scheduling is implemented, SPC provides the best curve, as expected, but no significant differences are seen when comparing REF to DAC, meaning that these scheduling approaches are less influenced by an increase of load due to changing traffic mix conditions.

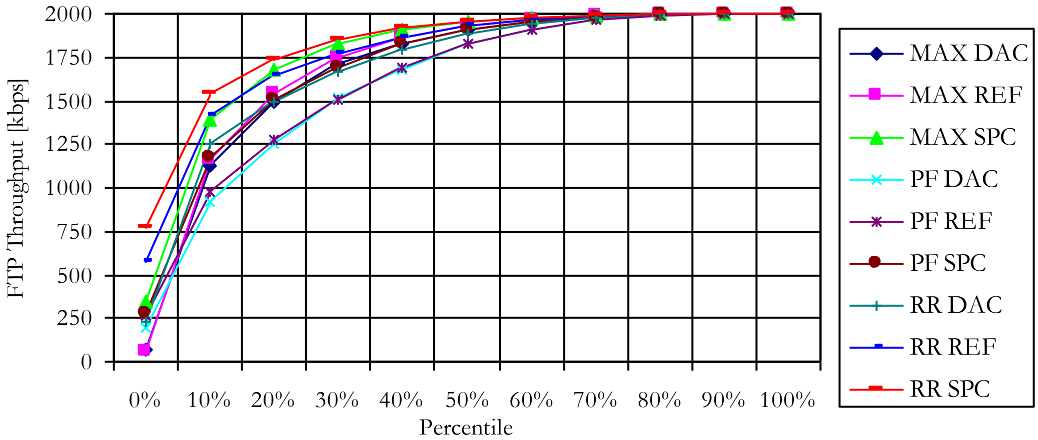


Figure 4.27 – FTP throughput cumulative distribution for variable per user traffic mix – Dense Urban.

In conclusion, the choice of the scheduling algorithm must not only consider the operators's objectives in terms of conciliation of QoS requirements for the different applications, but also the fact that the usage of the different applications can change dramatically alongside a single day period. The conclusions issued in Sub-section 4.2.3 highlighted the tradeoff decisions that must be done between priotising between real-time or BE services, highlighting the fact that if, e.g., it is mandatory to provide full accessibility to services like VoIP, then, services like WWW or FTP will be slower in terms of bitrate. A detailed analysis of different traffic mixes makes this problem even more complex, as the operator shall not look only at the particular traffic mix of one particular period, but rather to the different traffic mixes that are offered to the network at different periods of the day. Results summarised in the present section provide some useful information to this task, showing, for instance, that the RR approach is the one more subjected to the influence of the traffic mix, which is explained by the fact that there in no intelligence at all behind this algorithm. On the contrary, both the MAX and PF approaches are the ones that provide more security to an operator when considering the variable traffic mix problem.

System capacity is also affected by the number of simultaneously addressed users, as the MAP efficiency decreases when more users are served at the same time. This is particulary important, since if continuous and short burst applications, such as VoIP, are dominant, then, the associated MAP overhead is higher and the overall system capacity is reduced, whereas when heavy and short term applications, such as WWW, prevail, MAP efficiency is higher, and so is system capacity. Thus, a WiMAX operator must account for such differences, as it is expected that higher system capacity is available during data services peak hours, while the same capacity is decreased during speech ones. The scheduling policy also has an effect at this level, since algorithms like MAX, although maximising user throughput, result in lower system efficiency, as a consequence of allowing for more users to be simultaneously addressed. The same trend occurs when PF is applied, since this approach prioritises VoIP services, allowing for more users to be addressed simultaneously due to the low capacity requirements of VoIP.

PF, due to its QoS mechanism that prioritises real-time services over BE ones, provides the best accessibility and reliability for both VoIP and Video Call. This scheduling approach is also best suited to cope with changing traffic mix conditions alongside a full 24 hours day period. On the contrary the RR mechanism provides the worse real-time services performance, being highly subjected to traffic mix variations, and strongly penalised when there is an increase of the weight of more demanding applications, i.e., data ones. The same trend is applied for BE services performance: all scheduling approaches perform better from the FTP throughput point of view in the case of SPC traffic mix, but the RR approach is the one that is more negatively affected when the weight of data services increases (due to system load increase).

4.4 Variable Network Configuration Scenarios

In this section, one analyses the system response to variable network configuration figures. Again, simulations are carried out for the 4 considered radio environents (present section only addresses Dense Urban, with results for Urban, Suburban and Rural being summarised in Annex I and J) and for each of these scenarios, results are obtained when using individually each of the three implemented scheduling algorithms. For each set of simulations, the number of users covered by the BS is kept unchanged, based on the results of Section 4.2 (5250 for DU, U and SU scenarios, 8000 for RU scenarios), as well as the traffic mix, set to REF as summarised in Table 4.1. The effect of the channel bandwidth (5 vs. 10 MHz) is assessed. The TDD split (2:1 vs. 3:1) theoreticall impact is briefly explained without detailed analysis. The section finishes by again considering a 10 MHz channel bandwidth, and by assessing the impact of different Admission Control and Congestion Control thresholds for RTPS calls when using the PF approach.

4.4.1 Channel Bandwidth Impact and TDD Split

Results summarised in Sections 4.2 and 4.3 considered a channel bandwidth of 10 MHz. As detailed in Sub-section 2.2.1, WiMAX is flexibile enough so that different channel bandwidths can be considered. In the present section, results for a 5 MHz channel bandwidth are presented, considering the REF scenario analysed in Section 4.2, and a total number of users covered by the BS of 5250, i.e., the number assumed for Section 4.3 that resulted in a VoIP Refused Call Rate around 1%. Cell radius is not changed, i.e., the calculated cell radius for 10 MHz is kept when 5 MHz is considered.

By comparing Figure 4.28 with Figure 4.13, the reduction of the total number of simultaneously served users is clear. With a 10 MHz channel bandwidth, around 21 simultaneous users are served, while with 5 MHz this figure varies between 15 to 20, depending on the scheduling algorithm. The reduction is simply due to the fact that the system does not have enough resources to handle the same number of users, thus, some of these are blocked by Admission Control. The exception is only for the PF scheduling, as this policy implements the necessary QoS mechanism that handles VoIP users in first hand.

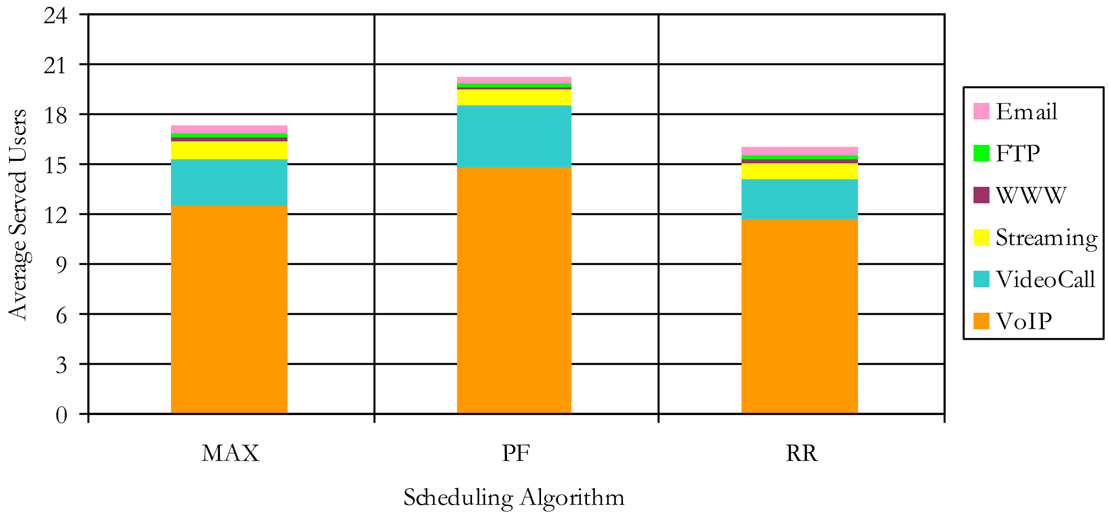


Figure 4.28 – Average number of served users for 5MHz channel bandwidth – Dense Urban.

As illustrated by Figure 4.29, the system is completely loaded, as the average load is above 80%, which explains the lack of capacity of the system to handle the same number of users as for a 10 MHz channel bandwidth.

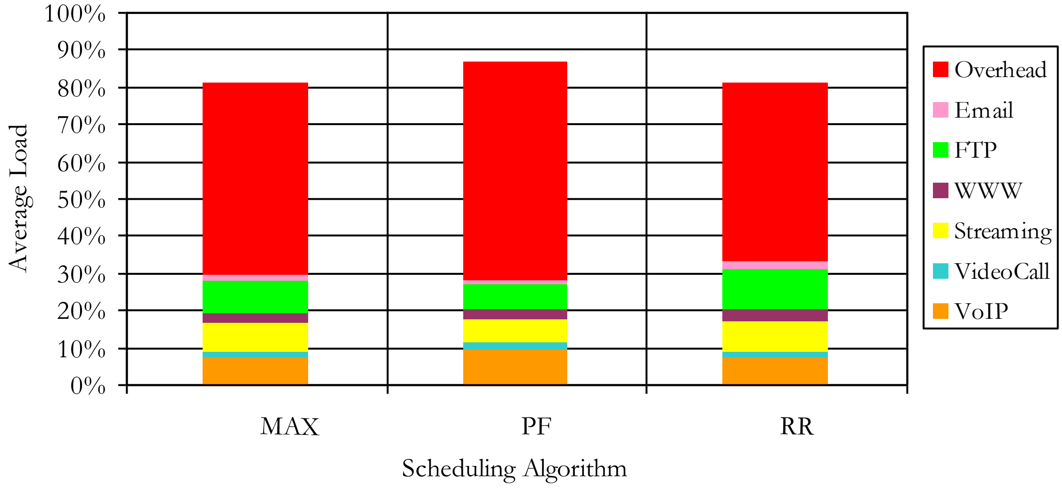


Figure 4.29 – System load for 5MHz channel bandwidth – Dense Urban.

A consequence of the channel bandwidth reduction is the link budget improvement. Since the cell radius is kept unchanged when comparing 10 to 5 MHz channel bandwidth performance, system noise is reduced by 3 dB, as explained in AnnexA. This results in a shift in the AMC usage distribution towards better AMCs. Figure 4.17 shows that for 10 MHz, QPSK1/2 use is near 40%, while it decreases to less than 30% when 5 MHz is used, as illustrated by Figure 4.30.

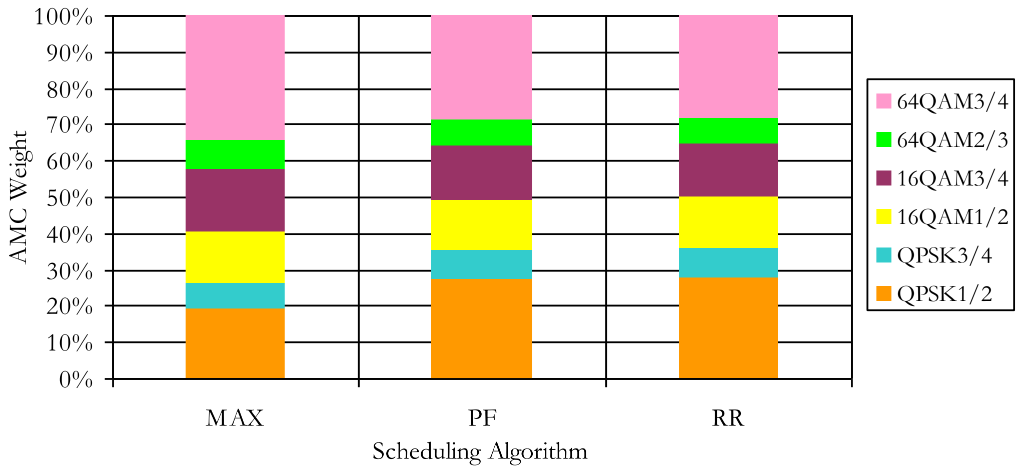
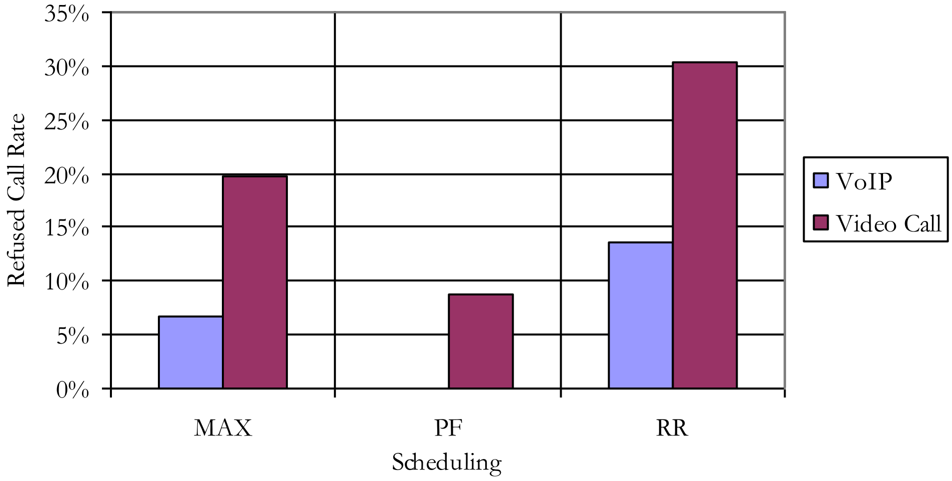
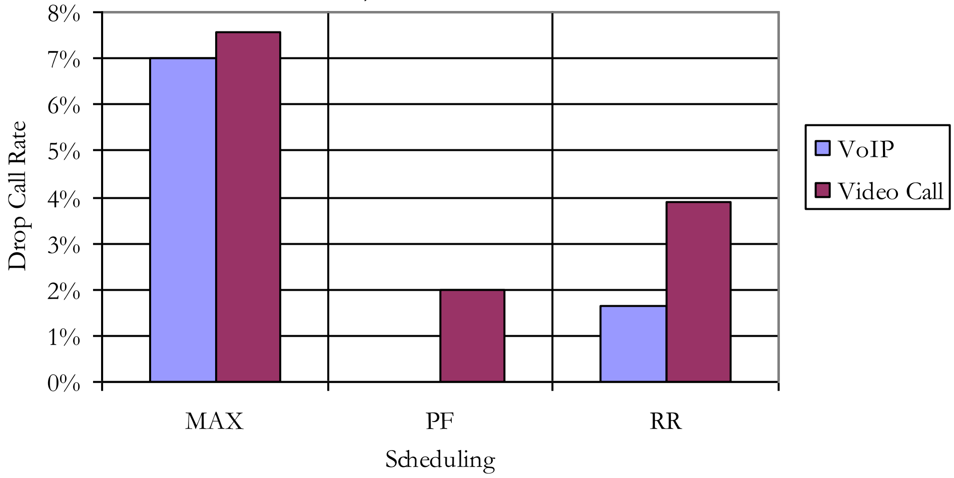


Figure 4.30 – AMC distribution for 5MHz channel bandwidth – Dense Urban.

The negative impact on the main performance indicators of real-time aplications is clear, although when the PF mechanism is considered still no blocking is seen for VoIP. For MAX and RR, whereas a blocking around 1% is achieved with 10 MHz, this figure is raised to values between 6 to 14% when 5 MHz is used. Since the available system resources is proportional to the channel bandwidth in use, this means that there is not a linear relation between the number of users that the system can address and system capacity, otherwise one would expect to see the Refused Call Rate doubling and not growing 6 to 14 times more.



a) Refused Call Rate



b) Drop Call Rate.

Figure 4.31 – RTPS main performance indicators for 5MHz channel bandwidth – Dense Urban.

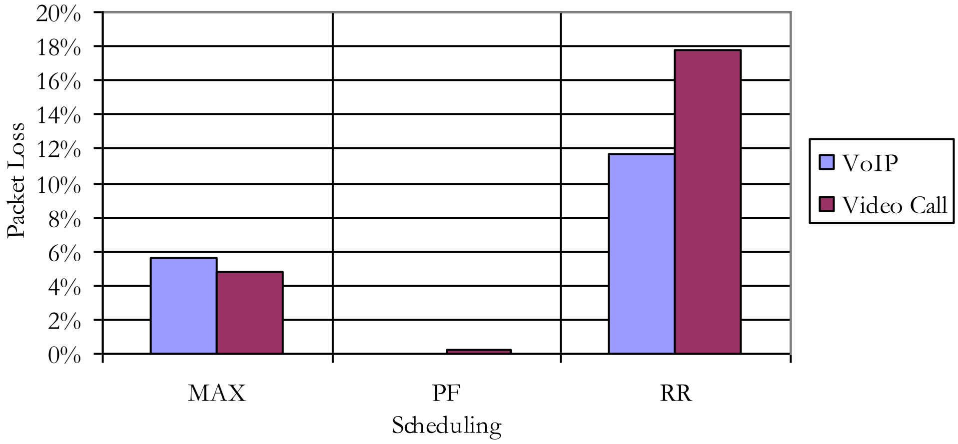


Figure 4.32 – RTPS packet loss for 5MHz channel bandwith.

Naturally, BE services are also severely affected. With MAX scheduling, which provides the lower percentage of delayed packets, there are around 10% of packets delayed, while with 10 MHz only 5% of packets are delayed, according to Figure 4.22. The degradation with PF and RR is even worse, with the percentage of delayed packets growing from 15% to 55 and 40%, respectively. In the same way, for delayed packets, the average delay is also increased significantly.

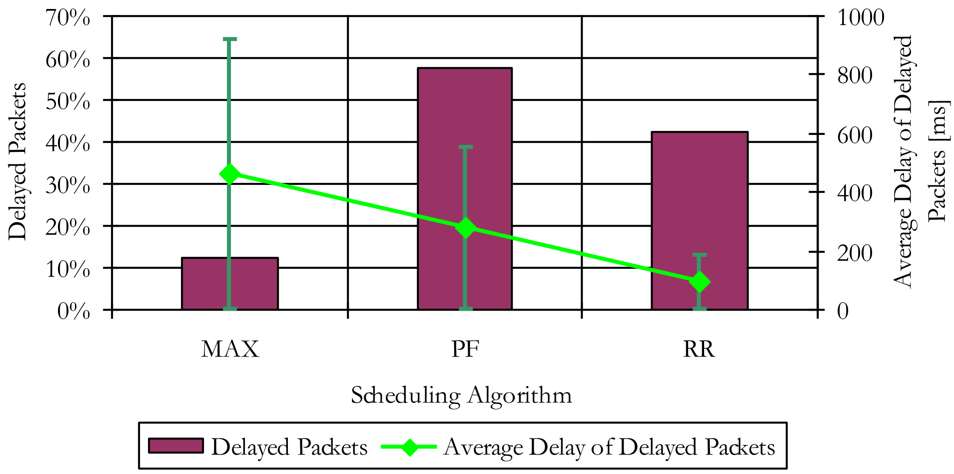


Figure 4.33 – BE packet delay for 5MHz channel bandwidth – Dense Urban.

One of the consequences in the increase of delay is the degradation of the average FTP throughput. With a 10MHz channel bandwidth, the average FTP throughput is around 1600 to 1750 kbps, Figure 4.26, but with 5 MHz, the best case is only around 1000 kbps. The cumulative distribution for FTP throughput, illustrated in Figure 4.34, shows that, for instance, for a RR scheduling, only 30% of users have a throughput above 1750 kbps, while this is the average throughput value with a 10MHz channel bandwidth. With PF, FTP throughput gets really low, as the few system resources available are mainly allocated for VoIP users, leaving very few resources for the remaining applications, in particular, for FTP: only 10% of users have throughputs above 1 Mbps.

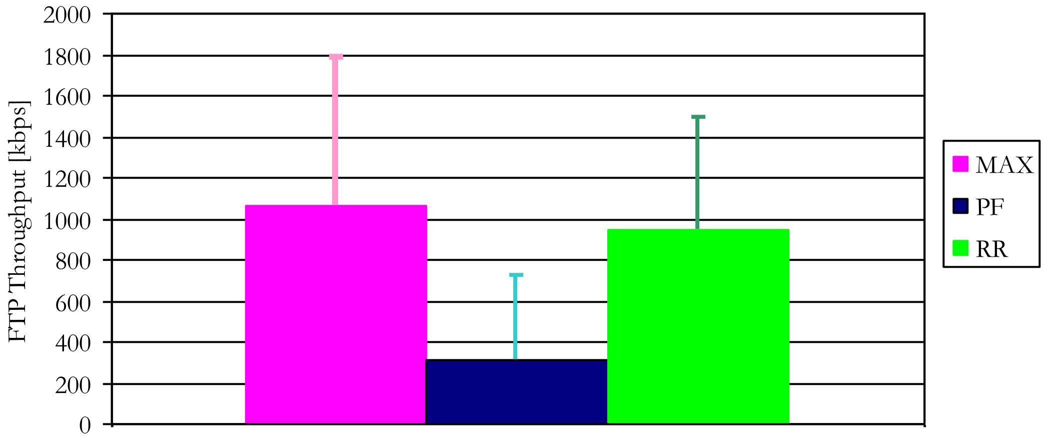


Figure 4.34 – Average FTP throughput for 5MHz channel bandwidth – Dense Urban.

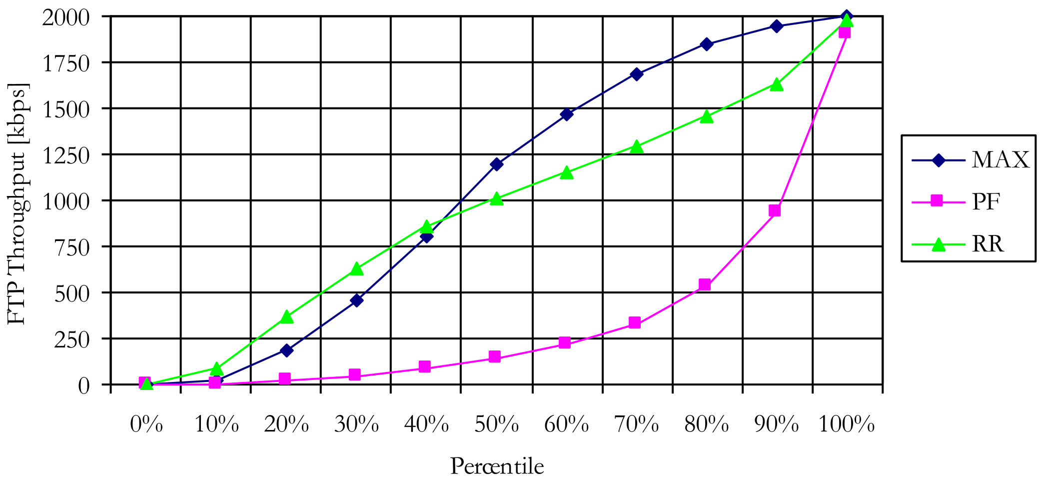


Figure 4.35 – Cumulative per user FTP throughput distribution for 5MHz channel bandwidth – Dense Urban.

Concerning TDD split, as detailed in Sub-section 2.2.2, 802.16e foresees the possibility of different splits to be used. The decision between a 1:1, 2:1 or 3:1 splits mostly depend on the type of traffic carried by a network:

- Networks dominated by services such as VoIP or Video Call will most likely tend to use a 1:1 approach, as this type of traffic is symmetric.
- Networks dominated by data applications such as WWW, Streaming or FTP will most likely use splits like 2:1 or even 3:1.
- More recently, new applications, such as Peer-to-Peer, although being data applications like for instance FTP, are based on the simultaneous reception and transmission of data files (movies, music, etc.), demanding for resources in both UL and DL, i.e., 1:1 kind of splits.

Generally, the adoption of, e.g., a 1:1 TDD split instead of the 2:1 one considered in the presented results results in a lower availability of DL resources and in a higher availability of UL ones. Thus, the impact on DL is similar to the one resulting from the adoption of a 5 MHz channel bandwidth instead of 10 MHz. and results have the same trend as the ones already presented in Sub-section 4.4.1, while the impact of UL is outside the scope of the present thesis. Consequently, since no detailed results on the commitment between the need of UL and DL resources can be presented, and in order not to repeat results that would be somehow redundant to the ones presented in Sub-section 4.4.1, no quantitative results on the influence of TDD split are presented.

4.4.2 Admission/Congestion Control Configuration for PF Scheduling

The previous results taking the PF approach considered 100% thresholds for both VoIP and Video Call Admission Control and Congestion Control algorithms. Results show that this approach performs very well from an RTPS service point of view, with very low Refused Call Rate and Drop Call Rate figures, but is worse concerning BE services. In order to find out a better commitment between these two types of applications, Admission Control and Congestion Control may be configured diferently, so that a tolerable VoIP and Video Call packet loss and/or blocking can be achieved, releasing resources to be used for the remaining applications, while still satisfying basic QoS requirements of RTPS calls. Configurations summarised in Table 4.5 were simulated.

Table 4.5 – Admission Control and Congestion Control configurations analysed.

	VoIP		Video Call	
Configuration	**Admission Control [%]**	**Congestion Control [%]**	**Admission Control [%]**	**Congestion Control [%]**
#1	35	35	55	55
#2	40	40	60	60
#3	45	45	65	65
#4	50	50	70	70
#5	40	25	60	35
#6	40	30	60	40
#7	40	35	60	45
#8	40	40	60	50
#9	100	100	100	100

Generally:

- Configurations #1 to #4 have the same thresholds for both Admission Control and Congestion Control, being only different between VoIP and Video Call. These thresholds are lower for configuration #1 and higher for configuration #4.
- Configurations #5 to #8 have an unique Admission Control thresholds for VoIP (40%) and Video Call (60%), while the respective Congestion Control thresholds grow from 25% to 40% and 35% to 50%, respectively.
- Configuration #9 is the one considered for PF results presented in Section 4.2 (REF traffic mix), with 100% thresholds considered for the two algorithms and the two services as well.

The impact of different configurations for either Admission or Congestion Control on the number of simultaneously served VoIP and Video Call users is clear, Figure 4.36. For configurations #1 to #3, there is an increase trend of served VoIP users, as more users are allowed by the system, while from configuration #3 to #4 there is no visible variation. Configurations #5 to #8 also shows the same growth trend, as expected since Congestion Control becomes gradually less restrictive. It is also visible that previous system load results (configuration #9) are, roughly, matched by either configuration #3, #4 or #8. Thus, indeed a proper setting of Admission and Congestion Control gives a WiMAX operator control over how to distribute resources between RTPS and BE services.

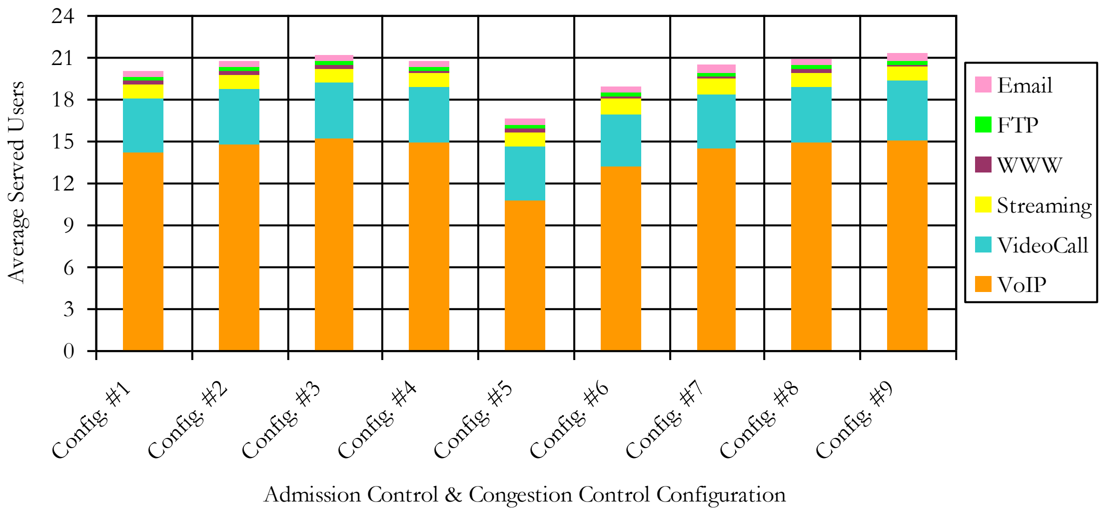


Figure 4.36 – Impact of different Admission and Congestion Control thresholds on average served users – Dense Urban.

As expected, the same trend is seen when considering network load, which is increased when the thresholds become less restrictive, Figure 4.37. An additional improvement coming from the adoption of limitations to the number of RTPS services concerns the reduction of overhead, mostly because of the lower number of VoIP users served.

Admission and Congestion Control configurations provide an important role on system capacity as well, since an operator can trade resources for RTPS services, which originate less traffic but consume a lot of resources due to overhead, by resources for BE ones, which originate more traffic. This is illustrated by Figure 4.38, which shows that there are configurations that allow more traffic to be carried than others. When Admission and Congestion Control are set too low, it is likely that some RTPS users are blocked, and, although maximum resources should be free for BE services, no maximum traffic volume is carried. This is clear, e.g., for configuration #5. On the other hand, too restrictive thresholds for Admission and Congestion Control allow for resources to be fully used by RTPS services, but this 'steals' capacity for BE ones, meaning that less traffic is carried. Thus, from a system capacity point of view, a tradeoff solution is the most recommended one.

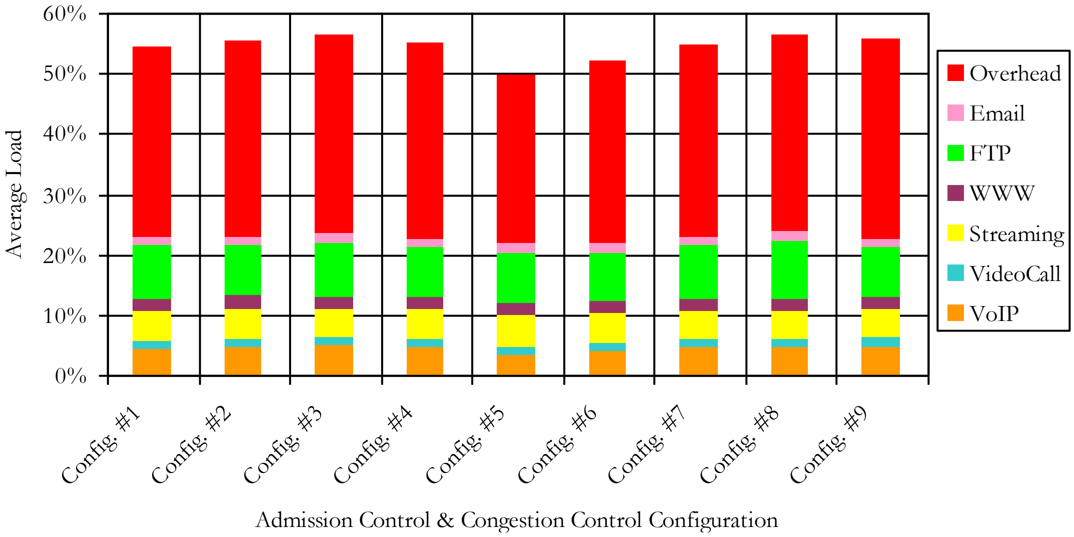


Figure 4.37 - Impact of different Admission and Congestion Control thresholds on network load – Dense Urban.

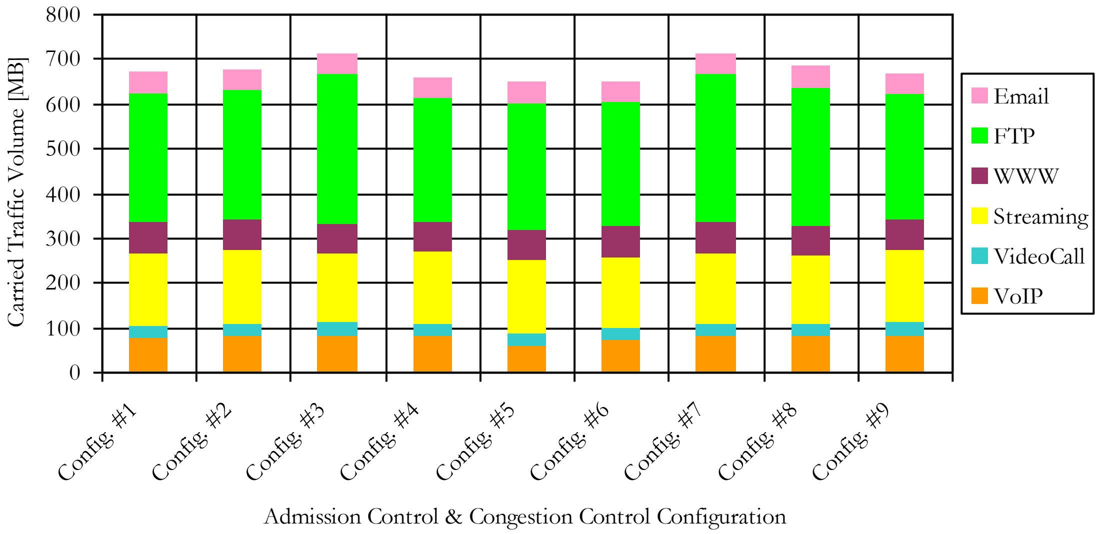


Figure 4.38 - Impact of different Admission and Congestion Control thresholds on carried traffic volume – Dense Urban.

As illustrated by Figure 4.39, AMC distribution is also affected by the configuration assumed for Admission and Congestion Control. More restrictive configurations as, e.g., configuration #5 have a higher probability of blocking RTPS users suffering from worse AMC, as these require more resources for the same data volume. Thus, RTPS users subject to worse AMC may not be served, leaving resources free for BE users that experience better AMC condititions, which results in an overall higher system capacity.

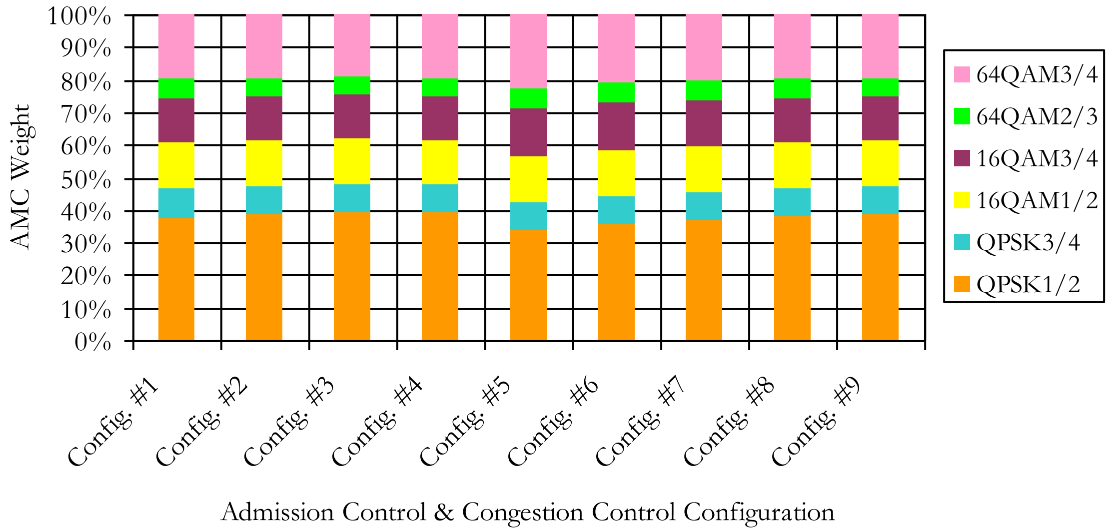


Figure 4.39 – Impact of different Admission and Congestion Control thresholds on AMC Distribution – Dense Urban.

As expected, there is a clear dependence between Admission Control thresholds and Refused Call Rate, as well as between Congestion Control and Drop Call Rate, Figure 4.40. For increasing Admission Control thresholds, Refused Call Rate becomes lower, while for increasing Congestion Control thresholds it is the Drop Call Rate that becomes lower. For an abstract target of more or less 1% for each, configurations #2 and 8# are enough.

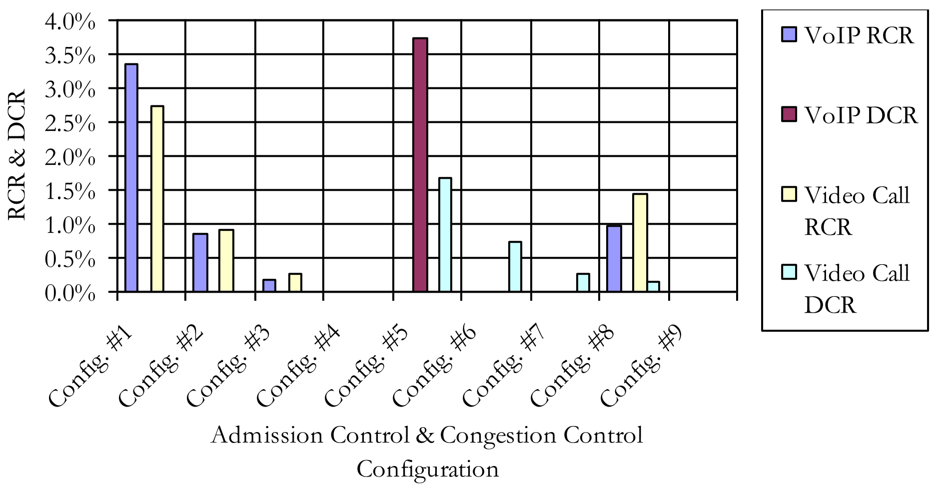


Figure 4.40 - Impact of different Admission and Congestion Control thresholds on RTPS services main performance indicators – Dense Urban.

Together with Figure 4.40, Figure 4.41 shows that Admission Control and Congestion Control thresholds cannot be set individually, as the impact of each is seen at a different level. For instance, with configuration #5 or #6, although VoIP Refused Call Rate figures are close to 0% due to the high enough Admission Control settings, many packets are lost due to a too restrictive

Congestion Control configuration, resulting in unnaceptable quality from an end-user perspective. Thus, it makes no sense to, e.g., allow many calls to come in if these are to be strongly restricted later on.

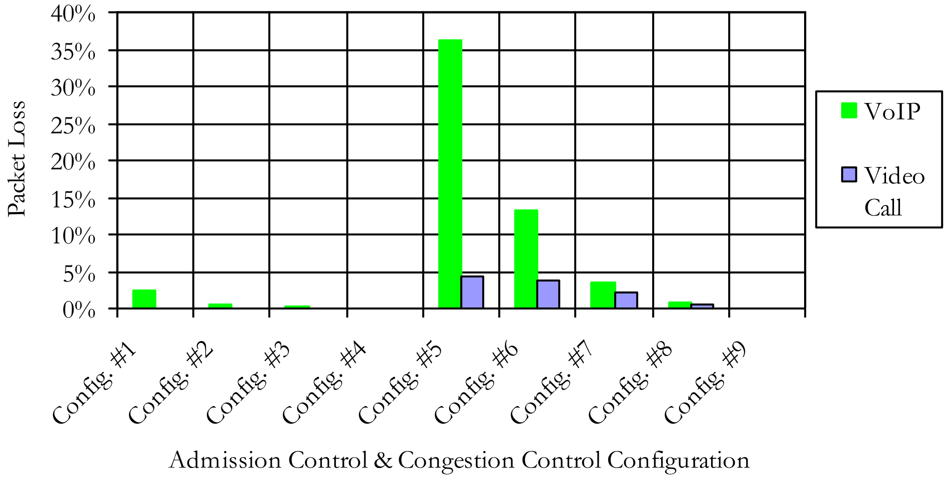


Figure 4.41 - Impact of different Admission and Congestion Control thresholds on RTPS services packet loss – Dense Urban.

Naturally, BE services performance has an opposite trend. Delay performance is illustrated by Figure 4.42, where it can be seen that the percentage of delayed packets becomes higher when the configuration for both Admission and Congestion Control algorithms becomes less restrictive. It is interesting to notice as well that, for configurations #5 and #6, although the percentage of delayed packets is lower due to more resources being available for BE applications, the average delay of delayed packets is higher. This is explained by the fact that these delayed packets tend to be always the ones coming from users suffering from worst AMC.

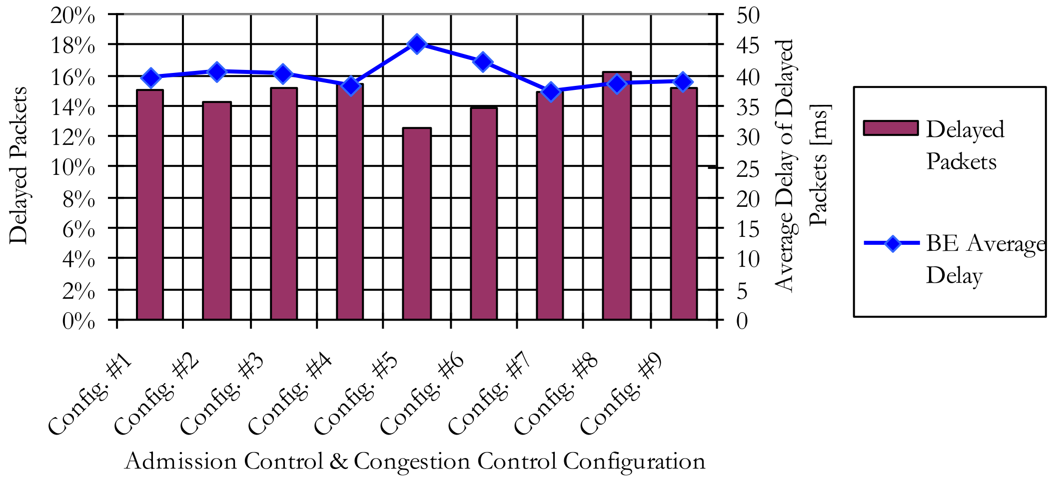


Figure 4.42 - Impact of different Admission and Congestion Control thresholds on BE services delay – Dense Urban.

FTP throughput figures are illustrated by Figure 4.43 and Figure 4.44. Configurations that are more restrictive for RTPS services provide higher FTP throughputs, and vice-versa. It is also important to notice that configuration #9, being the one that is the less restrictive of all for RTPS services, is the one that provides worse FTP throughput. Thus, according to Figure 4.40, it may be preferable to go for a configuration like #2, which still provides acceptable RTPS performance indicators, and is able to increase FTP throughput significantly, providing the needed tradeoff between the various types of applications.

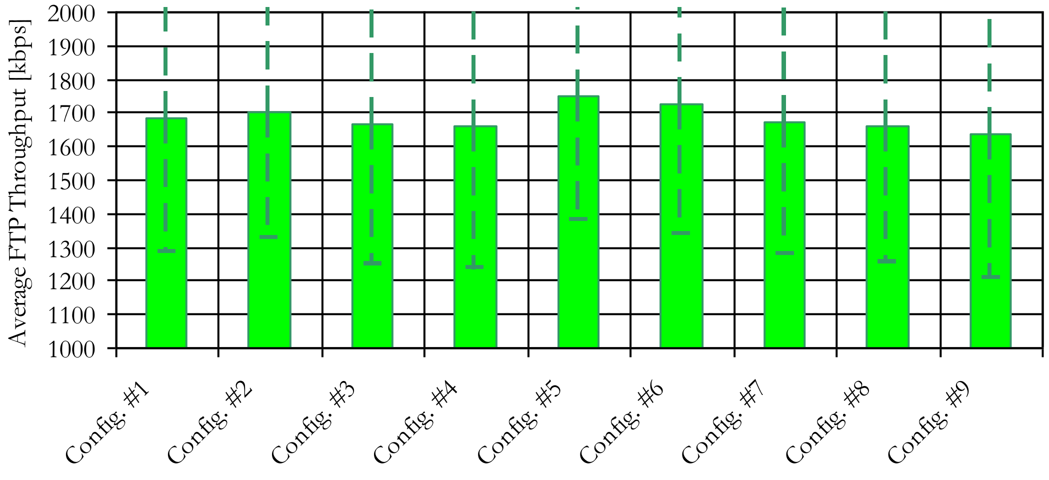


Figure 4.43 - Impact of different Admission and Congestion Control thresholds on average FTP throughput – Dense Urban.

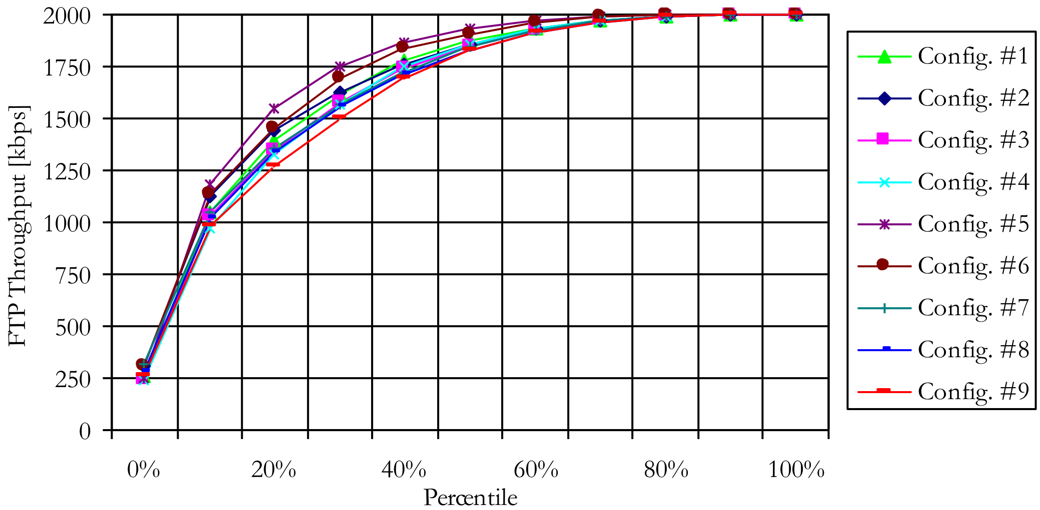


Figure 4.44 – Cumulative FTP throughput for different Admission and Congestion Control thresholds – Dense Urban.

In conclusion, WiMAX, being a flexible system based on the concept of SOFDMA, provides means for operators to tradeoff capacity by coverage, through the use of different channelisations. Higher channelisations provide more system capacity at the cost of less coverage and vice-versa, which is a completely new concept introduced by this standard. An operator can, for instance, start a deployment using lower channelisations, maximising the initial rollout coverage for a reduced number of users. With time, customer base trends to increase, and the operator can, later on, when the demand for capacity grows, change to a higher channelisation and deploy additional sites to fill in eventual coverage holes that appear due to the increase in channelisation.

Varying the TDD split is another concept included in WiMAX, but the rational behind its configuration is quite different. In this case, it is rather the type of traffic mix that rules its configuration, with the symmetrical kind of applications, like VoIP or Video Call, needing 1:1 splits, whereas appllications such as FTP or WWW need splits like 2:1 or even 3:1.

Concerning the configuration of Admission and Congestion Control thresholds, and following the remarks summarised in the end of Section 4.2, clearly algorithms that are able to provide the necessary QoS mechanism that ensure the necessary accessibility and reliability for RTPS services, like VoIP, not in a 'blind' way but rather allowing for some blocking or even droppping to occur in certain conditions and under acceptable levels, appear as being more interesting from a technical point of view. Instead of guaranteeing near perfect levels for this type of services, with a penalty for the remaining BE kind of services, at least basic performance levels are also guaranteed for BE services, at the cost of a slight degradation of RTPS indicators.

Chapter 5

Conclusions

This chapter summarises the main conclusions of the present work and ideas for future work.

This thesis investigates the impact of different scheduling approaches for providing both real-time and BE based services through 802.16e based systems. 802.16e has been developed with many objectives in mind, including support for flexible architecture (fully IP based system), dynamic optimisation of traffic mix through proper mechanisms, mobility, and high capacity.

802.16e based systems face the challenge of supporting an extensive variety of services with ever increasing demands for higher data rates, and with a wide diversity of QoS needs. Moreover, the expected increase for traffic demand drives operators to require larger system capacity to sustain their business growth.

Under that scenario, this M.Sc. thesis has assessed 802.16e as a potential radio access technology that can facilitate the evolution of the wireless communication market by satisfying the demands of end users and operators. Regarding BE services, the focus has been on Streaming, WWW, FTP and Email traffic, due to its enormous success on the Internet. Regarding Real Time services, the focus has been on VoIP and Video Call services, because they are still (especially voice) a relevant traffic type to be conveyed.

802.16e introduces a new concept that does not exist, so far, in current cellular networks. In UMTS, for instance, there is a clear split between data services, mostly carried over HSxPA, and conversational services (speech), carried over dedicated channels. Due to this division, the DL shared resource in UMTS, which is the BS power, is mostly divided between dedicated channels and HSxPA, with the latter usually resulting from the power not used for dedicated channels, i.e., working in a BE way. Thus, while for conversational services there are no significant problems concerning scheduling (a kind of fixed pipe is implemented for each active call), concerning HSxPA there are already different scheduling algorithms that rule scheduling, generally similar to the ones implemented in the developed simulator. However, UMTS/HSxPA currently is only used for data applications, such as FTP or WWW, whereas speech or video call are still carried by real-time dedicated radio bearers, thus, not subject to the same scheduling decisions, meaning that the main criteria behind HSxPA scheduling is related to the instantaneous channel conditions and long term throughput for all data applications, with no differentiation at all between applications. However, in the future, with the introduction of VoIP over UMTS/HSxPA systems, the same kind of decisions and algorithms will have to consider the existence of speech services as well, with all its requisites in terms of latency, packet loss, etc., meaning that UMTS/HSxPA scheduling will substancially become more complex. On the other hand, 802.16e scheduling algorithms must incorporate the existence of both voice and data based services from the beginning.

WiMAX, with its capability of differentiating services, and its flexibility in terms of AMC allocation depending on radio channel conditions, provides a significant margin for intelligent scheduling mechanisms to be implemented. A good algorithm may have input variables such as the instantaneous radio channel conditions felt by the different users, the medium/long term average rate of each user, the recent delay experimented by the different users, or the tolerable packet loss for real-time services, without jeopardising the overall quality. Thus, an equipment supplier clearly has room for differentiation, as it can develop algorithms that combine all these variables in the search for an algorithm that is capable of satisfying the different requirements associated to the different applications and be adapted by the operators according to their own goals in terms of QoS.

The performance of 802.16e is assessed for different scheduling approaches and for different traffic inputs, for which a simulator was studied and developed. Results clearly show that the system does not have infinite resources, and, in order to favour a given kind of applications, others are penalised. A system supplier or operator must then decide how the system shall be designed or configured, according to the defined targets in terms of overall QoS, knowing that, regardless of the approach assumed, there are always advanges and disadvanges associated to each option. The scheduling approach has a major role on this.

Results show that there is a clear tradeoff between the throughputs achievable by BE applications and the accessibility and retainability of RTPS ones. PF is the policy providing best performance for RTPS applications, due to its QoS mechanism combined with 100 % thresholds for Admission Control and Congestion Control of RTPS services, while the RR policy provides the worse performance from this point of view. However, the opposite occurst when considering BE services: RR provides the best results and PF the worse ones. MAX approach maximises BE throughput, but it can severely penalise users suffering from worse AMC. It is also clear that the PF implementation, although useful to guarantee real-time applications accessibility, is smart enough in order to discard part of the real-time traffic, as both a reduced packet loss and reduced blocking probability are tolerable.

Concerning VoIP Refused Call Rate, it is seen that for, e.g., a 1 % target blocking probability, the MAX approach can manage up to 26 % more users than the RR one (for a reference traffix mix), while PF provides no blocking at all for the same number of users. The price to pay concerns performance of BE services as, for instance, FTP achieves an average throughput of 1804 kbps when RR is used, against 1744 kbps of MAX and 1636 kbps of PF. The operator trades resources for RTPS services by resources for BE ones, knowing that if better accessibility figures are required for RTPS, BE services are subject to higher delays. However, when PF is implemented, thresholds used to block new call requests or to block resources for on-going calls can be changed so that tolerable losses are introduced for RTPS services, either by blocking some of these or by dropping some packets for on-going calls, which results in an overall improvement of BE services performance.

It is shown that the choice of the scheduling algorithm must not only consider the operators's objectives in terms of conciliation of QoS requirements for the different applications, but also the fact that the usage of the different applications can change dramatically alongside a single day period. A detailed analysis of different traffic mixes makes the scheduling issue even more complex, as the operator shall not look only at the particular traffic mix of one particular period, but rather to the different traffic mixes that are offered to the network at different periods of the day. Generally, if from a reference traffic mix VoIP weight increases 15 %, with subsequent decrease of remaining services' weight, throughput can increase up to 6 % and VoIP blocking can be improved by up to 50 %, depending on the scheduling algorithm. On the contrary, if the weight of VoIP is decreased by 15 %, with subsequent increase of the remaining services, throughput can be degraded up to 4 %, while VoIP blocking can be degraded by up to 66 %. It is seen that basic RR approaches are the ones more subject to the influence of the traffic mix, which is explained by the fact that there in no intelligence at all behind this algorithm. On the contrary, both the MAX and the PF approaches provide more security to an operator when considering the variable traffic mix problem. PF, due to its QoS mechanism that prioritises real-time services over BE ones, provides the best accessibility and reliability for both VoIP and Video Call, as this scheduling approach is the best suited to cope with changing traffic mix conditions alongside a full 24 hours day period. The same trend is applied for BE services performance: all scheduling approaches perform better from the FTP throughput point of view, in the case of SPC traffic mix, but the RR approach is the one that is more negatively affected when the weight of data services increases (due to system load increase).

Results also highlight the MAP efficiency issue of WiMAX, as system capacity is affected by the number of simultaneously addressed users, with MAP efficiency decreasing when more users are served at the same time. This is particulary important, since if continuous and short burst applications, such as VoIP, are dominant, then the associated MAP overhead is higher and the overall system capacity is reduced, whereas when heavy and short term applications, such as

WWW, prevail, MAP efficiency is higher, and so is system capacity. Thus, an operator must account for such differences, as it is expected that higher system capacity is available during data services peak hours, while the same capacity is decreased during speech ones. The scheduling policy also has an effect at this level, since algorithms like MAX, although maximising user throughput, result in lower system efficiency, since they allow for more users to be simultaneously addressed. The same trend occurs when PF is applied, since this approach prioritises VoIP services, allowing for more users to be addressed simultaneously due to the low capacity requirements of VoIP.

Beyond the type of scheduling approach adopted, 802.16e, being a flexible system based on the concept of SOFDMA, provides means for operators to tradeoff capacity by coverage, through the use of different channelisations. Higher channelisations provide more system capacity at the cost of less coverage, and vice-versa, which is a completely new concept introduced by this standard. An operator can, for instance, start a WiMAX deployment using lower channelisations, maximising the initial rollout coverage for a reduced number of users. With time, the customer base trends to increase, and the operator can later on, when the demand for capacity grows, change to a higher channelisation and deploy additional sites to fill in eventual coverage holes that appear due to the increase in channelisation. Varying TDD split is another concept included in WiMAX, but the rational behind its configuration is quite different. In this case, it is rather the type of traffic mix that rules its configuration, with the symmetrical kind of applications, like VoIP or Video Call, needing for 1:1 splits, whereas appllications such as FTP or WWW need splits like 2:1 or even 3:1.

Concerning the configuration of Admission and Congestion Control thresholds, these algorithms are able to provide the necessary QoS mechanism that ensure the necessary accessibility and reliability for RTPS services, like VoIP, not in a 'blind' way but rather allowing for some blocking or even droppping to occur in certain conditions and under acceptable levels, and seem to be more interesting from a technical point of view. Instead of guaranteeing near perfect levels for this type of services, with a penalty for the remaining BE kind of services, at least basic performance levels are also guaranteed for BE services, at the cost of a slight degradation of RTPS indicators.

Despite all the advantages and disadvantages highlighted in the previous paragrpahs, the decision among a RR, MAX or PF approach is not merely technical, as an operator can, for instance, decide to implement a cheaper algorithm, like RR, instead of buying a more intelligent algorithm, such as MAX or PF, and decide to deploy BS closer to each other, limiting the number of users covered by the BS, thus, mitigating the accessibility problems for RTPS applications that are seen with RR. On the contrary, another operator may wish to be guaranteed in terms of real-time services QoS, and decide to use a PF scheduling algorithm, which certainly has an additional cost, since network vendors usually charge for the use of such algorithms. But then the operator has a margin to deploy less BSs, reducing the need for investment, since more users can be served by one BS if PF is implemented, without jeopardising RTPS services accessibility and retainability.

As a final conclusion, regardless of the points of view previously stated, 802.16e, with its capability of differentiating services and its flexibility in terms of AMC allocation depending on radio channel conditions, provides a significant margin for intelligent scheduling mechanisms to be implemented. A good algorithm may have input variables such as the instantaneous radio channel conditions felt by the different users, the medium/long term average rate of each user, the recent delay experimented by the different users, or the tolerable packet loss for real-time services, without jeopardising the overall quality. Thus, an equipment supplier clearly has room for differentiation, as it can develop algorithms that combine all these variables in the search for an algorithm that is capable of satisfying the different requirements associated to the different applications and be adapted by operators according to their own goals in terms of QoS.

For future work, there are many 802.16e related concepts that have a key role on scheduling functions, and that are not considered for the present thesis. One subject concerns the reassembly and segmentation of upper layer packets onto the lower layers: for the present thesis packets sizes are considered to be static throughout a whole session, but theoretically packets can be segmented into small size ones, more easily fitable into the PHY layer, which requires functions of numbering, buffering or reordering at both BS and ST. Another subject concerns the coexistence of UL and DL traffic, as many applications have a more or less linear relation between the achievable DL and UL throughputs. The present thesis only concerns DL but, in reality, the room for UL scheduling must be accounted as well, since it affects the existing DL data rate. Another area that is not detailed in the present thesis concerns the different performances achievable through the use of the different subchannel allocation modes as PUSC, FUSC, AMC, etc.. The thesis only focuses on the mandatory PUSC, but it is known that each of these options has its own advantages and disadvantages, depending on the radio environment in which the system is deployed. However, for this area to be properly evaluated, precise channel model characterisation for OFDMA is needed. Another key area included in the 802.16e standard concerns the subchannelisation concept, which is used to tradeoff coverage by capacity. In the present work, it was not considered, but this is clearly an option that requires proper evaluation as it may be useful to compensate for the high path loss linked to the propagation in the 3.5 GHz band.

The radio channel model is another topic that is not covered by the present work. Since 802.16 is based on OFDMA, it is of major interest to have a correct characterisation of the radio channel, so that the particular radio conditions affecting each of the subcarriers of each of the active users in terms of RSSI or CINR can be correctly characterised, which allows for more reliable scheduling decisions by the BS. Still concerning radio channel characterisation, the lack of existence of propagation models suited for 3.5 GHz bands in all type of environments is another issue that could be worked out, as there is an increasing trend of using these bands for cellular-like wireless networks. Other key 802.16e aspects not covered by the present thesis concern the support for smart antenna technologies, like MIMO or beamforming.

Annex A

Link Budget

In this annex, detailed link budget calculations are presented and described.

The total propagation attenuation is calculated individually for both UL and DL using the following expression [Corr03]:

$$L_{p\,[\mathrm{dB}]} = EIRP_{[\mathrm{dBm}]} + G_{ar[\mathrm{dBi}]} - L_{r[\mathrm{dB}]} + G_{r[\mathrm{dB}]} - P_{r[\mathrm{dBm}]} \tag{A.1}$$

where:

- $EIRP$ is the equivalent isotropic radiated power, given by (A.2):

$$EIRP_{[\mathrm{dBm}]} = P_{t[\mathrm{dBm}]} + G_{t[\mathrm{dB}]} - L_{t[\mathrm{dB}]} + G_{at[\mathrm{dBi}]} \tag{A.2}$$

- P_t is the power at the exit of the transmitting radio unit. Values differ between BS and ST, being given by Table 2.3.
- G_t represents the additional gains, e.g. due to the use of multiple transmitting antennas or use of power boosters.
- L_t is the attenuation between the transmitting radio unit and the transmitting antenna due to cable loss, connector loss, etc.. In WiMAX, typical commercial solutions allow the radio units to be installed quite close to the antennas, so minimum losses are caused by such causes. A value of 0.7 dBi is assumed.
- G_{at} is the gain of the transmitter antenna. Typical values are:
  1. For BS antennas, 17 dBi for 65º horizontal aperture antennas, plus additional 3 dB derived from considering an array of two transmit antennas in the BS, [WiMA06a].
  2. For ST antennas, 0 dBi.
- G_{ar} is the gain of the receiving antenna.
- L_r is the attenuation between the receiving antenna and the receiving radio unit, due to cable loss, connector loss, etc.. For the DL, there are no cables at the receiving side, thus, this value is set to 0 dB.
- G_r represents the additional gains, e.g. due to the use of receive diversity. Due to received diversity, 3 dB gain is usually assumed.
- P_r is the power received at the receiving antenna.

In OFDM, the receiver's sensitivity per subcarrier can be calculated through:

$$P_{r\,\min/subcarrier[\mathrm{dBm}]} = SNR_{req[\mathrm{dB}]} + N_{tot[\mathrm{dBm}]} \tag{A.3}$$

where SNR_{req} is the required Signal-to-Noise Ratio for the signal, given by Table 2.4, and N_{tot} is the total noise power, given by (A.4):

$$N_{tot[\mathrm{dBm}]} = F_{N[\mathrm{dB}]} - 174 + 10 \cdot \log(\Delta f_{[\mathrm{Hz}]}) \tag{A.4}$$

where:

- F_N is the receiver's noise figure:
  - For BS equipment, typical noise figure is around 4.0 dB [WiMA06a].
  - For ST equipment, typical noise figure is around 7.0 dB [WiMA06a].
- Δf is the channel band-width, in this case related to the subcarrier. According to Table 2.1, it equals 10.94 kHz.

The composite receiver sensitivity, i.e., considering all subcarriers is given by:

$$P_{rx\min[\mathrm{dBm}]} = P_{rx\min/subcarrier[\mathrm{dBm}]} + 10 \cdot \log\left(N_{usedsubcarriers}\right) \tag{A.5}$$

where $N_{usedsubcarriers}$ stands for the number of used subcarriers (Table 2.2).

The total system gain is given by:

$$L_{ptot\,[\mathrm{dB}]} = EIRP_{[\mathrm{dBm}]} + G_{ar[\mathrm{dBi}]} - L_{r[\mathrm{dB}]} + G_{r[\mathrm{dB}]} - P_{rxmin[\mathrm{dBm}]} \tag{A.6}$$

However, since additional margins must be considered to account for the effect of fading (both fast and slow-fading), interference (caused by other active users or stations) and penetration losses (e.g. due to walls for indoor users), the total maximum allowed path loss is given by:

$$L_{ptotmax\,[\mathrm{dB}]} = L_{ptot[\mathrm{dB}]} - M_{SF[\mathrm{dB}]} - M_{FF[\mathrm{dB}]} - L_{pen[\mathrm{dB}]} - M_{int[\mathrm{dB}]} \tag{A.7}$$

where:

- M_{SF} is the margin to account for slow fading.
- M_{FF} is the margin to account for fast fading.
- L_{pen} stands for the penetration losses.
- M_{int} is the margin to account for interference.

Fast fading is normally characterised by a Rayleigh distribution, [Sale98], the probability for the received power to be below a given threshold is given by:

$$P\left(P_{[\mathrm{dBm}]} \le P_{0[\mathrm{dBm}]}\right) = 1 - 2^{-\frac{P_{0[\mathrm{dBm}]}}{P_{m[\mathrm{dBm}]}}} \tag{A.8}$$

where:

- P_m is the median of the received power.
- P_o is the received power threshold.

For the present work, a value of 6 dB is assumed, as proposed in [WiMA06a].

Slow fading is typically characterised through a log-normal distribution with a standard deviation between 4 and 8 dB, depending on variables such as the path distance, environment, weather conditions, etc., [Sale98]. Under this assumption, the probability for the received power to be below a given threshold is given by:

$$P\left(P_{[\mathrm{dBm}]} \leq P_{0[\mathrm{dBm}]}\right) = \frac{1 - erf\left(\dfrac{\bar{P}_{[\mathrm{dBm}]} - P_{0[\mathrm{dBm}]}}{\sqrt{2} \cdot \sigma_{[\mathrm{dB}]}}\right)}{2} \tag{A.9}$$

where σ is the standard-deviation of the received power, measured in dB and function of the propagation model and frequency, and $\bar{P}$ is the average received power.

If the received power has the following behaviour:

$$\bar{P}_{[\mathrm{dBm}]} = \bar{P}_{R[\mathrm{dBm}]} - 10 \cdot n \cdot \log\left(\frac{r}{R}\right) \tag{A.10}$$

, the coverage probability inside a circle of radius *R* is given by:

$$F_{area} = \frac{1 + erf(a) + e^{(2a+1)/b^2} \cdot \left[1 - erf\left(\dfrac{a \cdot b + 1}{b}\right)\right]}{2} \tag{A.11}$$

where:

- $a = \dfrac{P_{R[\mathrm{dBm}]} - P_{\min[\mathrm{dBm}]}}{\sqrt{2} \cdot \sigma_{[\mathrm{dB}]}}$ (A.12)
- $b = \dfrac{10 \cdot n \cdot \log e}{\sqrt{2} \cdot \sigma_{[\mathrm{dB}]}}$ (A.13)
- σ depends on the propagation model used and for the present thesis the following figures are assumed:
    - σ = 6 dB for Walfisch-Ikegami propagation model [Gonç98].
    - σ = 10.5 dB for the COST231 propagation model [Corr03].
- P_R is the received power at a distance R from the BS.

Variable *n* included in equation (A.10) depends on the propagation model considered. For the COST 231 - Hata model (Annex B), *n* is given by:

$$n = \frac{44.90 - 6.55 \cdot \log(h_{BS})}{10} \tag{A.14}$$

For the Walfisch-Ikegami model (Annex B), n is equal to 3.8 since in the present work $L_{rts}+L_{msd} > 0$ and $h_{BS}>h_{roof}$.

In order to simplify these calculations, and according to [WiMA06a], a slow fading margin of 5.56 dB assures a 75% coverage probability at the cell edge and 90% coverage probability over the entire area. This value is considered for cell radius estimation.

Indoor losses for the 3.5 GHz band can be extrapolated from existing studies for lower bands, particularly for GSM900, GSM1800 and UMTS. According to [SLCM05], two classifications should be made in order to characterise the indoor penetration losses: deep indoor and light indoor. The figure below illustrates the CDF of the attenuation due to penetration into buildings on the GSM900:

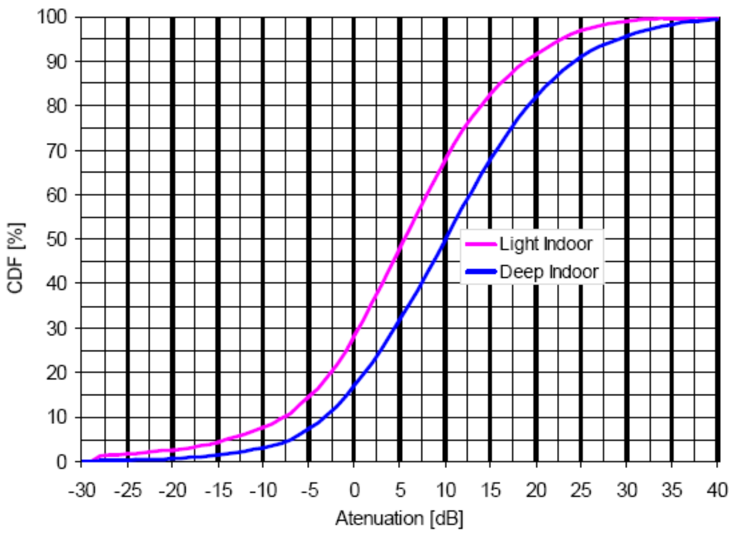


Figure A.1 - CDF of the indoor penetration losses for the GSM900 band (extracted from [SLCM05]).

According to [SLCM05], the shift when moving from the GSM900 band to either the GSM1800 or the UMTS band is 1.9 dB, with the CDF curve keeping the same trend. Since 802.16e is being targeted to work in either the 2.5 GHz or the 3.5 GHz band, by keeping a linear factor when using higher frequency bands, it is assumed in the present work that, regarding indoor penetration losses, the shift from GSM900 to the 2.5 GHz band equals 2.5 dB and the shift from GSM900 to the 3.5 GHz band equals 3.5 dB.

Thus, by setting the desired indoor coverage probability, the indoor penetration loss can be extracted from Figure A.1 and, by adding the necessary shift according to the adopted 802.16e band, i.e., 2.5 dB if 2.5 GHz is used or 3.5 dB if the 3.5 GHz band is used, one can estimate the indoor penetration losses in 802.16e.

For indoor users, to cope with the influence of the height of the user inside the building, i.e., the building floor, the following formula is used [SLCM05], where n_{fl} stands for the floor:

$$L_{pen[\mathrm{dB}]} = -0.8629 n_{fl} + 9.4064 \tag{A.15}$$

As stated previously, a 2.5 or 3.5 dB shift must be added, case the system operates at 2.5 or at 3.5 GHz, respectively. This formula represents the average penetration loss for all deep indoor, indoor window and indoor daylight, and it provides either positive (for lower floors) or negative (for higher floors) values. This shift is related to street level, at ground floor height, and the reason why there can be negative values for higher floors is due to the fact that users located at higher floors can be out of the shadow caused by neighbour buildings.

Annex B

Propagation Models

In this annex, detailed propagation model formulas are presented.

B.1 COST231-Hata Propagation Model

The COST231-Hata model [DaCo9] has been chosen to calculate the cell radius in Rural and Axial areas. This model is an extension of the Hata [Hata80] model for higher frequency range [1500, 2000] MHz, based on Okumura measurements performed in the [150, 2000] MHz band.

For a ST height of 1.5 m, the propagation attenuation is obtained by (B.1).

$$L_{CH[\mathrm{dB}]} = \mathrm{K}_1 + \mathrm{K}_2 \log(d) - a(h_{ST}) \quad \text{(B.1)}$$

where

- $K_1 = 46.3 + 33.9\log_{10}(f_{[\mathrm{MHz}]}) - 13.82\log(h_{BS[\mathrm{m}]})$ (B.2)
- $K_2 = 44.9 - 6.55\log_{10}(h_{BS[\mathrm{m}]})$ (B.3)
- f is the frequency
- d is the distance between the BS and the ST
- $a(h_{ST})$ is the correction factor which takes into account the antenna height variation of the ST, given by Table B.1.

Table B.1 – h_{MT} values per environment for COST 231 - Hata propagation model calculations.

Environment	$\boldsymbol{a(h_{ST})}$
Dense urban	$[1.1\log_{10}(f_{[\mathrm{MHz}]}) - 0.7]h_{ST[\mathrm{m}]} - 1.56\log_{10}(f_{[\mathrm{MHz}]}) + 0.8 -3$
Urban	$[1.1\log_{10}(f_{[\mathrm{MHz}]}) - 0.7]h_{ST[\mathrm{m}]} - 1.56\log_{10}(f_{[\mathrm{MHz}]}) + 0.8$
Suburban	$[1.1\log_{10}(f_{[\mathrm{MHz}]}) - 0.7]h_{ST[\mathrm{m}]} - 1.56\log_{10}(f_{[\mathrm{MHz}]}) + 0.8$
Rural	0

- h_{ST} (mobile antenna height)=1.8 m for indoor and pedestrian, and 1.2 m for vehicular.
- h_{BS} is antenna height of the base station. 30 m is assumed in the present work.

Some correction factors are introduced given the morphology. For rural environments, the corrections summarised in Table B.2 are applied.

Table B.2 – Correction factors due to morphology to be applied for COST 231 - Hata propagation model calculations.

Environment	$a(h_{ST})$
Suburban	$-2\log^2_{10}(f_{[MHz]}/28) - 5.4$
Rural	$-2\log^2_{10}(f_{[MHz]}/28) - 4.2$ (correction forest)
Rural	$-4.78\,[\log_{10}(f_{[MHz]})]^2 + 18.33\log_{10}(f_{[MHz]}) - 35.94$ (correction semi-open)

This model can be applied under the following conditions:

- $1 \le d \le 20$ km
- $30 \le h_{BS} \le 200$ m
- $1500 \le f \le 2000$ MHz
- $1 \le h_{ST} \le 10$ m

B.2 Walfisch-Ikegami Propagation Model

The COST231-Walfisch Ikegami model [DaCo9] has been chosen to calculate the cell radius in Dense Urban, Urban and Suburban areas. In the previous revisions of this document model COST231-Hata was used independently of the environments under consideration. However, COST231-Walfisch Ikegami is the most suitable theoretical propagation model for urban environments and presents the advantage of having a wide range of application, from 800 MHz to 2000 MHz.

If line-of-sight does not exist between the BS and the ST (common condition) pathloss is given by:

$$L_{CWI[\mathrm{dB}]} = \begin{cases} L_{0[\mathrm{dB}]} + L_{rts[\mathrm{dB}]} + L_{msd[\mathrm{dB}]} \\ L_{0[\mathrm{dB}]}\,,\ \text{for } L_{rts[\mathrm{dB}]} + L_{msd[\mathrm{dB}]} \le 0 \end{cases} \tag{B.4}$$

where:

- L_0 represents the free-space loss.
- L_{rts} the "roof-top-to-street" diffraction and scatter loss.
- L_{msd} the "multi-screen loss".

The free-space loss is given by:

$L_{0[\mathrm{dB}]} = 32.4 + 20\log_{10}(d_{[\mathrm{km}]}) + 20\log_{10}(f_{[\mathrm{MHz}]})$ (B.5)

where:

- d is the distance ST – BS.
- f is the frequency.

The "rooftop-to-street" diffraction and scatted loss is calculated with:

$L_{rts[\mathrm{dB}]} = -16.9 - 10\log_{10}(w_{[\mathrm{m}]}) + 10\log_{10}(f_{[\mathrm{MHz}]}) + 20\log_{10}(dh_{ST[\mathrm{m}]}) + L_{ori[\mathrm{dB}]}$ (B.6)

where:

- w is the street width (in meter). See the two figures presented at the end of model's description.
- L_{ori} is given by (B.7):

$$L_{ORI[\mathrm{dB}]} = \begin{Bmatrix} -10 + 0.354 \times \varphi_{[^\circ]} & \text{for } 0 < \varphi < 35 \\ 2.5 + 0.075 \times (\varphi_{[^\circ]} - 35) & \text{for } 35 \le \varphi \le 55 \\ 4.0 - 0.114 \times (\varphi_{[^\circ]} - 55) & \text{for } 55 < \varphi \le 90 \end{Bmatrix} \quad \text{(B.7)}$$

- φ is the angle measured between the street direction and the incident wave direction. In the absence of a particular value, a 90º angle is commonly assumed.
- $dh_{ST[\mathrm{m}]} = h_{roof[\mathrm{m}]} - h_{ST[\mathrm{m}]}$ (B.8)
- $dh_{BS[\mathrm{m}]} = h_{BS[\mathrm{m}]} - h_{roof[\mathrm{m}]}$ (B.9)
- h_{roof} is the average height of the buildings' roofs.
- h_{ST} is the ST height.
- h_{BS} is the height of the BS antenna.

The multi-screen diffraction loss is given by:

$L_{msd[\mathrm{dB}]} = L_{bsh[\mathrm{dB}]} + k_{a[\mathrm{dB}]} + k_{d[\mathrm{dB}]}\log_{10}(d_{[\mathrm{m}]}) + k_{f[\mathrm{dB}]}\log(f_{[\mathrm{MHz}]}) - 9\log_{10}(b_{[\mathrm{m}]})$ (B.10)

where:

- b is the inter-buildings distance.
- L_{bsh} is given by:

$$L_{bsh[\mathrm{dB}]} = \begin{Bmatrix} -18 + \log_{10}(1 + dh_{BS[\mathrm{m}]})_i & \text{for } h_{\mathrm{BS}} > h_{roof} \\ 0 & \text{for } h_{\mathrm{BS}} \le h_{roof} \end{Bmatrix} \quad \text{(B.1)}$$

- K_a is given by:

$$K_{a[\mathrm{dB}]} = \begin{cases} 54 & \text{for dh}_{\mathrm{BS}} > h_{roog} \\ 54 - 0.8 \times (dh_{BS[\mathrm{m}]}) & \text{for } d \geq 0.5 \text{ km and } h_{BS} \leq h_{roof} \\ 54 - 0{,}8 \times (dh_{BS[\mathrm{m}]}) \times \dfrac{d_{[\mathrm{m}]}}{0.5} & \text{for } d < 0.5 \text{ km and } h_{BS} \leq h_{roof} \end{cases} \tag{B.2}$$

- K_d is given by:

$$K_{d[\mathrm{dB}]} = \begin{cases} 18 & \text{for } h_{\mathrm{BS}} > h_{roof} \\ 18 - 15\dfrac{dh_{BS[\mathrm{m}]}}{h_{roof[\mathrm{m}]}} & \text{for } h_{\mathrm{BS}} \leq h_{roof} \end{cases} \tag{B.3}$$

- K_f is given by:

$$K_{f[\mathrm{dB}]} = -4 + \begin{cases} 0.7 \times \left(\dfrac{f_{[\mathrm{MHz}]}}{925} - 1 \right) & \text{for medium sized cities or centres with moderate tree density} \\ 1.5 \times \left(\dfrac{f_{[\mathrm{MHz}]}}{925} - 1 \right) & \text{for metropolitan centres} \end{cases} \tag{B.4}$$

Figure B.1 and Figure B.2 illustrate some of the parameters used in the COST231-Walfisch-Ikegami model.

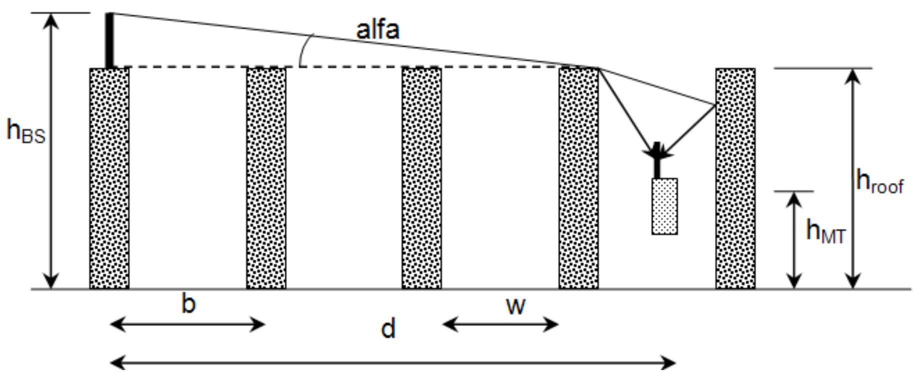


Figure B.1 – Parameters used in COST231-Walfisch-Ikegami model.

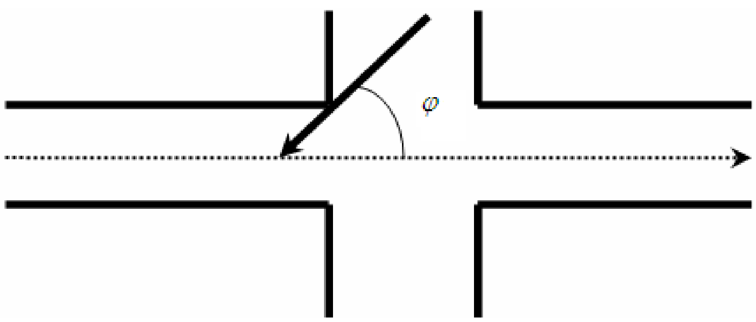


Figure B.2 – Definition of Street Orientation for COST231-Walfisch-Ikegami model.

More simply, when propagation occurs along the street ($\varphi = 0°$) path loss can be computed using

$$L_{CWI[dB]} = 42.6 + 26\log_{10}\left(d_{[km]}\right) + 20\log_{10}\left(f_{[MHz]}\right), \quad d > 0.02\,\text{km} \qquad \text{(B.5)}$$

In the absence of concrete data the following values are recommended:

- $b \in [20, 50]$ m.
- $W = b/2$.
- $\varphi = 90^{\circ}$.
- $H_B = 3 \cdot$ (Number of floors) + H_{roof}.
- $H_{roof} = 3$ if tilted roof or 0 if plain roof.

The COST-231 Walfisch Ikegami is restricted to the following ranges:

- 0.02km $\leq d \leq$ 5km
- $4 \leq h_{BS} \leq 50$ m
- $800 \leq f \leq 2000$ MHz
- $1 \leq h_{MT} \leq 3$ m

Annex C

Traffic Models

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

C.1 VoIP

VoIP services typically present 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 intercaleted 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 C.1.

Table C.1 – Typical VoIP codecs (extracted from [Nune02]).

Codec	Frame Duration [ms]	Frame Size [bytes]	Bitrate [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 bitrate, a final bitrate 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 C.2.

Table C.2 – 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
Transmission Time Interval [ms]	20
Mean Call Duration [s]	120

C.2 Video Telephony

Several codecs exist for the purpose of transmisting video, some examples being ([Agui03]) ITU-T H.261, ITU-T H.263, MPEG-1 (Moving Picture Expert Group-1), MPEG-2 or MPEG-4. Within the present work, ITU-T H.263 is considered for video phone calls. Although there is no specific mandatory codec for WiMAX, this codec has been standartised by 3GPP for mobile terminals supporting conversational multimedia services over packet switching, which is the case of WiMAX.

[Heym97] presents a model for video based on observed video traffic statistical features for video conference services. This model is represented as the Gamma Beta Auto-Regressive (GBAR) Model, which is a first order auto-regressive process that relies on two statistical features observed in H.261 and H.263 ([LaGD00]) Variable Bit Rate (VBR) video conferencing traffic: (1) the marginal distribution follows a Gamma distribution; (2) the auto-correlation function is geometric.

The model, being auto-regressive, only requires a previous sample for generating the next one, i.e., the process is used as a source model by generating non-integer values from (C.1) and then rounded to the nearest integer. In this context,

$$X_n = A_n X_{n-1} + B_n \tag{C.1}$$

defines a stationary stochastic process $\{X_n\}$ with a marginal $Ga(\beta,\lambda)$ distribution, where:

- X_{n-1} is a $Ga(\beta,\lambda)$ distribution.
- A_n is BE($\alpha,\beta-\alpha$) distribution.
- B_n is $Ga(\beta-\alpha,\lambda)$ distribution.

These three distributions are mutually independent [Agui03]. $G_a(\beta,\lambda)$ denotes a random variable with a Gamma distribution with shape parameter β and scale parameter λ, with its PDF (Probability Density Function) being given by:

$$p_G(x) = \frac{\lambda(\lambda x)^{\beta}}{\Gamma(\beta+1)} e^{-\lambda x}, \quad x > 0 \tag{C.2}$$

where $\Gamma(\cdot)$ is the Gamma function. Similarly, $B_e(p,q)$ denotes a random variable with a Beta distribution with parameters p and q, with its PDF being given by:

$$p_B(x) = \frac{\Gamma(p+q)}{\Gamma(p+1)\Gamma(q+1)} x^{p-1}(1-x)^{q-1}, \quad 0 < x < 1 \tag{C.3}$$

where p and q are both larger than –1.

[FiRe01] provides values for the characterisation of the above distributions, which were calculated from real traces from VBR videos with 176 by 144 pixels (which is the expected screen definition for next generation wireless handsets [FrNg00]) and a fixed frame rate of 25 frames per second, Table C.3.

Table C.3 - Parameter values for characterisation of the GBAR model for video call.

VBR Video Sequence	Λ	β	α
Office Camera	0.008437	7.625225	7.190587

β/λ and β/λ^2 are the mean and the variance of a $Ga(\beta,\lambda)$ distribution, respectively.

C.3 Streaming

Audio and Video Streaming services typically present an asymmetrical nature, as they refer mostly to the download of large files. Contrary to VoIP, streaming is more flexible to end-to-end transmission delays and to its variations, although the implementation of buffering is required to accommodate those fluctuations. Although video traffic can be transported either with a CBR or with a VBR, VBR is one expected to be the prime example, as it has several potential advantages over CBR, particularly the possibility for implementing statistical multiplexing, allowing for improved channel allocation.

IEEE has not standartised a specific codec for video streaming yet. On the contrary, 3GPP has specified the use of both MPEG-4 and H.263 codecs for video streaming services [3GPP02b]. Under the present work, a model for MPEG-4 VBR video streaming is considered. MPEG-4 is a coded based on objects, while legacy codecs, such as MPEG-1, are frame based. Under MPEG-4, each scene consists of a set of Video Objects (VOs) that are individually encoded. Additionally, each individual VO may have several scalability layers, which are referred to as Video Object Layers (VOLs). Finally, every VOL consists of an ordered sequence of snapshots in time, referred to as Video Object Planes (VOPs). There are 3 types of VOPs: intracoded (*I*), forward predicted (*P*), and bidirectionally predicted (*B*), with I and P also named anchor VOPs. The *I*, *P*, and *B* VOPs are arranged in a periodic pattern referred to as a Group Of Pictures (GOP). This periodic pattern is usually described as a (*N*,*M*) cyclic GOP, where *N* is the spacing between successive *I* VOPs, and *M* is the spacing between successive anchor VOPs. For instance, a typical GOP structure for a 25 Hz frame rate video on-demand is *IBBPBBPBBPBB* ((12,3) cyclic GOP).

The model adopted under the present work is the GOP GBAR model [FrNg00], which is a generalisation of the GBAR model for video conferencing presented in Section 0. Under this model, the size X_k of the k^{th} frame in an MPEG-encoded video sequence starting with an *I*-frame, using a (*N*,*M*) cyclic GOP given by

$$X_k = \begin{cases} \lambda_1 Z_{1k} + \lambda_2 Z_{2k} + \lambda_3 Z_{3k}, & \text{if } k = 1 \bmod N, \\ \lambda_1 Z_{1k} + \lambda_2 Z_{2k}, & \text{if } k \neq 1 \bmod N \text{ but } k = 1 \bmod M, \\ \lambda_1 Z_{1k,} & \text{otherwise} \end{cases} \tag{C.4}$$

where $\{Z_{1k} = 0, 1, \ldots\}$ = GBAR (α_1,ρ_1), $\{Z_{2k} = 0, 1, \ldots\}$ = GBAR (α_2,ρ_2) and $\{Z_{3k} = 0, 1, \ldots\}$ = GBAR (α_3,ρ_3) are independent stationary GBAR processes. The nine parameters needed to characterise the model can be estimated from statistical parameters taken from real traces. [FiRe01] presents the results from some video sequences data for two known movies (with digital format 176 by 144, using (12,3) cyclic GOP), with two levels of quality: (1) low quality, corresponding to quantisation parameters of 10 for *I* frames (VOPs), 14 for *P* frames, and 18 for *B* frames; (2) high quality, corresponding to quantisation parameters of 4 for all three frame types. The resulting parameters to be applied for characterisation of the model are summarised in Table C.4.

Table C.4 - Parameter values for characterisation of the GBAR model for video streaming.

Film	Quality	λ_1 [byte]	λ_2 [byte]	λ_3 [byte]	α_1	α_2	α_3	ρ_1	ρ_2	ρ_3
Jurassic Park I	Low	263.8	562.2	476.6	188.42	1.12	5.62	0.93600	0.79228	24.06809
	High	1012.5	1404.3	-149.6	2813.49	0.91	-18.17	0.95500	0.74627	-64.83057
Silence of the Lambs	Low	464.1	986.5	1099.6	90.04	0.40	1.63	0.96000	0.86443	34.78203
	High	1532.8	2671.3	609.3	1529.25	0.29	3.72	0.97900	0.85008	90.68861

C.4 Non-Real Time Applications

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.

Various 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 C.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 C.1.

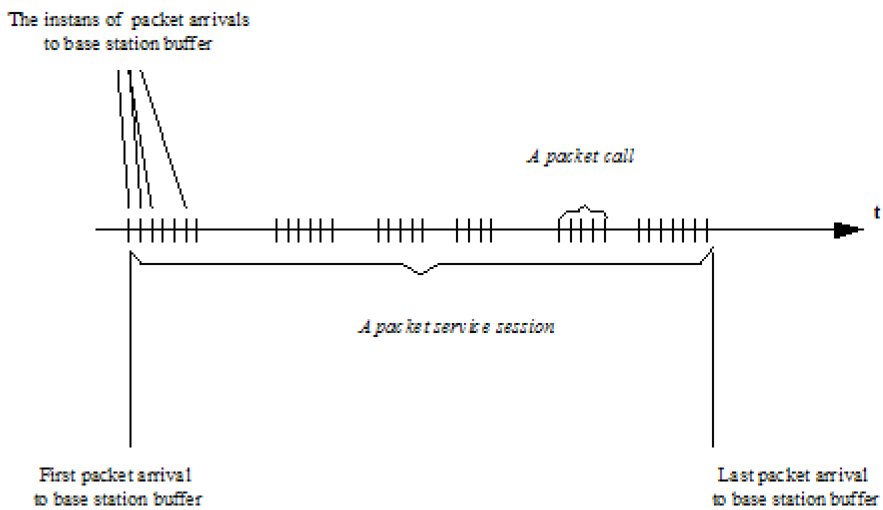


Figure C.1 – Typical Packet Service Session (extracted from [Agui03]).

A packet service session 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 modelled [ETSI98]:

- Number of packet calls per session, N_{pc}
- Reading time between packet calls, D_{pc}
- Number of packets within a packet call, N_d
- Interarrival time between packets (within a packet call), D_d
- Size of a packet, S_d.

The session length is implicitly modelled by the number of events during the session. [ETSI98] specifies how the above events should be modelled as well:

- **Number of packet call requests per session**, N_{pc}: This is a geometrically distributed random variable, with mean μ_{Npc}, where the probability of generating one packet call, P_{PC} is given by:

  $$P_{PC} = \frac{1}{\mu_{Npc}} \tag{C.5}$$

  The probability, $p_{Ge}[n]$, of generating n packet call requests during a session is given by the following probability function:

  $$p_{Ge}[n] = P_{PC} \cdot (1 - P_{PC})^{n-1} \tag{C.6}$$

- **Reading time between two consecutive packet call requests in a session**, D_{pc}: This is a geometrically distributed random variable, with mean μ_{Dpc}. The reading time starts when the last packet of the packet call is completely received by the user, and it ends when the user makes a request for the next packet call.
- **Number of packets in 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 μ_{Nd}.
- **Time interval between two consecutive packets inside a packet call**, D_d: This should be a geometrically distributed random variable, with mean μ_{Dd}. If there is only one packet in a packet call (e.g., FTP), this variable is not needed. However, it is assumed under the

present work that the packet inter-arrival time matches the radio interface frame size (5ms) through TCP window size adjustment.

- **Packet size**, S_d: The packet size is defined as $S_d = \min(P_a, M_{ps})$, where P_a is a Pareto distributed random variable and M_{ps} is the maximum allowed packet size (66 666 bytes). The packet size distribution model is based on a Pareto distribution with cut-off. The normal Pareto PDF (without cut-off) is given by [ETSI98] (k corresponds to the minimum packet size):

$$p_P(x) = \begin{cases} 0, x < k \\ \dfrac{\alpha_p \cdot k^{\alpha_p}}{x^{\alpha_p+1}}, x \geq k \end{cases} \tag{C.7}$$

Table C.5 summarises the mean values for the distributions of typical WWW services. According to the values for α_p and k in the Pareto distribution, the average packet size μ is 480 bytes. The average requested file size is μ_{Nd} by μ = 25 by 480 bytes ~ 12 kBytes. The interarrival time is set to 5ms, assuming that the TCP window size adjustment follows the radio interface frame size.

Table C.5 – Parameters for HTTP traffic modelling (extracted from [ETSI98]).

Packet based information types (WWW Browsing)	μ_{Npc}	μ_{Dpc} [s]	μ_{Nd}	μ_{Dd} [s]	Parameters for S_d distribution
UDD 2048 kbit/s	5	4-12	25	0.00195	k = 81.5 a_p = 1.1

Other non-real time applications are the commonly called 'background' type services. These correspond typically to data transfer applications that do not demand a quick response from the user, such as e-mail or FTP. To handle these applications, the present report bases their modelling on the results from the project summarised in [KlLL01], where an intense IP traffic characterisation was performed at the Internet Service Provider (ISP) dial-in modem/Integrated Services Digital Network (ISDN) link of the University of Dortmund, based on the measured traffic data; extrapolations for the bit rates allowed by UMTS Release 99 networks were performed. This traffic model uses the notion that a user running background applications follows a characteristic usage pattern and that a single user may run different applications that may be concurrently active, e.g., WWW browsing while downloading files from FTP servers.

Different applications are described differently by their own specific statistical properties, which comprise of an alternating process of ON and OFF periods with some application specific length or data volume distribution. Within each ON-period the packet arrival process is completely captured by the packet interarrival times and the corresponding packet sizes, Figure C.2.

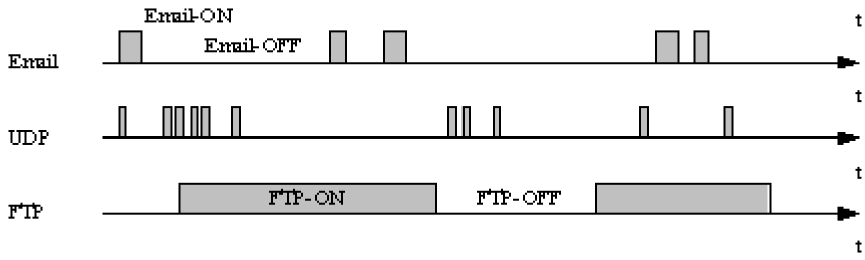


Figure C.2 – Characterisation of session level user traffic (extracted from [Agui03]).

Within this model, different levels must be analysed in order to fully characterise traffic. Particularly, the characterisation of FTP and e-mail applications for the present work is summarised in Table C.6 and Table C.7.

- Connection-level: describes the corresponding distribution of connection interarrival times and connection data volume for each individual application, Table C.6.
- Packet-level: characterises the packet interarrival time distribution and the packet size distribution within the application specific connections. For the present thesis, it is assumed that the packet inter-arrival time matches the radio interface frame size through TCP window adjustment, meaning that a packet interarrival time of 5ms is assumed for FTP and Email. Table C.7summarises packet size related figures.

Table C.6 - Statistical properties for connection level in bytes and seconds (extracted from [KlLL01] and [Ljun06]).

Application	Quantity	Distribution	Parameters
Email	Interarrival time	Pareto (k; a_p)	(16.0229; 2.1223)
	Data Volume	Lognormal (μ_{LN}; σ_{LN}^2)	(8.4124; 3.6439)
FTP	Interarrival time	Single connection within a session	
	Data Volume	Uniform(*min*; *max*)	(100000; 5000000)

Unfortunately for e-mail and FTP modelling, [KlLL01] only presents results from UMTS Release 99 radio bearers, which go up only to 384 kbps, while WiMAX can provide throughputs around some Mbps. Nevertheless, and since no literature was found about specific modelling of these applications for WiMAX, the modelling summarised in [KlLL01] for UMTS Release 99 is extended to the present work.

Table C.7 - Percentage of different packet sizes in bytes in overall traffic (extracted from [KlLL01]).

Application	Fractions of packets in overall traffic [%]			
	Packet size 40 bytes	**Packet size 576 bytes**	**Packet size 1500 bytes**	**Other packet sizes**
Email	38.25	25.98	9.51	26.26
FTP	40.43	18.08	9.33	32.16

Annex D

Statistical Distributions Validation

In this annex, the basic validation of the statistical distributions used within the simulator is presented.

Statistical functions are used throughout the present thesis for numerous purposes, such as traffic models generation. For the implementation of statistical models, RNGs are crutial. The following different RNGs were necessary: Uniform, Beta, Gamma, Geometrical, Normal, LogNormal, Pareto, Exponential, Triangular and Poisson. The selected Uniform RNG is the basis for the remaining RNG algorithms, which are based on transformations of this Uniform RNG.

Validation of each statistical function is based on a comparison between the respective RNG (10000 points generated) and the respective theoretical CDFs. Additionally, for each distribution, the CDF correlation is calculated based on discrete generated samples and the theoretical distribution expressions. The correlation figures given by Microsoft's Excel CORREL function, are very close to 100%, Table D.1.

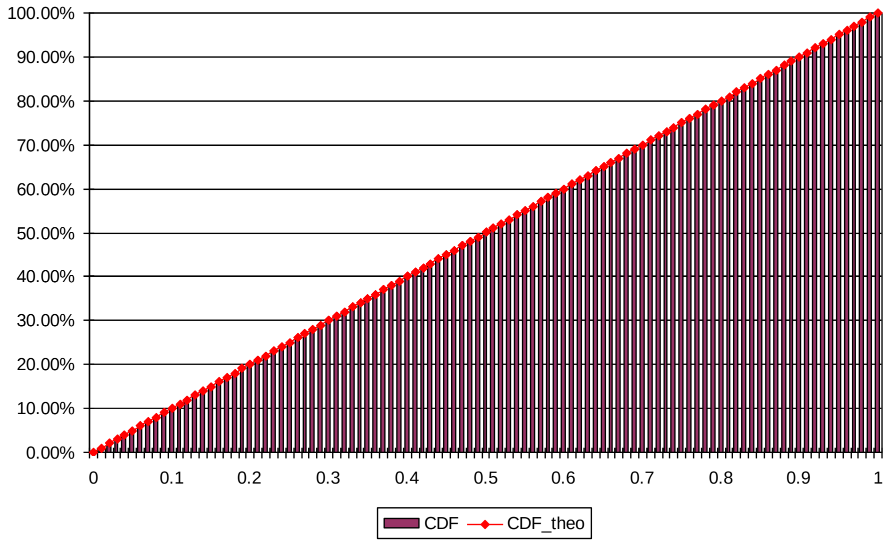


Figure D.1 – Theoretical and generated CDF for Uniform distribution (between 0 and 1).

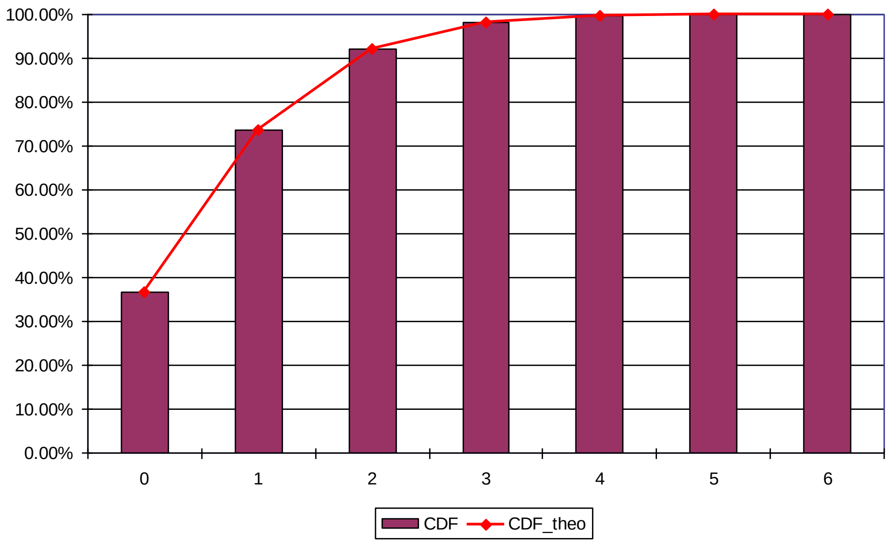


Figure D.2 - Theoretical and generated CDF for Poisson distribution (mean = 1).

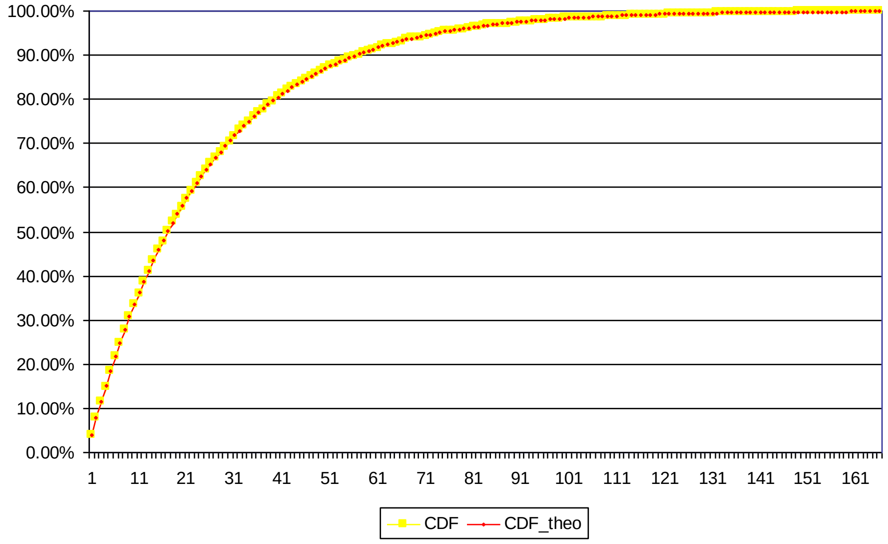


Figure D.3 - Theoretical and generated CDF for Geometric distribution ($p = 1/25$).

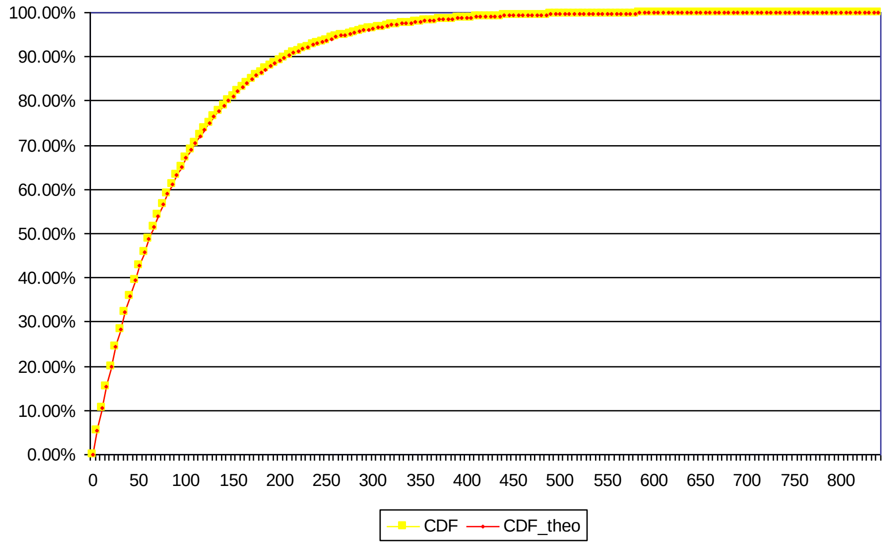


Figure D.4 - Theoretical and generated CDF for Exponential distribution ($\lambda = 90$).

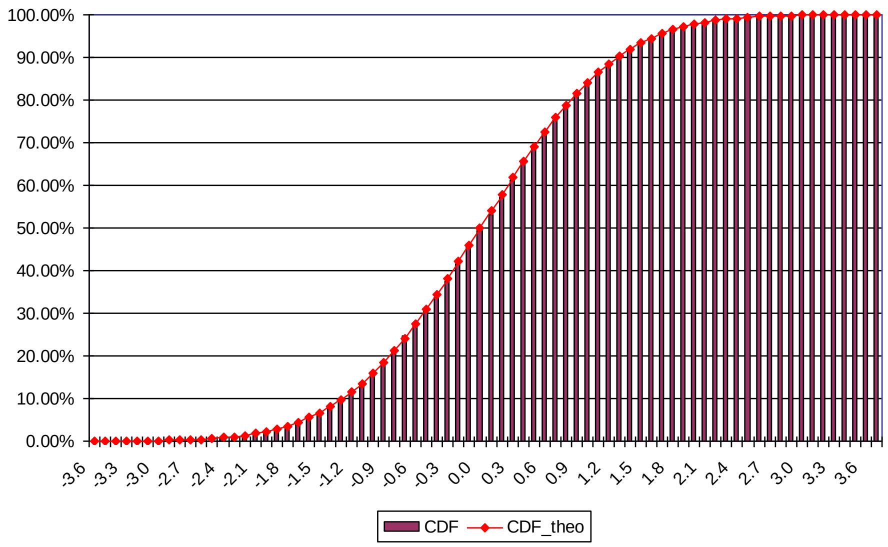


Figure D.5 - Theoretical and generated CDF for Normal distribution (mean=0, variance=1).

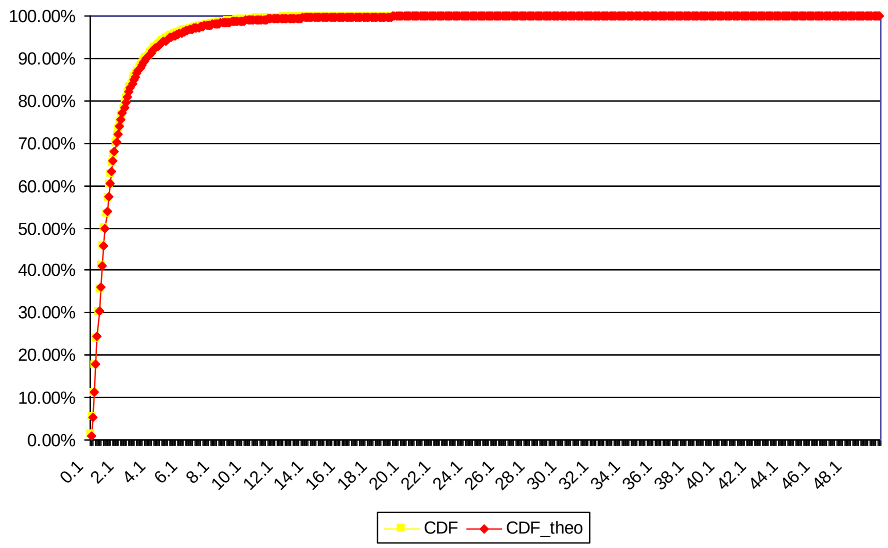


Figure D.6 - Theoretical and generated CDF for Log-Normal distribution (mean=0, variance=1).

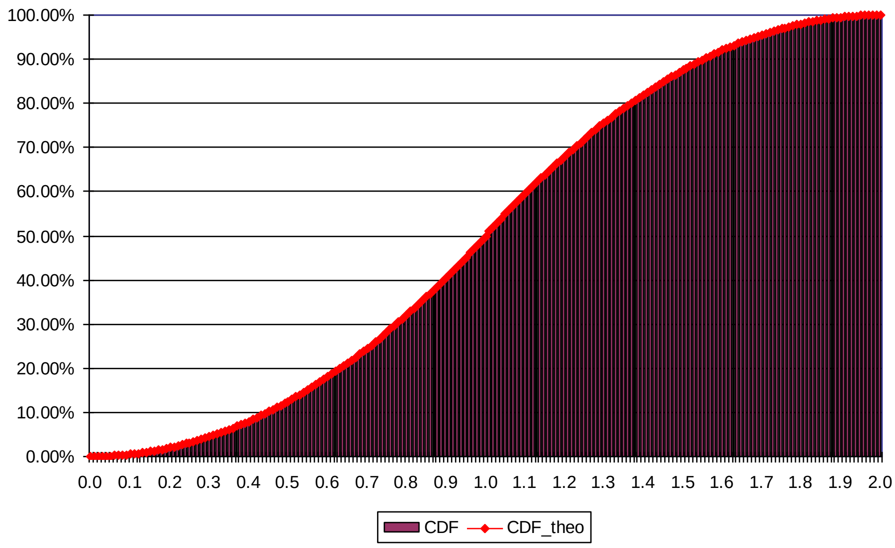


Figure D.7 - Theoretical and generated CDF for Triangular distribution (a=0, b=2, c=1).

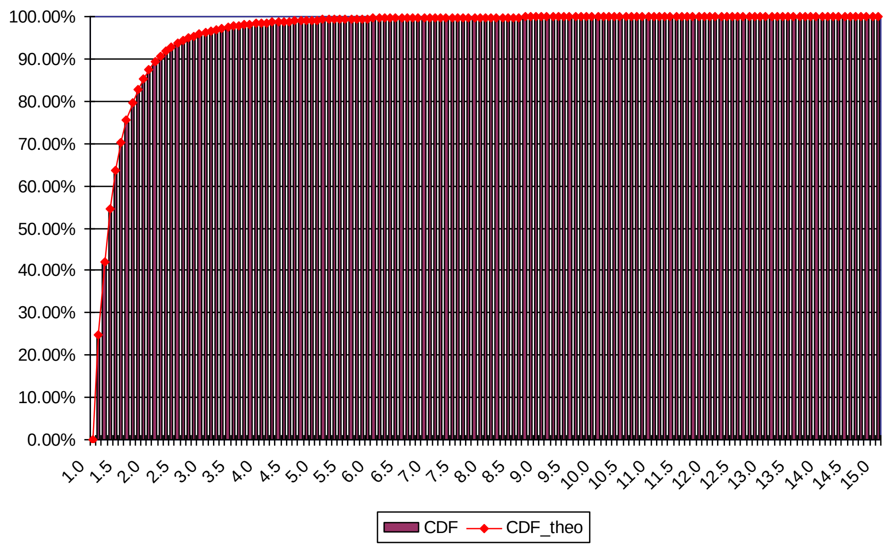


Figure D.8 - Theoretical and generated CDF for Pareto distribution (x_m=1, k=3).

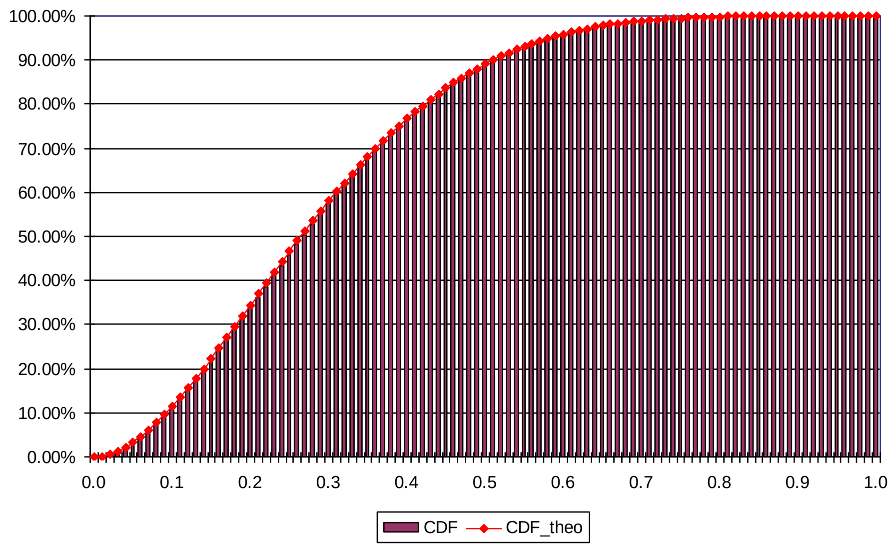


Figure D.9 - Theoretical and generated CDF for Beta distribution (α=2, β =5).

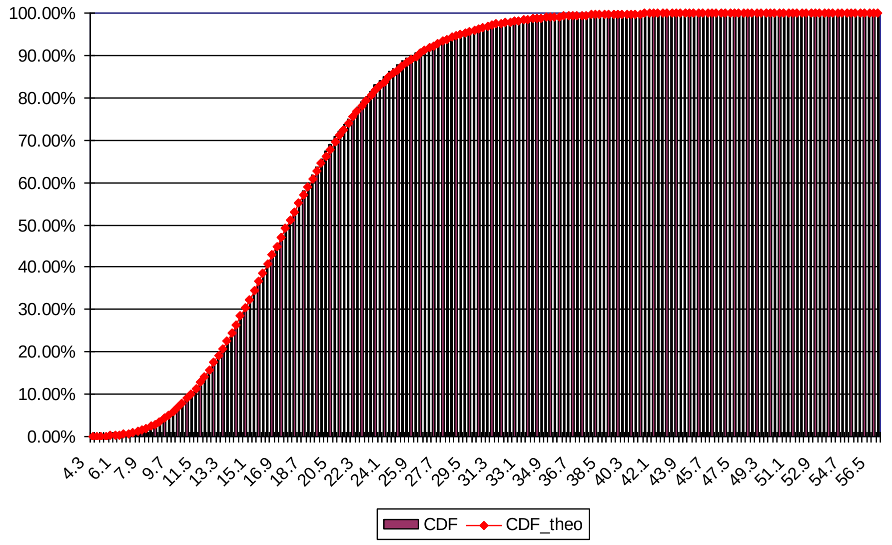


Figure D.10 - Theoretical and generated CDF for Gamma distribution (α=9, β =2).

Table D.1 – Correlation between theoretical and generated CDF for the different statistical distributions.

Distribution Type	CDF Correlation
Uniform	100.0000%
Poisson	100.0000%
Geometric	99.9993%
Exponential	100.0000%
LogNormal	99.9969%
Normal	99.9986%
Triangular	100.0000%
Pareto	99.9945%
Beta	99.9972%
Gamma	99.9943%

Annex E

Input Files

In this annex, the input files used by the simulator are summarised.

The simulator takes several inputs, which can be extracted from comma separated type files.

- Generic System Configuration: this file contains generic information as the number of users, the clutter type, the operating frequency, etc., Table E.1

Table E.1 – System Configuration Input Table.

Parameter / Header Entry	Unit	Type	Example
Number of Users	N/A	int	5000
Clutter	N/A	string	Dense Urban
Cell Range	M	int	1000
Operating Frequency	MHz	int	3500
Channel Bandwidth	MHz	int	10
BS DL Transmit Power	dBm	float	43
BS DL Additional Gain	dB	float	3
BS Cable Losses	dB	float	2
BS Antenna Gain at Main Lobe	dB	float	17
ST Antenna Gain	dB	float	0
ST Losses	dB	float	2
ST DL Additional Gain	dB	float	2
ST Noise Factor	dB	float	7
QPSK1/2 Min CINR	dB	float	5
QPSK3/4 Min CINR	dB	float	8
16QAM1/2 Min CINR	dB	float	10.5
16QAM3/4 Min CINR	dB	float	14
64QAM2/3 Min CINR	dB	float	18
64QAM3/4 Min CINR	dB	float	20
BS Logging Period	N/A	int	100

- Users Information Summary: this file contains information concerning the distribution of users within the different segments, penetration losses associated to each user environment, etc. Due to the significant header size, header entries concerning SOHO and Mass Market users is not present in Table E.2 but they follow the same structure as the presented one for Business users.

Table E.2 – Users Information Summary Table.

Parameter / Header Entry	Unit	Type	Example
Number of Users	N/A	int	5000
Indoor User Average Speed	km/h	int	0
Pedestrian User Average Speed	km/h	int	3
Incar User Average Speed	km/h	int	50
Indoor User Penetration Losses at Ground Floor	dB	int	20
Pedestrian User Penetration Losses	dB	int	0
Incar User Penetration Losses	dB	int	5
Business Penetration	%	int	30%
SOHO Penetration	%	int	40%
Mass Market Penetration	%	int	30%
VoIP Percentage Within Business Segment	%	int	18%
Video Telephony Percentage Within Business Segment	%	int	68%
Streaming Percentage Within Business Segment	%	int	69%
FTP Percentage Within Business Segment	%	int	49%
WWW Percentage Within Business Segment	%	int	51%
Email Percentage Within Business Segment	%	int	36%
Indoor Percentage Within VoIP Business Segment	%	int	65%
Pedestrian Percentage Within VoIP Business Segment	%	int	12%
Incar Percentage Within VoIP Business Segment	%	int	59%
Indoor Percentage Within Video Telephony Business Segment	%	int	70%
Pedestrian Percentage Within Video Telephony Business Segment	%	int	14%
Incar Percentage Within Video Telephony Business Segment	%	int	21%
Indoor Percentage Within Streaming Business Segment	%	int	0%
Pedestrian Percentage Within Streaming Business Segment	%	int	79%
Incar Percentage Within Streaming Business Segment	%	int	8%
Indoor Percentage Within FTP Business Segment	%	int	23%
Pedestrian Percentage Within FTP Business Segment	%	int	88%
Incar Percentage Within FTP Business Segment	%	int	64%
Indoor Percentage Within WWW Business Segment	%	int	88%
Pedestrian Percentage Within WWW Business Segment	%	int	56%
Incar Percentage Within WWW Business Segment	%	int	81%
Indoor Percentage Within Email Business Segment	%	int	62%
(...)			

* Services Information Summary: this file contains information concerning the characterisation of the different services considered by the simulator, which is used for the respective statistical distributions, Table E.3.

Table E.3 - Services Information Summary Table.

Parameter / Header Entry	**Unit**	**Type**	**Example**
VoIP Mean Holding Time	seconds	int	5000
FTP Mean Data Volume	Bytes	float	8.8409
FTP Data Volume Variance	Bytes	float	4.3343
FTP Mean Packet Interarrival Time	seconds	float	0.06375
Email Mean Data Volume	Bytes	float	11.6795
Email Data Volume Variance	Bytes	float	14.4979
Email Mean Session Interarrival Time (Pareto Index)	seconds	float	16.0229
Email Session Interarrival Time Variance (Pareto Location)	seconds	float	3.1223
Email Session Mean Data Volume	Bytes	float	8.4124
Email Session Data Volume Variance	Bytes	float	3.6439
Email Mean Packet Interarrival Time	seconds	float	0.104
WWW Mean Session Arrival Time	seconds	float	5
WWW Mean Reading Time Between Packet Calls	seconds	float	8
WWW Mean Number of Packets Within Packet Call	N/A	float	25
WWW Mean Interarrival Time Between Packets Within Packet Call	seconds	float	0.0104
WWW Mean Packet Size Within Packet Call (Pareto Index)	Bytes	float	81.5
WWW Mean Packet Size Within Packet Call (Pareto Location)	Bytes	float	1.1

* Propagation Model Configuration: this file contains information concerning the parameters used for propagation model calculations, Table E.4.
* Antenna Radiation Pattern: this file contains information concerning the radiation pattern of the BS antenna, Table E.5.

Table E.4 – Propagation Model Configuration Table.

Parameter / Header Entry	Unit	Type	Example
Okumura-Hata BS Height	m	int	35
Okumura-Hata Environmnent	N/A	int	1 = Dense Urban and Large Cities 2 = Urban or Small City 3 = Suburban with Correction Applied 4 = Axial and Rural Forest 5 = Axial and Rural Semi-Open / Countryside
Walfisch-Ikegami BS Height	m	int	30
Walfisch-Ikegami Street Width	m	int	30
Walfisch-Ikegami Inter-Building Distance	m	int	60
Walfisch-Ikegami Angle Between Street Direction and Incident Wave Direction	deg	int	45
Walfisch-Ikegami Buildings' Roof Height	m	int	23
Walfisch-Ikegami City Type	N/A	int	1 = medium sised cities and suburban centres. 2 = metropolitan centres

Table E.5 – Antenna Radiation Pattern Table.

Parameter / Header Entry	Unit	Type	Example
Azimuth	deg	int	0, 1, 2, ..., 359
Losses Relative to Azimuth 0º	dB	float	2

Annex F

Output Files

In this section, the output files produced by the simulator are summarised.

As result of the simulator's execution, several output files are produced concerning the outcome of the simulations. These are comma separated type files, easily processed in 'common' tools such as Microsoft's Excel.

- VoIP Sessions Summary: this file summarises the outcome of each attempted VoIP session, Table F.1.

Table F.1 – VoIP Session Summary.

Parameter / Header Entry	Unit	Type	Example
User Number	N/A	int	2
Session Start	seconds	int	14
Original Session Duration	seconds	int	90
Final Session Duration	seconds	int	60
Session Blocked	N/A	boolean	FALSE
Session Dropped	N/A	boolean	TRUE
Average Number of Consumed Slots	N/A	int	5
Maximum Number of Consumed Slots	N/A	int	8
Average AMC	N/A	string	16QAM ½
Best AMC	N/A	string	64QAM 2/3
Worse AMC	N/A	string	QPSK ½
Number of Received Frames	N/A	int	50000
Number of Lost Frames	N/A	int	750
Percentage of Lost Frames	%	float	1.5%
Maximum Number of Consecutive Lost Frames	N/A	int	400
Traffic Volume	kBytes	int	3940

- Video Telephony Sessions Summary: this file summarises the outcome of each attempted video telephony session, Table F.2.
- Streaming Sessions Summary: this file summarises the outcome of each attempted streaming session, Table F.3.
- FTP Sessions Summary: this file summarises the outcome of each attempted FTP session, Table F.4.
- WWW Sessions Summary: this file summarises the outcome of each attempted WWW session, Table F.5.

Table F.2 – Video Telephony Session Summary.

Parameter / Header Entry	Unit	Type	Example
User Number	N/A	int	5
Session Start	seconds	int	20
Original Session Duration	seconds	int	100
Final Session Duration	seconds	int	70
Session Blocked	N/A	boolean	FALSE
Session Dropped	N/A	boolean	TRUE
Average Number of Consumed Slots	N/A	int	10
Maximum Number of Consumed Slots	N/A	int	15
Average AMC	N/A	string	16QAM ¾
Best AMC	N/A	string	64QAM 3/4
Worse AMC	N/A	string	QPSK ½
Number of Received Frames	N/A	int	50000
Number of Lost Frames	N/A	int	750
Percentage of Lost Frames	%	float	1.5%
Maximum Number of Consecutive Lost Frames	N/A	int	400
Traffic Volume	kBytes	int	7940

Table F.3 – Streaming Session Summary.

Parameter / Header Entry	Unit	Type	Example
User Number	N/A	int	2
Session Start	seconds	int	15
Session End	seconds	float	70
Session Quality	N/A	string	LQ
Session Data Volume	bytes	long	400000
Session Average Packet Size	bytes	int	1000
Packet Interarrival Time	ms	float	5
Number of Packets	N/A	long	500
Number of Packets Delayed	N/A	long	2
Percentage of Delayed Packets	%	float	1.00%
Total Delay	ms	long	10
Average Delay of Delayed Packets	ms	int	5
Session Average Throughput	kbps	float	300
Average Number of Consumed Slots	N/A	int	16
Maximum Number of Consumed Slots	N/A	int	20
Average AMC	N/A	string	16QAM ¾
Best AMC	N/A	string	64QAM 2/3
Worse AMC	N/A	string	QPSK ¾

Table F.4 – FTP Session Summary

Parameter / Header Entry	Unit	Type	Example
User Number	N/A	int	2
Session Start	seconds	int	12
Session End	seconds	float	20
Session Data Volume	bytes	long	400000
Session Packet Size	bytes	int	1000
Packet Interarrival Time	ms	float	5
Number of Packets	N/A	long	500
Number of Packets Delayed	N/A	long	5
Percentage of Delayed Packets	%	float	1.00%
Total Delay	Ms	long	50
Average Delay of Delayed Packets	Ms	int	10
Session Average Throughput	kbps	float	400
Average Number of Consumed Slots	N/A	int	15
Maximum Number of Consumed Slots	N/A	int	18
Average AMC	N/A	string	16QAM 1/2
Best AMC	N/A	string	64QAM 2/3
Worse AMC	N/A	string	QPSK 1/2

Table F.5 – WWW Session Summary.

Parameter / Header Entry	Unit	Type	Example
User Number	N/A	int	1
User's Session Number	N/A	int	3
Session Start	seconds	int	12
Session End	seconds	float	20
Session Data Volume	bytes	long	400000
Packet Size	bytes	int	1000
Packet Interarrival Time	ms	float	6,1
Number of Packets	N/A	long	500
Number of Packets Delayed	N/A	long	5
Percentage of Delayed Packets	%	float	1,00%
Total Delay	ms	long	50
Average Delay of Delayed Packets	ms	int	10
Session Average Throughput	kbps	float	400
Session Maximum Throughput	kbps	float	1200
Average Number of Consumed Slots	N/A	int	12
Maximum Number of Consumed Slots	N/A	int	14
Average AMC	N/A	string	16QAM 1/2
Best AMC	N/A	string	64QAM 2/3
Worse AMC	N/A	string	QPSK 1/2

- Email Sessions Summary: this file summarises the outcome of each attempted Email session, Table F.6.

Table F.6 – Email Session Summary.

Parameter / Header Entry	Unit	Type	Example
User Number	N/A	int	1
User's Session Number	N/A	int	3
Session Start	seconds	int	12
Session End	seconds	float	20
Session Data Volume	bytes	long	400000
Reading Time	seconds	int	100
Packet Size	bytes	int	1000
Packet Interarrival Time	ms	float	5
Number of Packets	N/A	long	500
Number of Packets Delayed	N/A	long	5
Percentage of Delayed Packets	%	float	1.00%
Total Delay	ms	long	50
Average Delay of Delayed Packets	ms	int	10
Session Average Throughput	kbps	float	400
Session Maximum Throughput	kbps	float	1200
Average Number of Consumed Slots	N/A	int	13
Maximum Number of Consumed Slots	N/A	int	16
Average AMC	N/A	string	16QAM 1/2
Best AMC	N/A	string	64QAM 2/3
Worse AMC	N/A	string	QPSK 1/2

- BS Log: this file records every 1 second for each BS sector the usage of BS resources, Table F.7.
- Network Counters: this file stores the main indicators regarding network behaviour, including load, number of users served, BS resources usage, signal strength, etc., Table F.8.

Table F.7 – BS Log Table.

Parameter / Header Entry	Unit	Type	Example
BS Sector	N/A	int	1
Frame	N/A	int	300
Served VoIP Users	N/A	int	10
Slots Used for VoIP	N/A	int	30
VoIP Instantaneous Throughput	kbps	float	122
Served Video Telephony Users	N/A	int	2
Slots Used for Video Telephony	N/A	int	20
Video Telephony Instantaneous Throughput	kbps	float	128
Served Streaming Users	N/A	int	1
Slots Used for Streaming	N/A	int	13
Streaming Instantaneous Throughput	kbps	float	64
Served FTP Users	N/A	int	3
Slots Used for FTP	N/A	int	50
FTP Instantaneous Throughput	kbps	float	1500
Served WWW Users	N/A	int	5
Slots Used for WWW	N/A	int	100
WWW Instantaneous Throughput	kbps	float	1500
Served Email Users	N/A	int	3
Slots Used for Email	N/A	int	60
Email Instantaneous Throughput	kbps	float	720
Overhead Slots	N/A	int	30
Free Slots	N/A	int	404
System Load	%	string	75%
Total Instantaneous Throughput	kbps	float	4034

Table F.8 – Network Counters Table.

Parameter / Header Entry	Unit	Type	Example
BS Sector	N/A	int	1
Average Served VoIP Users	N/A	float	10
Max Served VoIP Users	N/A	int	12
Average Served Video Telephony Users	N/A	float	2
Max Served Video Telephony Users	N/A	int	3
Average Served Streaming Users	N/A	float	1
Max Served Streaming Users	N/A	int	2
Average Served FTP Users	N/A	float	3
Max Served FTP Users	N/A	int	4
Average Served WWW Users	N/A	float	5
Max Served WWW Users	N/A	int	6
Average Served Email Users	N/A	float	3
Max Served Email Users	N/A	int	4
Average Slots Used for VoIP	N/A	float	30
Max Slots Used for VoIP	N/A	int	33

Average Slots Used for Video Telephony	N/A	float	20
Max Slots Used for Video Telephony	N/A	int	23
Average Slots Used for Streaming	N/A	float	13
Max Slots Used for Streaming	N/A	int	16
Average Slots Used for FTP	N/A	float	50
Max Slots Used for FTP	N/A	int	53
Average Slots Used for WWW	N/A	float	100
Max Slots Used for WWW	N/A	int	103
Average Slots Used for Email	N/A	float	60
Max Slots Used for Email	N/A	int	63
Average Overhead Slots	N/A	float	30
Max Overhead Slots	N/A	int	33
Average Free Slots	N/A	float	200
Minimum Free Slots	N/A	int	100
Average System Load	%	float	40%
Max System Load	%	string	70%
Average VoIP Throughput	kbps	float	122
Max VoIP Throughput	kbps	float	146.4
Average Video Telephony Throughput	kbps	float	128
Max Video Telephony Throughput	kbps	float	192
Average Streaming Throughput	kbps	float	64
Max Streaming Throughput	kbps	float	128
Average FTP Throughput	kbps	float	1500
Max FTP Throughput	kbps	float	2000
Average WWW Throughput	kbps	float	1500
Max WWW Throughput	kbps	float	1800
Average Email Throughput	kbps	float	720
Max Email Throughput	kbps	float	960
VoIP Traffic Volume	kBytes	long	1500
Video Telephony Traffic Volume	kBytes	long	2000
Streaming Traffic Volume	kBytes	long	5000
FTP Traffic Volume	kBytes	long	2000
WWW Traffic Volume	kBytes	long	10000
Email Traffic Volume	kBytes	long	3000
Percentage of Lost VoIP Frames	%	float	0.50%
Percentage of Lost Video Frames	%	float	1.00%
Percentage of Streaming Delayed Packets	%	float	1.00%
Average Delay of Streaming Delayed Packets	ms	int	15
Standard Deviation Delay of Streaming Delayed Packets	ms	int	5
Percentage of Best-Effort Delayed Packets	%	float	1.00%
Average Delay of Best-Effort Delayed Packets	ms	int	15
Standard Deviation Delay of Best-	ms	int	5

Effort Delayed Packets			
Average RSSI	dBm	float	-85
Max RSSI	dBm	float	-60
Min RSSI	dBm	float	-105
Average CINR	dB	float	17
Max CINR	dB	float	28
Min CINR	dB	float	10
QPSK1/2 Weight	%	float	20%
QPSK3/4 Weight	%	float	20%
16QAM1/2 Weight	%	float	15%
16QAM3/4 Weight	%	float	15%
64QAM2/3 Weight	%	float	15%
64QAM3/4 Weight	%	float	15%

- Call Counters: this file stores the main indicators regarding number of call attempts, blocking and dropping probability.

Table F.9 – Call Counters Table.

Parameter / Header Entry	Unit	Type	Example
VoIP Calls Attempted	N/A	int	500
Video Telephony Calls Attempted	N/A	int	5
Streaming Calls Attempted	N/A	int	20
FTP Sessions Attempted	N/A	int	10
WWW Sessions Attempted	N/A	int	50
Email Sessions Attempted	N/A	int	30
VoIP Blocking Probability	%	float	2.00%
VoIP Drop Probability	%	float	1.00%
Video Blocking Probability	%	float	3.00%
Video Drop Probability	%	float	1.50%

- Simulation Generic Information: this file summarises the start and end time of the simulation.

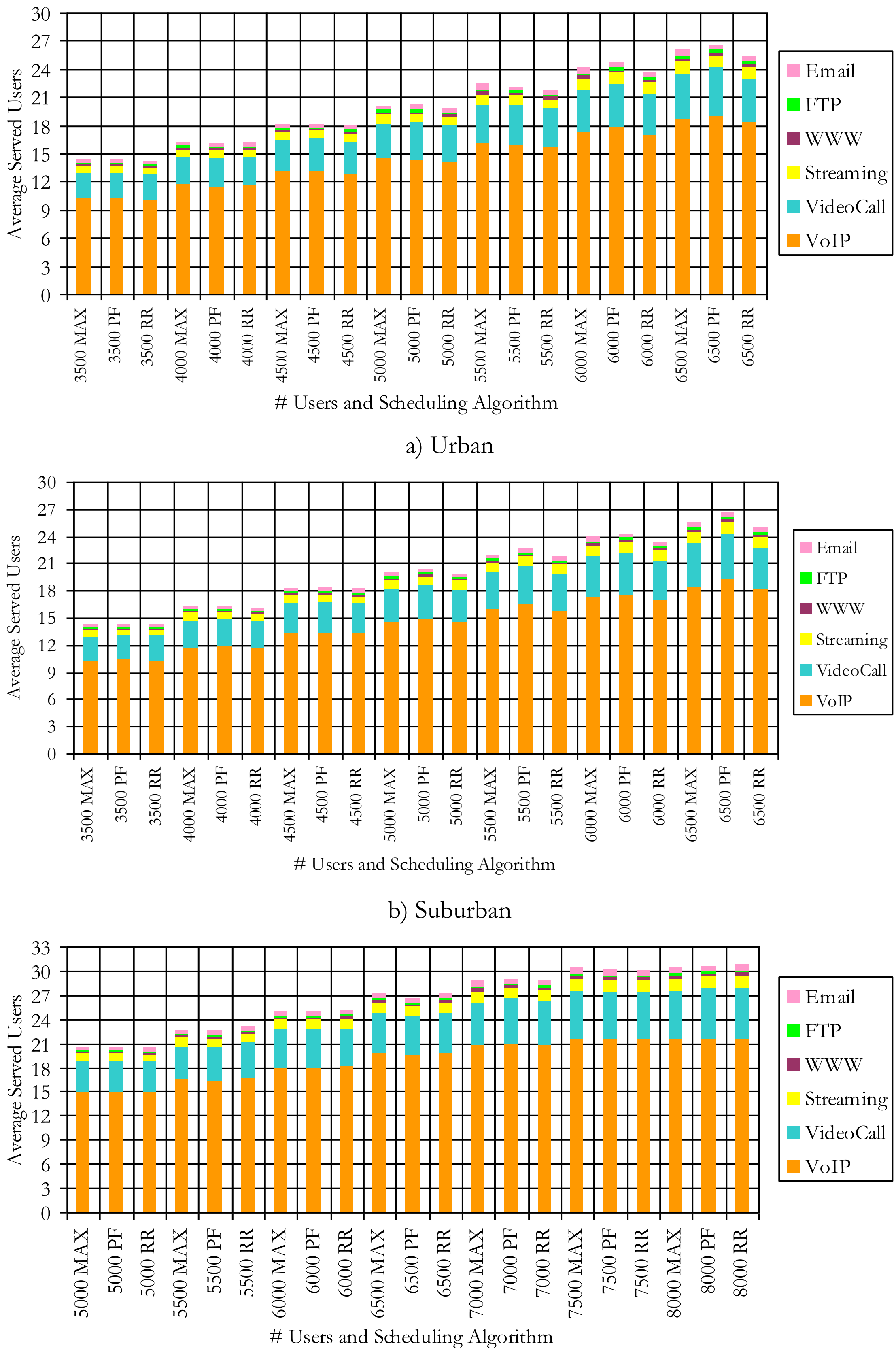
Table F.10 – Simulation Generic Information Table.

Parameter / Header Entry	Unit	Type	Example
Start Time	N/A	time	14:50:42
End Time	N/A	time	15:15:20

Annex G

Results for Variable Number of Users

This annex summarises the results obtained when varying the number of users covered by the BS without changing the traffic profile and the system configuration, for Urban, Suburban and Rural.

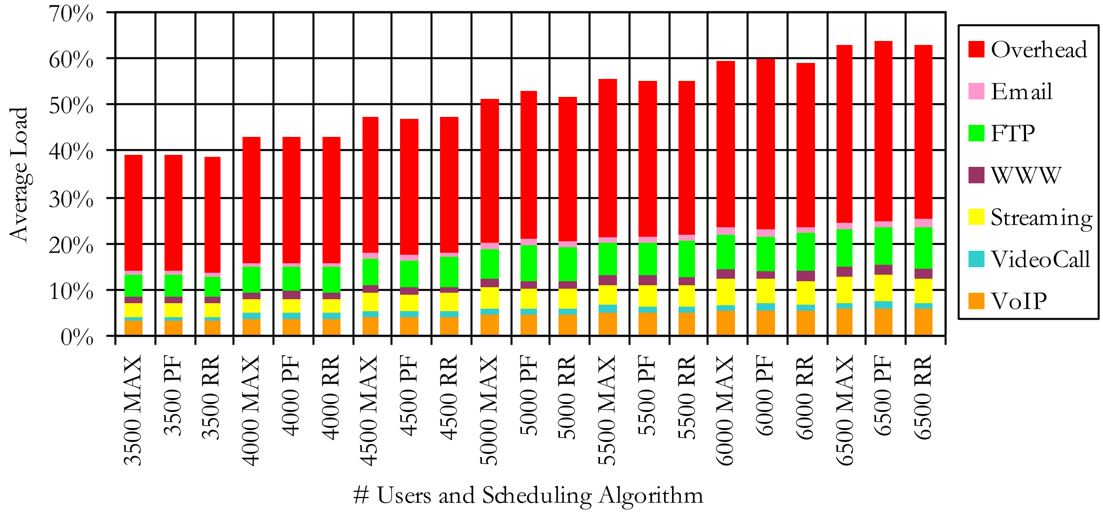


a) Urban

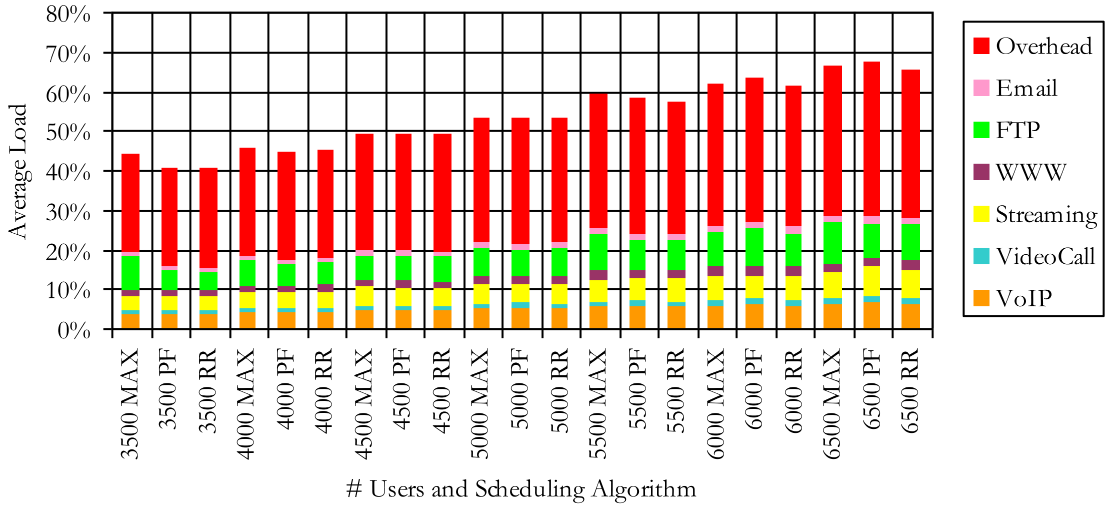
b) Suburban

c) Rural

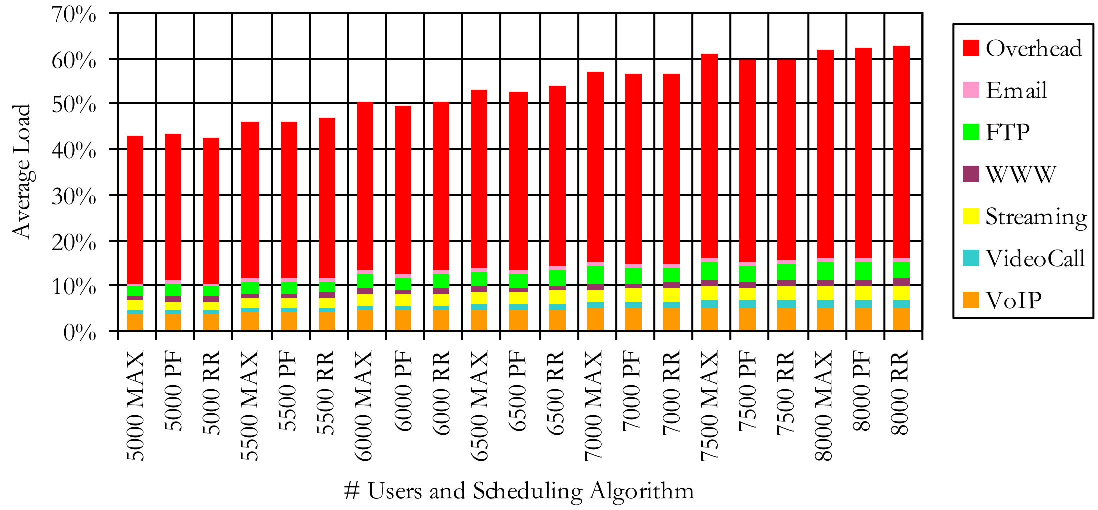
Figure G.1 – Average number of users served per application for variable number of users covered by the BS.



a) Urban

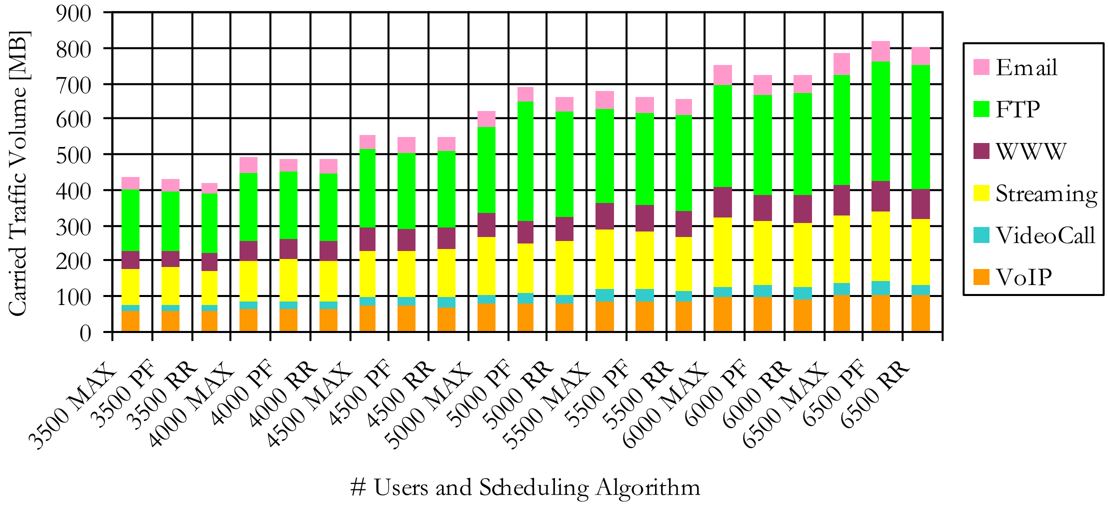


b) Suburban

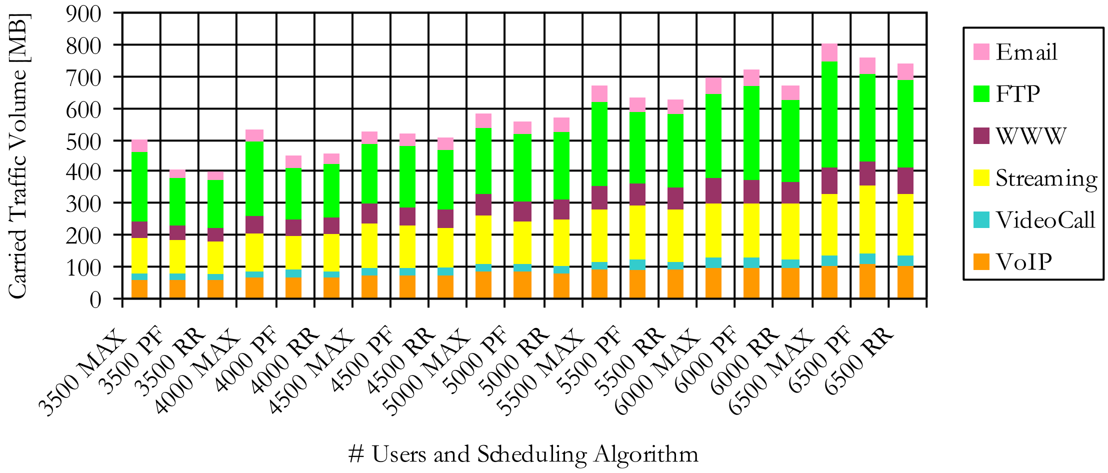


c) Rural

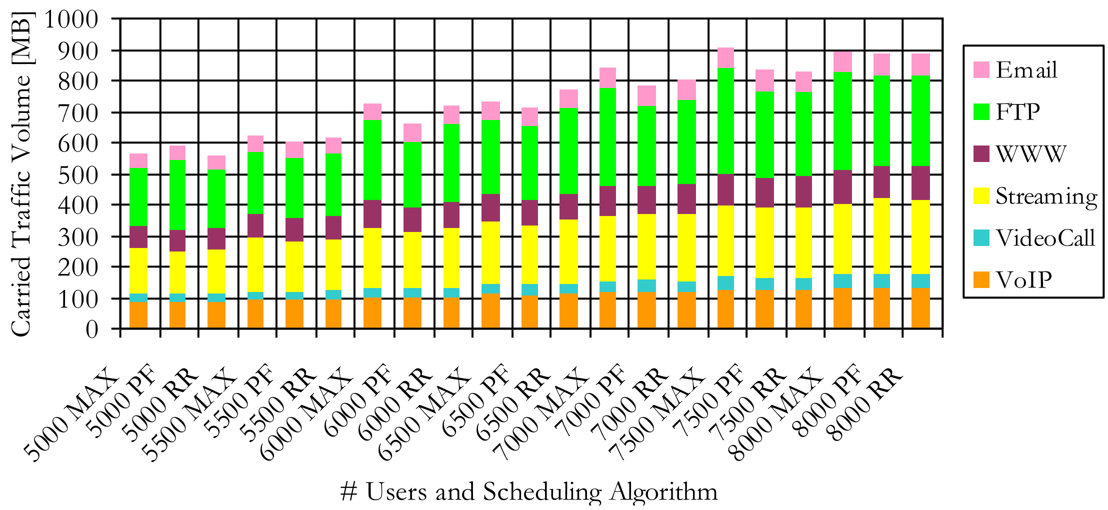
Figure G.2 - Average load for variable number of users covered by the BS.



a) Urban

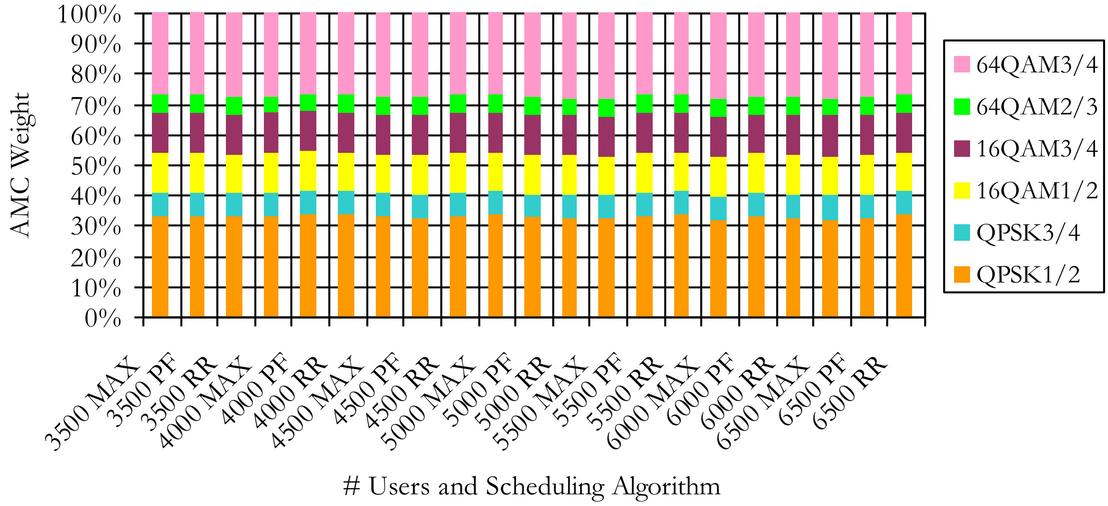


b) Suburban

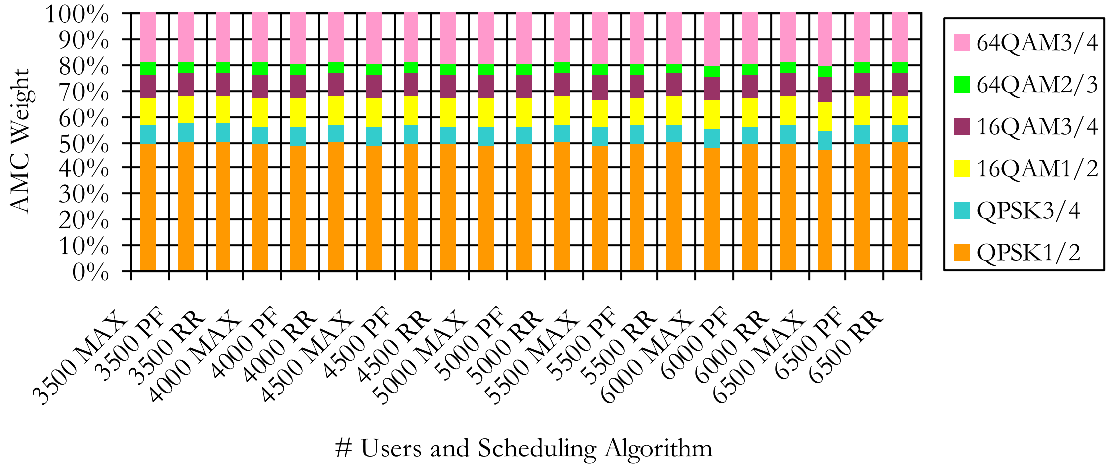


c) Rural

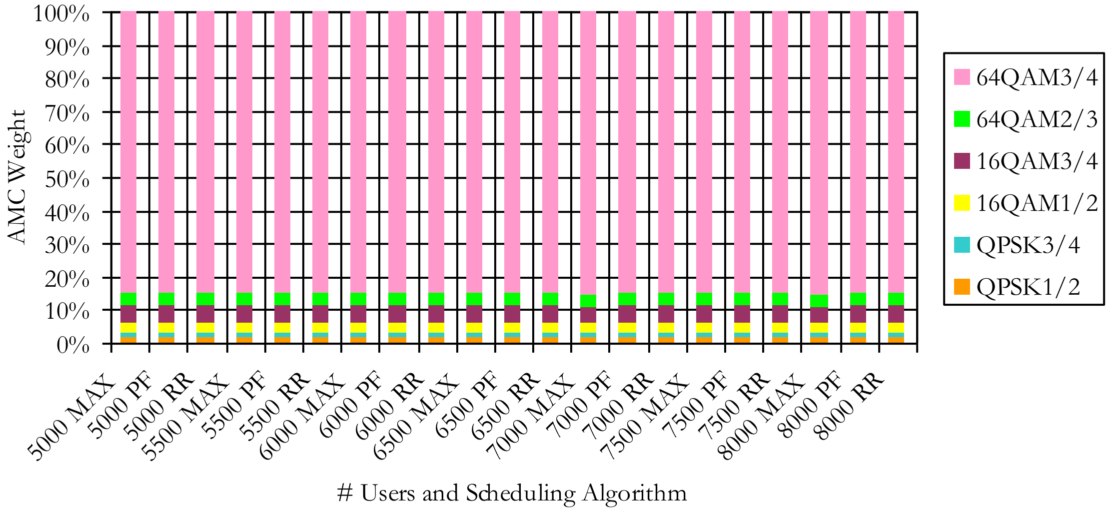
Figure G.3 – Traffic volume carried by the system.



a) Urban

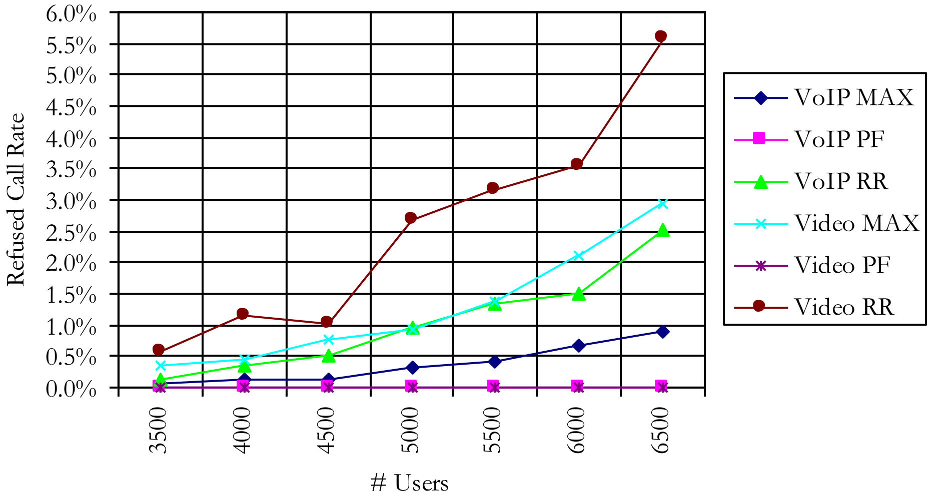


b) Suburban

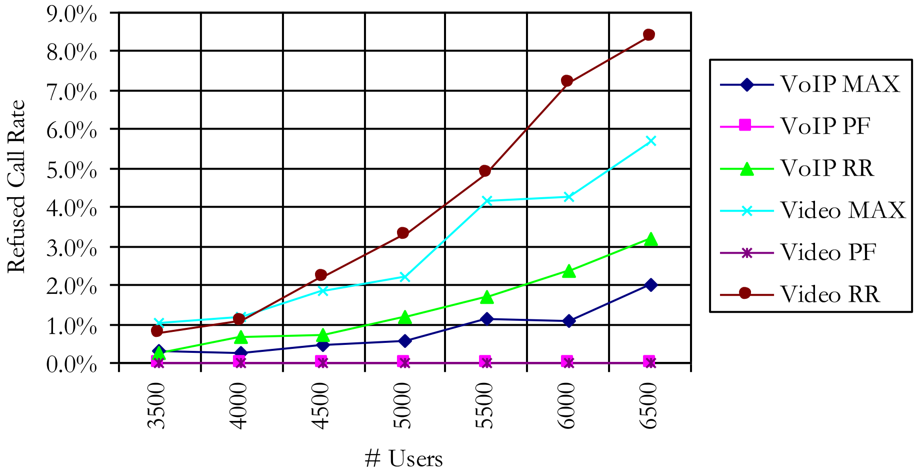


c) Rural

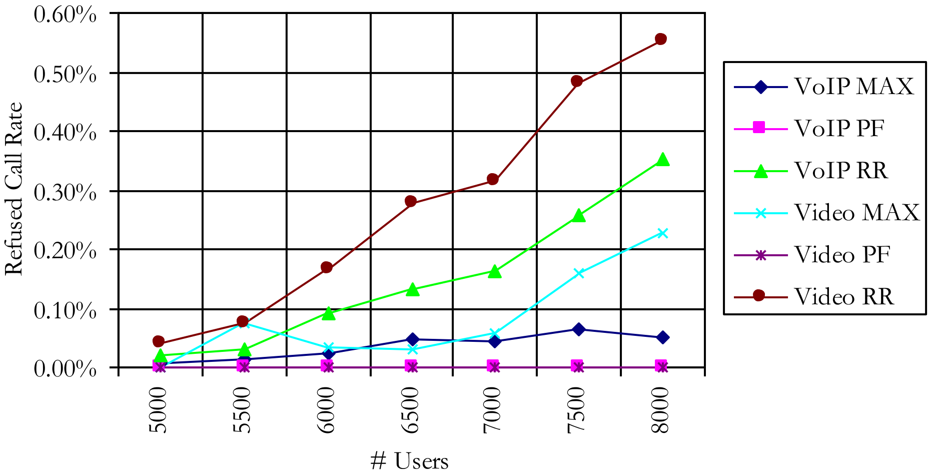
Figure G.4 – AMC distribution.



a) Urban

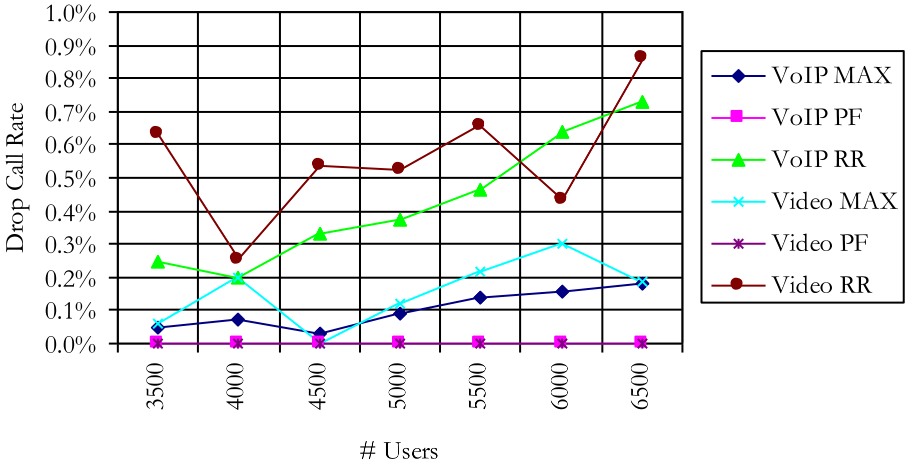


b) Suburban

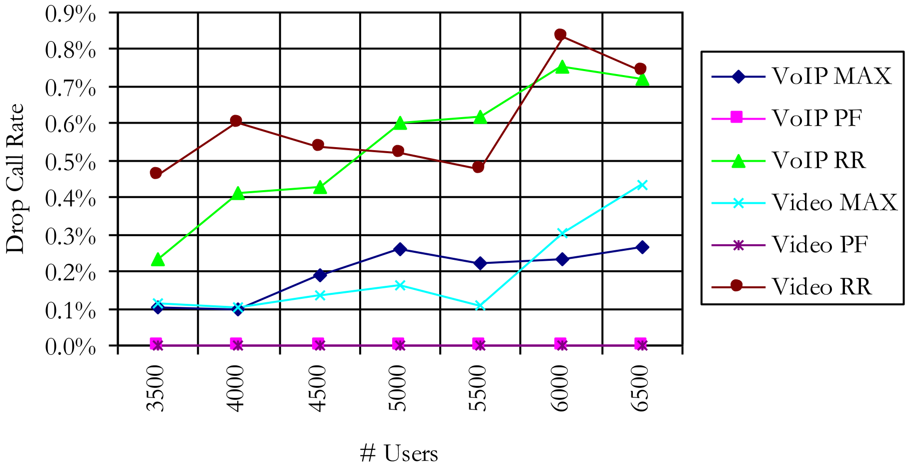


c) Rural

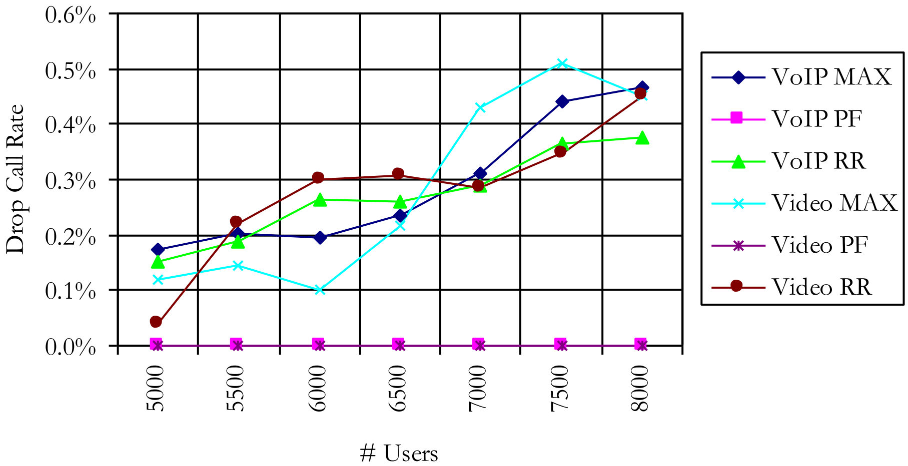
Figure G.5 – RTPS Refused Call Rate.



a) Urban

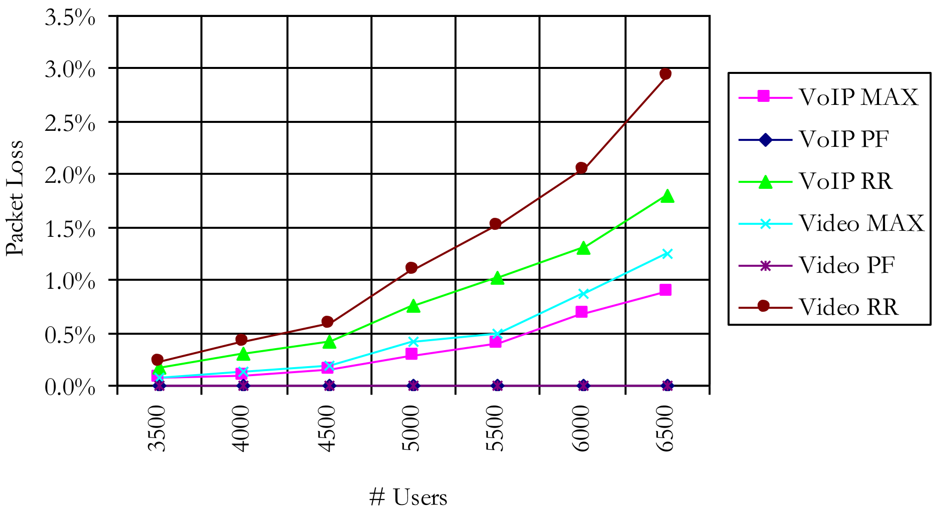


b) Suburban

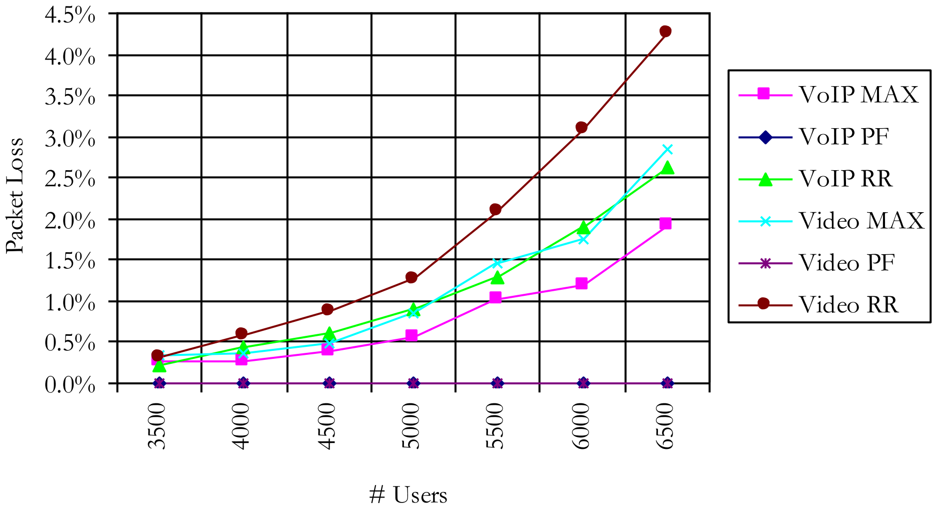


c) Rural

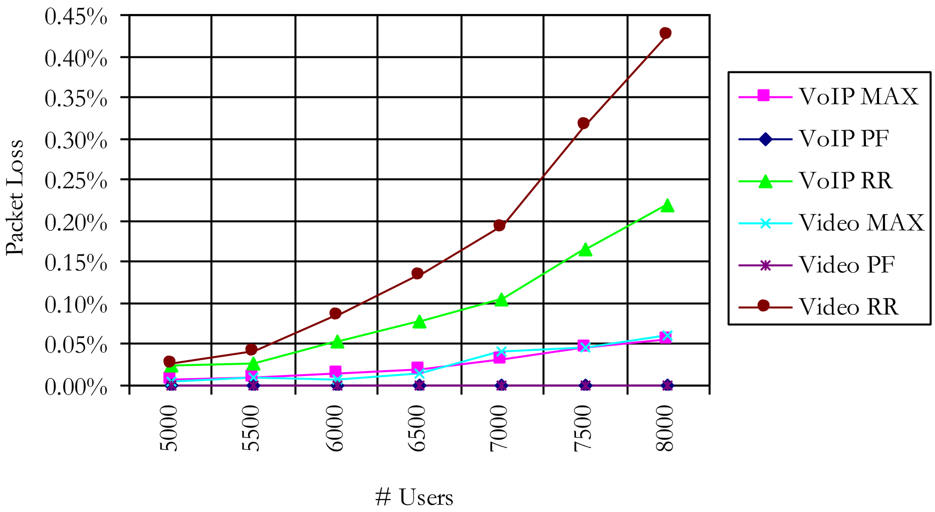
Figure G.6 – RTPS Drop Call Rate.



a) Urban

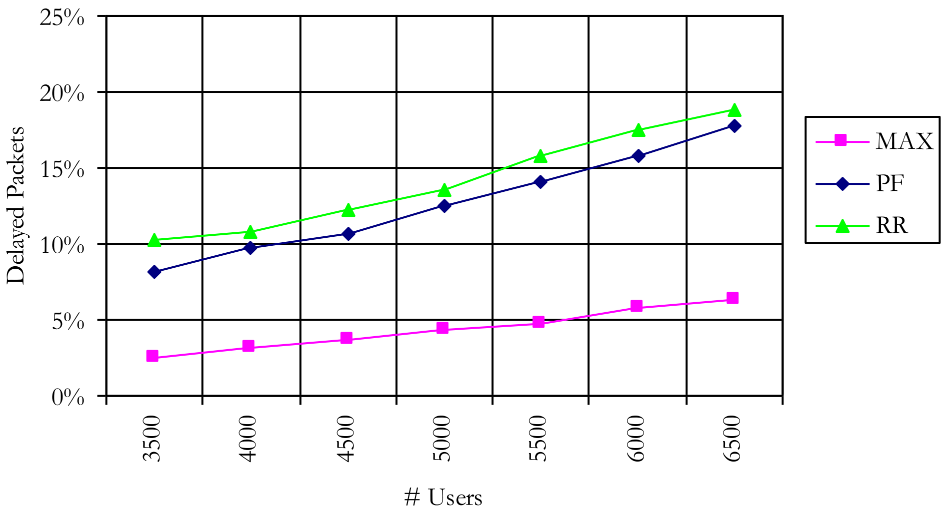


b) Suburban

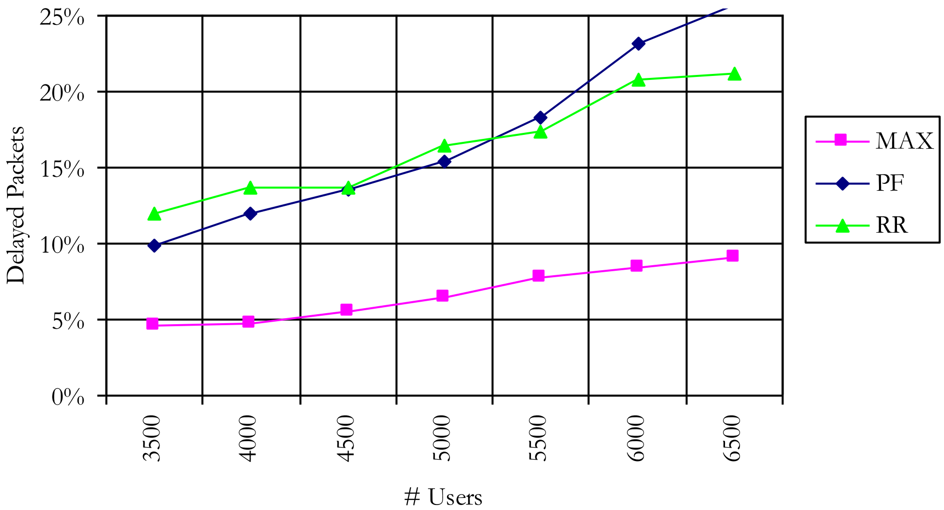


c) Rural

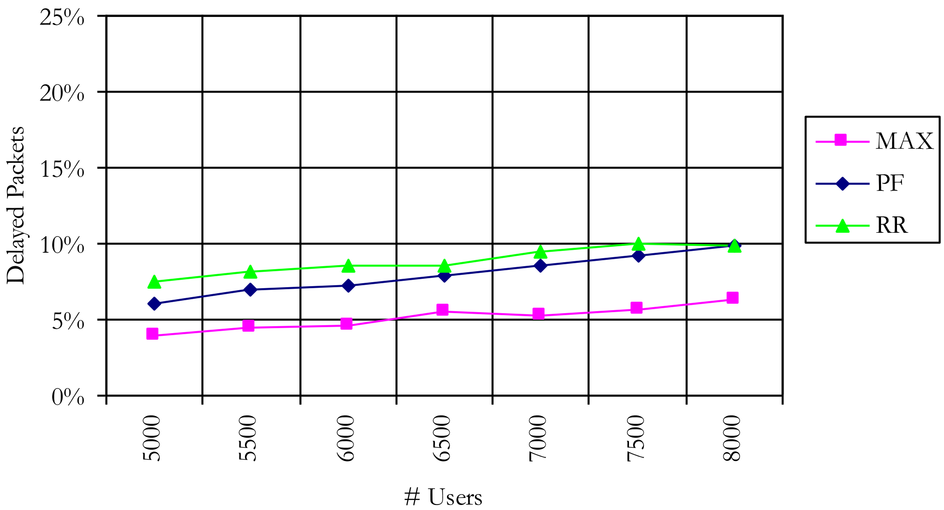
Figure G.7 – RTPS Packet Loss.



a) Urban

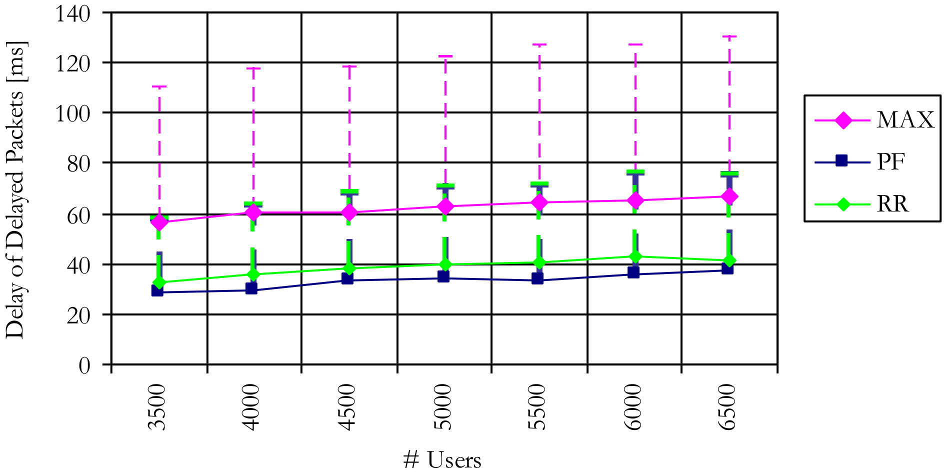


b) Suburban

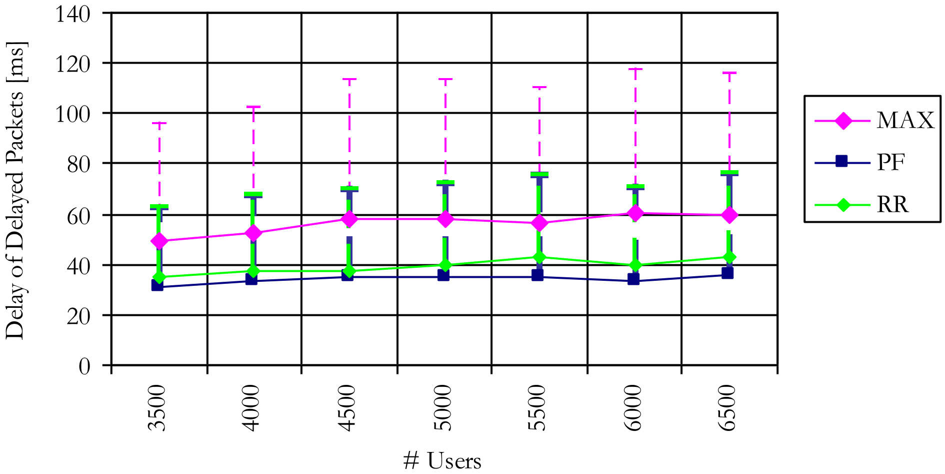


c) Rural

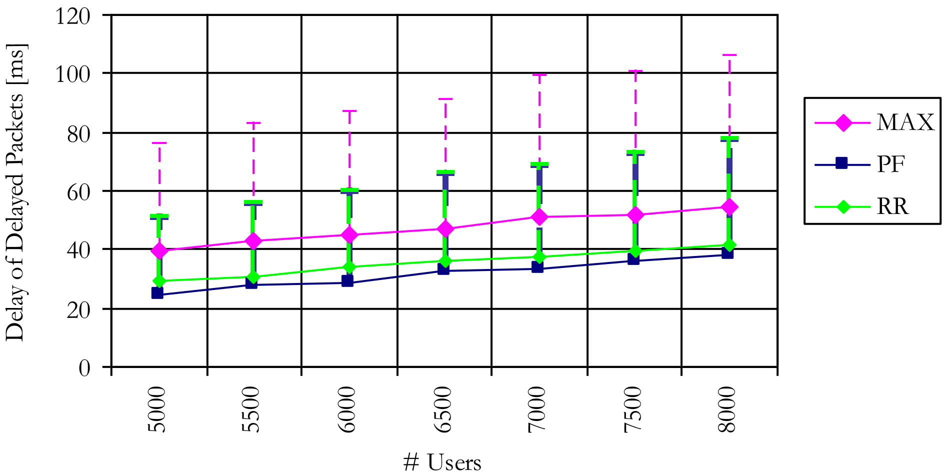
Figure G.8 – BE percentage of delayed packets.



a) Urban

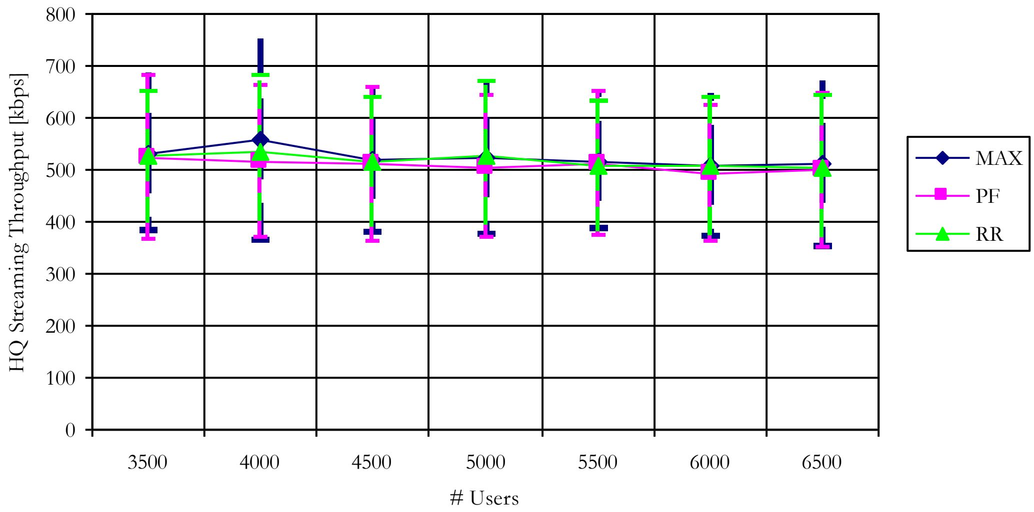


b) Suburban

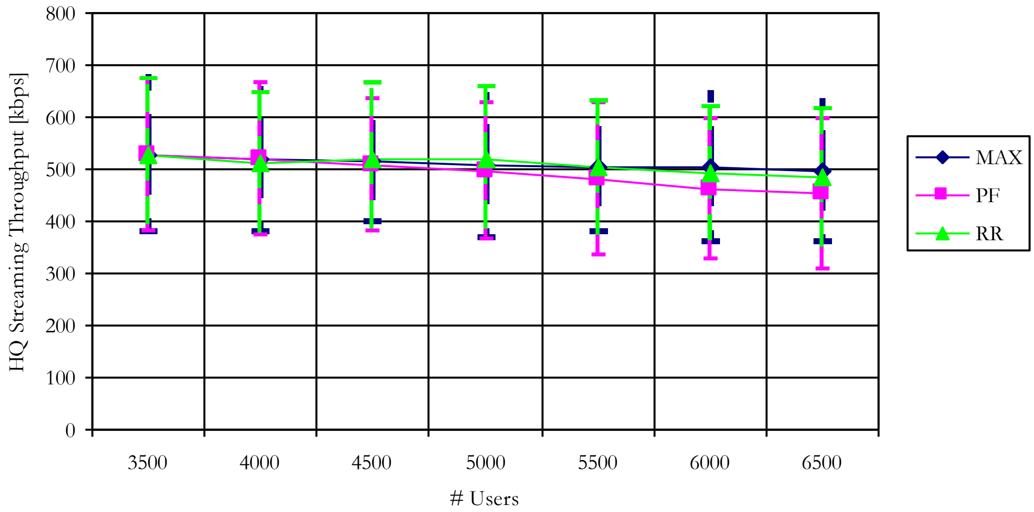


c) Rural

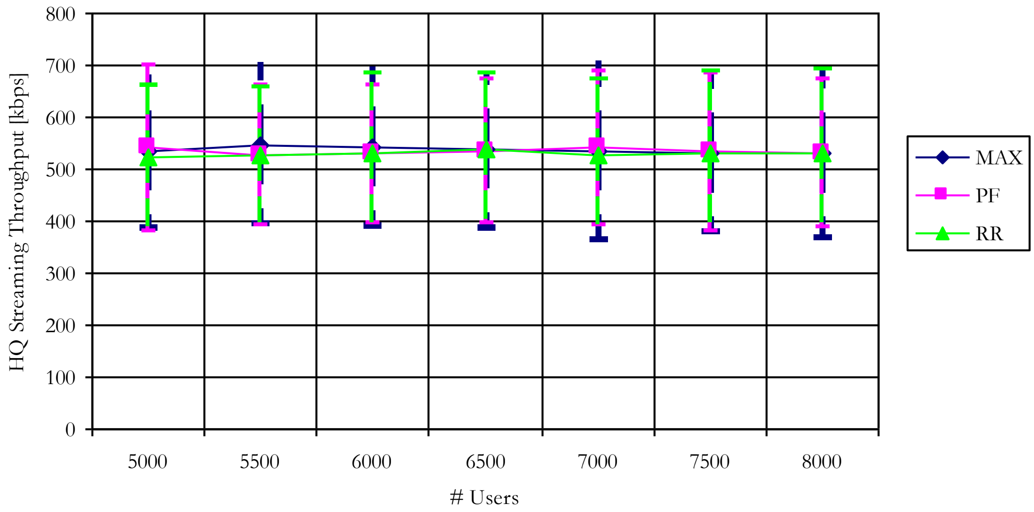
Figure G.9 – BE average delay of delayed packets.



a) Urban

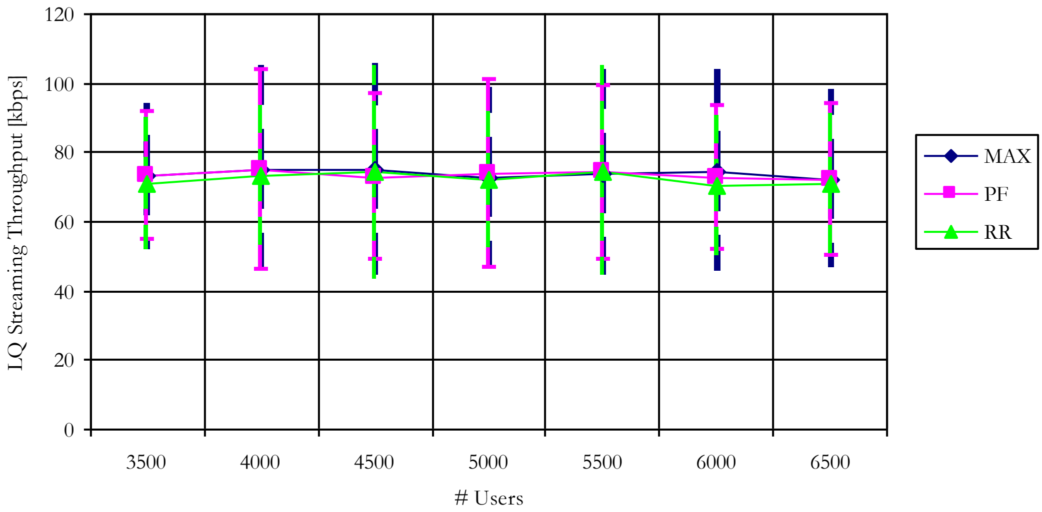


b) Suburban

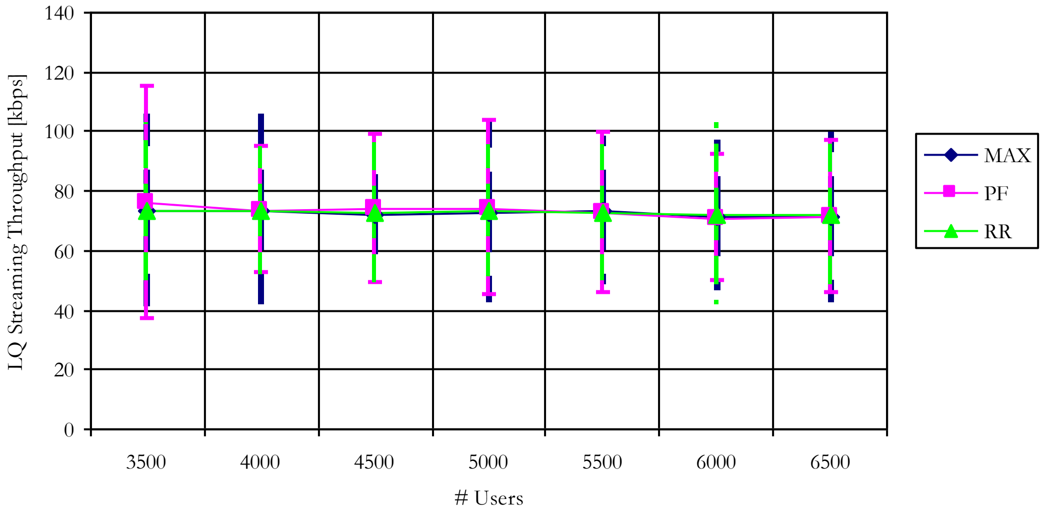


c) Rural

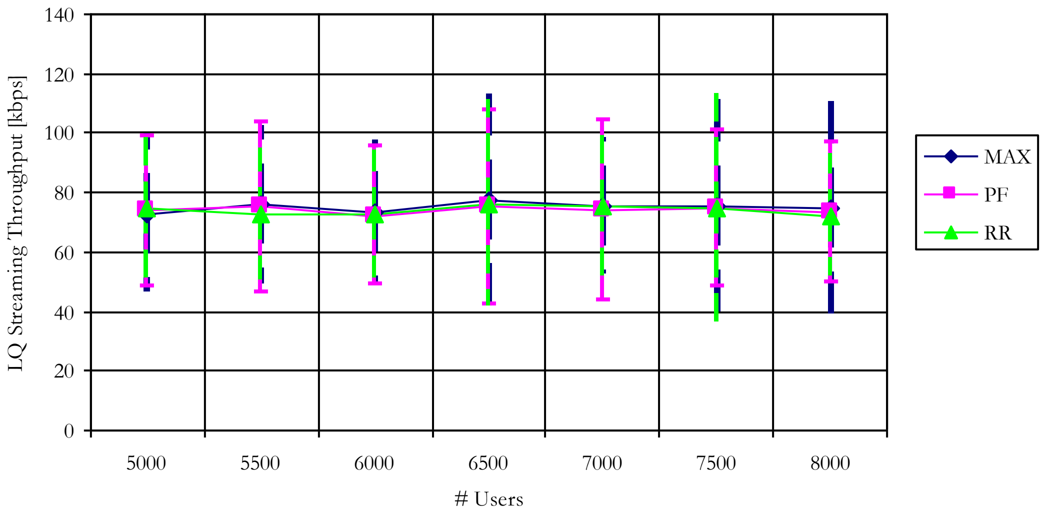
Figure G.10 – Average streaming throughput and standard deviation for increasing number of users served by BS in case of high quality video.



a) Urban

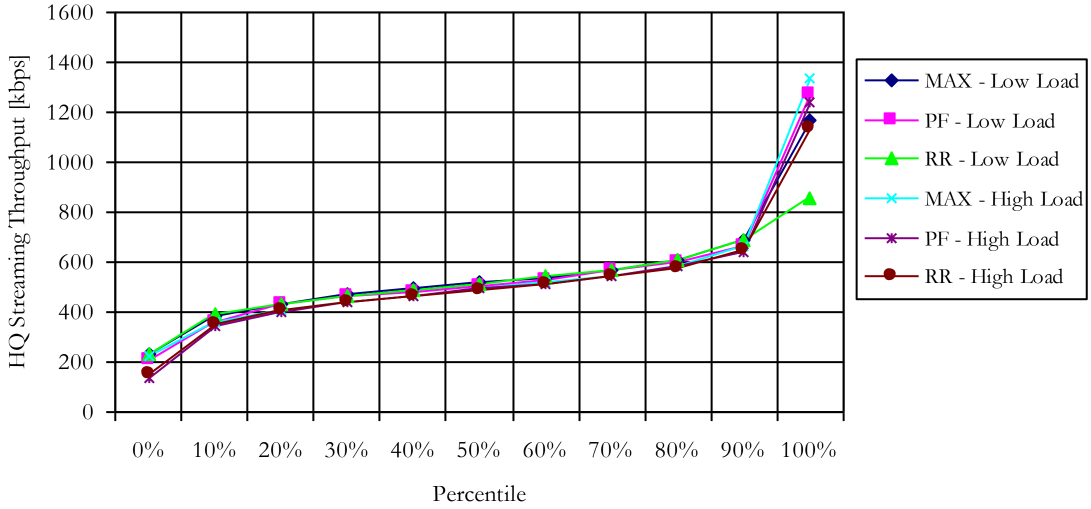


b) Suburban

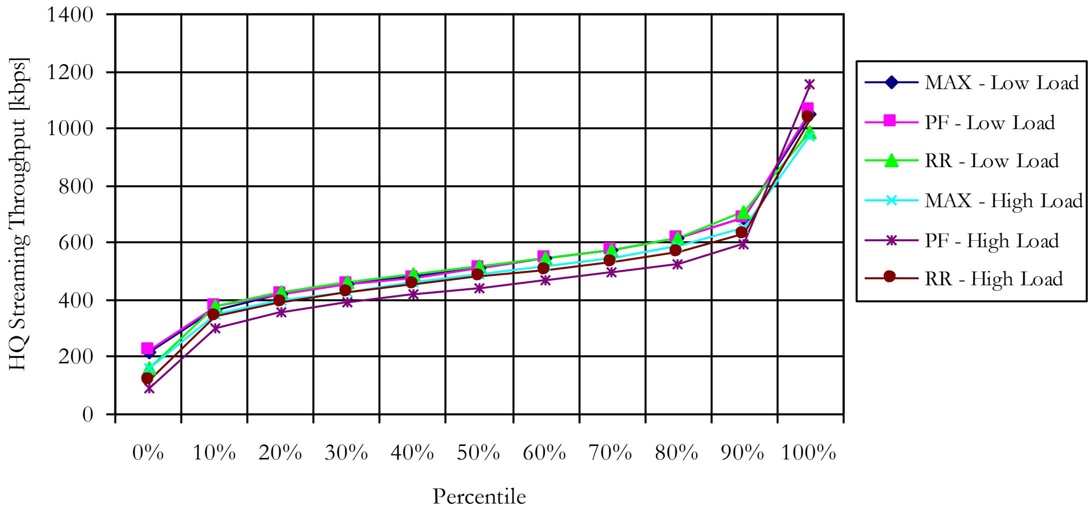


c) Rural

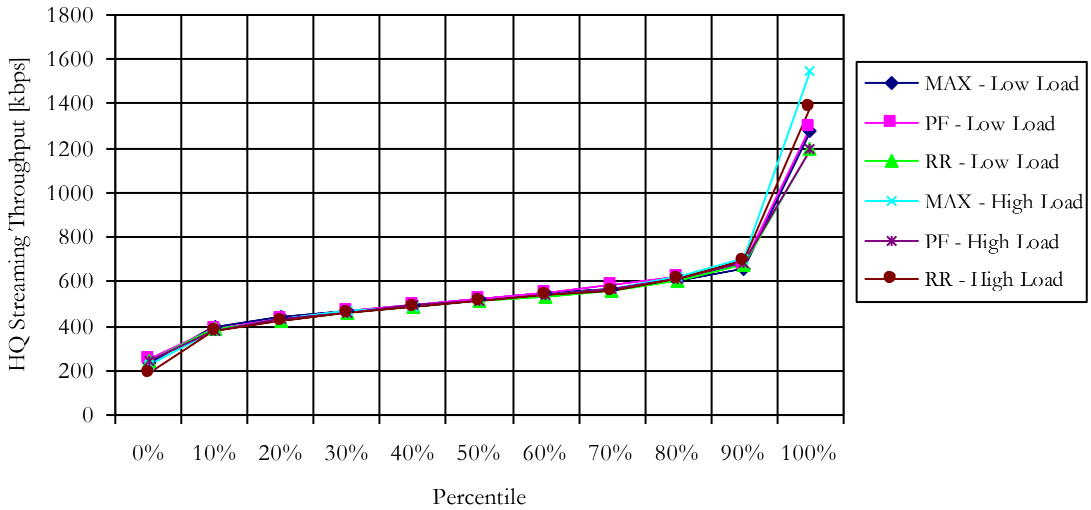
Figure G.11 – Average streaming throughput and standard deviation for increasing number of users served by BS in case of low quality video.



a) Urban

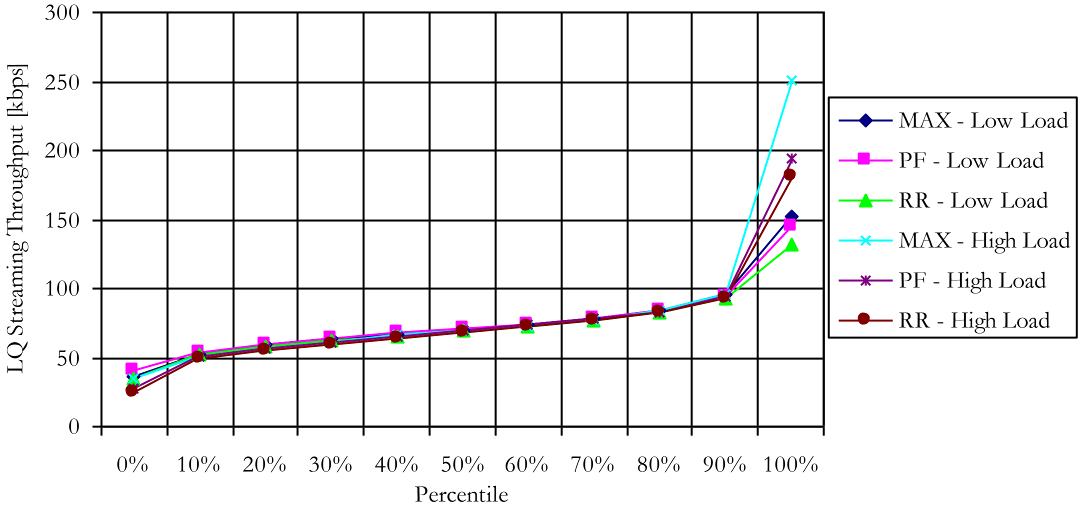


b) Suburban

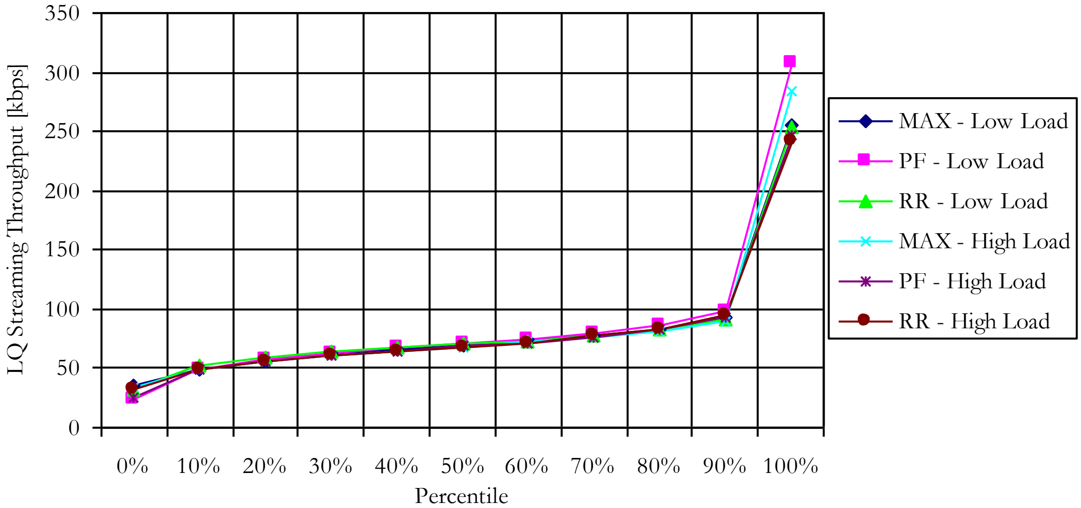


c) Rural

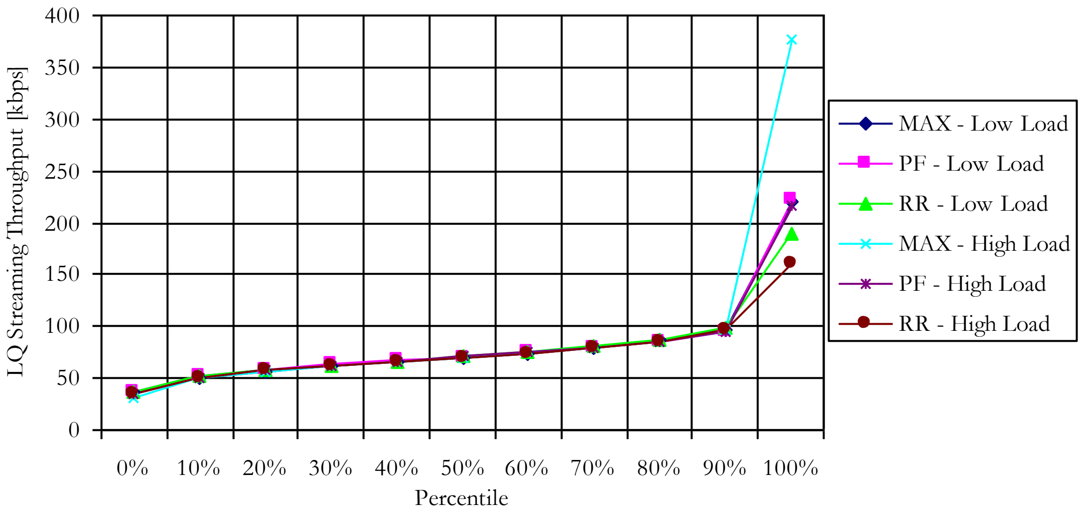
Figure G.12 - Streaming throughput cumulative distribution in case of high quality video.



a) Urban

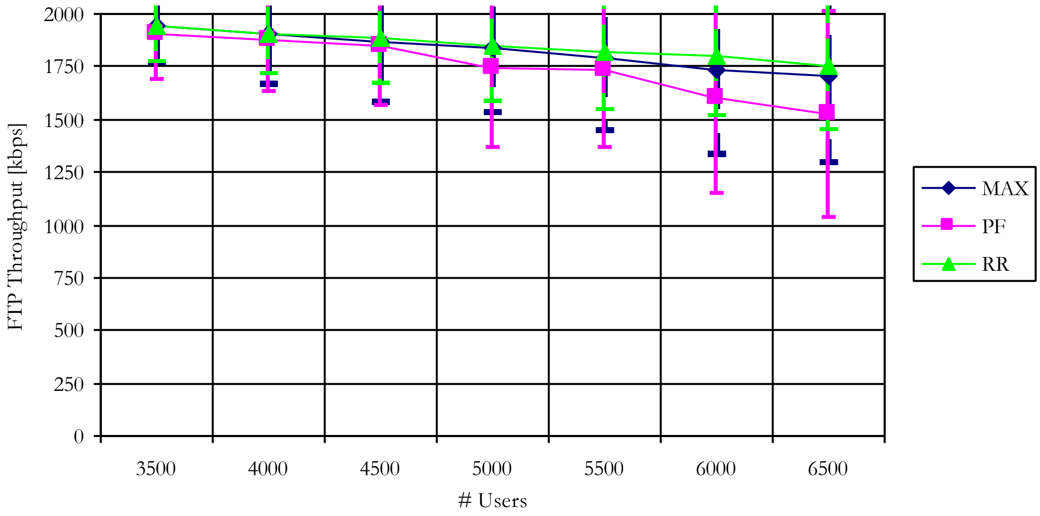


b) Suburban

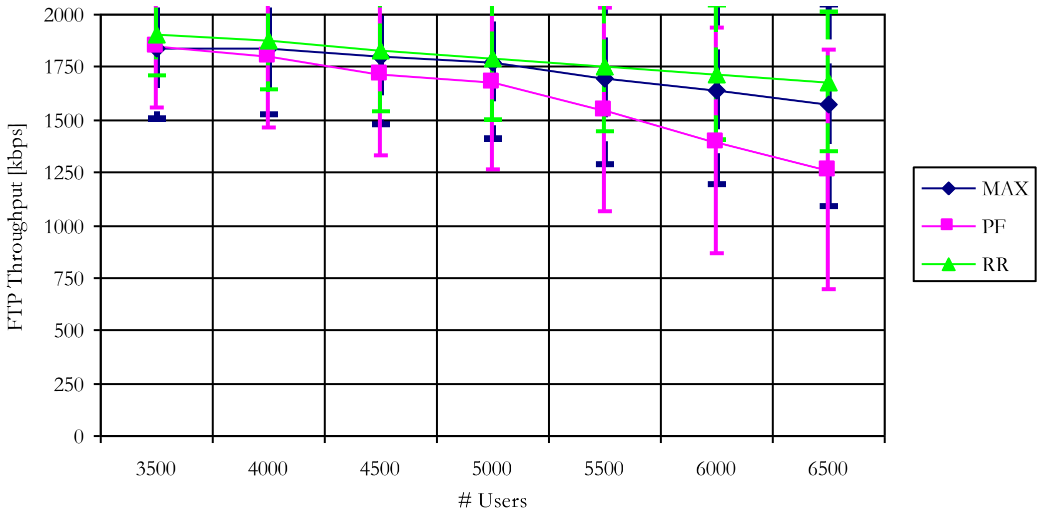


c) Rural

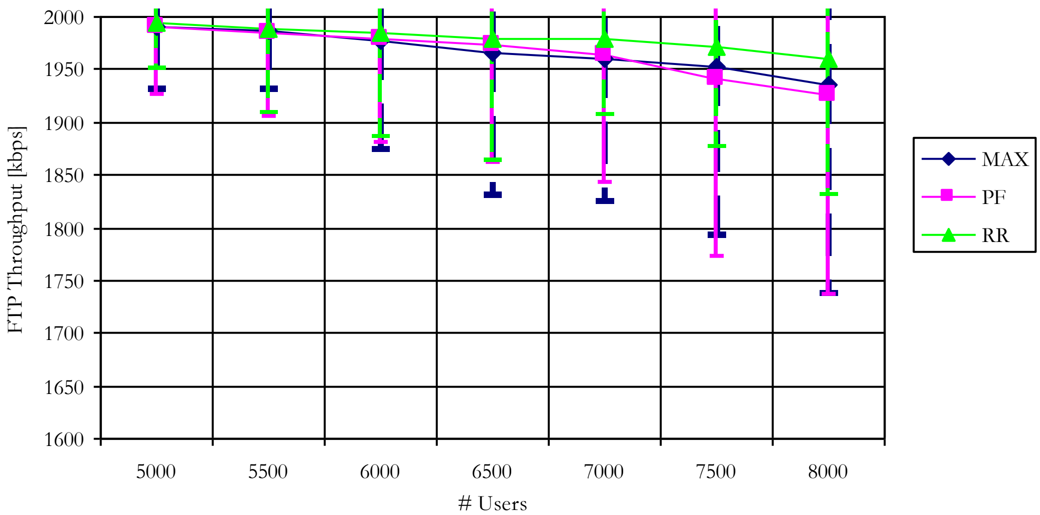
Figure G.13 - Streaming throughput cumulative distribution for low load scenario in case of low quality video.



a) Urban

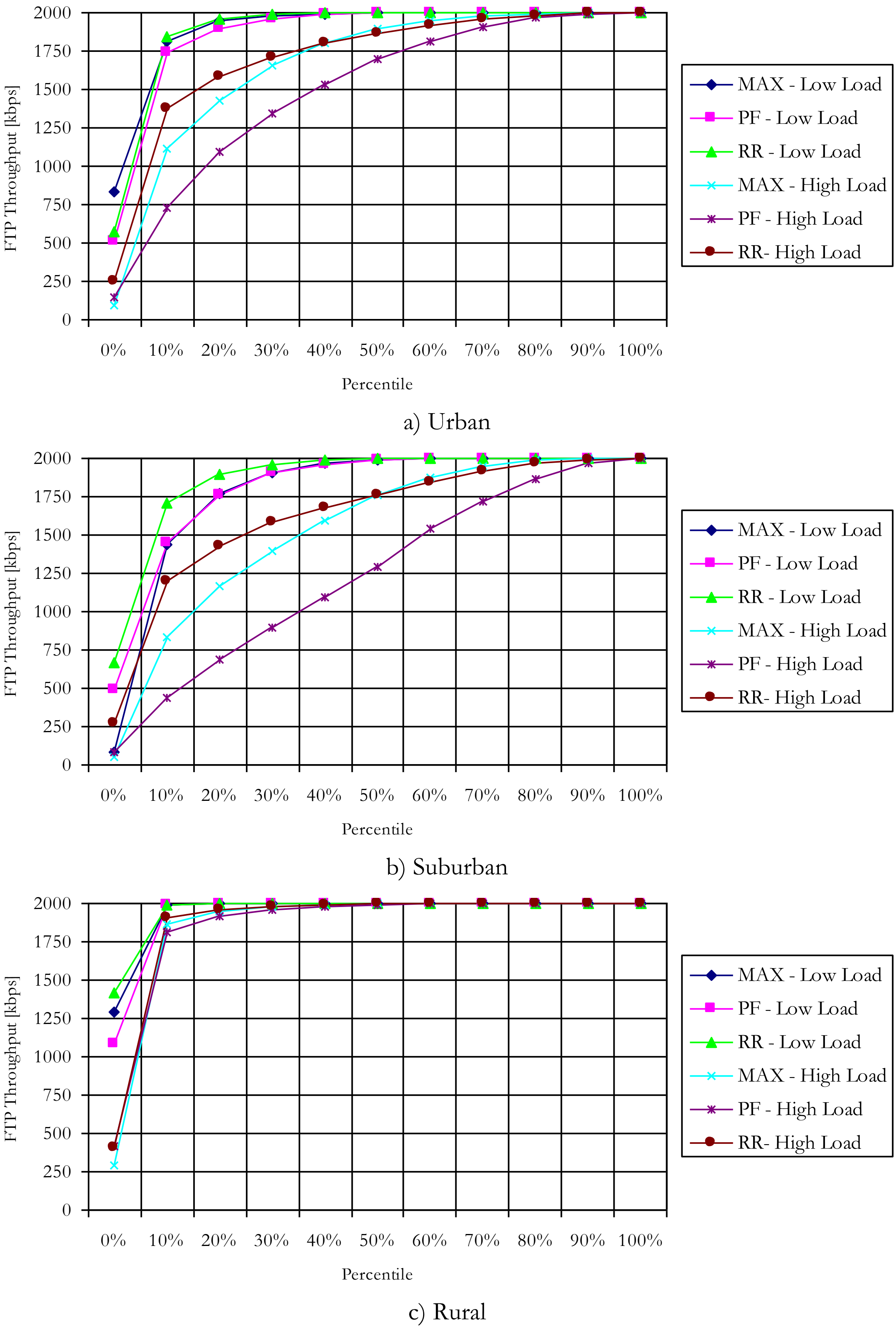


b) Suburban



c) Rural

Figure G.14 – Average per user FTP throughput and standard deviation for increasing number of users served by BS.



a) Urban

b) Suburban

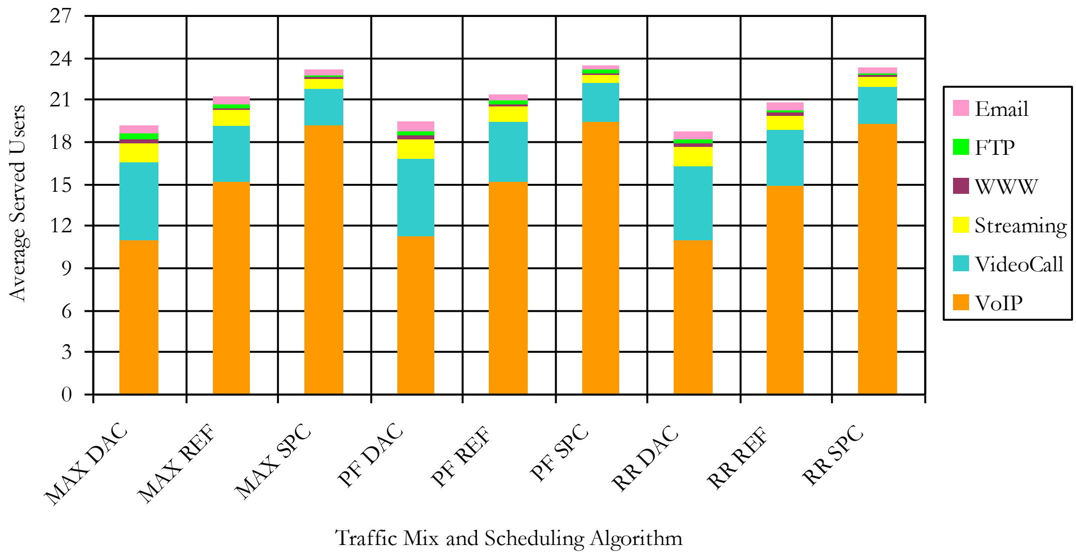
c) Rural

Figure G.15 - FTP throughput cumulative distribution.

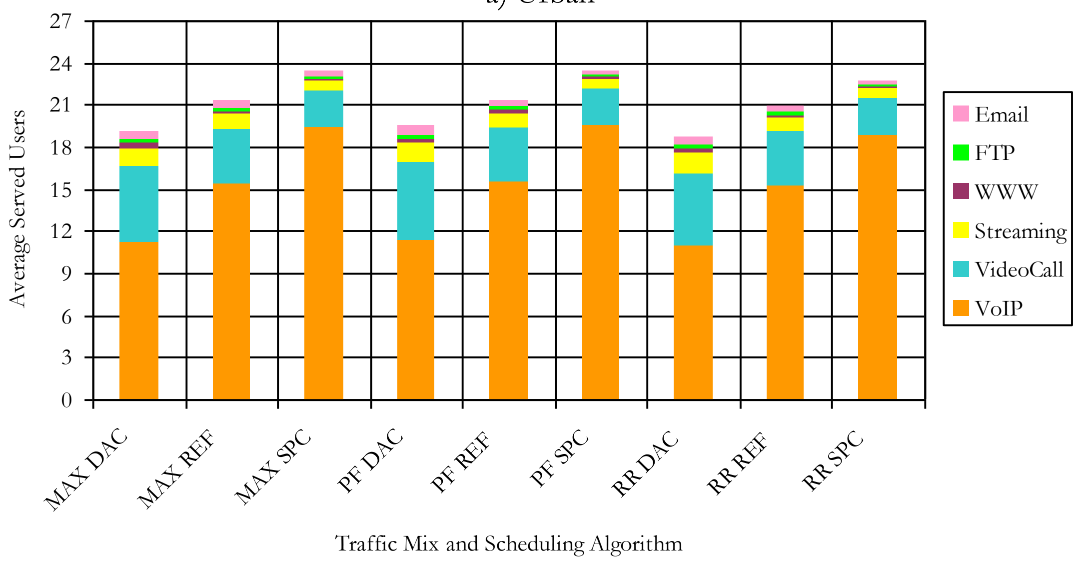
Annex H

Results for Variable per User Traffic Mix

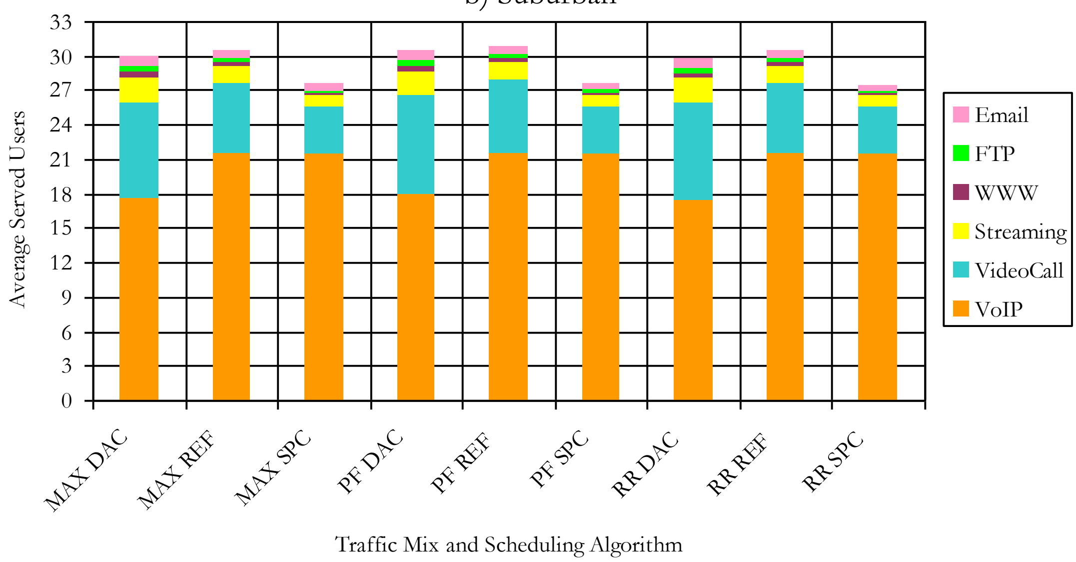
This annex summarises the results obtained when varying the per user traffic mix without changing the number of users and the system configuration, for Urban, Suburban and Rural.



a) Urban

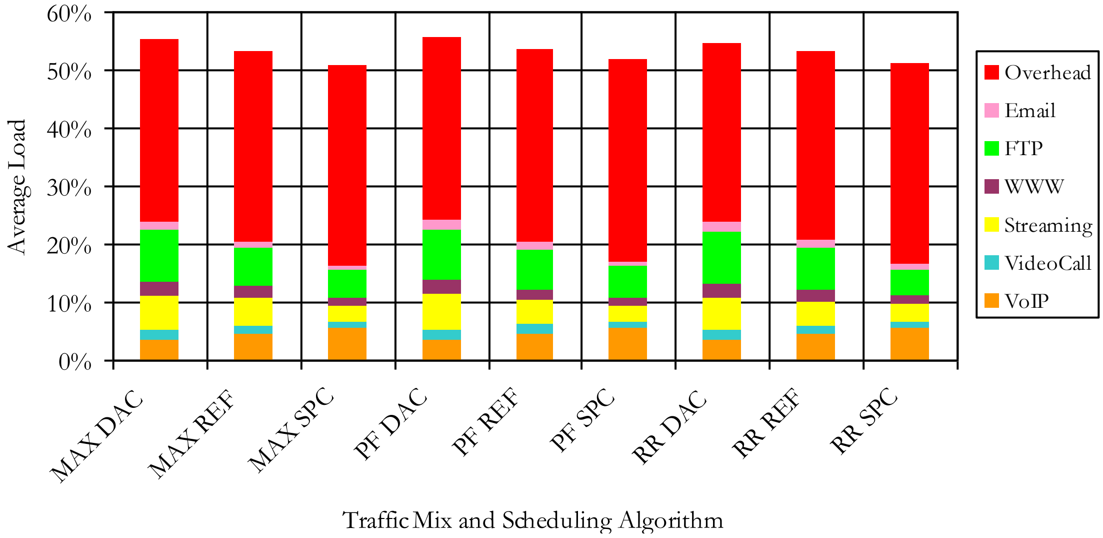


b) Suburban

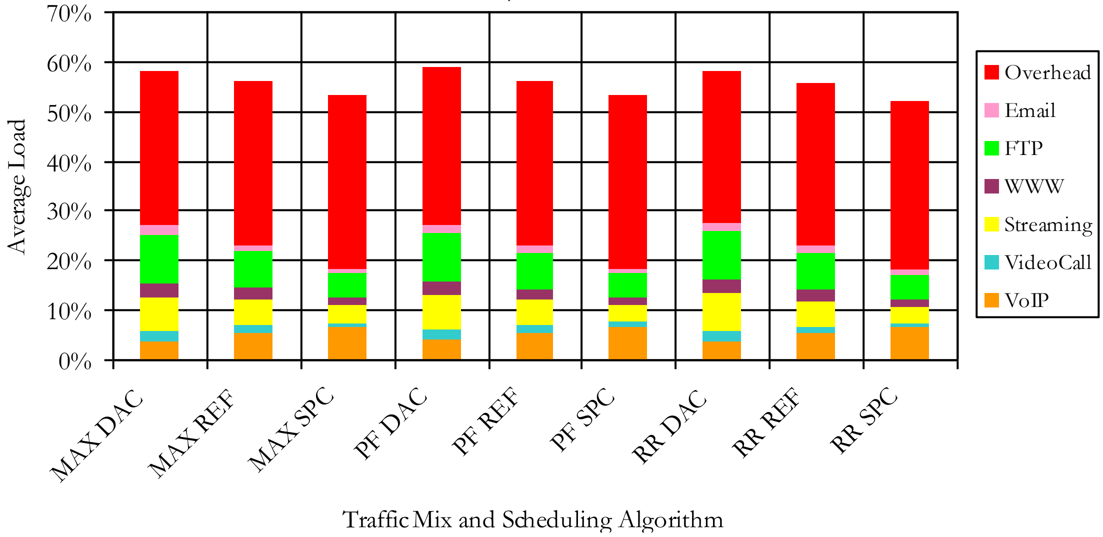


c) Rural

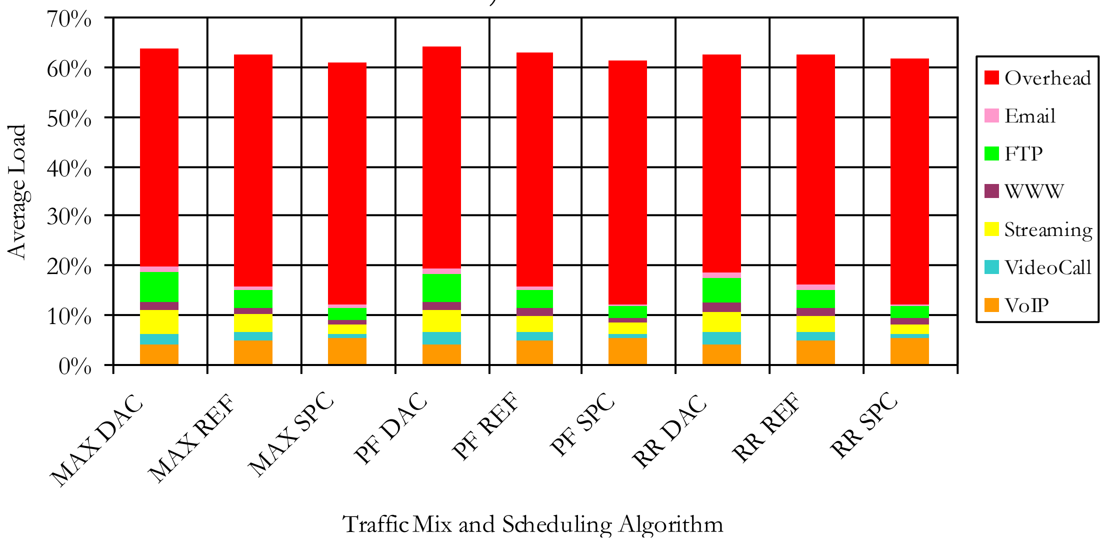
Figure H.1 – Average number of users served per application for variable per user traffic mix.



a) Urban

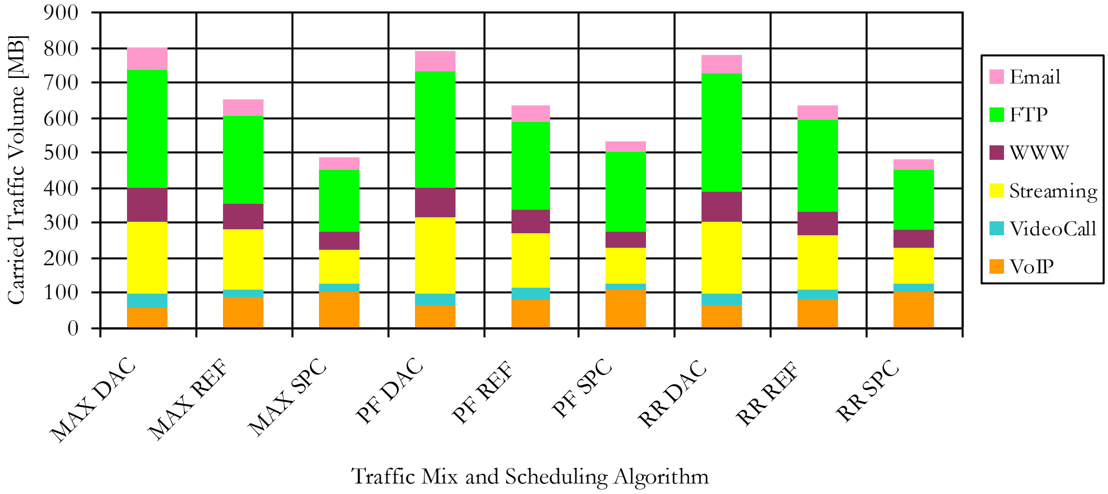


b) Suburban

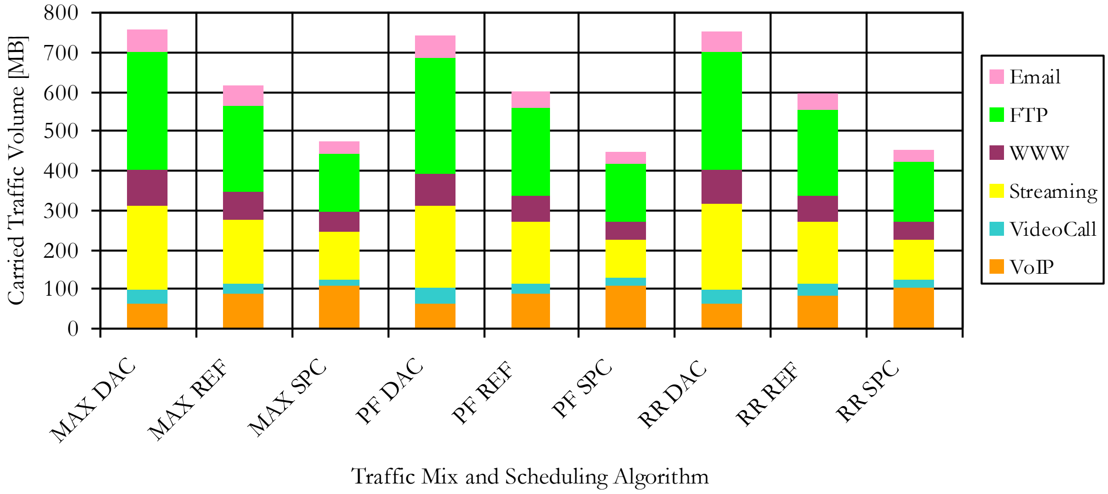


c) Rural

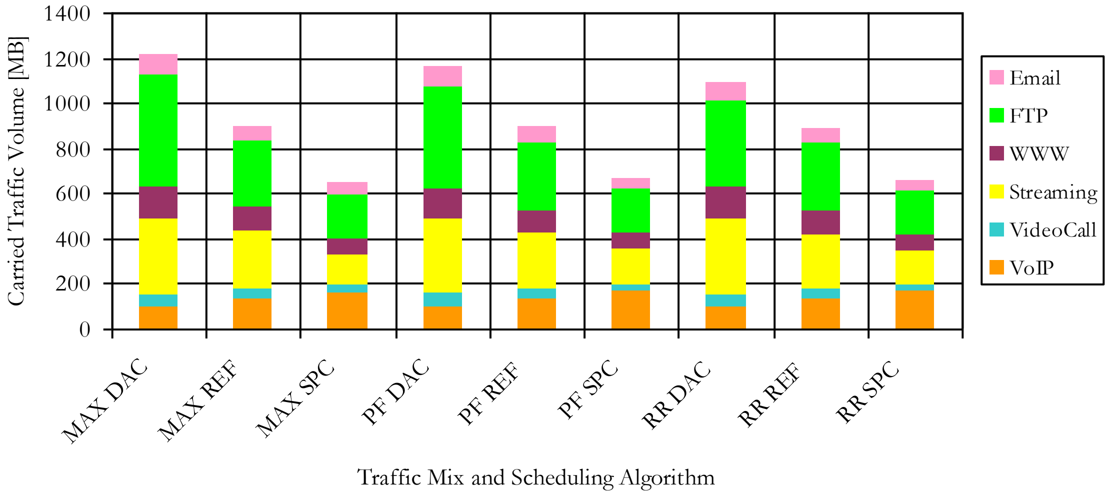
Figure H.2 - Average load for variable per user traffic mix.



a) Urban

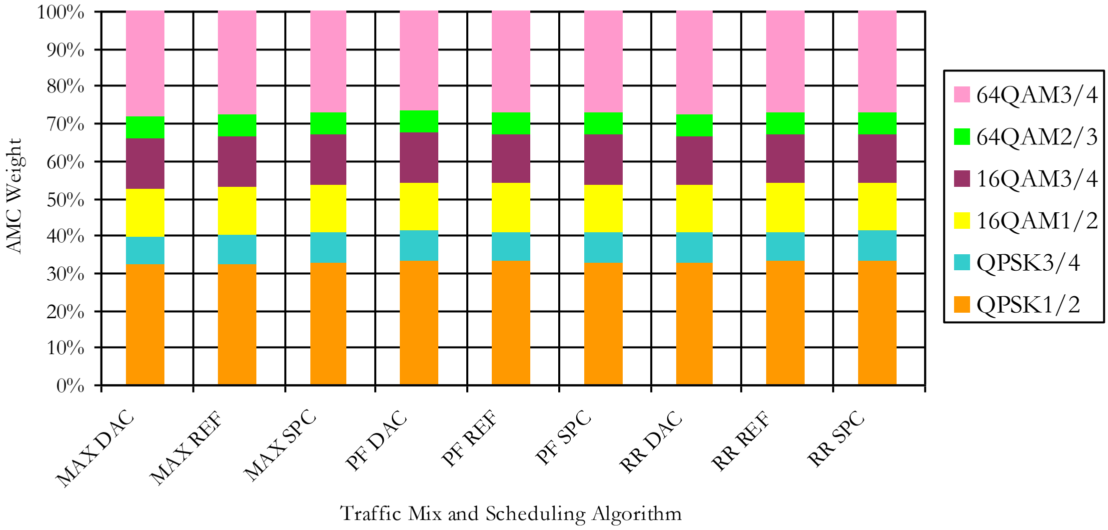


b) Suburban

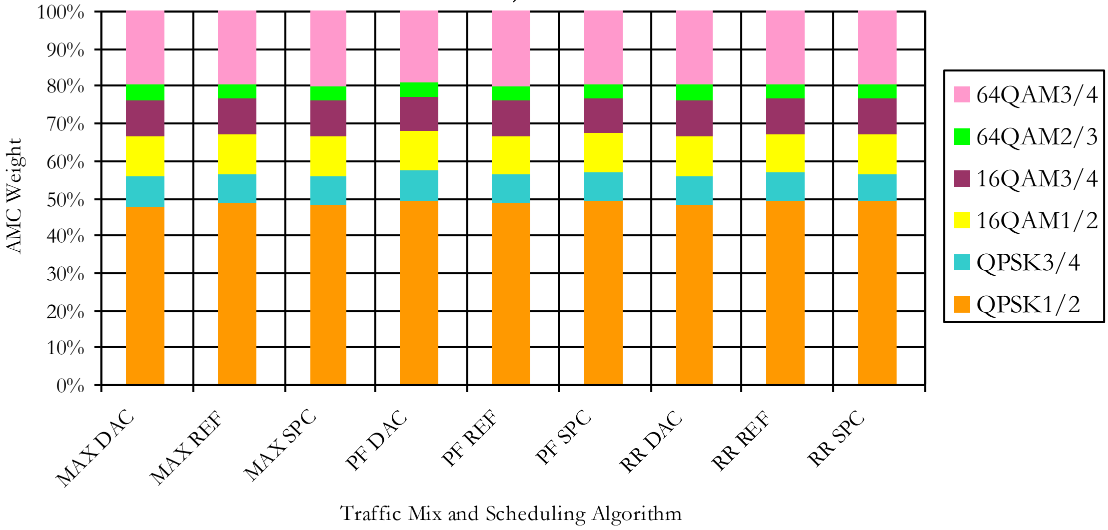


c) Rural

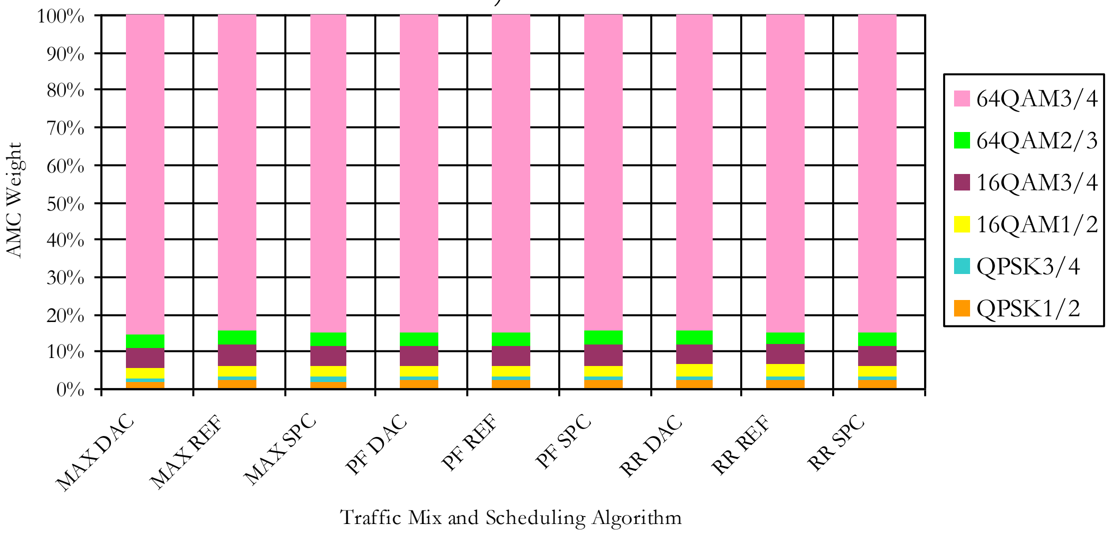
Figure H.3 – Traffic volume carried by the system for variable traffic mix per user.



a) Urban

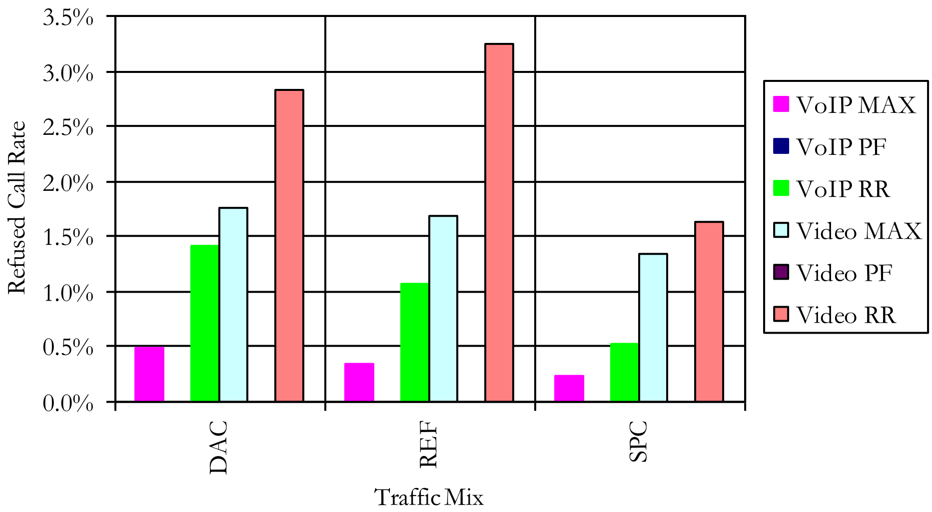


b) Suburban

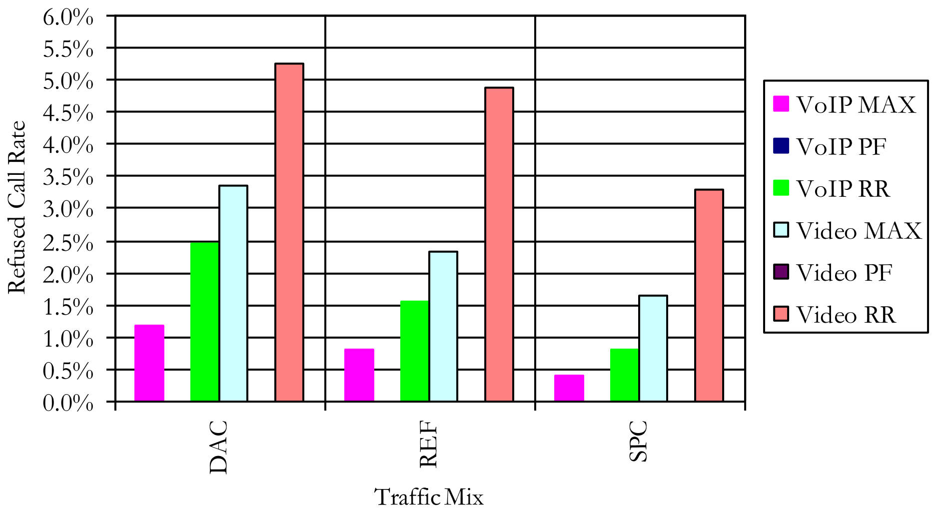


c) Rural

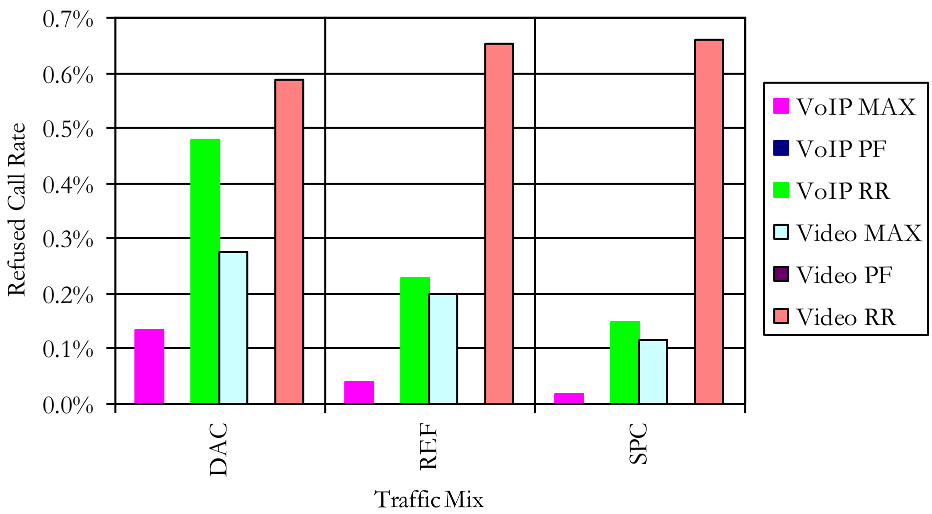
Figure H.4 – AMC distribution per user distribution.



a) Urban

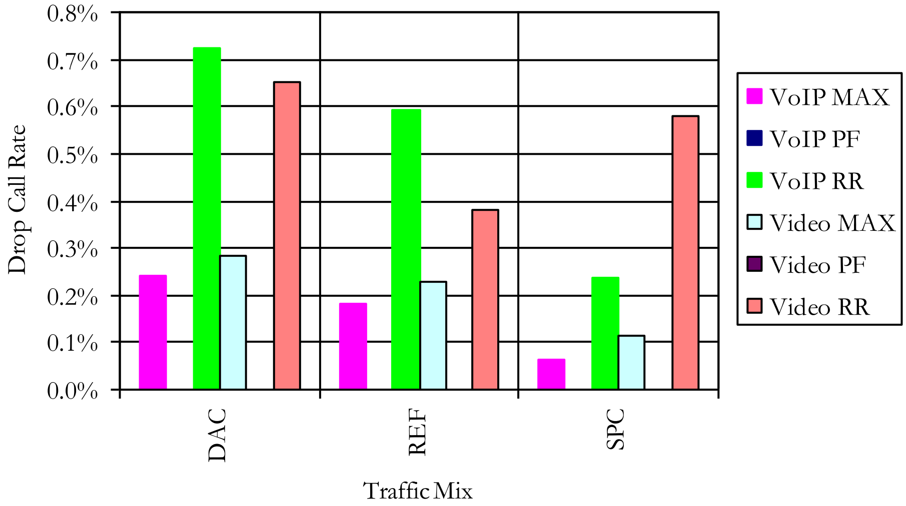


b) Suburban

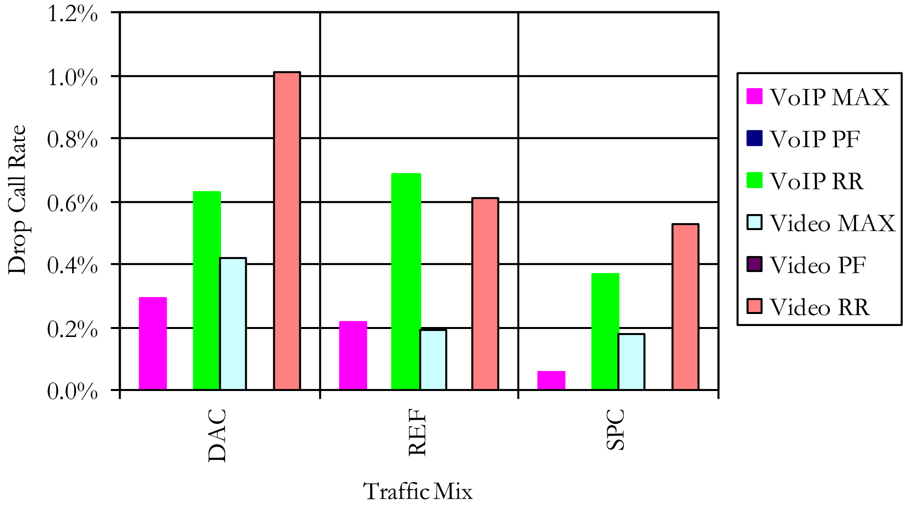


c) Rural

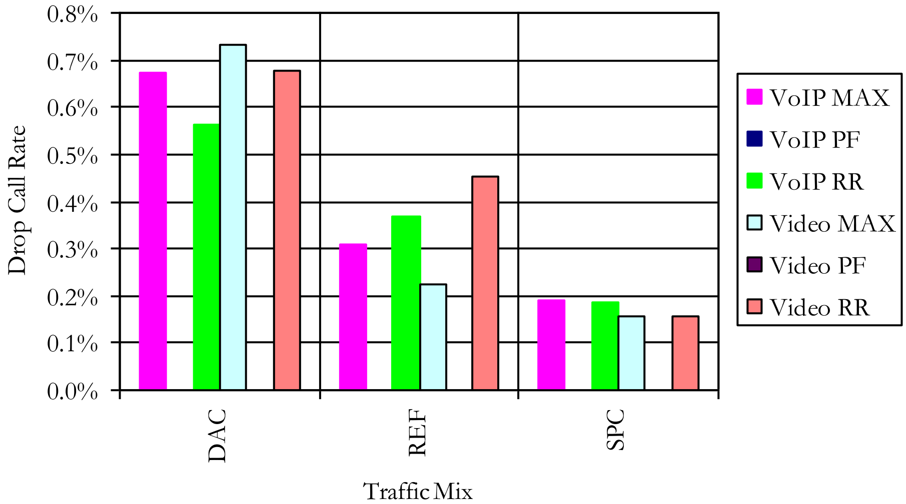
Figure H.5 – RTPS Refused Call Rate for variable per user traffic mix.



a) Urban

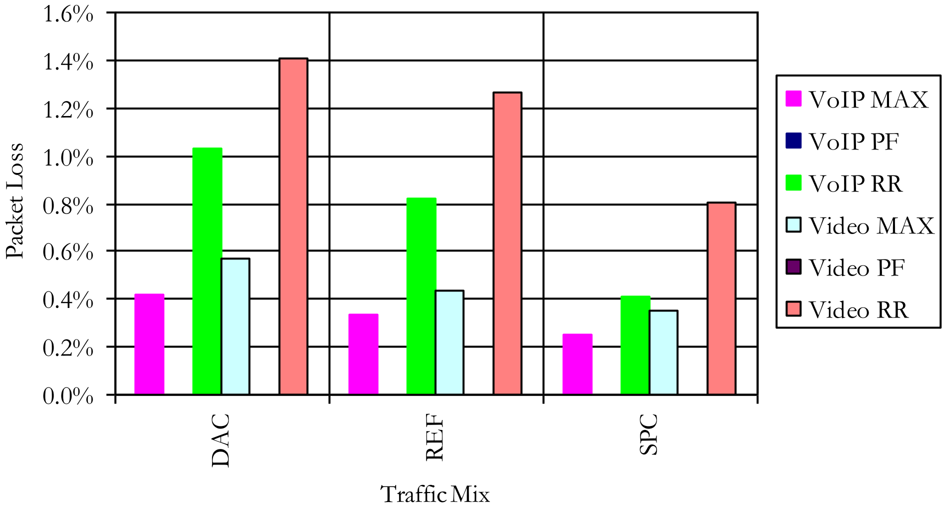


b) Suburban

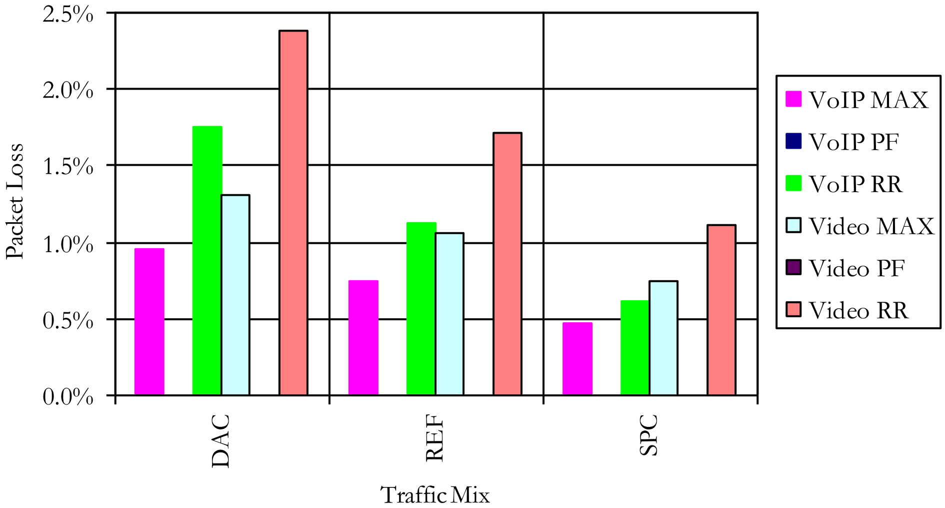


c) Rural

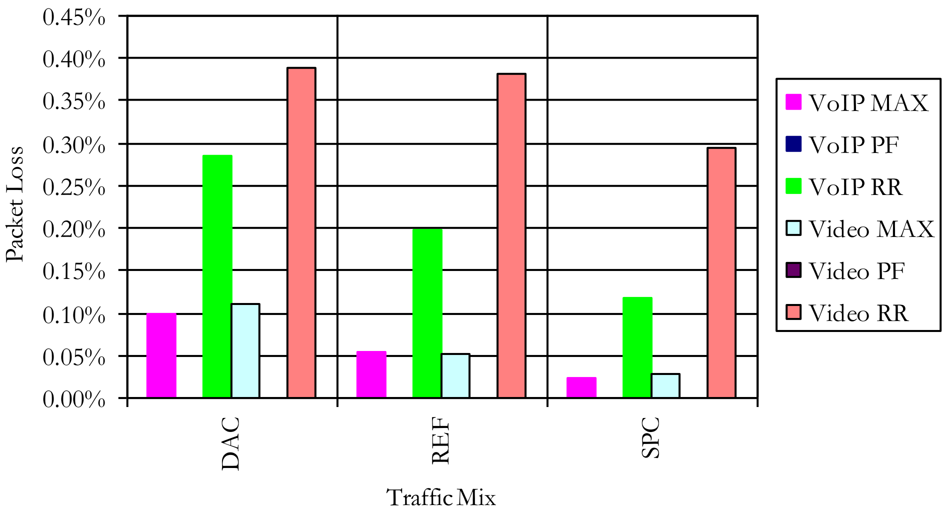
Figure H.6 – RTPS Drop Call Rate for variable per user traffic mix.



a) Urban

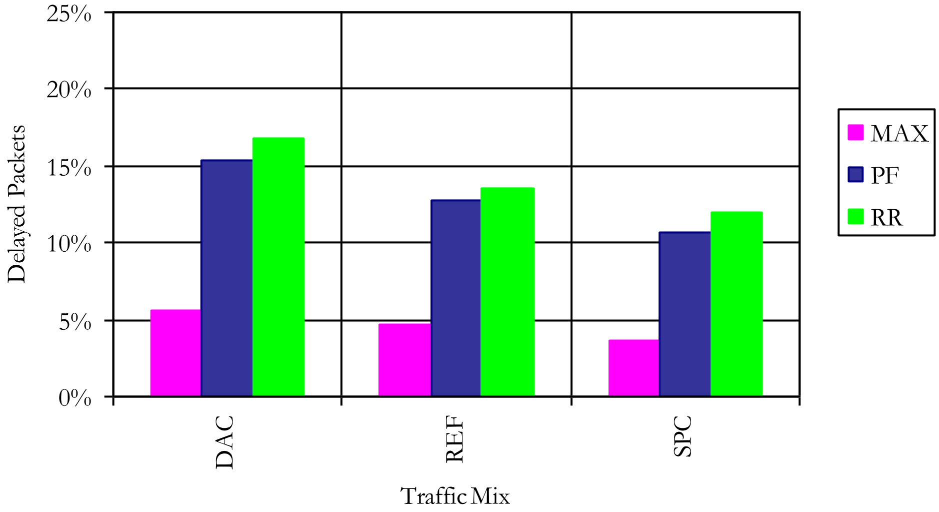


b) Suburban

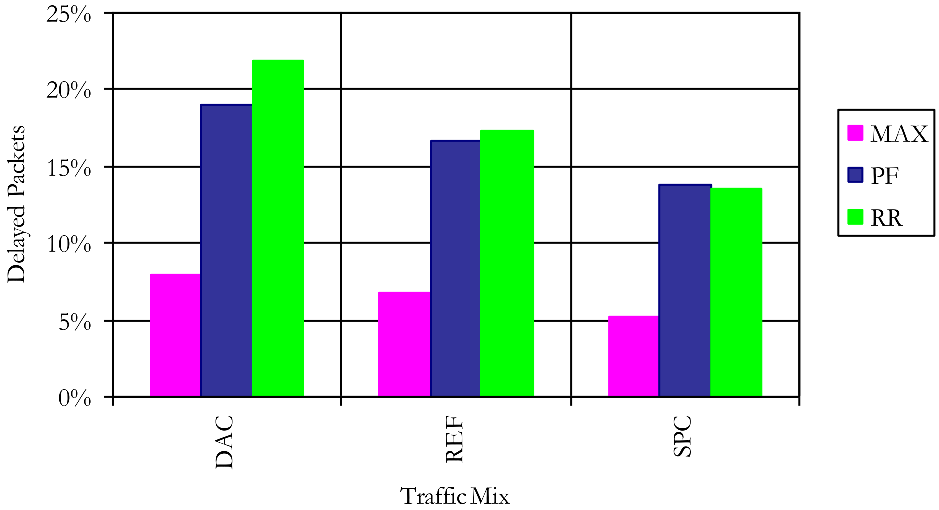


c) Rural

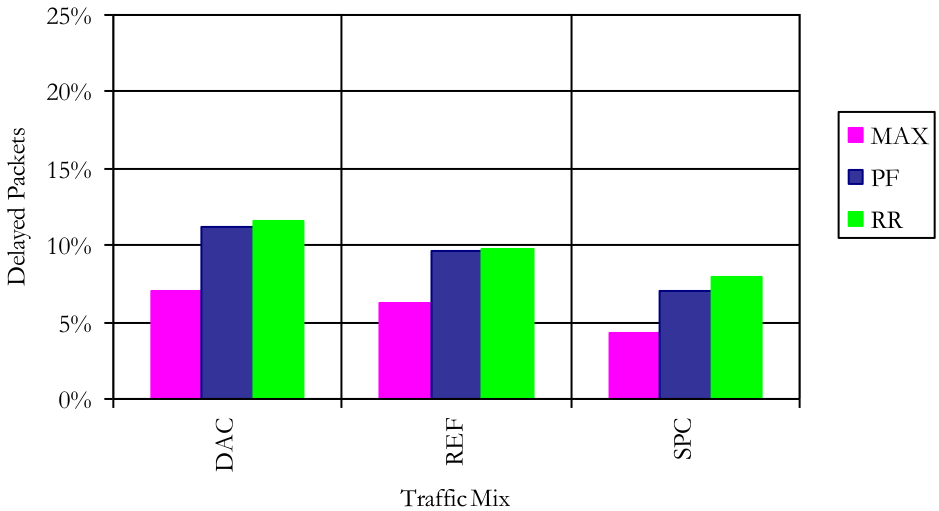
Figure H.7 – RTPS Packet Loss for variable per user traffic mix.



a) Urban

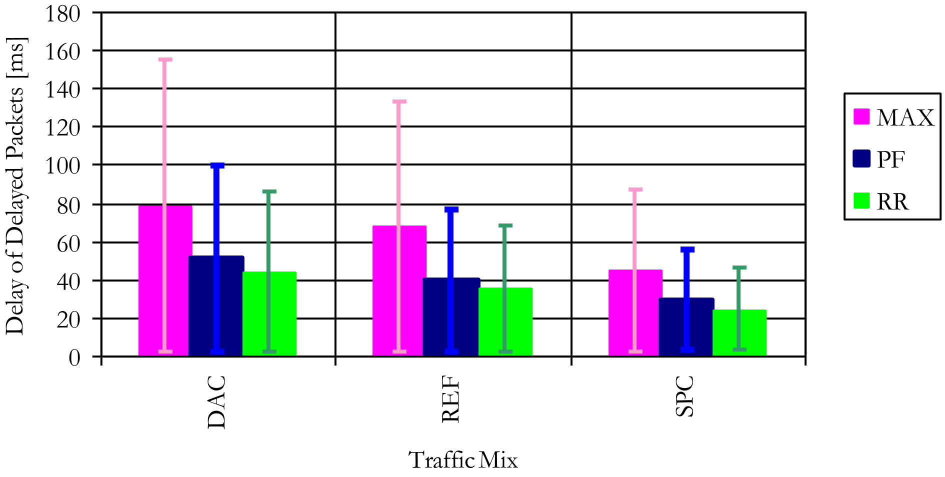


b) Suburban

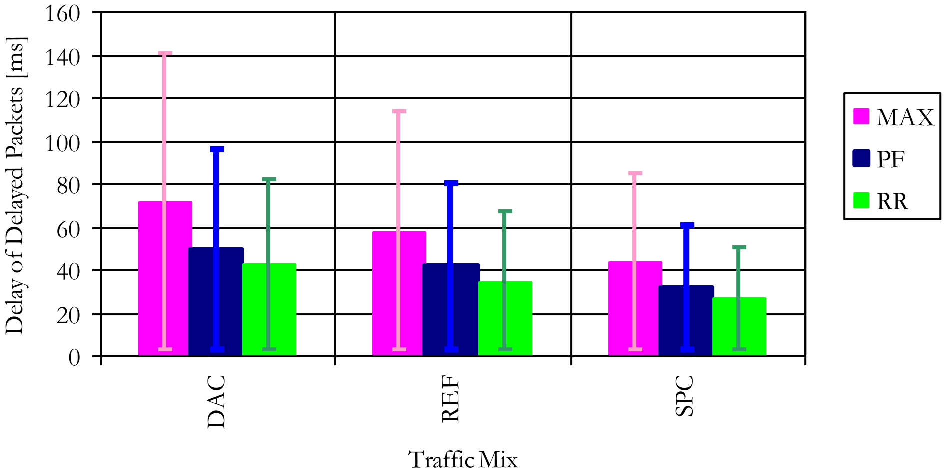


c) Rural

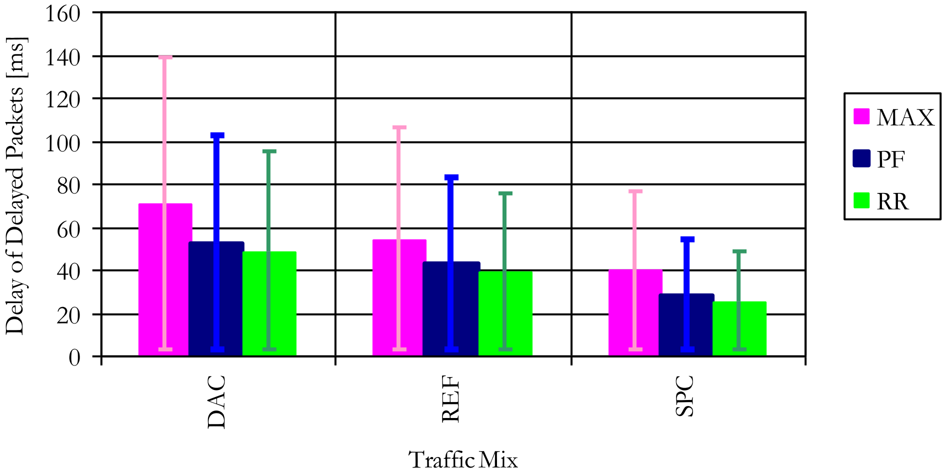
Figure H.8 – BE percentage of delayed packets for variable per user traffic mix.



a) Urban

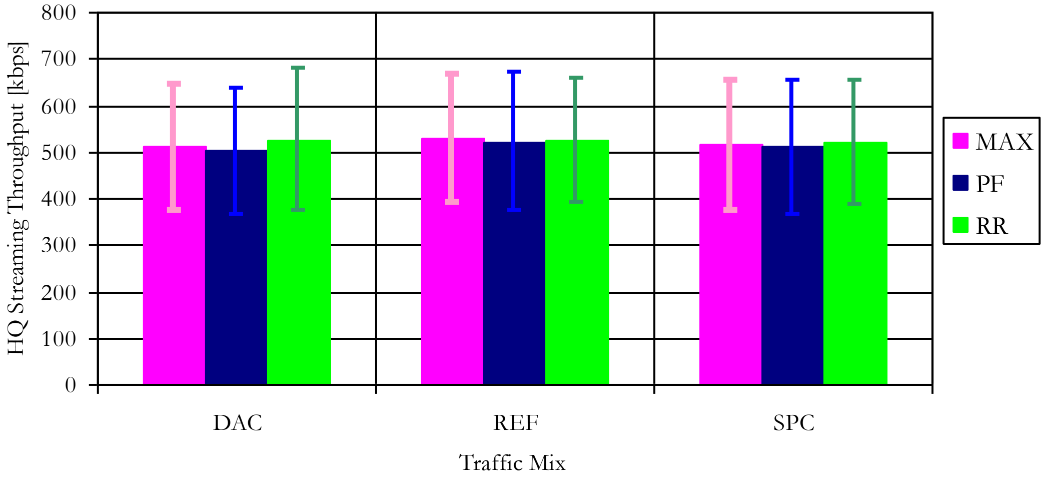


b) Suburban

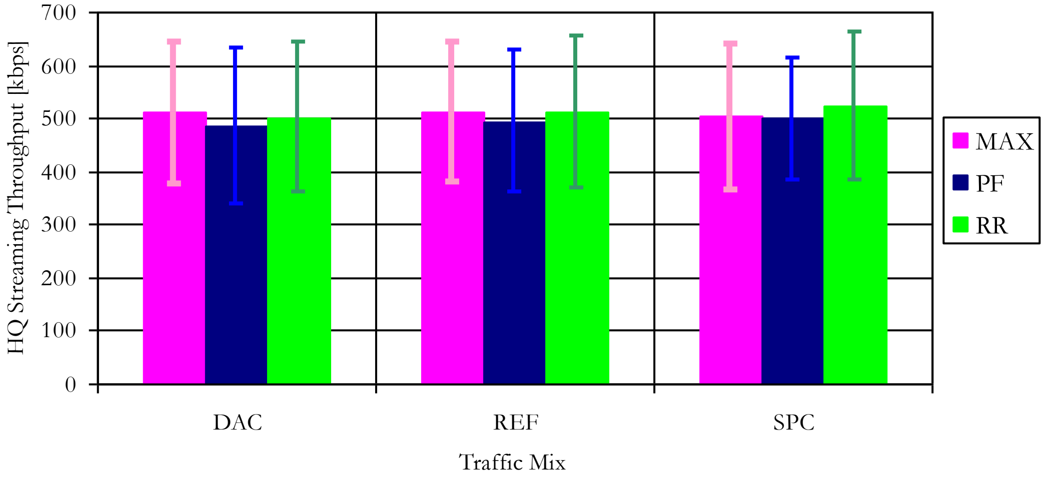


c) Rural

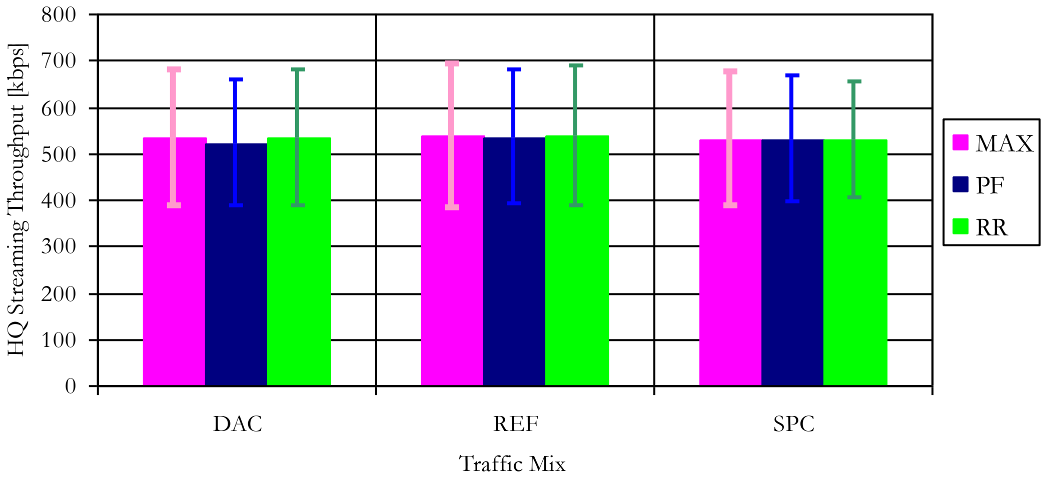
Figure H.9 – BE average delay of delayed packets for variable per user traffic mix.



a) Urban

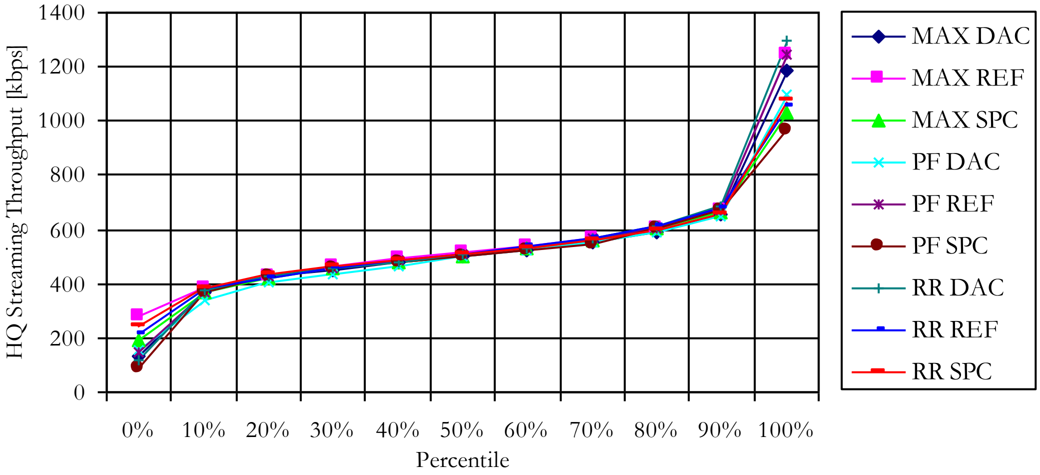


b) Suburban

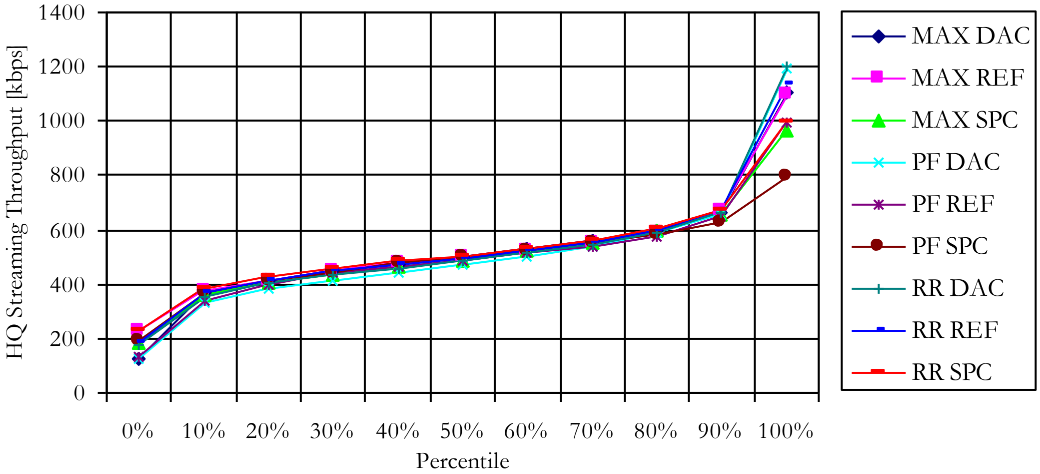


c) Rural

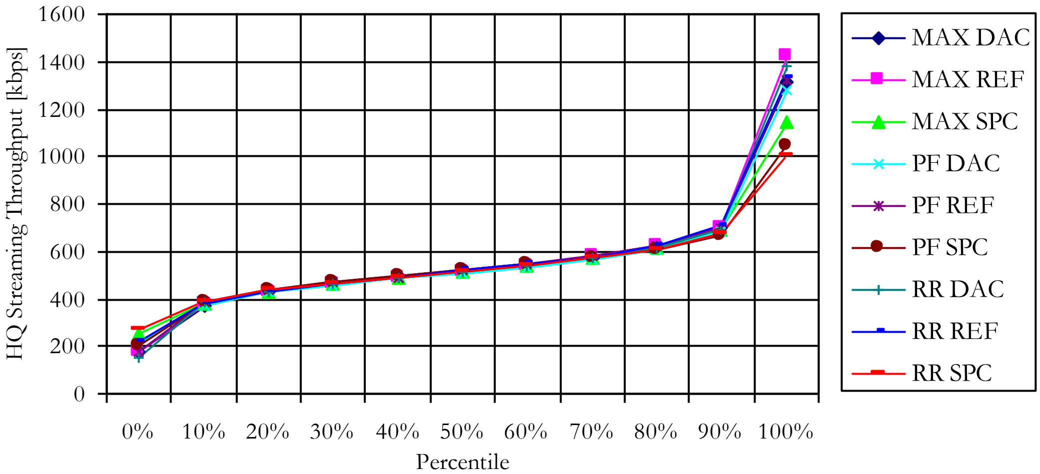
Figure H.10 – Average high quality streaming throughput and standard deviation for variable per user traffic mix.



a) Urban

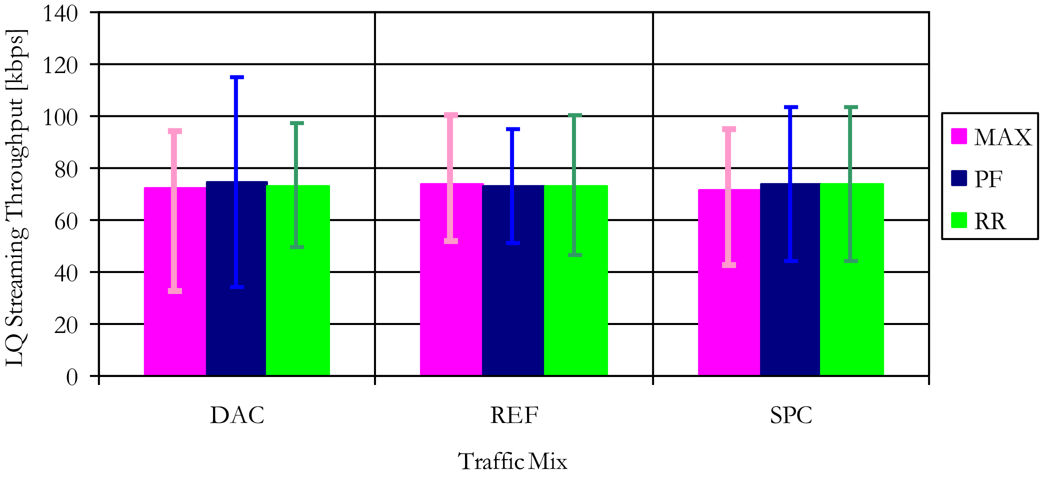


b) Suburban

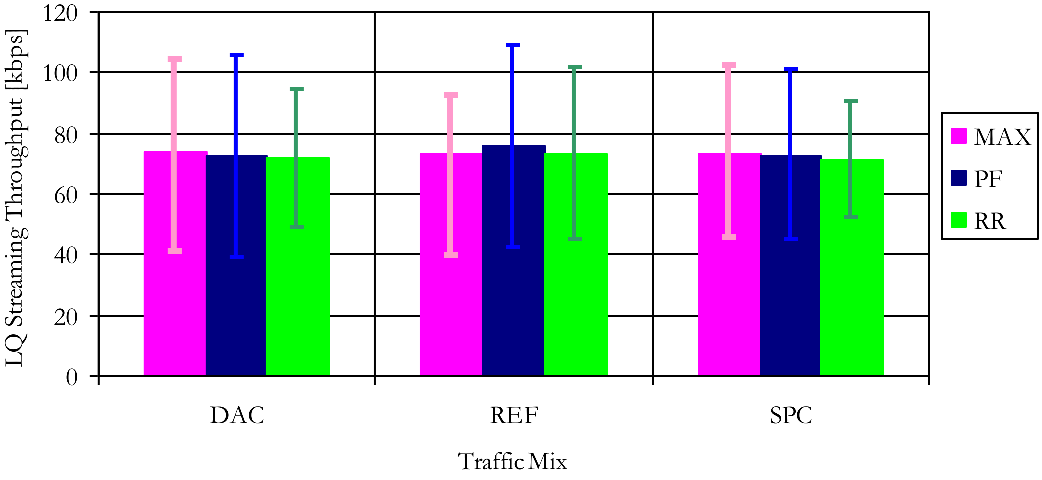


c) Rural

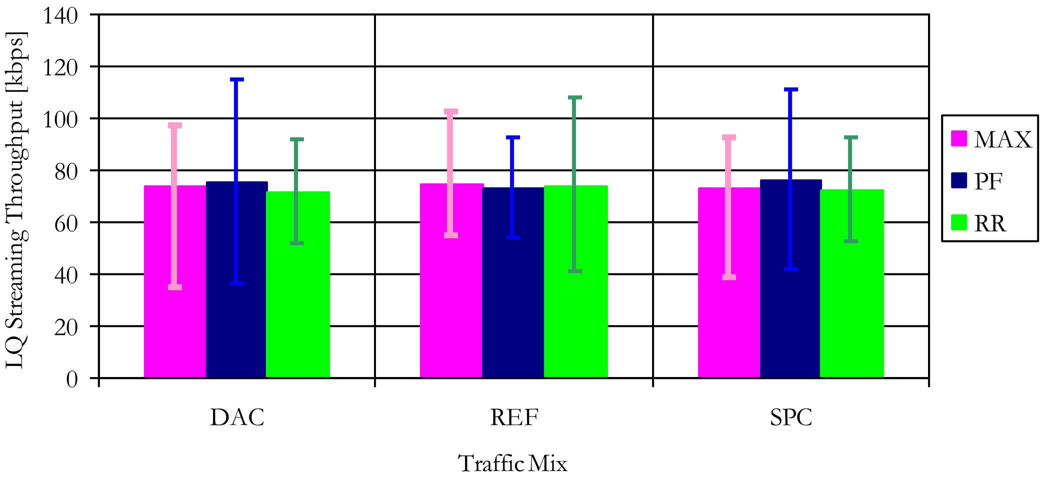
Figure H.11 – Cumulative distribution of high quality streaming throughput for variable per user traffic mix.



a) Urban

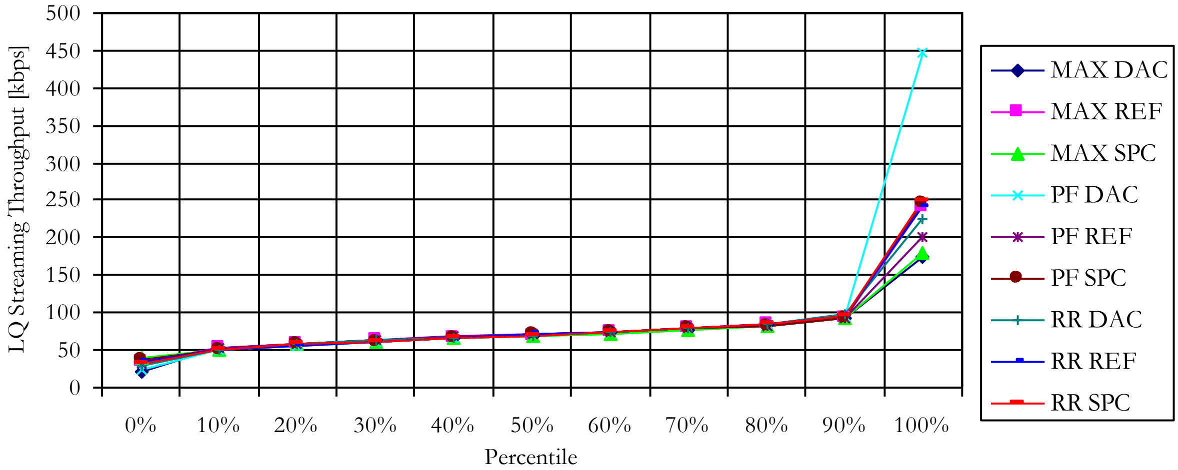


b) Suburban

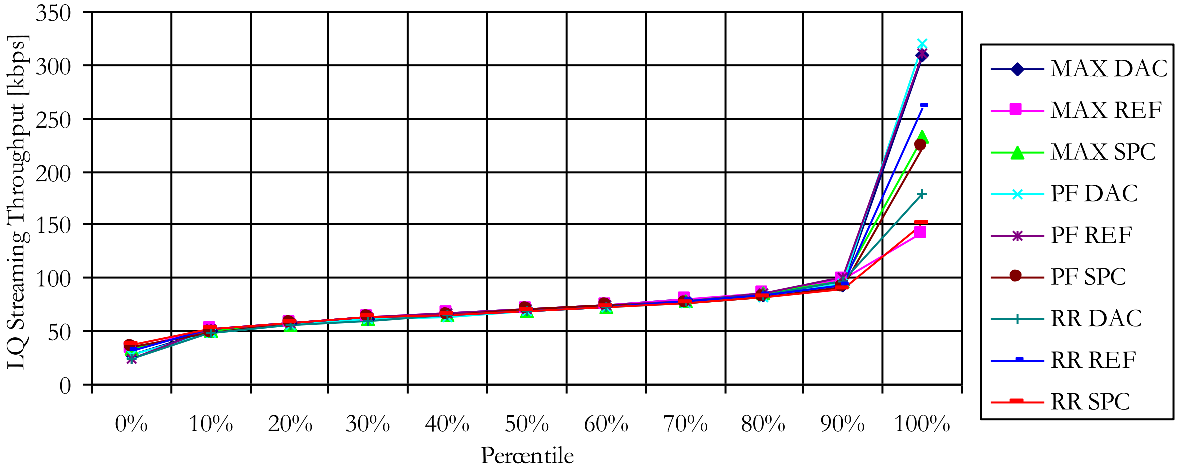


c) Rural

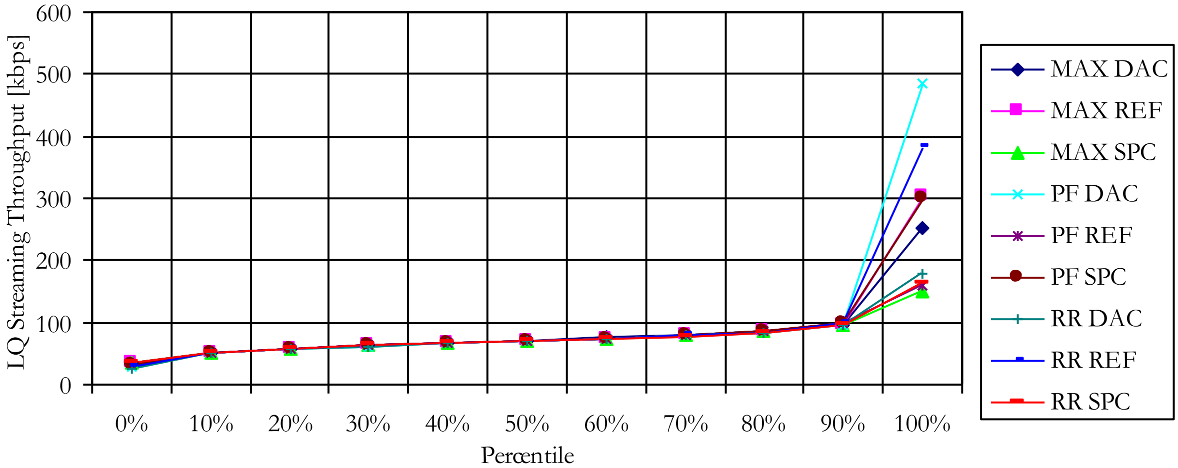
Figure H.12 - Average low quality streaming throughput and standard deviation for variable per user traffic mix.



a) Urban

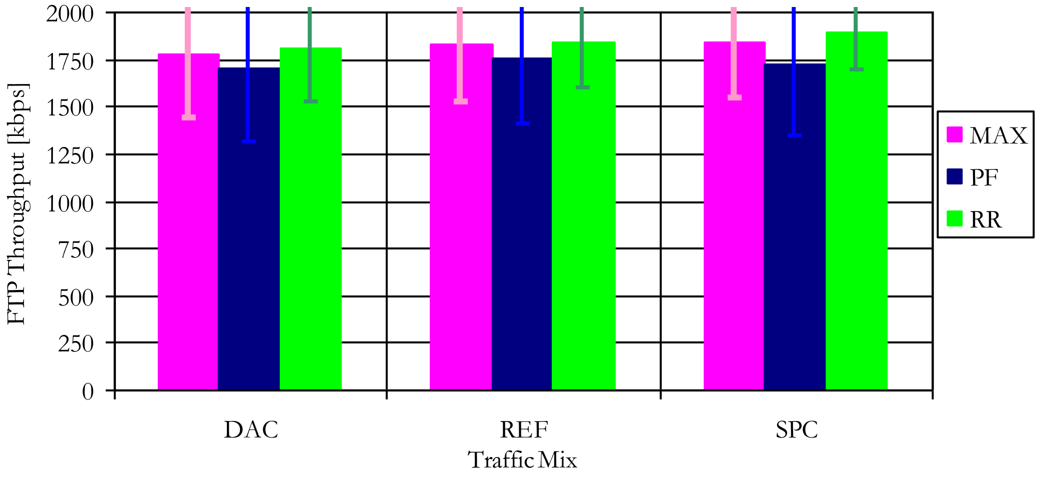


b) Suburban

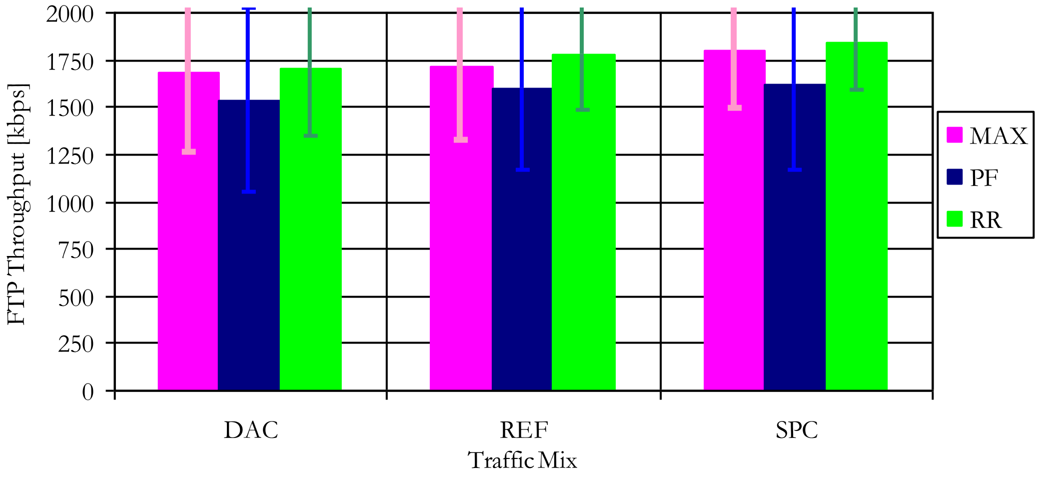


c) Rural

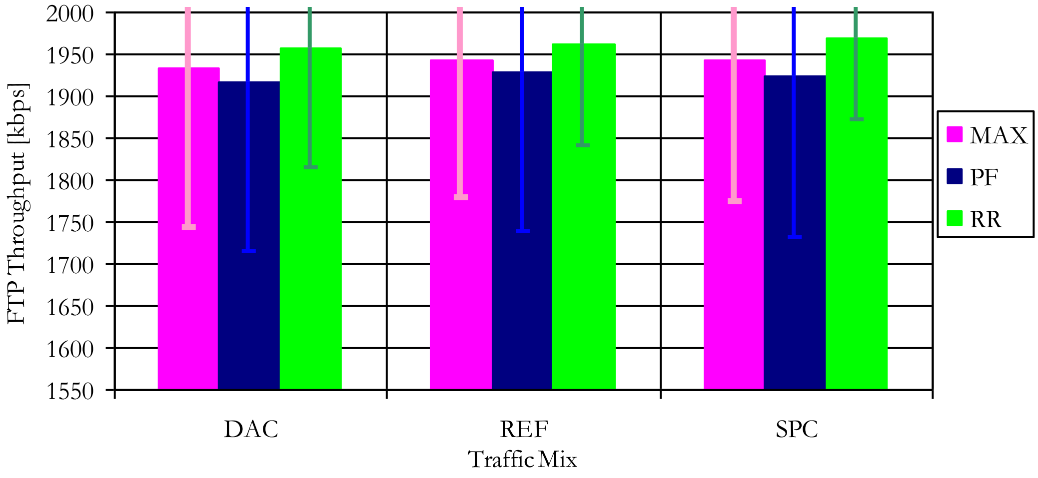
Figure H.13 - Cumulative distribution of low quality streaming throughput and standard deviation for variable per user traffic mix.



a) Urban

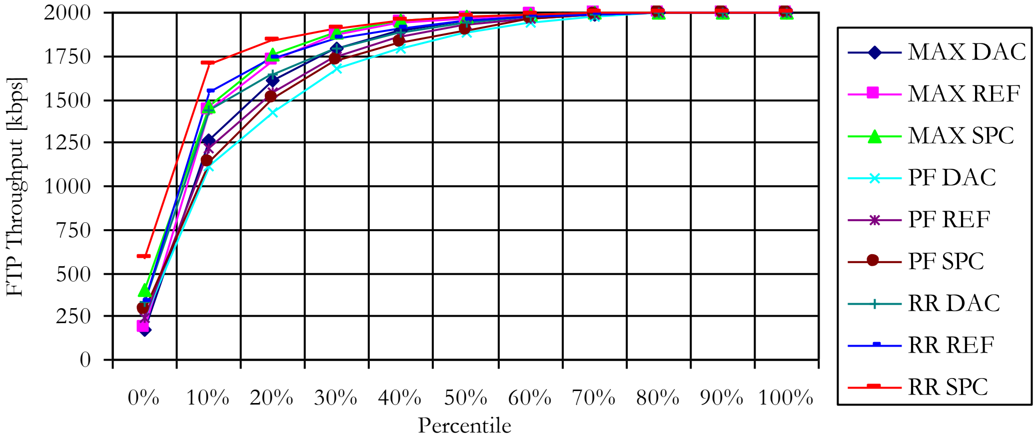


b) Suburban

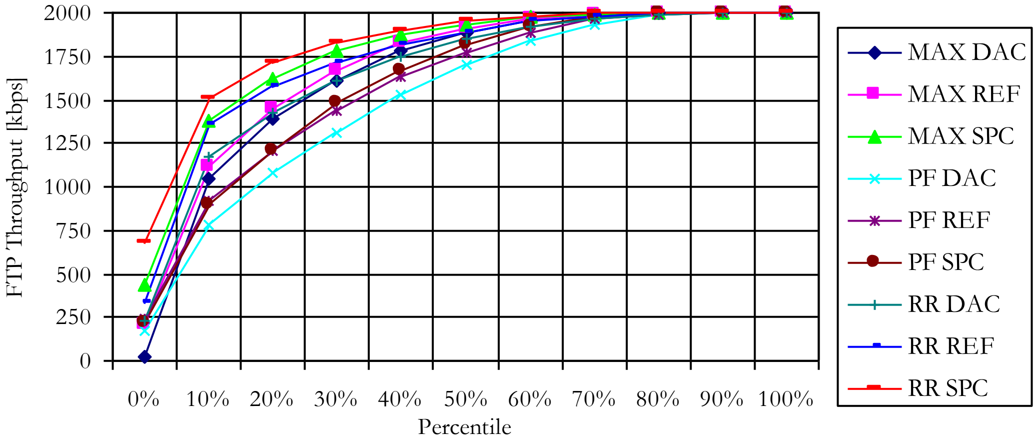


c) Rural

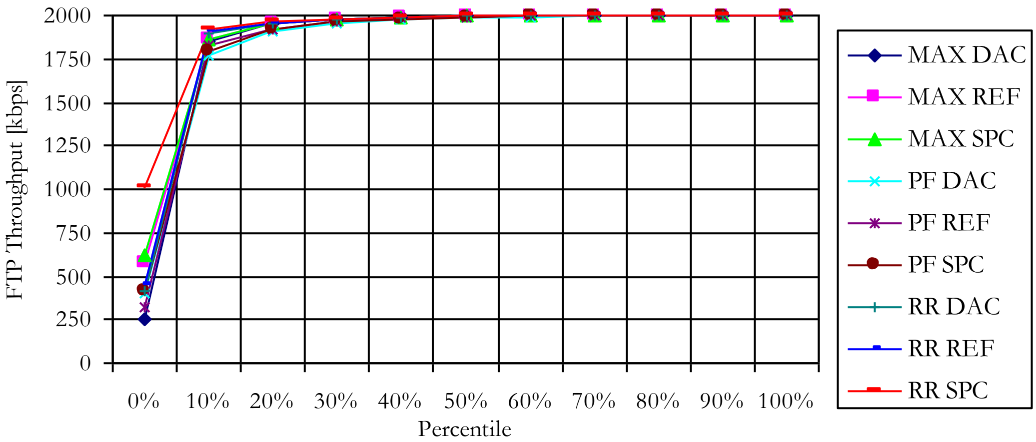
Figure H.14 – Average per user FTP throughput and standard deviation for variable per user traffic mix.



a) Urban



b) Suburban



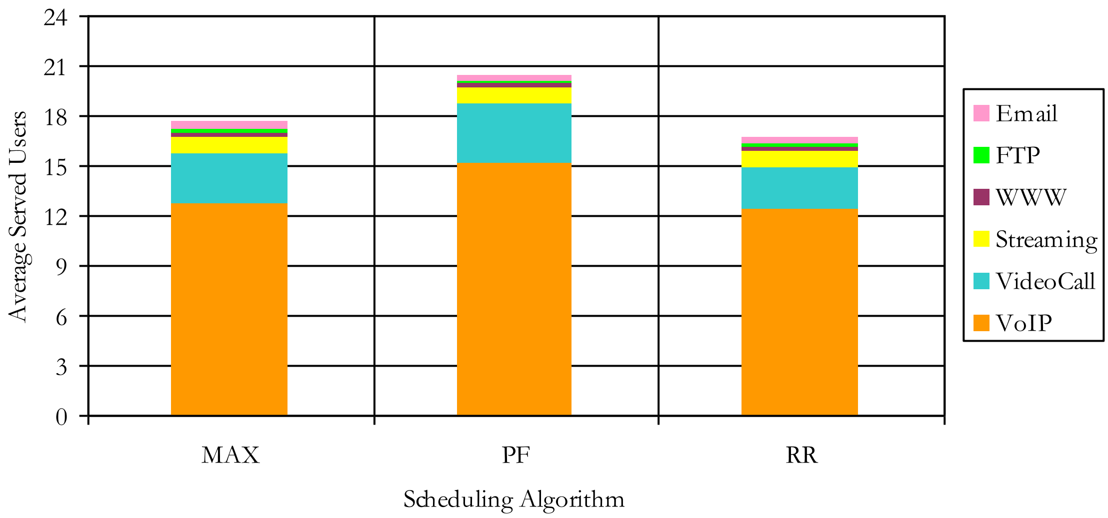
c) Rural

Figure H.15 - FTP throughput cumulative distribution for variable per user traffic mix.

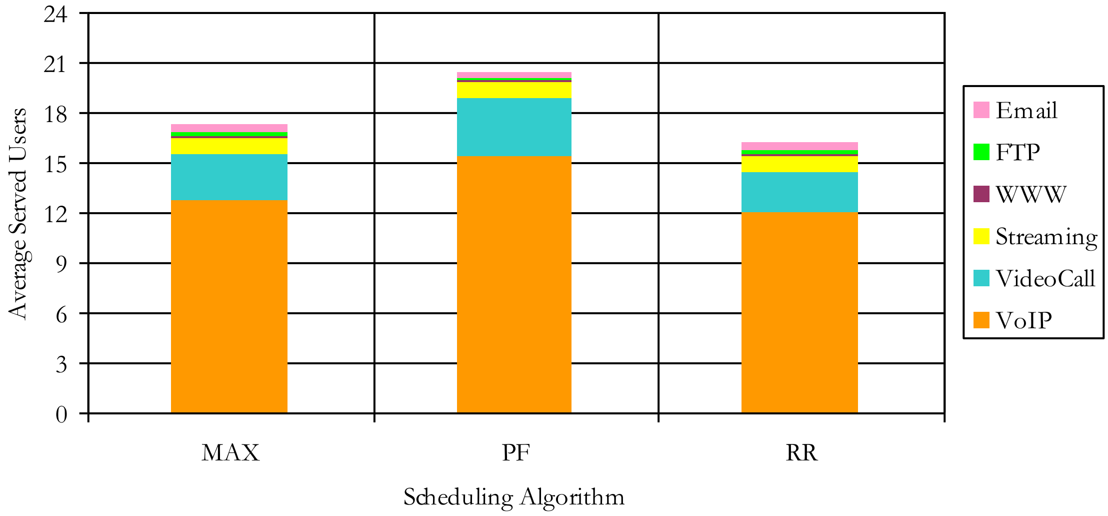
Annex I

Results for Variable Channel Bandwidth

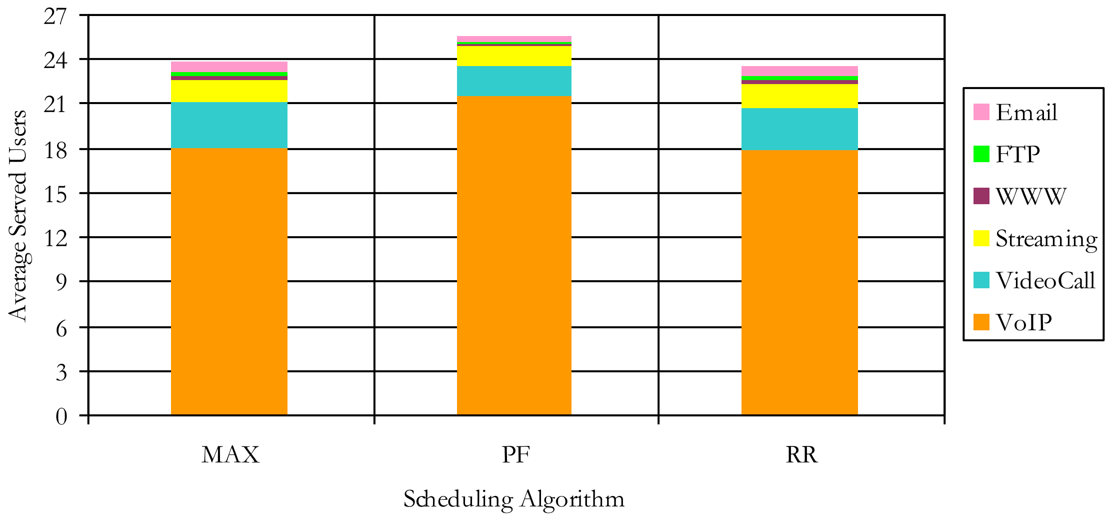
This annex summarises the results obtained when varying the system configuration, particularly channel bandwidth, for Urban, Suburban and Rural.



a) Urban

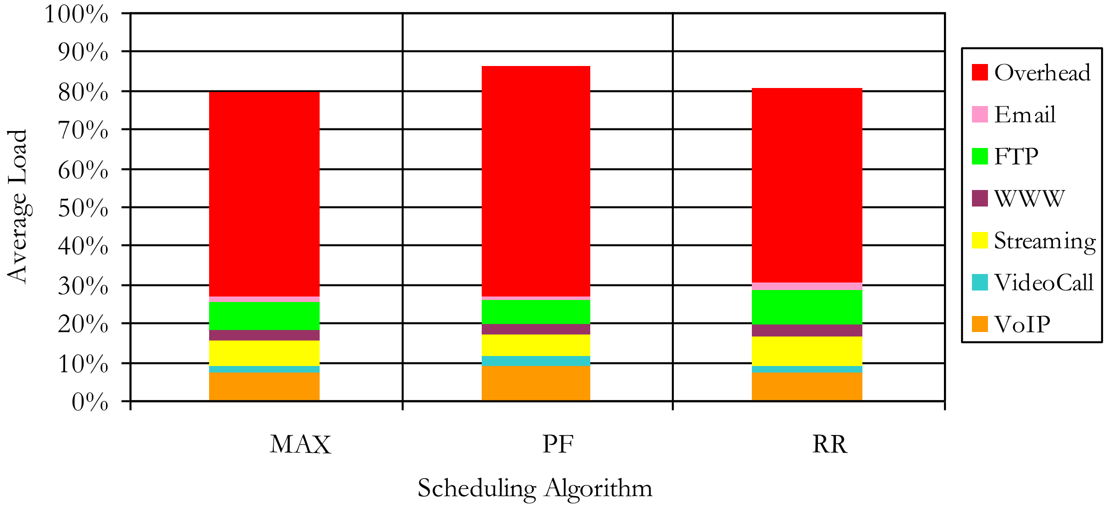


b) Suburban

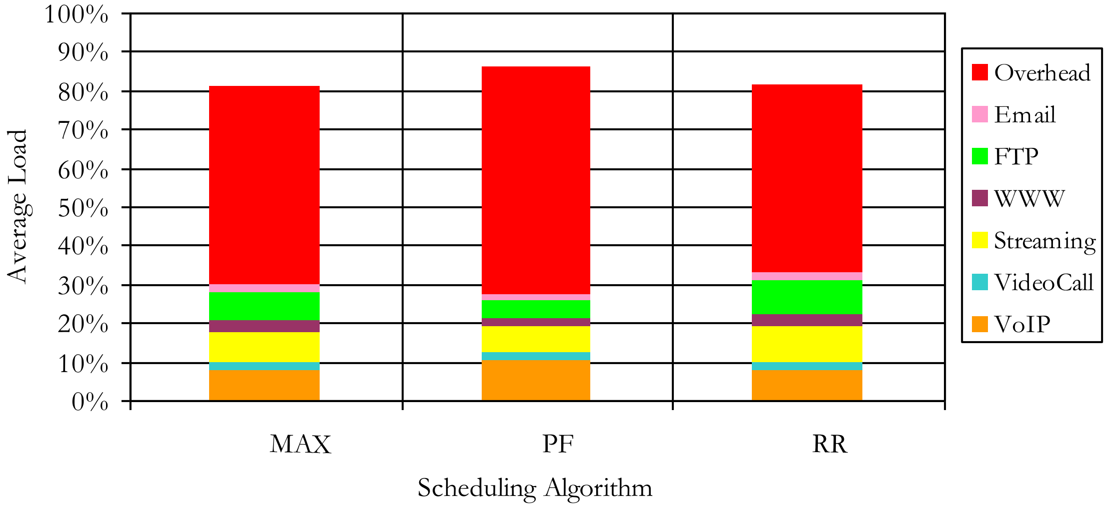


c) Rural

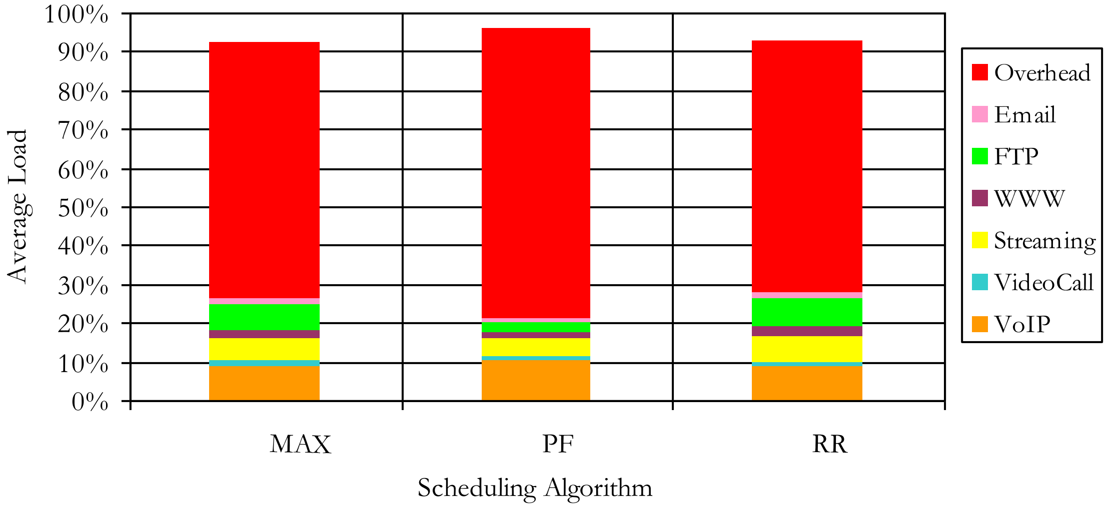
Figure I.1 - Average number of served users for 5MHz channel bandwidth.



a) Urban

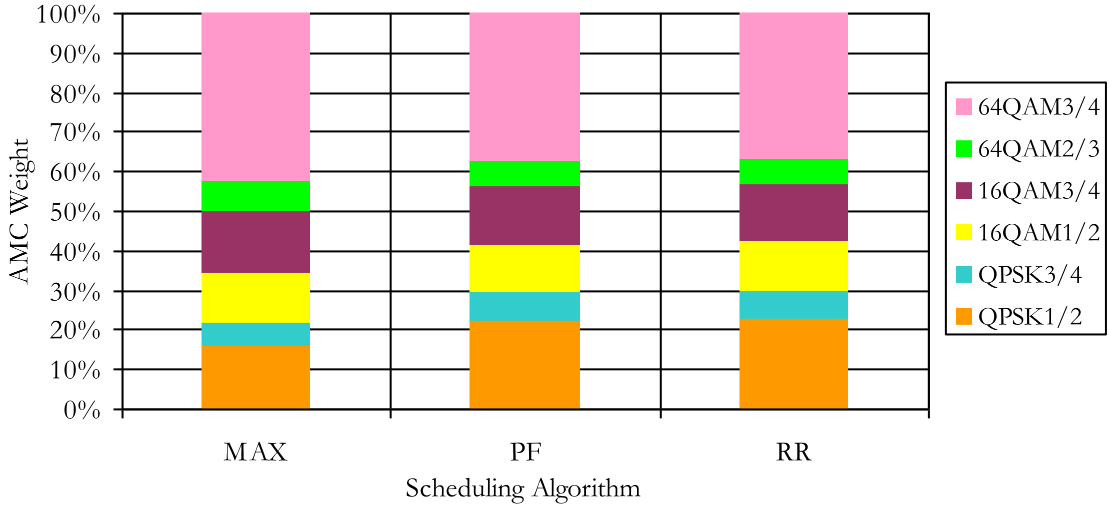


b) Suburban

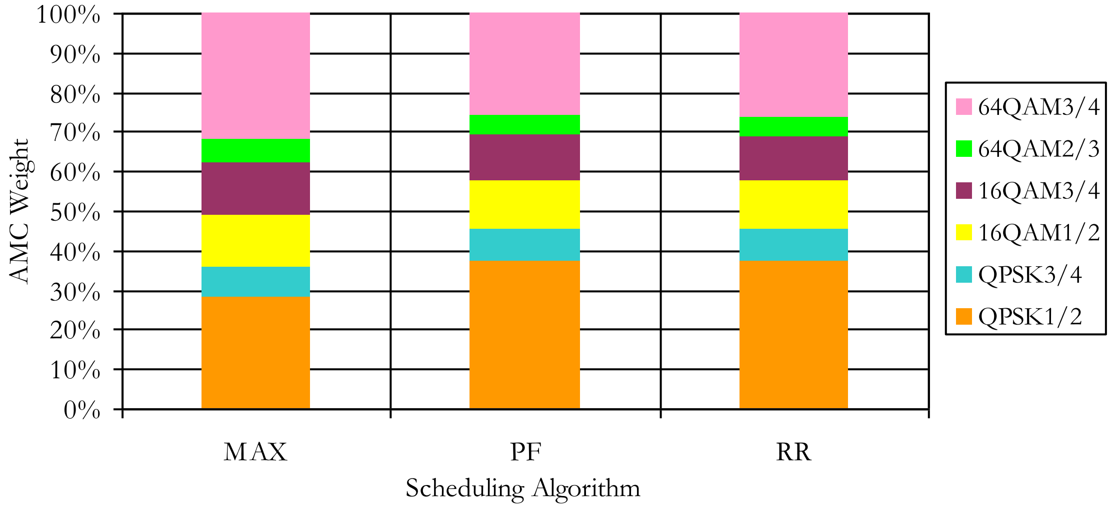


c) Rural

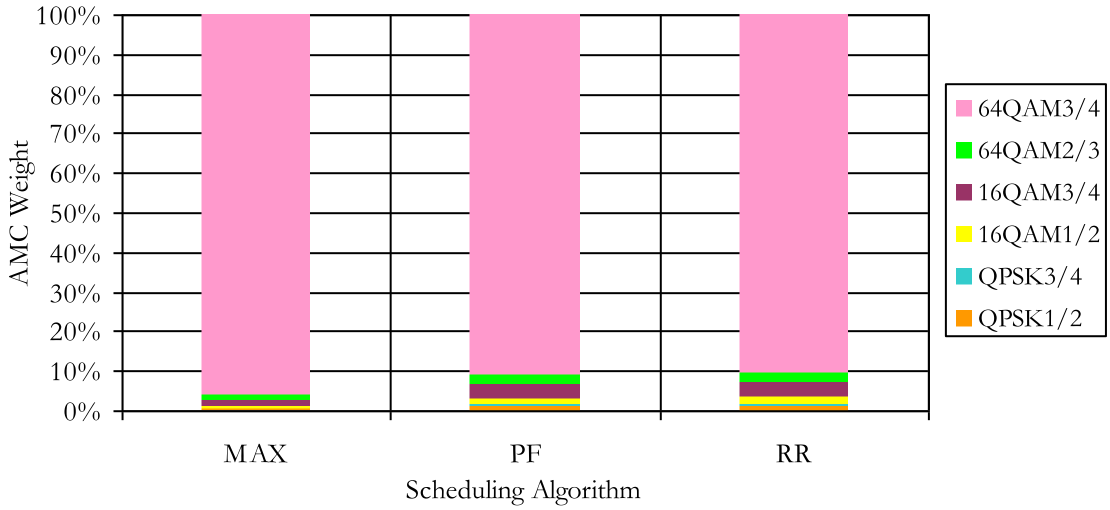
Figure I.2 - Network load for 5MHz channel bandwidth.



a) Urban

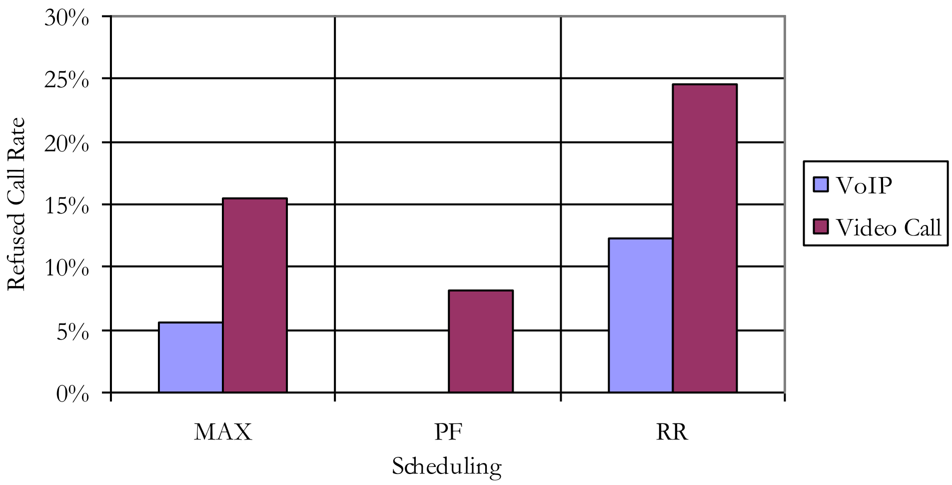


b) Suburban

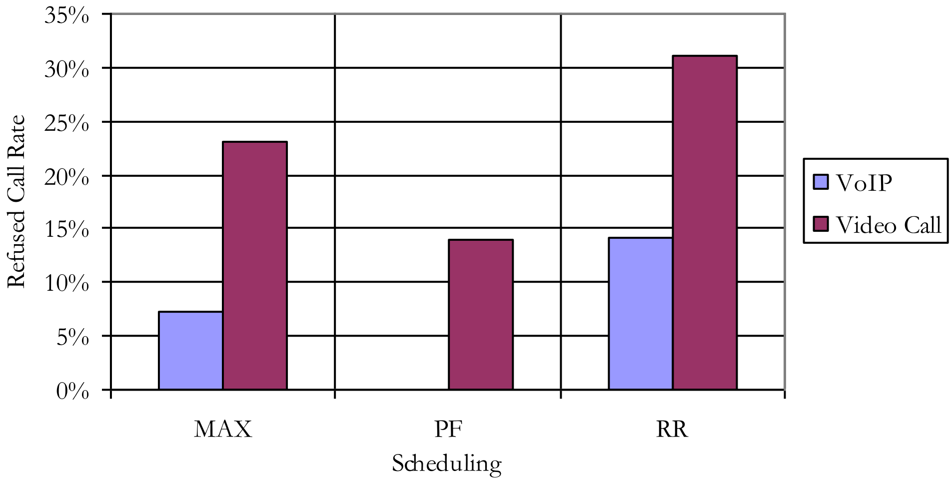


c) Rural

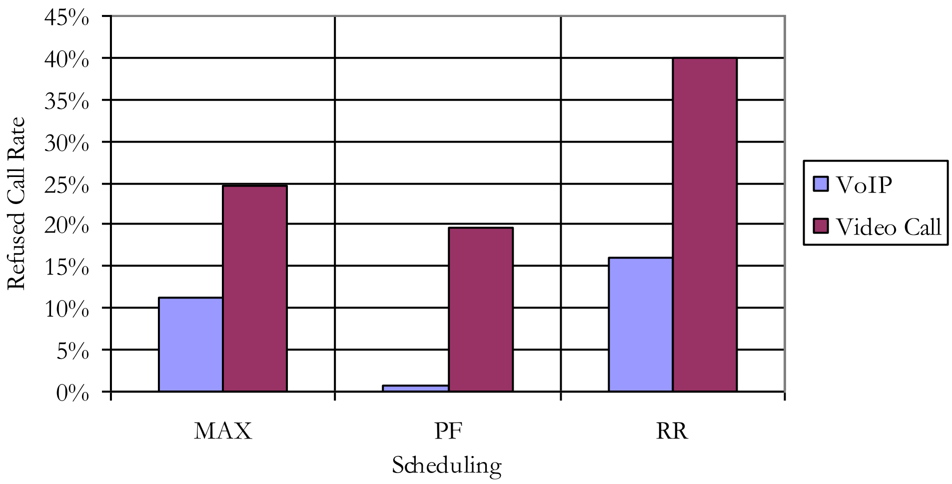
Figure I.3 - AMC distribution for 5MHz channel bandwidth.



a) Urban

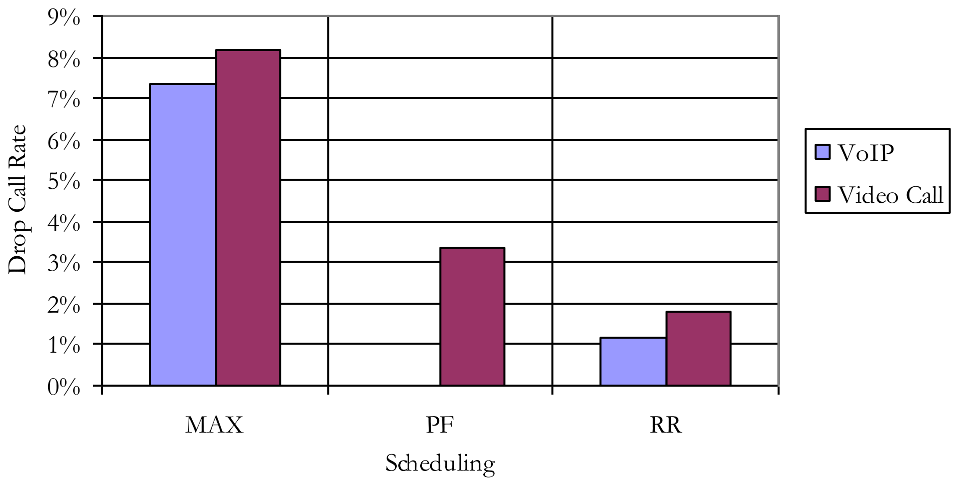


b) Suburban

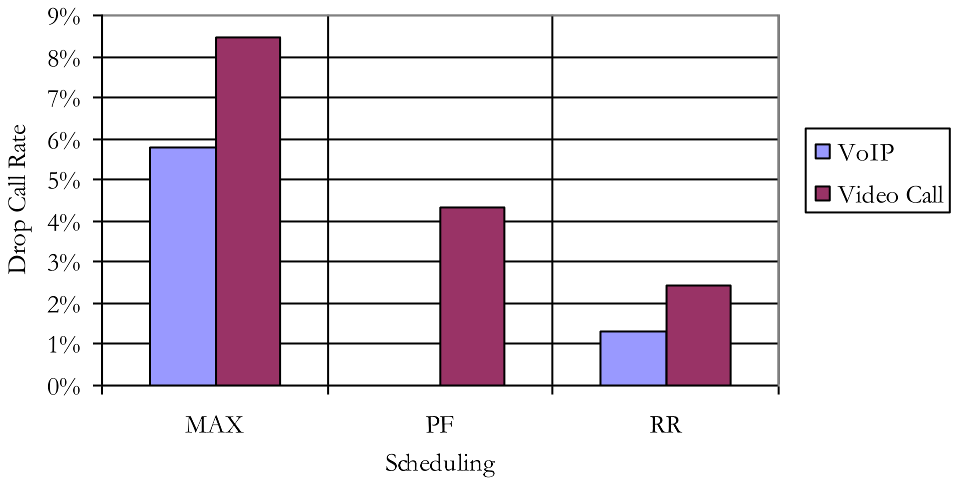


c) Rural

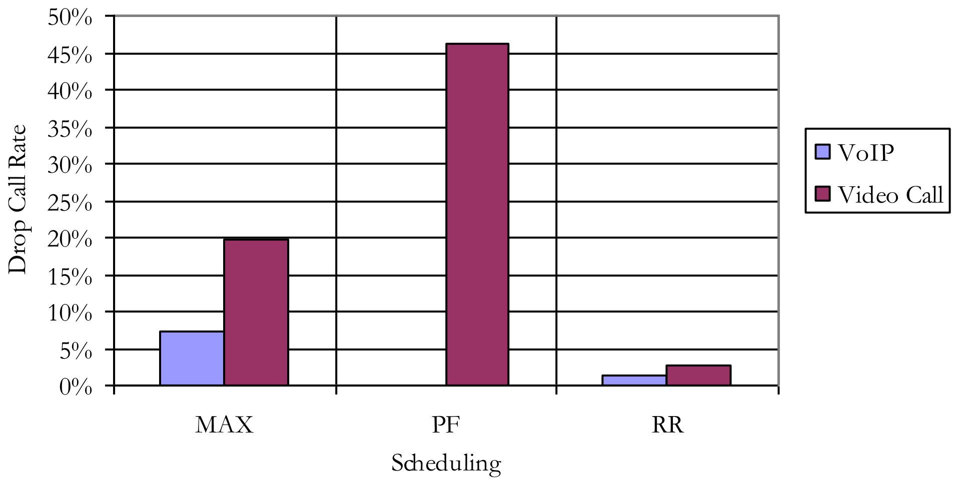
Figure I.4 - VoIP and Video Call Refused Call Rate for 5MHz channel bandwidth.



a) Urban

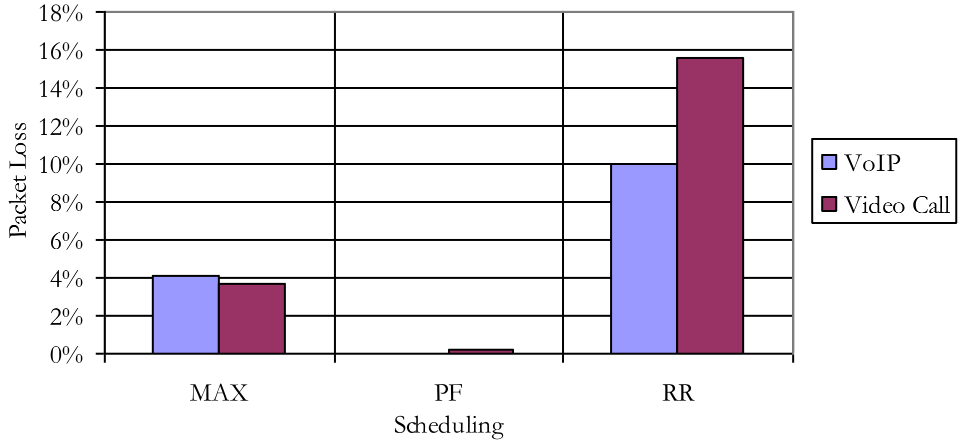


b) Suburban

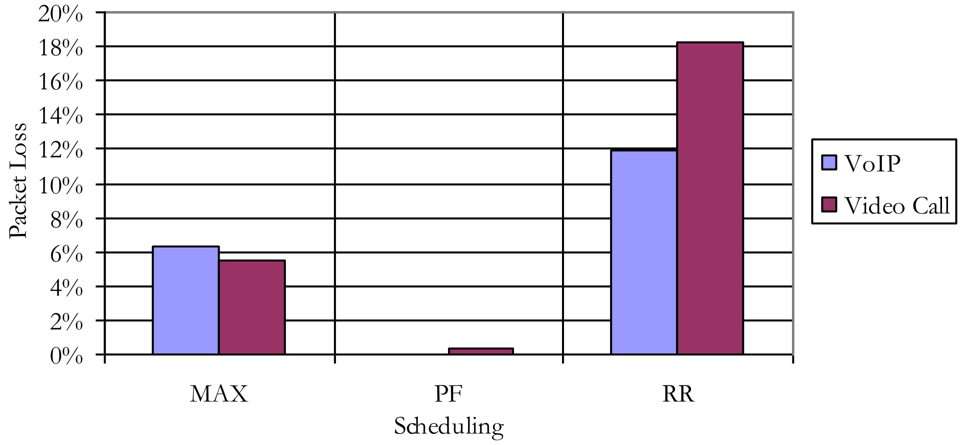


c) Rural

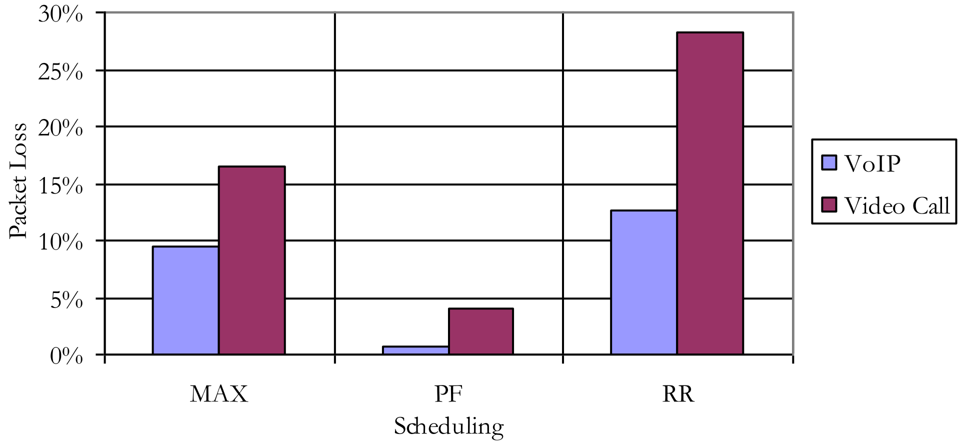
Figure I.5 - VoIP and Video Call Drop Call Rate for 5MHz channel bandwidth.



a) Urban

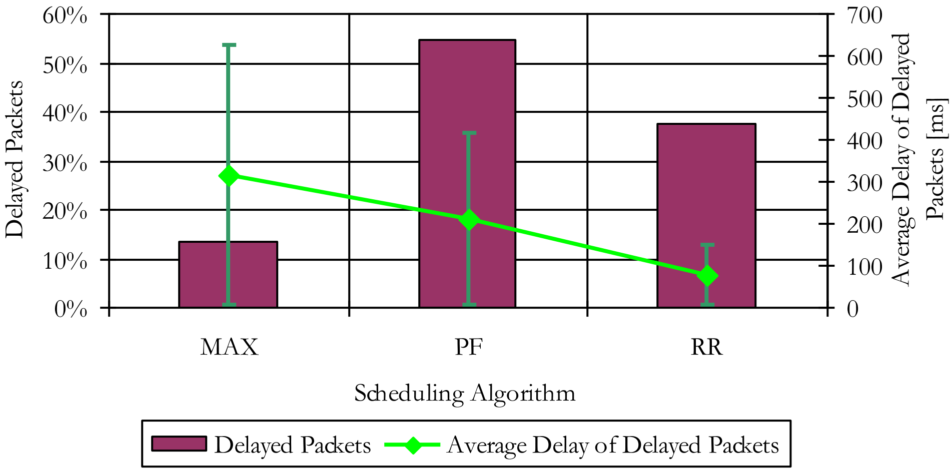


b) Suburban

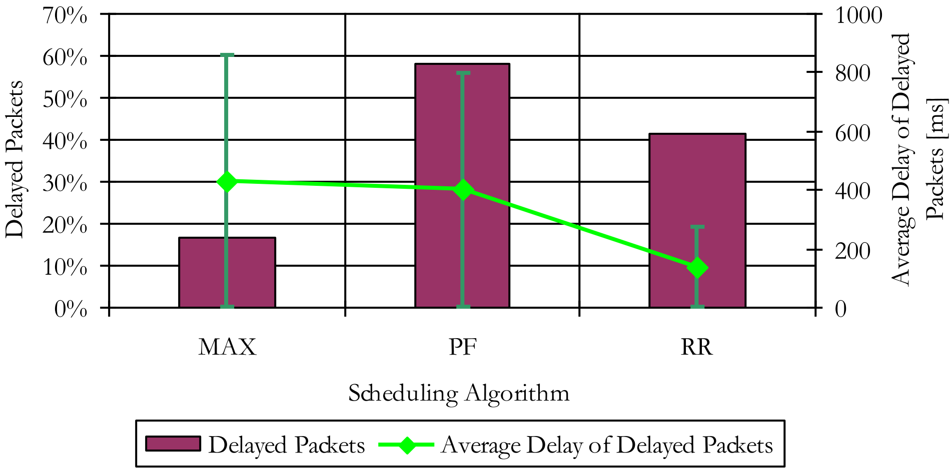


c) Rural

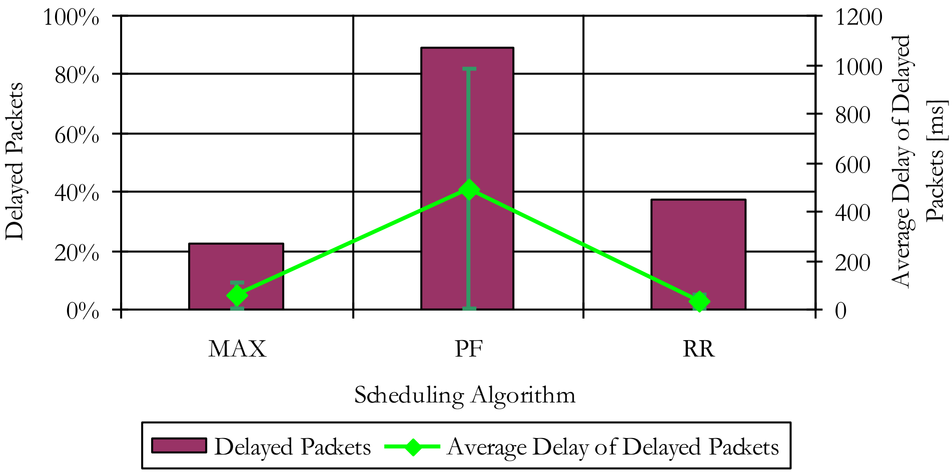
Figure I.6 - VoIP and Video Call packet loss for 5MHz channel bandwith.



a) Urban

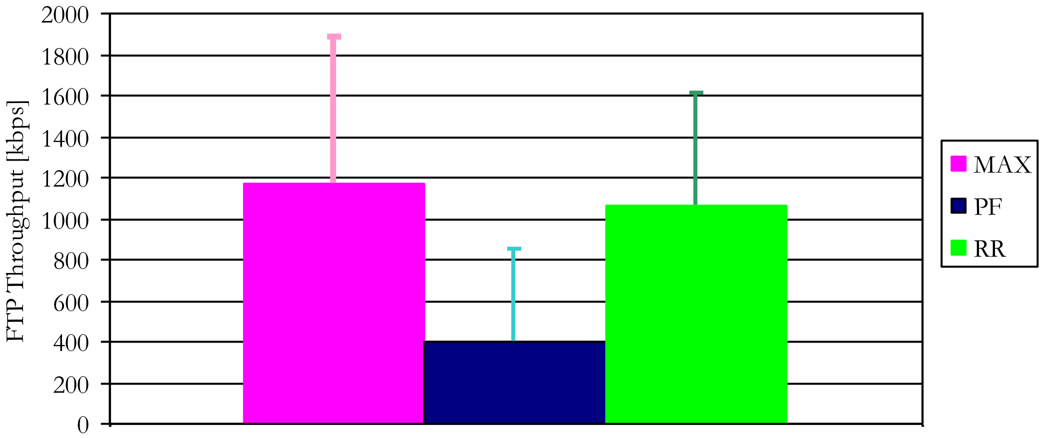


b) Suburban

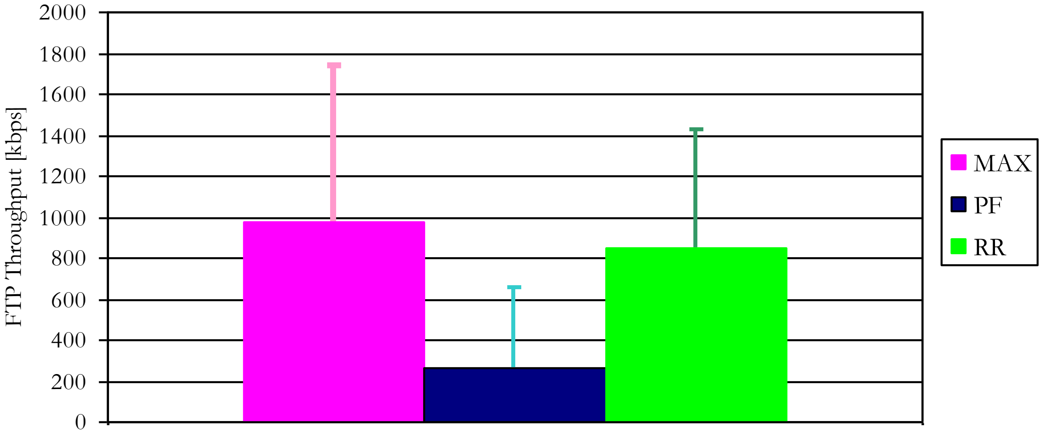


c) Rural

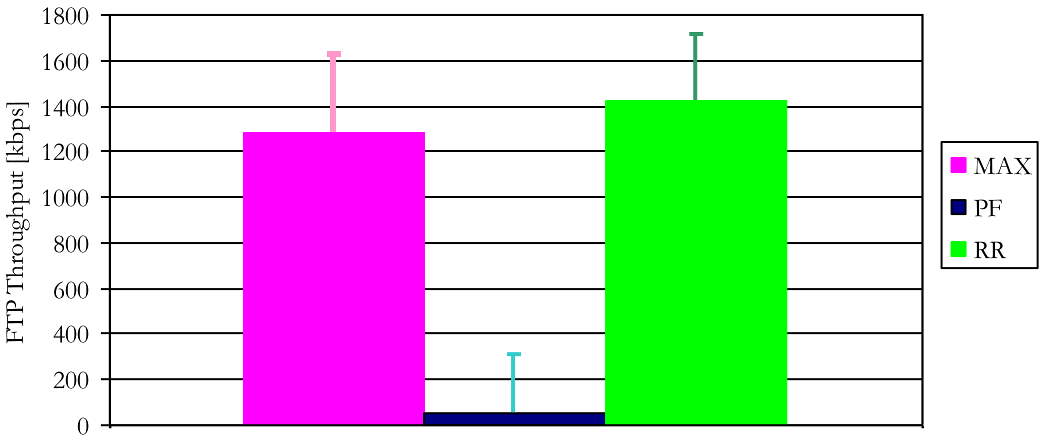
Figure I.7 - BE packet delay for 5MHz channel bandwidth.



a) Urban

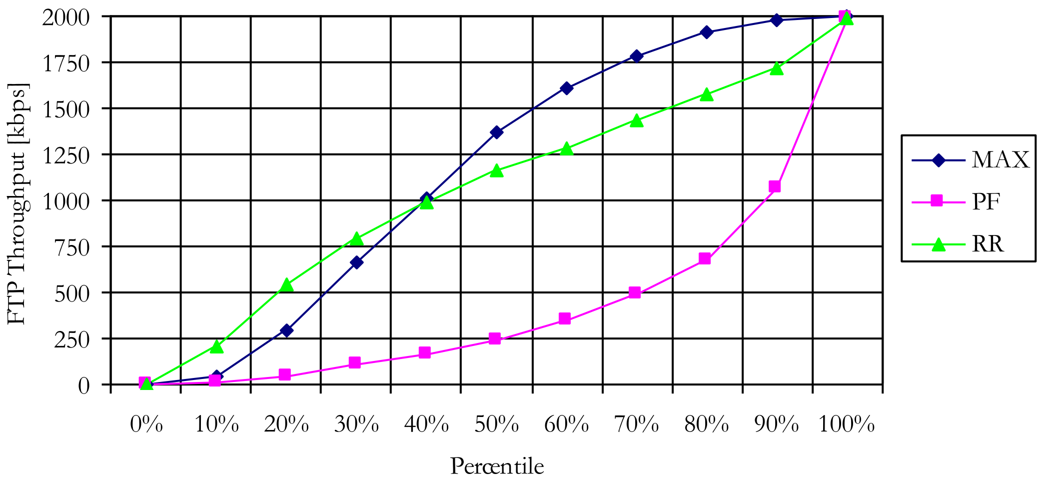


b) Suburban

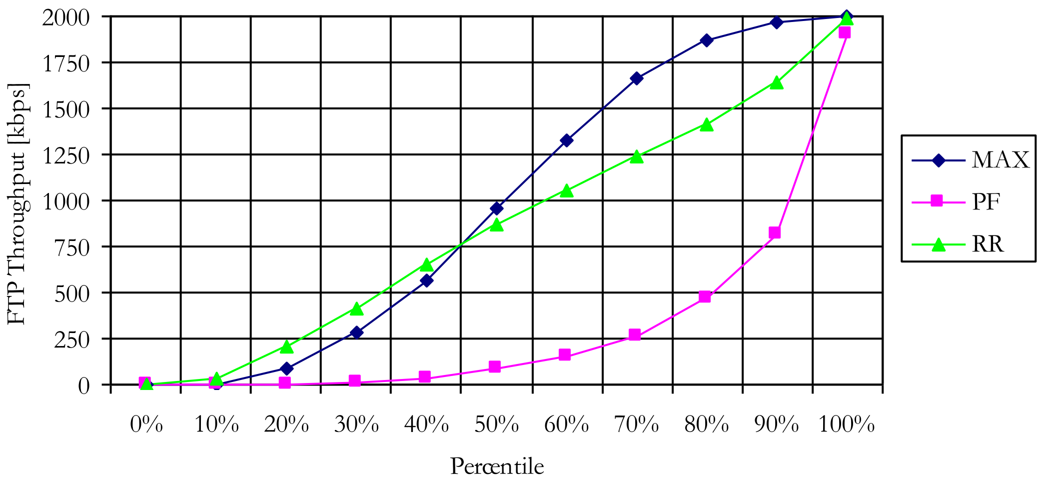


c) Rural

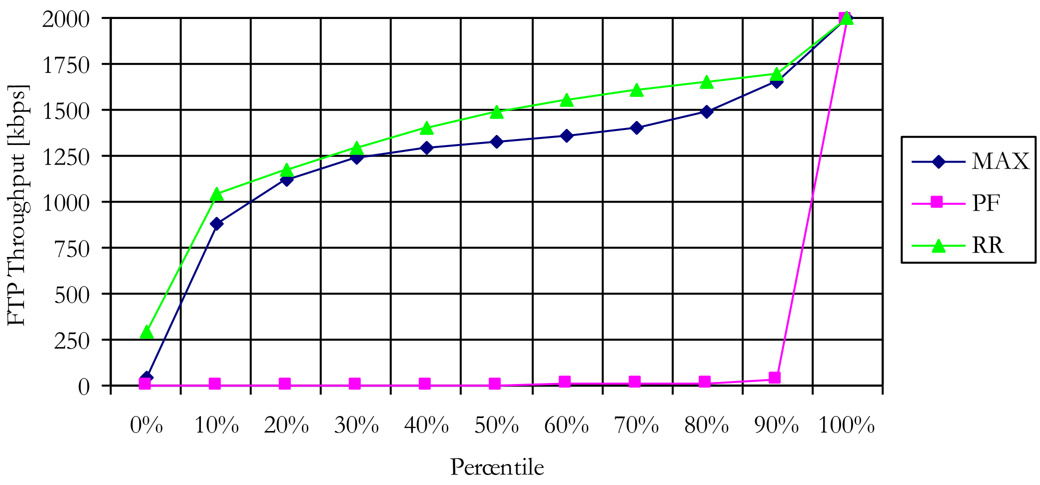
Figure I.8 - Average FTP throughput for 5MHz channel bandwidth.



a) Urban



b) Suburban



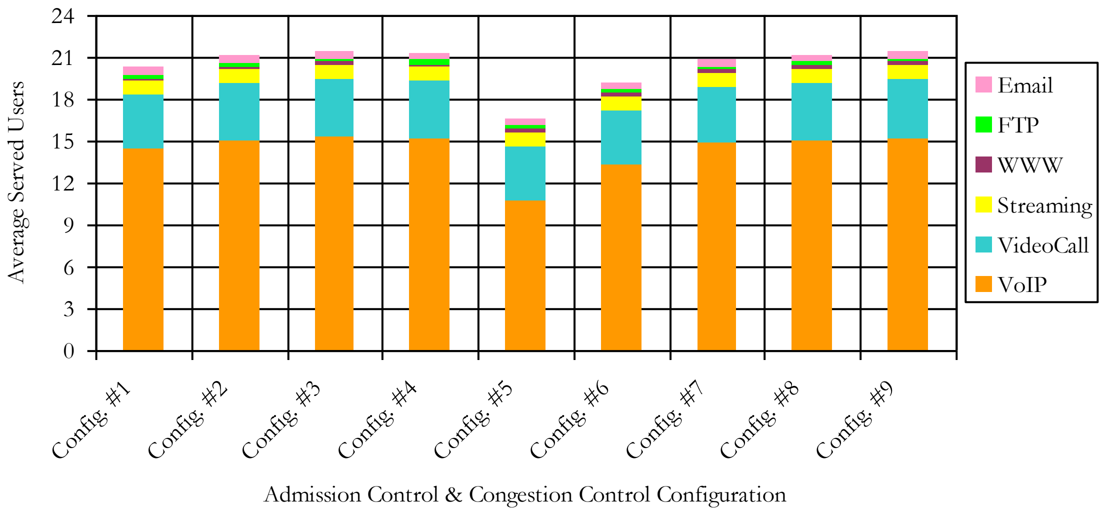
c) Rural

Figure I.9 - Cumulative per user FTP throughput distribution for 5MHz channel bandwidth.

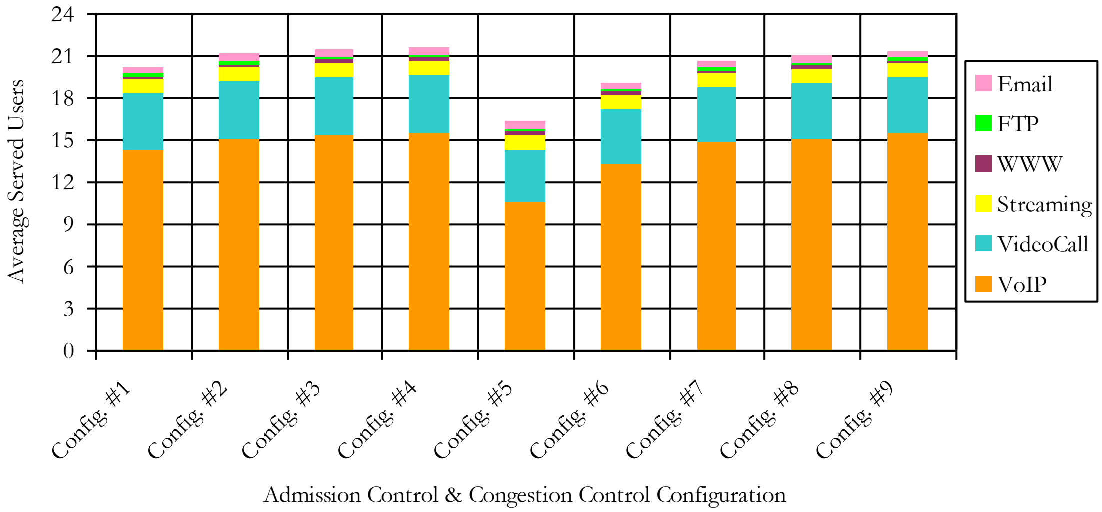
Annex J

Results for Variable Admission/Congestion Control Thresholds

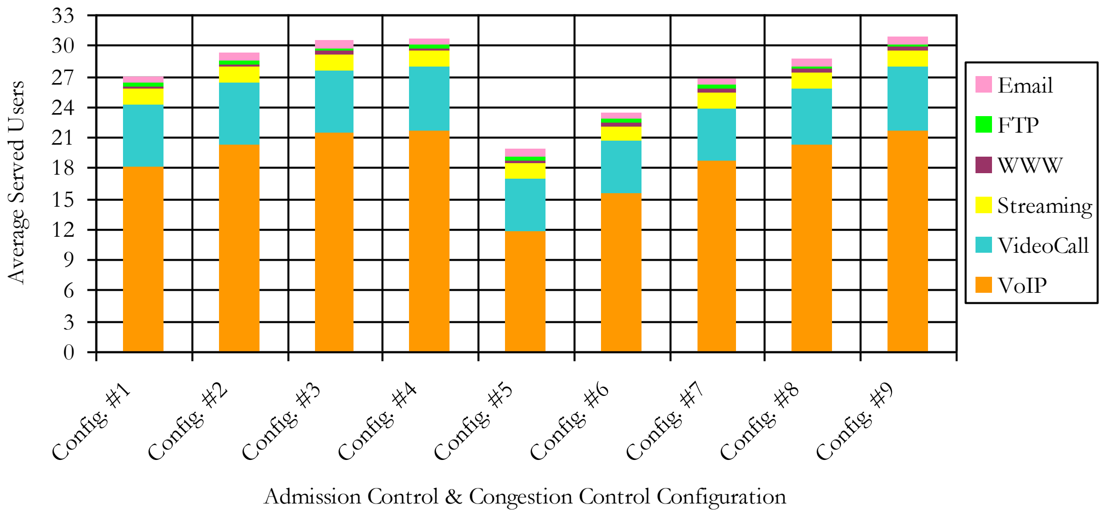
This annex summarises results for different Admission/Congestion Control thresholds.



a) Urban



b) Suburban



c) Rural

Figure J.1 - Impact of different Admission and Congestion Control thresholds on average served users.

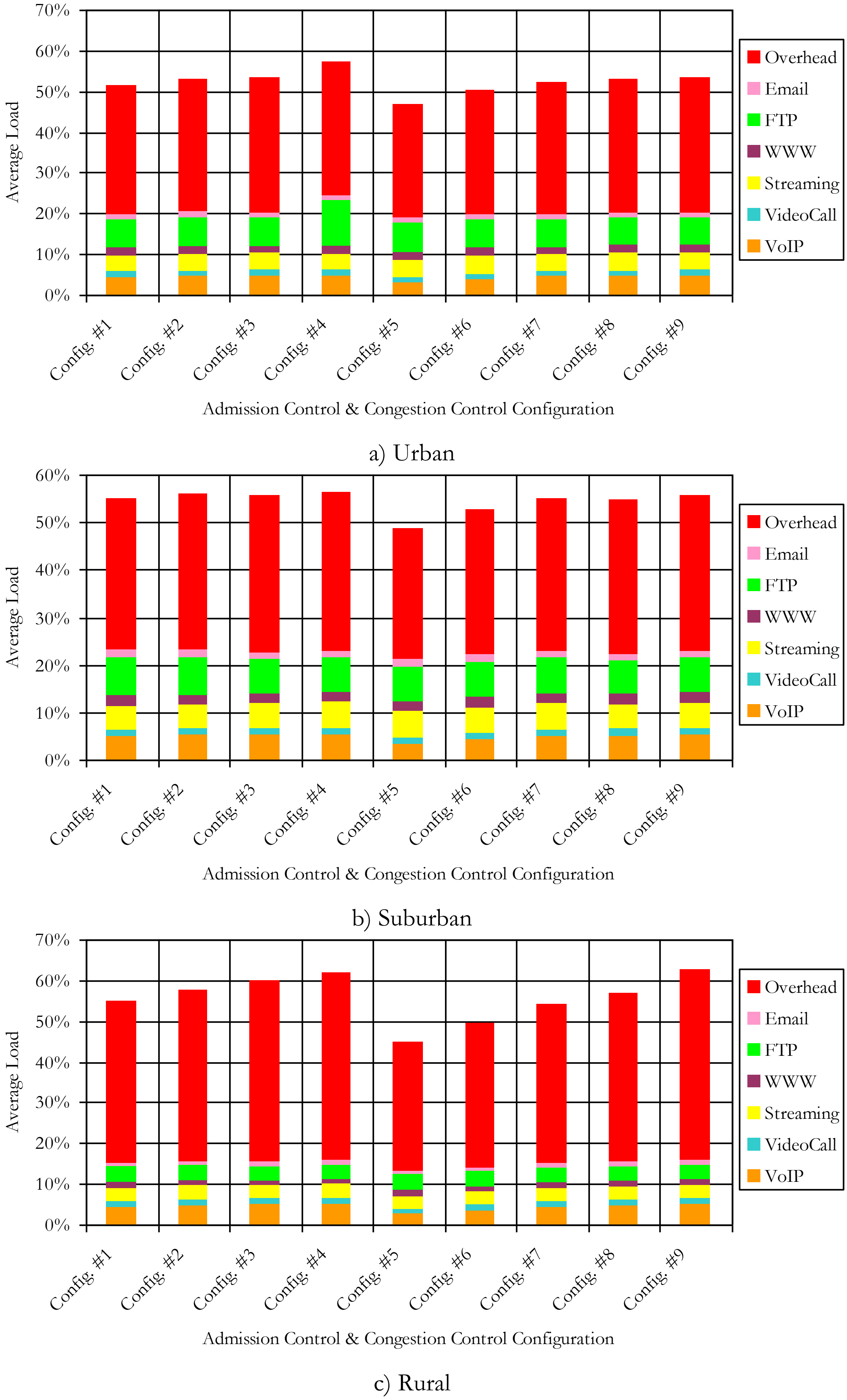
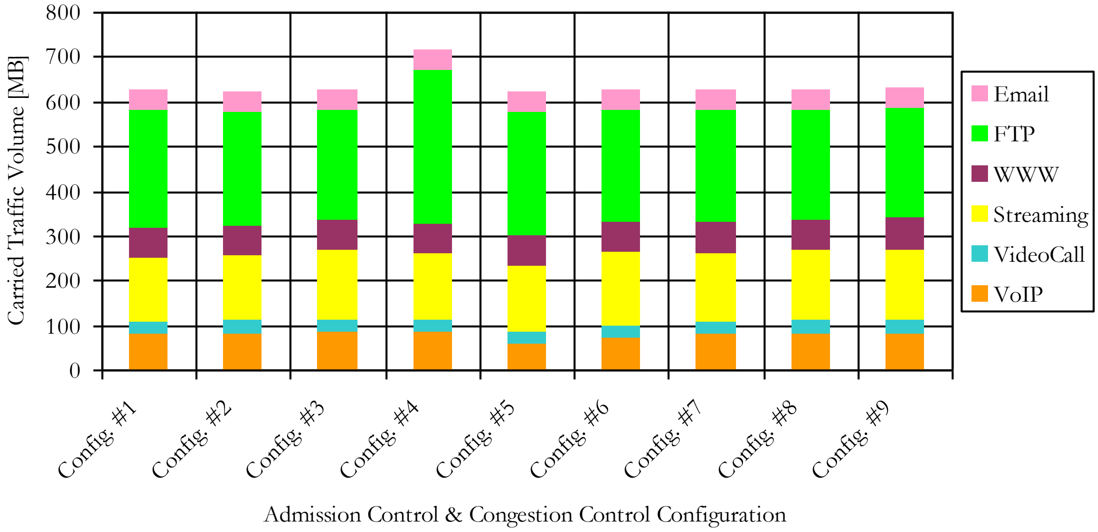
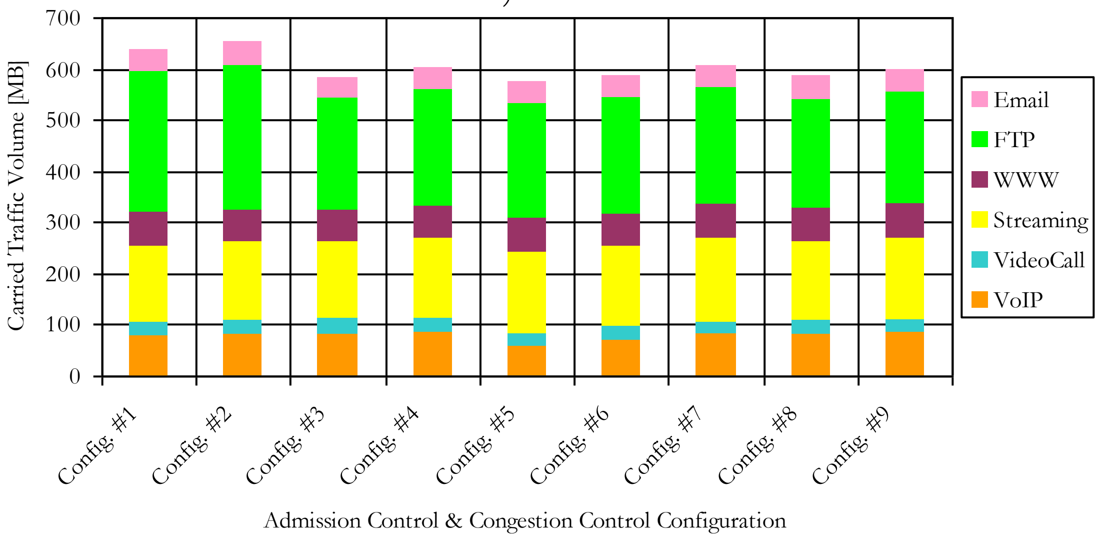


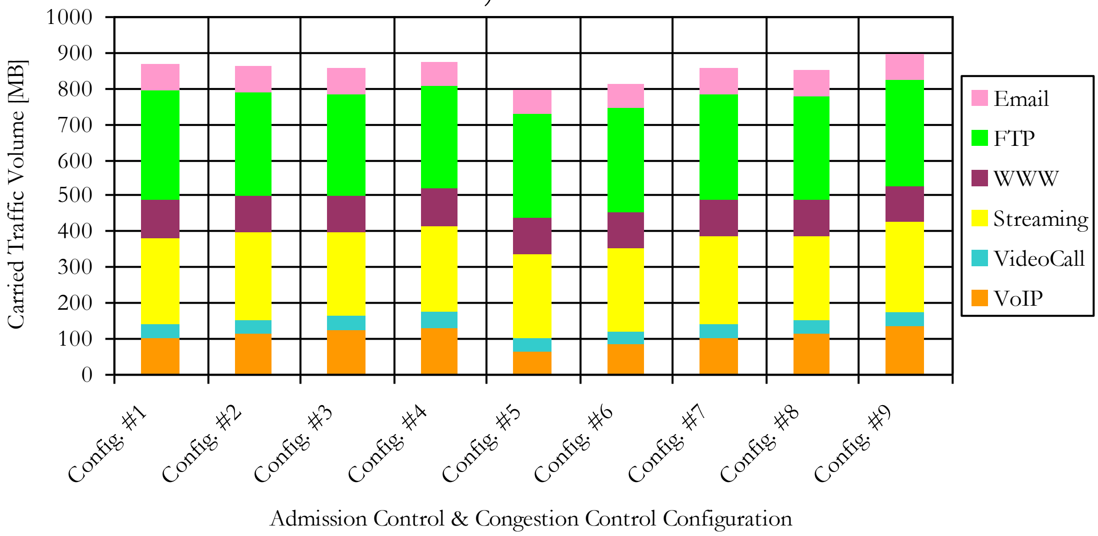
Figure J.2 – Impact of different Admission/Congestion Control thresholds on network load.



a) Urban

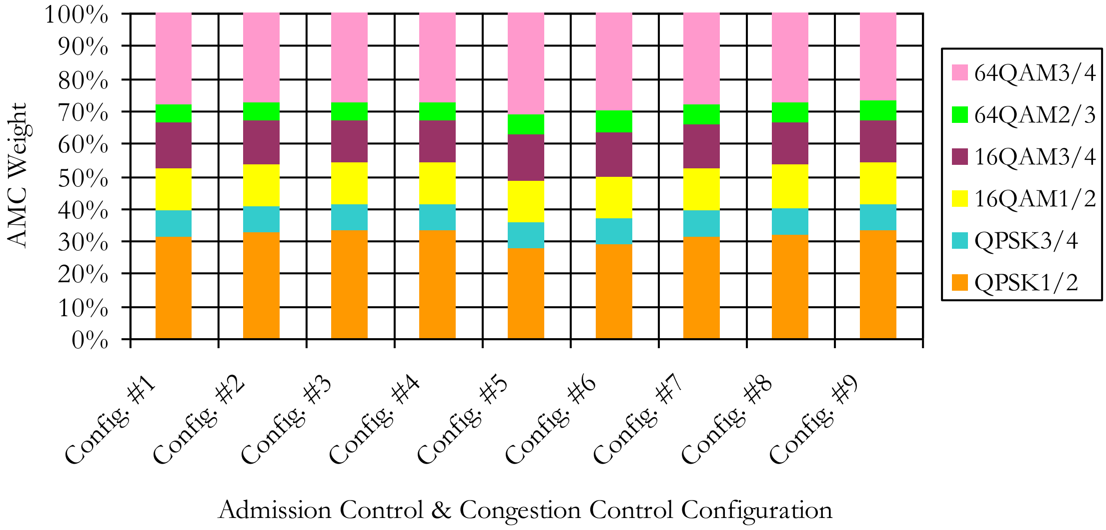


b) Suburban

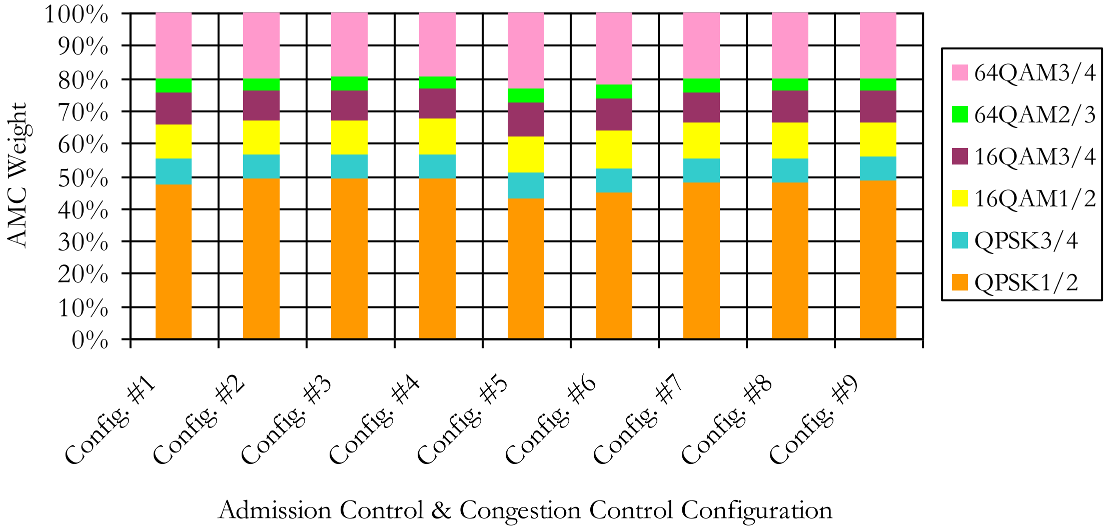


c) Rural

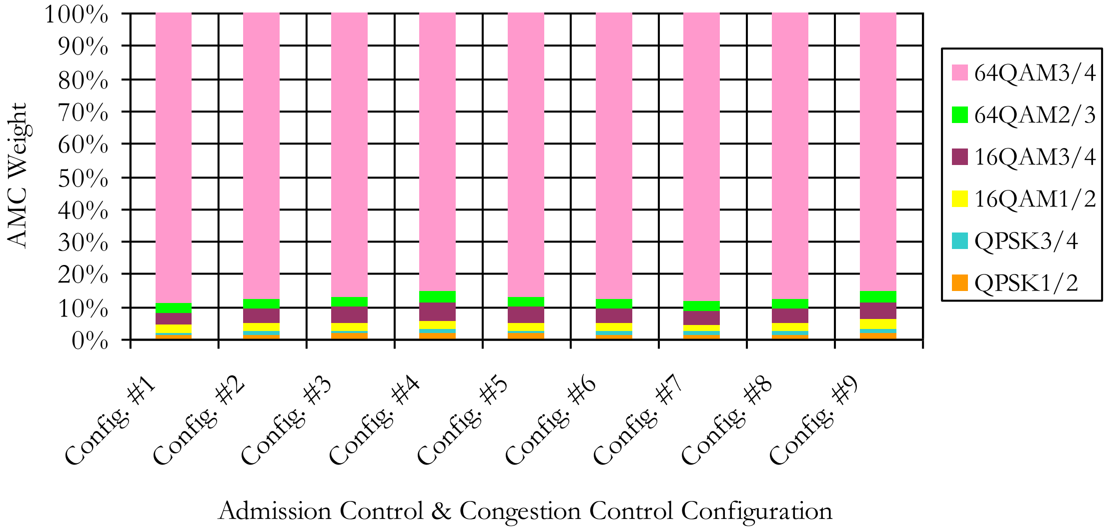
Figure J.3 - Impact of different Admission and Congestion Control thresholds on carried traffic volume.



a) Urban

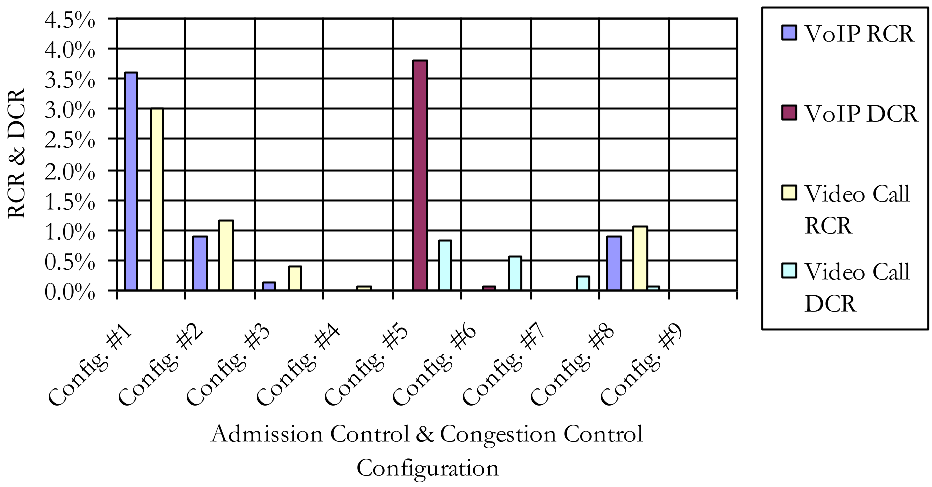


b) Suburban

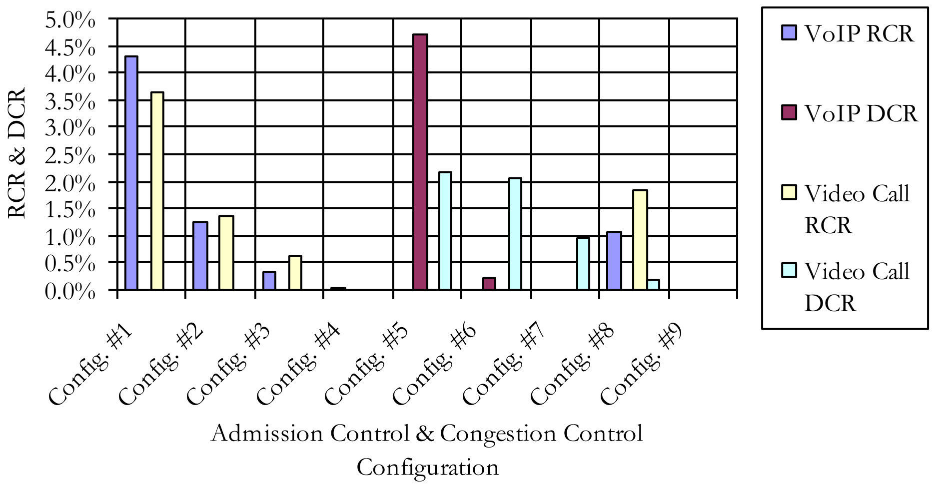


c) Rural

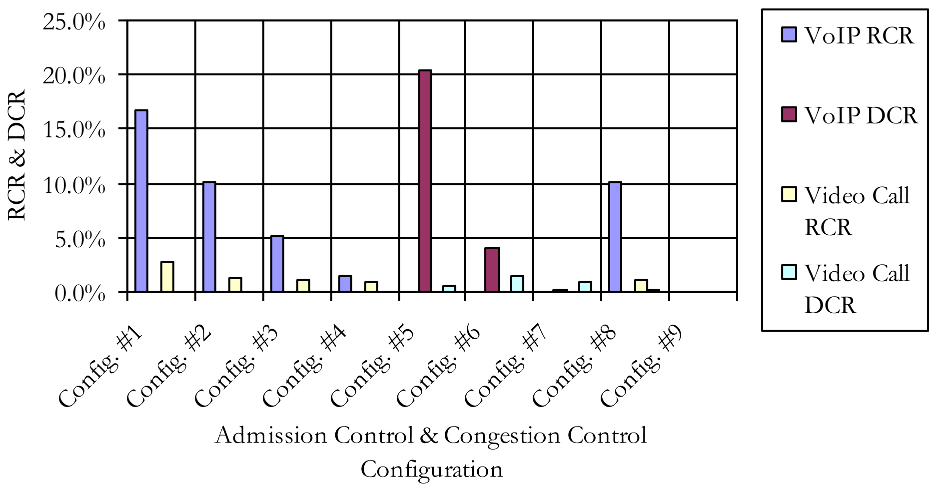
Figure J.4 - Impact of different Admission/Congestion Control thresholds on AMC Distribution.



a) Urban



b) Suburban



c) Rural

Figure J.5 - Impact of different Admission and Congestion Control thresholds on RTPS services main performance indicators.

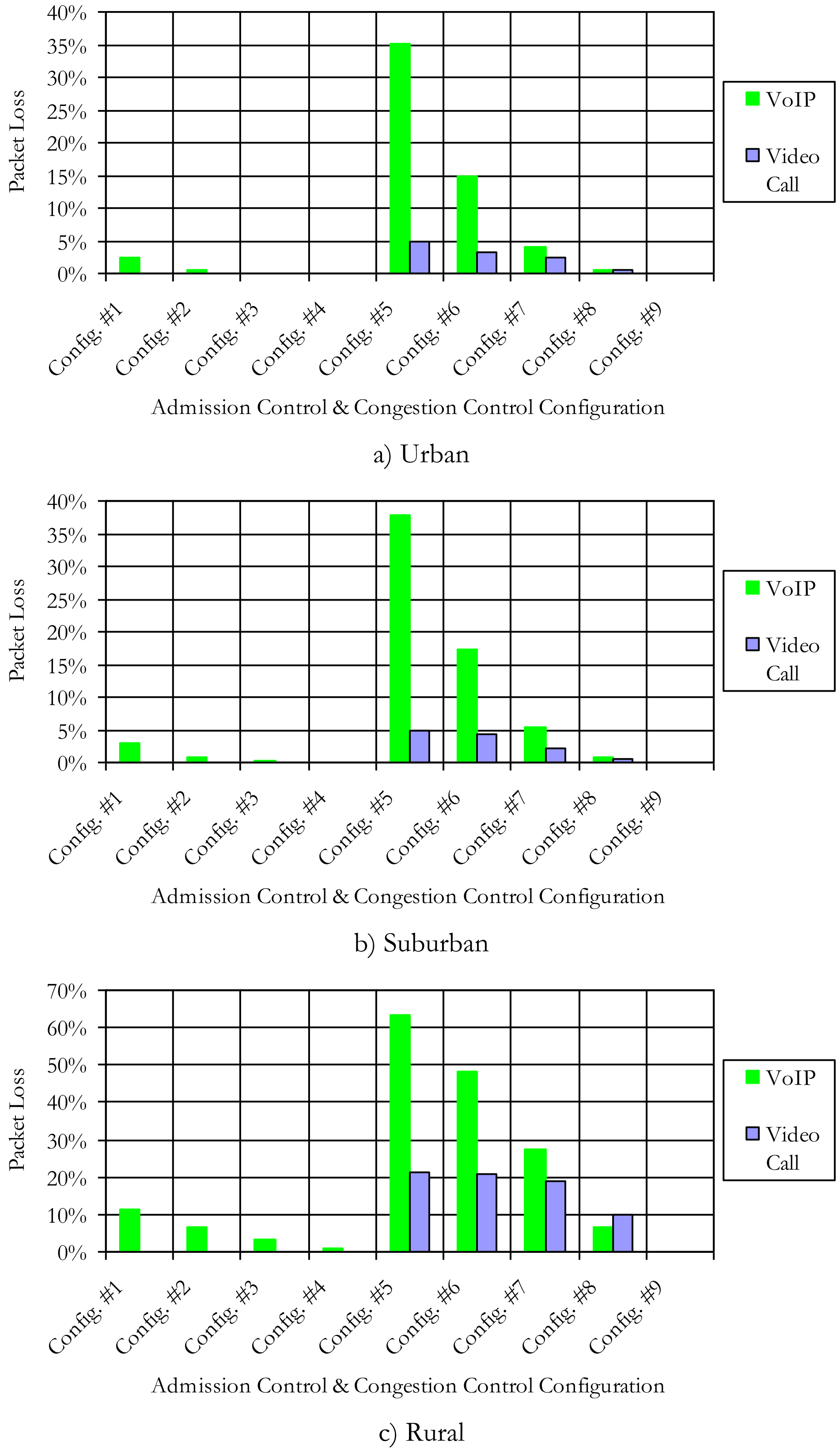


Figure J.6 - Impact of different Admission and Congestion Control thresholds on RTPS services packet loss.

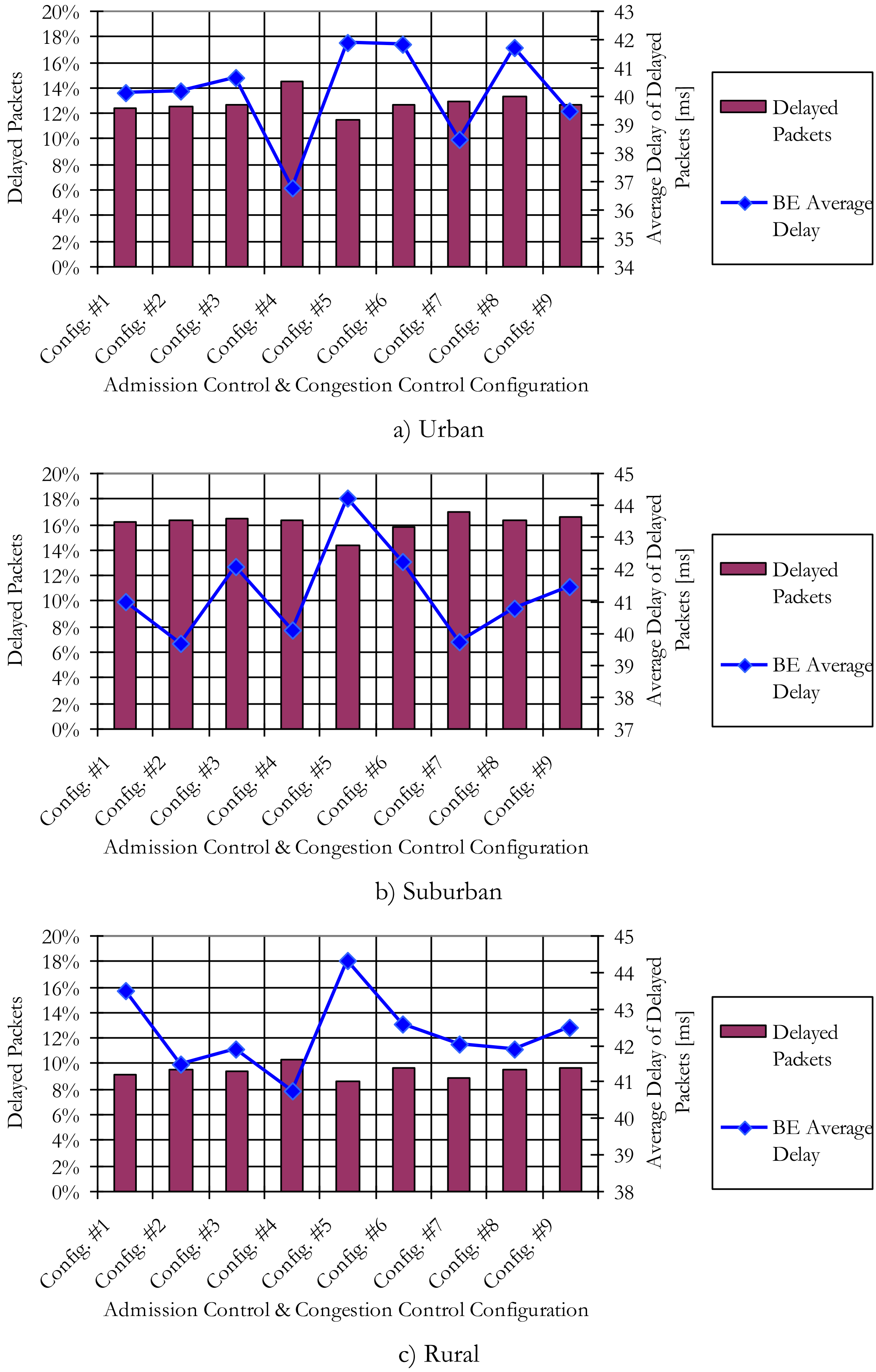
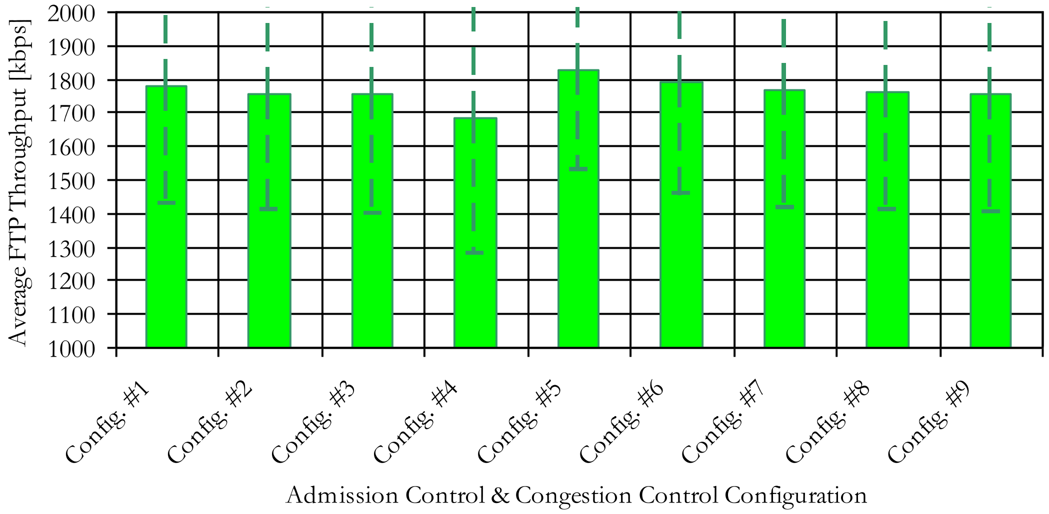
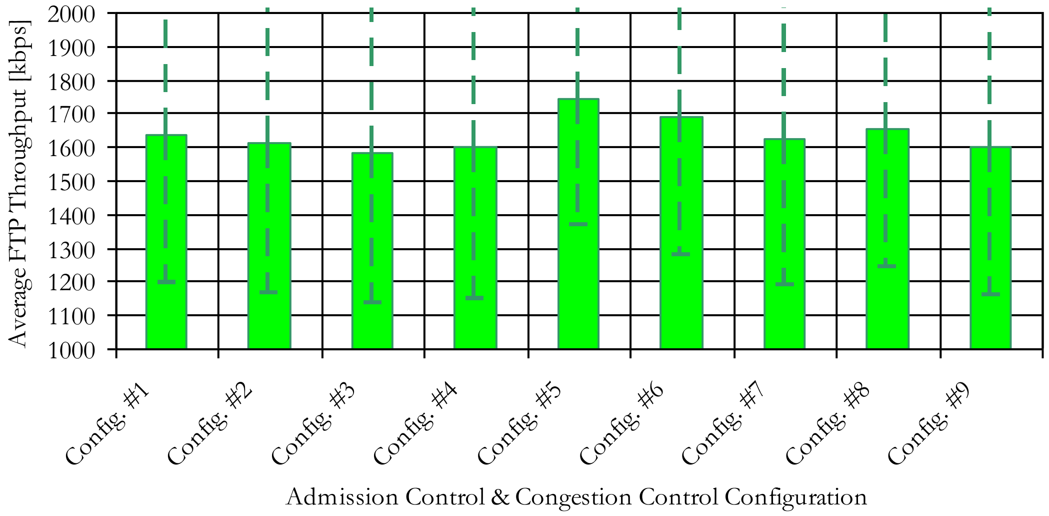


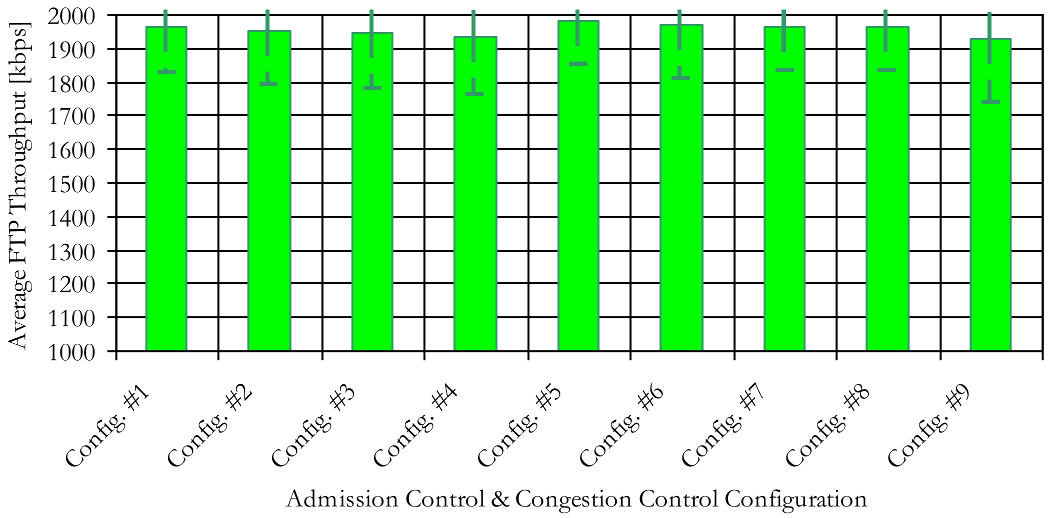
Figure J.7 - Impact of different Admission and Congestion Control thresholds on BE services delay.



a) Urban

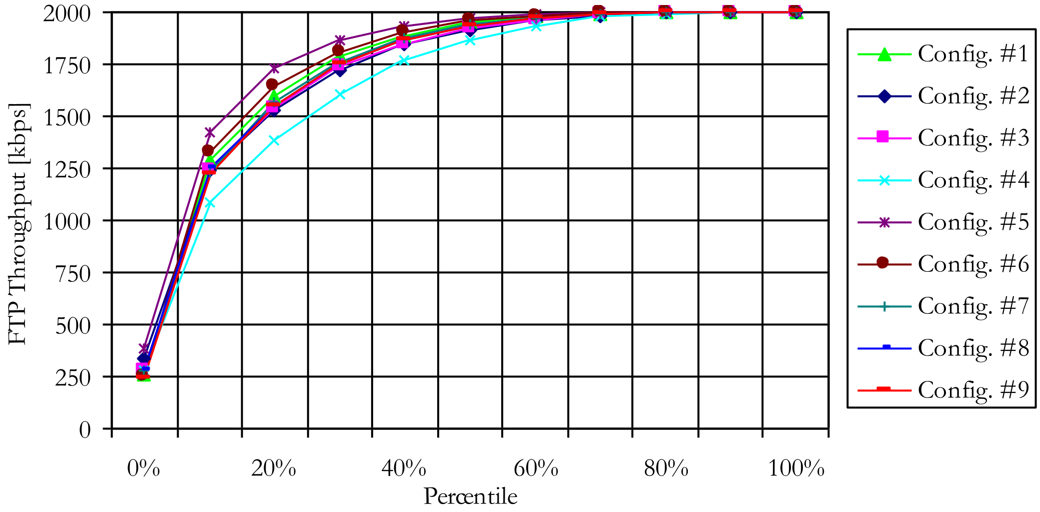


b) Suburban

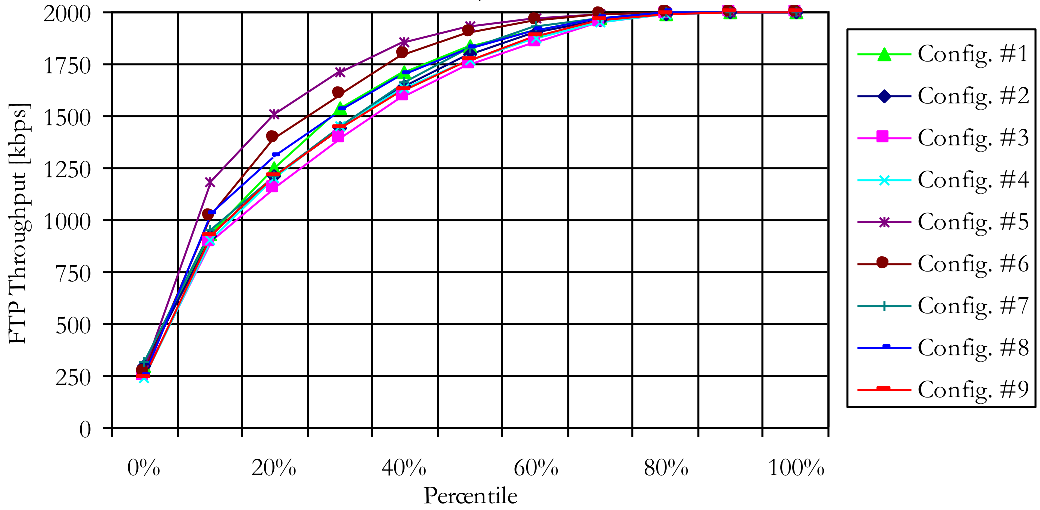


c) Rural

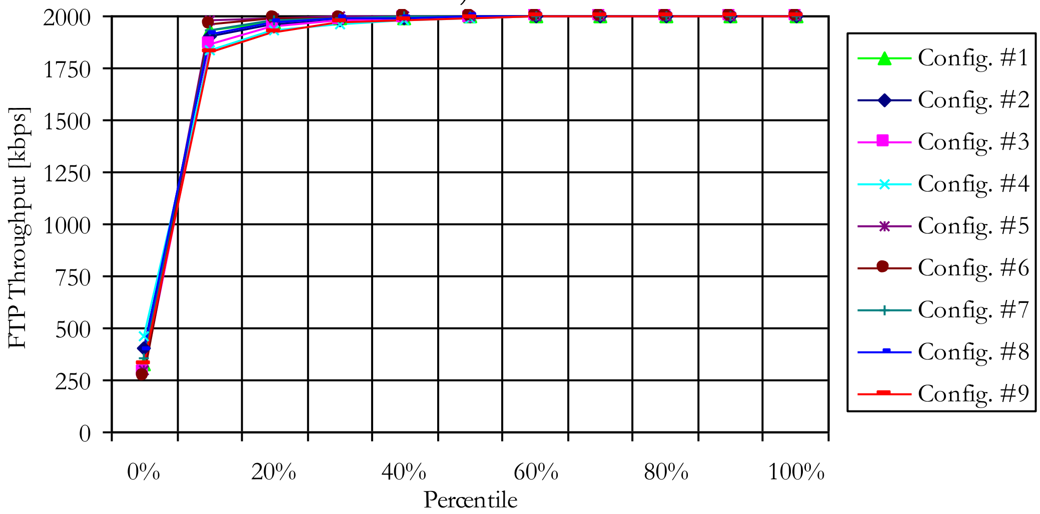
Figure J.8 - Impact of different Admission/Congestion Control limits on FTP throughput.



a) Urban



b) Suburban



c) Rural

Figure J.9 - Cumulative FTP throughput for different Admission/Congestion Control thresholds.

References

[3GPP02a] 3GPP, *3GPP system to Wireless Local Area Network (WLAN) Interworking; System description (Release 6)*, 3GPP Technical Specification No. 23.234 Ver. 1.0.0, Valbonne, France, Aug. 2002 (www.3gpp.org).

[3GPP02b] 3GPP, *Transparent end-to-end packet switched streaming service (PSS); Protocols and codecs (Release 4)*, 3GPP Technical Specification No. 26.234 Ver. 4.3.0, Valbonne, France, Mar. 2002. (www.3gpp.org)

[3GPP06] 3GPP, *3GPP System to Wireless Local Area Network (WLAN) interworking; WLAN User Equipment (WLAN UE) to network protocols; Stage 3 (Release 6)*, 3GPP Technical Specification No. 24.234 Ver 6.7.0, Valbonne, France, Sep. 2006 (www.3gpp.org).

[3GPP07a] 3GPP, *3GPP system to Wireless Local Area Network (WLAN) interworking; Stage 3 (Release 6)*, 3GPP Technical Specification No. 29.234 Ver 6.11.0, Valbonne, France, June 2007 (www.3gpp.org).

[3GPP07b] 3GPP, *3G Security; Wireless Local Area Network (WLAN) interworking security (Release 6)*, 3GPP Technical Specifcation No. 33.234 Ver. 6.9.0, Valbonne, France, Mar. 2007 (www.3gpp.org).

[Agui03] Aguiar,J., *Traffic Analysis at the Radio Interface in Converging Mobile and Wireless Communication Systems*, M.Sc. Thesis, Instituto Superior Técnico, Lisbon, Portugal, 2003.

[Alam06] Alamouti,S.M., WhyMAX? Mobile WiMAX Overview & Performance, *in Proc. of India Broadband Wireless Summit*, India, 2006.

[Alva06] Alvarion, *Understanding the Radio Technologies of Mobile WiMAX,* Alvarion, Israel, 2006 (http://www.alvarion.com/upload/contents/291/Radio_Mobile_WiMAX_wp.pdf).

[Corr03] Correia,L.M., *Mobile Communication Systems*, Course Notes, IST, Lisbon, Portugal, 2003.

[DaCo99] Damasso,E. and Correia,L.M. (eds), *Digital Mobile Radio Towards Future Generation Systems COST231 Final Report*, COST Office, Brussels, Belgium, 1999.

[EMPS06] Eklund,C., Marks,R.B., Ponnuswamy,S., Stanwood,K.L. and van Waes,N.J.M., *Wireless MAN – Inside the IEEE 802.16 Standard for Wireless Metropolitan Networks*, IEEE Press, Piscataway, NJ, USA, 2006.

[Eric06] Ericsson, *WiMAX – Copper in the Air*, Ericsson, Sweden, 2006 (http://www.ericsson.com/technology/whitepapers/3058_WiMAX_A.pdf).

[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.

[FiRe01] Fitzek,F.H.P. and Reisslein,M., "MPEG-4 and H.263 Video Traces for Network Performance Evaluation", *IEEE Network*, Vol. 15, No. 6, Nov./Dec. 2001, pp. 40-54.

[FrNg00] Frey,M. and Nguyen-Quang,S., "A Gamma-Based Framework for Modeling Variable-Rate MPEG Video Sources: The GOP GBAR Model", *IEEE/ACM Transactions on Networking*, Vol. 8, No. 6, Dec. 2000, pp. 710-719.

[Gonç98] Gonçalves,N.C., *Propagation Models for Microcellular Systems in the UHF Band,* M.Sc.. Thesis, Instituto Superior Técnico, Lisbon, Portugal, 1998.

[Guti03] Gutiérrez,P.J.A., *Packet Scheduling and Quality of Service in HSDPA*, Ph.D. Thesis, Aalborg University, Aalborg, Denmark, 2003.

[Hara03] Hara.,S. and Prasad,R., *Multicarrier Tecnhniques for 4G Mobile Communications*, Artech House, Norwood, MD, USA, 2003.

[Hata80] Hata,M., "Empirical Formula for Propagation Loss in Land Mobile Radio Services", *IEEE Transactions on Vehicular Technology*, Vol. VT-29, pp. 317-325, 1980.

[Heym97] Heyman,D., "The GBAR Source Model for VBR Videoconferences", *IEEE/ACM Transactions on Networking*, Vol. 5, No. 4, Aug. 1997, pp. 554-560.

[ITUR97] ITU-R, *Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000*, ITU-R Recommendation M.1225, Geneva, Switzerland, 1997

[IEEE02] IEEE, *IEEE Std. 802.16-2001 IEEE Standard for Local and Metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE, New York, NY, USA, Apr. 2002 (http://standards.ieee.org/getieee802/802.16.html).

[IEEE03] IEEE, *IEEE Standard for Local and metropolitan area networks --- Part 16: Air Interface for Fixed Broadband Wireless Access Systems--- Amendment 2: Medium Access Control Modifications and Additional Physical Layer Specifications for 2-11 GHz*, IEEE, New York, NY, USA, Apr. 2003 (http://standards.ieee.org/getieee802/802.16.html).

[IEEE04] IEEE, *802.16 IEEE Standard for Local and metropolitan area networks – Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE, New York, NY, USA, Oct. 2004 (http://standards.ieee.org/getieee802/802.16.html).

[IEEE06] IEEE, *802.16 IEEE Standard for Local and metropolitan area networks – Part 16: Air Interface for Fixed Broadband Wireless Access Systems – Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1*, IEEE, New York, NY, USA, Feb. 2006 (http://standards.ieee.org/getieee802/802.16.html).

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

[IkYU84] Ikegami,F., Yoshida,S. and Umehira,M., "Propagation Factors Controlling Mean Field Strength on Urban Streets", *IEEE Trans. on Antennas and Propagation*, Vol. 32, No. 8, Aug. 1984, pp. 822-829.

[KlLL01] Klemm,A., Lindemann,C. and Lohmann,M., "Traffic Modeling and Characterisation for UMTS Networks", in *Proc. of GLOBECOM'01 - IEEE Global Telecommunications Conference*, San Antonio, TX, USA, Nov. 2001.

[Kine07] Kineto Wireless, *The Complementary Roles of UMA and IMS in Fixed-Mobile Convergence*, Internal Report, Milpitas, CA, USA, 2007 (http://www.kinetowireless.com/products/downloads/kineto_wp_UMA_IMS_2007.pdf)

[Kirk06] Kirkaldy, N., *Wireless Broadband Solutions",* NTRA, Egypt, May 2006 (http://www.tra.gov.eg)

[LaGD00] Lazaro,O., Girma,D. and Dunlop,J., "Statistical analysis and evaluation of modelling techniques for self-similar video source traffic", in *Proc. of PIMRC'2000 - 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, London, UK, Sep. 2000.

[Ljun06] Ljung,R. (ed.) *et al*., *Target Scenarios Specification: Vision at Project Stage 1,* IST-AROMA Project, Del. D05, EC-IST Office, Brussels, Belgium, Mar. 2006, (http://www.aroma-ist.upc.edu)

[Mark03] Marks,R., "The IEEE 802.16 WirelessMAN Standard for Broadband Wireless Metropolitan Area Networks", in *IEEE Computer Society Distinguished Visitors Program*, IEEE C802.16-03/06, CO, USA, April 2003

[NCMP04] Nair,G., Chou,J., Madejski,T., Perycz,K., Putzolu and D., Sydir,J., "IEEE 802.16 Medium Access Control and Service Provisioning", *Intel Technology Journal,* Vol. 08, Issue 03, Aug. 2004, pp. 2-16.

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

[Nune03] Nunes. M. S., *Multi-Service Access Networks –Wireless Access Technologies (in Portuguese)* AEIST, Lisbon, Portugal, 2003.

[Orr04] Orr, J., *802.16 and 802.11: The Right Technology in the Right Place*, CEENet – Central and Eastern European Networking Association, Poland, 2004 (http://www.ceenet.org/workshops/lectures2004/Richard_Perlman/additional_materials_(wimax)/proxim-wimax.ppt - 256,1,802.16 and 802.11: The right technology in the right place).

[Pras98] Prasad,R., *Universal Wireless Personal Communications*, Artech House, Norwood, MD, USA, 1998.

[Sale98] Salema, C., *Hertzian Beams (in Portuguese)*, IST Press, Lisbon, Portugal, 1998.

[Stal00] Stallings, W., *Data & Computer Communications*, 6th ed., Prentice Hall, NJ, USA, 2000

[SLCM05] Sebastião,D., Ladeira,D., Carpinteiro,G., Martins,G., Esteves,H., Ferreira,L., Pereira,M., Kuipers,M., Costa,P., Correia,S. and Correia,L.M., *Characterisation of signal penetration into buildings for GSM900, GSM1800 and UMTS-FDD*, v3.0, Instituto de Telecomunicações, Lisbon, Portugal, Dec. 2005.

[WaBe88] Walfisch, J. and Bertoni, H. L., "A Theoretical Model of UHF Propagation in Urban Environments", *IEEE Trans. on Antennas and Propagation*, Vol. 36, No. 12, Dec. 1988, pp. 1788-1796.

[WiMA04a] WiMAX Forum, *Regulatory Position and Goals of the WiMAX Forum*, Internal Report, Beaverton, OR, USA, 2004 (http://www.wimaxforum.org/technology/downloads/WiMAX_Forum_Regulatory_Whitepaper_v08092004.pdf).

[WiMA04b] WiMAX Forum, *WiMAX's technology for LOS and NLOS environments,* Internal Report, Beaverton, OR, USA, 2004 (http://www.wimaxforum.org/technology/downloads/WiMAXNLOSgeneral-versionaug04.pdf)

[WiMA05] WiMAX Forum, *Can WiMAX Address Your Applications*, Internal Report, Beaverton, OR, USA, 2005, (http://www.wimaxforum.org/technology/downloads/Can_WiMAX_Address_Your_Applications_final.pdf).

[WiMA06a] WiMAX Forum, "Mobile WiMAX – Part I: A Technical Overview and Performance Evaluation", Internal Report, Beaverton, OR, USA, 2006, (http://www.wimaxforum.org/technology/downloads/Mobile_WiMAX_Part1_Overview_and_Performance.pdf).

[WiMA06b] WiMAX Forum, "Mobile WiMAX – Part II: A Comparative Analysis", Internal Report, Beaverton, OR, USA, 2006 (http://www.wimaxforum.org/technology/downloads/Mobile_WiMAX_Part2_Comparative_Analysis.pdf).

[WiMA06c] WiMAX Forum, "WiMAX End-to-End Network Systems Architecture - Stage 2: Architecture Tenets, Reference Model and Reference Points – Part 1", Internal Report, Beaverton, OR, USA, 2004 (http://www.wimaxforum.org/technology/documents/WiMAX_End-to-End_Network_Systems_Architecture_Stage_2-3_Release_1.1.0.zip).

[WiMA06d] WiMAX Forum, "WiMAX End-to-End Network Systems Architecture - Stage 2: Architecture Tenets, Reference Model and Reference Points – 3GPP WiMAX Interworking", Internal Report, Beaverton, OR, USA, 2004 (http://www.wimaxforum.org/technology/documents/WiMAX_End-to-End_Network_Systems_Architecture_Stage_2-3_Release_1.1.0.zip).

[WiMA06e] WiMAX Forum "Mobile WiMAX: The Best Personal Broadband Experience!", Internal Report, Beaverton, OR, USA, 2006 (http://www.wimaxforum.org/technology/downloads/MobileWiMAX_PersonalBroadband.pdf)

[WZEA04] Wong,I.C., Zukang,S., Evans,B.L. and Andrews,J.G., "A low Complexity Algorithm for Proportional Resource Allocation in OFDMA Systems", Dept. of Electrical and Computer Engineering, The University of Texas, Austin, TX, USA

[Yagh04] Yaghoobi,H., "Scalable OFDMA Physical Layer in IEEE 802.16 WirelessMAN", *Intel Technology Journal*, Vol. 8, Issue 3, Aug. 2004