
**UNIVERSIDADE FEDERAL DO CEARÁ**
**CENTRO DE TECNOLOGIA**
**DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA**
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA**

**ROBERTO PINTO ANTONIOLI**

# ADAPTIVE RADIO RESOURCE ALLOCATION ALGORITHM FOR USER SATISFACTION MAXIMIZATION IN MULTIPLE SERVICES WIRELESS NETWORKS

**FORTALEZA**

**2017**

ROBERTO PINTO ANTONIOLI

# ADAPTIVE RADIO RESOURCE ALLOCATION ALGORITHM FOR USER SATISFACTION MAXIMIZATION IN MULTIPLE SERVICES WIRELESS NETWORKS

Dissertação apresentada ao Curso de Mestrado em Engenharia de Teleinformática da Universidade Federal do Ceará, como parte dos requisitos para obtenção do Título de Mestre em Engenharia de Teleinformática. Área de concentração: Sinais e Sistemas

Orientador: Prof. Dr. Tarcisio Ferreira Maciel

Co-Orientador: Prof. Dr. Emanuel Bezerra Rodrigues

FORTALEZA

2017

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

A64a Antonioli, Roberto Pinto.
Adaptive Radio Resource Allocation Algorithm for User Satisfaction Maximization in Multiple Services Wireless Networks / Roberto Pinto Antonioli. – 2017.
101 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2017.
Orientação: Prof. Dr. Tarcisio Ferreira Maciel.
Coorientação: Prof. Dr. Emanuel Bezerra Rodrigues.

1. Utility Theory. 2. Multiple Services. 3. Radio Resource Allocation. 4. Quality of Service Provision. 5. User Satisfaction Maximization. I. Título.

CDD 621.38

ROBERTO PINTO ANTONIOLI

# ADAPTIVE RADIO RESOURCE ALLOCATION ALGORITHM FOR USER SATISFACTION MAXIMIZATION IN MULTIPLE SERVICES WIRELESS NETWORKS

Dissertation presented to the Master Program in Teleinformatics Engineering at the Federal University of Ceará, as part of the requirements for obtaining the Master's Degree in Teleinformatics Engineering. Concentration area: Signal and Systems.

Approved in: 10/08/2017.

EXAMINATION BOARD

---

Prof. Dr. Tarcisio Ferreira Maciel (Advisor)
Federal University of Ceará

---

Prof. Dr. Emanuel Bezerra Rodrigues (Co-Advisor)
Federal University of Ceará

---

Prof. Dr. Yuri Carvalho Barbosa Silva
Federal University of Ceará

---

Prof. Dr. Carlos Héracles Morais de Lima
São Paulo State University

## ACKNOWLEDGEMENTS

Firstly and above all, I would like to thank God for giving me health and guiding me to successfully complete one more step in my life.

I am extremely grateful to my parents, brother and all family members for their support, prayers, love, caring and sacrifices for providing a good education to form the person I am now. In special, I would like to thank my mother, Lenilce, for being such a hardworking and honorable person who raised me with love providing me with everything I needed to get to where I am now. I also thank my beloved Fernanda for the love, support, understanding and encouragement during the moments I most needed.

I would also like to thank very much my advisor Prof. Dr. Tarcisio F. Maciel and to my co-advisor Prof. Dr. Emanuel Bezerra Rodrigues, for receiving me at the Wireless Telecommunications Research Group (GTEL) and for the support and guidance during this master's journey. I also thank Prof. Dr. Fco. Rodrigo P. Cavalcanti for giving me the opportunity to be part of the GTEL team.

Also, lot of thanks to all my friend at the UFC and GTEL for their sincere friendship. Special thanks goes to Diego Sousa, who helped me a lot every time I was in need.

I also thank Prof. Dr. Yuri Carvalho Barbosa Silva and Prof. Dr. Carlos Heracles Morais de Lima, whose comments helped me to greatly improve the quality of this thesis.

Finally, I acknowledge the technical and financial support from FUNCAP, Ericsson Research and Ericsson Innovation Center, Brazil, under EDB/UFC.40 Technical Cooperation Contract.

Fortaleza, August 2017.

Roberto Pinto Antonioli

## ABSTRACT

The enriched service scope, the steep increase in mobile traffic volume, and the ever increasing number of connected devices in mobile networks coupled with the scarcity of electromagnetic spectrum have raised the importance of designing flexible and ingenious means to guarantee high user satisfaction levels. Therefore, in order to capture and maintain a representative share of the wireless communication market, effective ways to manage the scarce physical resources of cellular networks are fundamental for cellular network operators. The Radio Resource Allocation (RRA) algorithms are responsible for performing such a relevant and arduous task. The efficiency of such algorithms is essential so that there exists a fair resource allocation among users and the Quality of Service (QoS) requirements of each individual user are met, thus guaranteeing high user satisfaction levels.

The recent scenarios of cellular networks are composed of a wide range of available services for mobile users, which demand conflicting QoS requirements. In order to achieve the objective of user satisfaction maximization in such networks, we formulate a utility-based cross-layer optimization problem targeted at maximizing the user satisfaction in multi-service cellular networks. The optimal solution of the proposed problem is very hard to be found. Thus, we mathematically manipulate the problem and derive a low complexity suboptimal solution from which we design an adaptive RRA technique. Our technique is composed of user weights and an innovative service weight that is adapted to meet the satisfaction target of the most prioritized service chosen by the network operator. Furthermore, the proposed algorithm is scalable to several classes of service and can be employed in the current and future generations of wireless systems.

The performance evaluation of the proposed algorithm was conducted by means of system-level simulations in various scenarios. The evaluation was performed considering different multi-service scenarios. Then, the performance was evaluated considering imperfect Channel State Information (CSI) estimation at the transmitter. Significant gains in the overall system capacity were obtained in comparison with four benchmarking algorithms from the literature, demonstrating that the adaptability and service prioritization of the proposed algorithm are effective towards the objective of simultaneously maximizing the user satisfaction for multiple services.



**Keywords**: Utility Theory, Multiple Services, Radio Resource Allocation, Quality of Service Provision, User Satisfaction Maximization.

## RESUMO

O escopo enriquecido de serviços, o aumento acentuado do volume de tráfego móvel e o número cada vez maior de dispositivos conectados nas redes móveis, acompanhado pela escassez do espectro eletromagnético, aumentaram a importância de projetar meios flexíveis e engenhosos para garantir altos níveis de satisfação dos usuários. Portanto, para capturar e manter uma participação representativa no mercado das comunicações sem fio, mecanismos efetivos para gerenciar os recursos físicos escassos das redes celulares são fundamentais para as operadoras das redes celulares. Os algoritmos de alocação dos recursos de rádio (do inglês, *Radio Resource Allocation* (RRA)) são os responsáveis por executar essa tarefa tão relevante e árdua. A eficiência desses algoritmos é essencial para que exista uma alocação justa de recursos entre os usuários e os requisitos individuais de qualidade de serviço (do inglês, *Quality of Service* (QoS)) de cada usuário sejam atendidos, garantindo assim altos níveis de satisfação dos usuários.

Os cenários atuais das redes celulares são compostos por uma ampla gama de serviços disponíveis para usuários móveis, que exigem requisitos de QoS conflitantes. Para alcançar o objetivo de maximizar a satisfação dos usuários nessas redes, formulamos um problema de otimização baseado na teoria da utilidade considerando múltiplas camadas que visa maximizar a satisfação dos usuários em redes celulares com múltiplos serviços. A solução ótima do problema proposto é muito difícil de ser encontrada. Dessa forma, nós manipulamos matematicamente o problema e derivamos uma solução subótima de baixa complexidade a partir da qual nós desenvolvemos um mecanismo adaptativo de RRA. Nosso mecanismo é composto por prioridades relacionadas aos usuários e uma inovadora prioridade relacionada ao serviço que é adaptada para atender um objetivo de satisfação dos usuários de um serviço com maior prioridade escolhido pela operadora da rede. Além disso, o algoritmo proposto é escalável para várias classes de serviço e pode ser empregado nas gerações atuais e futuras de sistemas celulares.

A avaliação de desempenho do algoritmo proposto foi realizada por meio de simulações sistêmicas em vários cenários. A avaliação foi realizada considerando diferentes cenários com múltiplos serviços. Então, o desempenho foi avaliado considerando estimativa imperfeita da informação do estado de canal (do inglês, *Channel State Information* (CSI)) no transmissor. Ganhos significativos foram obtidos na capacidade total do sistema em comparação com quatro algoritmos encontrados da literatura, demonstrando que a adaptabilidade e priorização do serviço feita pelo algoritmo proposto são eficazes para atingir o objetivo de maximizar simultaneamente a satisfação dos usuários para múltiplos serviços.



**Palavras-chaves**: Teoria da Utilidade, Múltiplos Serviços, Alocação de Recursos de Rádio, Provisão de Qualidade de Serviço, Maximização da Satisfação do Usuário.

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| 1G | 1st Generation |
| 3GPP | 3rd Generation Partnership Project |
| 3GPP2 | 3rd Generation Partnership Project 2 |
| 4G | 4th Generation |
| 5G | 5th Generation |
| AMC | Adaptive Modulation and Coding |
| AWGN | Additive White Gaussian Noise |
| BLER | BLock Error Rate |
| BS | Base Station |
| CBR | Constant Bit Rate |
| CDMA | Code-Division Multiple Access |
| CQI | Channel Quality Indicator |
| CRA | Capacity-driven Resource Allocation |
| CSI | Channel State Information |
| DFT | Discrete Fourier Transform |
| DRA | Dynamic Resource Assignment |
| DSM | Delay-based Satisfaction Maximization |
| eNB | Evolved Node B |
| EPA | Equal Power Allocation |
| EPS | Evolved Packet System |
| E-UTRAN | Evolved Universal Terrestrial Radio Access Network |
| EPC | Evolved Packet Core |
| EXP | Exponential |
| FDD | Frequency Division Duplex |
| FER | Frame Erasure Rate |
| FTP | File Transfer Protocol |
| GBR | Guaranteed Bit Rate |
| GSM | Global System for Mobile Communications |
| HARQ | Hybrid Automatic Repeat Request |
| HOL | Head Of Line |
| IP | Internet Protocol |
| ITU | International Telecommunication Union |
| JSM | Joint Satisfaction Maximization |
| LTE | Long Term Evolution |
| LTE-A | Long Term Evolution (LTE)-Advanced |
| MAC | Medium Access Control |

| | |
|---|---|
| MCS | Modulation and Coding Scheme |
| MDU | Max-Delay-Utility |
| MIMO | Multiple Input Multiple Output |
| MISO | Multiple Input Single Output |
| MLWDF | Modified Largest Weighted Delay First |
| MME | Mobility Management Entity |
| NRT | Non-Real Time |
| OFDM | Orthogonal Frequency Division Multiplexing |
| OFDMA | Orthogonal Frequency Division Multiple Access |
| PDCP | Packet Data Convergence Protocol |
| PDN | Packet Data Network |
| PF | Proportional Fair |
| P-GW | Packet Data Network Gateway |
| PHY | Physical |
| PLR | Packet Loss Ratio |
| QAM | Quadrature Amplitude Modulation |
| QCI | QoS Class Identifier |
| QHMLWDF | Queue-HOL-MLWDF |
| QoE | Quality of Experience |
| QoS | Quality of Service |
| QPSK | Quadrature Phase Shift Keying |
| QSM | Queue-based Satisfaction Maximization |
| RAT | Radio Access Technology |
| RB | Resource Block |
| RLC | Radio Link Control |
| RM | Rate Maximization |
| RRA | Radio Resource Allocation |
| RRC | Radio Resource Control |
| RRM | Radio Resource Management |
| RT | Real Time |
| SC-FDMA | Single Carrier - Frequency Division Multiple Access |
| S-GW | Serving Gateway |
| SINR | Signal to Interference-plus-Noise Ratio |
| SISO | Single Input Single Output |
| SNR | Signal to Noise Ratio |
| SORA-NRT | Satisfaction-Oriented Resource Allocation for Non-Real Time Services |
| SORA-RT | Satisfaction-Oriented Resource Allocation for Real Time Services |
| TDD | Time Division Duplex |
| TDL | Tapped Delay Line |

| | |
|---|---|
| TDMA | Time Division Multiple Access |
| TSM | Throughput-based Satisfaction Maximization |
| TTI | Transmission Time Interval |
| TU | Typical Urban |
| TV | Television |
| UE | User Equipment |
| UEPS | Urgency and Efficiency-based Packet Scheduling |
| UMTS | Universal Mobile Telecommunications System |
| VoIP | Voice over IP |
| VTMLWDF | Virtual Token MLWDF |
| Wi-Fi | Wireless Fidelity |
| ZMCSCG | Zero Mean Circularly Symmetric Complex Gaussian |

## LIST OF SYMBOLS

| | |
|---|---|
| $\mathcal{J}$ | Set of all UEs |
| $J$ | Total number of UEs |
| $j$ | Index representing an UE from $\mathcal{J}$ |
| $\mathcal{K}$ | Set of all RBs |
| $K$ | Total number of RBs |
| $k$ | Index representing a RB from $\mathcal{K}$ |
| $P_t$ | Total transmit power of a BS |
| $p_k$ | Power allocated for RB $k$ |
| $h_{j,k}$ | Channel coefficient of link between BS and UE $j$ on RB $k$ |
| $PL(d)$ | Distance-dependent path loss |
| $d$ | Distance between BS and UE |
| $\sigma_{\text{sh}}$ | Log-normal shadowing standard deviation |
| $G_h(\theta)$ | Horizontal antenna gain component |
| $\theta$ | Horizontal angle relative to the main beam positioning direction |
| $G_v(\phi)$ | Vertical antenna gain component |
| $\phi$ | Negative angle elevation to the horizontal plane |
| $\phi^{\text{tilt}}$ | Downtilt angle |
| $G(\theta, \phi)$ | Total antenna gain |
| $\gamma_{j,k}$ | SNR of UE $j$ on RB $k$ |
| $\sigma^2$ | Average AWGN power |
| $r_{j,k}$ | Maximum achievable rate of UE $j$ on RB $k$ |
| $f(\cdot)$ | Link adaptation function |
| $R_j$ | Total rate allocated to UE $j$ |
| $\mathcal{K}_j$ | Subset of RBs allocated to UE $j$ |
| $\hat{h}_{j,k}$ | Estimated channel coefficient of link between BS and UE $j$ on RB $k$ |
| $\psi$ | Degradation of channel estimation |
| $\eta$ | Channel estimation error |
| $n$ | Index of a given TTI |
| $\mathbb{E}\{\cdot\}$ | Expected value |

| | |
|---|---|
| $\Delta n$ | CSI reporting delay |
| $T_j$ | Throughput of UE $j$ |
| $\Phi_{\text{req}}^{\text{thr}}$ | Throughput requirement |
| $\text{FER}_j$ | FER of UE $j$ |
| $\kappa_j^{\text{desc}}$ | Number of discarded packets of UE $j$ |
| $\kappa_j^{\text{sucesso}}$ | Number of successfully transmitted packets of UE $j$ |
| $\text{FER}_{\text{req}}$ | FER requirement |
| $\Phi_{\text{req}}^{\text{delay}}$ | HOL packet delay requirement |
| $\Upsilon$ | Percentage of satisfied UEs |
| $J^{\text{sat}}$ | Number of satisfied UEs |
| $\mathbf{x}$ | Vector of a generic QoS metric |
| $x_j$ | Generic QoS metric |
| $F(\mathbf{x})$ | Jain's fairness index of a generic QoS metric |
| $\rho_{j,k}$ | Assignment variable indicating UE $j$ allocated on RB $k$ |
| $U(\cdot)$ | General user utility function |
| $V(\cdot)$ | Service utility function |
| $S$ | Total number of services |
| $s$ | Index representing one service |
| $\mathcal{J}_s$ | Set of UEs from service $s$ |
| $U_s(\cdot)$ | User utility function from service $s$ |
| $x_j^s$ | QoS metric of UE $j$ from service $s$ |
| $U_{\text{thr}}(\cdot)$ | Utility function of throughput-based services |
| $T_j[n]$ | Throughput of UE $j$ at TTI $n$ |
| $R_j[n]$ | Total rate allocated to UE $j$ at TTI $n$ |
| $U'_{\text{thr}}(\cdot)$ | First derivative of $U_{\text{thr}}(\cdot)$ |
| $w_j$ | General user marginal utility |
| $w_j^{\text{thr}}$ | User marginal utility of throughput-based services |
| $f_{\text{thru}}$ | Filtering constant for throughput calculation |
| $U_{\text{delay}}(\cdot)$ | Utility function of delay-based services |
| $d_j^{\text{hol}}[n]$ | HOL packet delay of UE $j$ at TTI $n$ |
| $U'_{\text{delay}}(\cdot)$ | First derivative of $U_{\text{delay}}(\cdot)$ |

| | |
|---|---|
| $w_j^{\text{delay}}$ | User marginal utility of delay-based services |
| $\lvert \cdot \rvert$ | Absolute value |
| $f_{\text{thru}}$ | Filtering constant for throughput calculation |
| $t_{\text{tti}}$ | TTI duration |
| $L$ | Packet arrival rate |
| $S_{\text{p}}$ | Packet size |
| $U_{\text{queue}}(\cdot)$ | Utility function of queue-based services |
| $\overline{Q}_j[n]$ | Average queue size of UE $j$ at TTI $n$ |
| $U'_{\text{queue}}(\cdot)$ | First derivative of $U_{\text{queue}}(\cdot)$ |
| $w_j^{\text{queue}}$ | User marginal utility of queue-based services |
| $f_{\text{queue}}$ | Filtering constant for queue size calculation |
| $Q_j[n]$ | Queue size of UE $j$ at TTI $n$ |
| $\alpha[n]$ | Amount of arrival bits of UE $j$ at TTI $n$ |
| $\omega_j$ | Source data rate of UE $j$ |
| $\mathcal{J}_{\text{thr}}$ | Set of UEs from throughput-based service |
| $\mathcal{J}_{\text{delay}}$ | Set of UEs from delay-based service |
| $\mathcal{J}_{\text{queue}}$ | Set of UEs from queue-based service |
| $w_j^s$ | Service marginal utility of UE $j$ using service $s$ |
| $j^\star$ | UE selected on RB $k$ at TTI $n$ |
| arg max$\{\cdot\}$ | Function that returns the maximum element from the input elements |
| $\sigma$ | Value that determines shape of user utility functions |
| $\mu$ | Value that defines if logistic utility function is decreasing or increasing |
| $x_j^{\text{req}}$ | General QoS requirement of UEs |
| $\rho$ | Proportion of general QoS metric |
| $\delta$ | Value of user logistic function for a given value of QoS metric |
| $\sigma_{\text{thr}}$ | Shape value for throughput-based user utility function |
| $\sigma_{\text{delay}}$ | Shape value for delay-based user utility function |
| $\sigma_{\text{queue}}$ | Shape value for queue-based user utility function |
| $z_j[n]$ | General values of a given user marginal utility |
| $\lambda$ | Value that determines shape of service utility function |
| $\Upsilon^1$ | Satisfaction target for service 1 |

| | |
|---|---|
| $w_j[n]$ | Average waiting time of UE $j$ at TTI $n$ |
| $\lambda_j$ | Satisfaction target for service 1 |
| $U_V(\mathrm{w})$ | Utility function for VoIP services |
| $U_S(\mathrm{w})$ | Utility function for streaming services |
| $U_B(\mathrm{w})$ | Utility function for best effort services |
| $U_{\mathrm{RT}}(\cdot)$ | Utility function for RT services |
| $U_{\mathrm{NRT}}(\cdot)$ | Utility function for NRT services |
| $\beta$ | Value that determines $U_{\mathrm{RT}}(\cdot)$ slope |
| $a$ | Value that determines $U_{\mathrm{NRT}}(\cdot)$ slope |
| $b$ | Value that determines $U_{\mathrm{NRT}}(\cdot)$ amplitude |
| $c$ | Value that determines $U_{\mathrm{NRT}}(\cdot)$ QoS requirement |
| $\alpha$ | Maximum allowed probability for packet exceeding delay requirement |
| $\overline{d_j^{\mathrm{hol}}}[n]$ | Mean HOL packet delay of all RT UEs at TTI $n$ |

**CONTENTS**

# 1 INTRODUCTION

This is an introductory chapter where the motivation and scope of this thesis are presented in sections 1.1 and 1.2, respectively. Then, the methods used for conducting the studies and performance evaluation of this thesis proposal are discussed in section 1.3. The main contributions os this Master's thesis are summarized in section 1.4. Section 1.5 depicts the thesis organization, and, finally, the main scientific production during the Master program are listed in section 1.6.

## 1.1 Motivation

Driven by the proliferation of mobile devices, the demand for more high-quality content has experienced a steep increase, resulting in a mobile data traffic growth of 4,000-fold over the past 10 years [1]. This avalanche of the mobile traffic volume has severely stretched the available wireless spectrum, leading the academy and industry to trigger an investigation of a new generation of cellular networks [2]. The 5$^{th}$ Generation (5G) of mobile communication systems is expected to be deployed beyond 2020, when it is foreseen that there will be a thousand times higher mobile traffic per unit area and 10-100 times higher user data rate [3].

The innovative mobile devices, such as smartphones and tablets, coupled with the improvements of mobile communication networks have produced an eruption of new services and applications with a variety of QoS requirements that need to be supported by recent and future mobile networks. Among all these services, mobile multimedia applications, such as streamed video viewing and IPTV, represent a big slice of mobile traffic, accounting for 55% ot the total mobile data traffic in 2015 [1]. Besides, it is predicted that approximately 75% of the world's mobile data traffic will be of video by 2021 [1, 4]. As an example of the mobile video traffic consumption growth, teens (aged 16-19) have increased smartphone TV/video viewing in 85% from 2011 to 2015 [4].

Streaming video services are characterized as Real Time (RT) services because they have low latency requirements between the communicating parts, namely the video servers and mobile video users. This short time response imposes requirements regarding packet delay, jitter and Frame Erasure Rate (FER). Besides that, the present-day video services are also characterized by requiring high throughput rates because of the current high resolution demands (e.g., 4K, 1080p and 720p). Other RT services, such as VoIP and on-line games, do not generate high data rates as video services; thus, these services are mainly characterized by requiring quick responses from the communicating parts because outdated information is not useful. On the other hand, there are the NRT services which do not have strict packet delay requirements, but high latencies are unacceptable. Examples of NRT services are Web browsing and File Transfer Pro-

tocol (FTP) services, which sometimes require high throughput demands in case of transfers of large files. Besides the aforementioned services, a very wide range of services is available for mobile users, which poses some challenges for the network operators when it comes to fulfilling these conflicting QoS requirements in scenarios with a limited number of resources.

Therefore, the complex recent scenarios are composed of a variety of services requiring more and more bandwidth in order to have their QoS requirements satisfied. This competitive scenario complicates the network operators task of efficiently managing the scarce and limited wireless resources since they need to guarantee the highest user satisfaction possible, even in scenarios with services that demand conflicting QoS requirements.

In this context, optimized RRA algorithms are required to support the unprecedented high demand for high-quality services. The RRA algorithms are responsible for selecting which User Equipments (UEs) will have access to the system resources, where the central objective is to efficiently manage the limited radio resources to maximize revenues and user satisfaction by meeting their QoS requirements. Accommodating all users using the restricted resources available in the network would be the best scenario, however this is not always possible in overload situations. As a consequence, the network operators may want to prioritize a given user or service class according to a certain criterion. Therefore, a desired feature that might be available in RRA frameworks is the flexibility for network operators to adjust the operating point reflecting strategic decisions or customers' trending [5].

Considering recent and complex scenarios, this thesis proposes an adaptive and low complexity RRA algorithm that targets the user satisfaction maximization for scenarios composed of multiple service classes with conflicting QoS requirements. Besides being normalized and unified across all service classes, the proposed algorithm employs an innovative service prioritization that is adapted to meet the satisfaction target of the most prioritized service. This specific feature allows network operators to flexibly define their strategy.

## 1.2 Thesis Scope

The main tool used in this thesis for the development of the proposed RRA framework is the Utility Theory. This theory has been widely used in the literature for designing RRA algorithms for cellular networks. For instance, this theory has been used for single service scenarios in [6, 7, 8, 9], and for multi-service scenarios in [10, 11, 12]. The Utility Theory was initially conceived for applications in the economics area, where it was applied to explain the consumers' behavior and help in the decision-taking process [13, 14]. However, this theory has also received some attention of the wireless communications research community over the last years [15].

The RRA algorithms for cellular networks aim at guaranteeing a trade-off between QoS, spectral efficiency and fairness in the Resource Blocks (RBs) allocation. In the economics field, the utility theory has been used to study the problem of providing a fair and efficient resource

allocation, where utility functions have been applied for quantifying the advantage of using particular resources [6]. A similar approach can be employed in the area of cellular networks, where an evaluation of how well the network is satisfying the users' applications requirements could be conducted by using metrics such as throughput, FER or outage probability [16].

Therefore, the utility theory emerges as a powerful tool for the conception of RRA algorithms since it allows us to quantify the user satisfaction levels for a given resource allocation. Thus, it is possible to design RRA algorithm capable of achieving different levels of fairness and user satisfaction in the resource allocation process [15, 11].

The work proposed in [9] is the basis of this thesis proposal; herein, we propose to extend the single-service cases analyzed in [9] for multi-service scenarios. Together with [9], the works in [10, 11, 12] were also used for designing this thesis proposal. Based on these works, this thesis proposes a RRA algorithm for the downlink of 4$^{th}$ Generation (4G) LTE cellular networks. In the downlink, the 4G cellular networks that follow the LTE standard employ Orthogonal Frequency Division Multiple Access (OFDMA), which is the multiple access scheme studied in this thesis. More details about the LTE standard, 4G cellular networks and OFDMA are given in Chapter 2.

The main objective of the proposed RRA algorithm is to maximize the user satisfaction in scenarios composed of multiples services. Satisfying a user is directly connected to meeting its QoS requirements; throughput, packet delay, jitter and FER are among the most common QoS requirements. The specific set of QoS requirements to be met depends on the service the user is accessing. In this thesis, we study scenarios composed of CBR, VoIP and video services, which have as their main QoS requirements to be met the throughput, FER and both throughput and FER, respectively. However, the amount of resource to be distributed among the users is limited. Thus, one can see that the studied scenarios are comprised of services demanding conflicting QoS requirements.

## 1.3 Research Method

The intrinsic characteristics of wireless communications systems, such as the fading in space and time domains as well as the time variability of the mobile radio channel, are very complex and difficult to be tackled mathematically. Thus, these characteristics make it intractable to conduct complete performance evaluations using an analytical/mathematical approach [17]. Additionally, the presence of a variety of different services modeled with random variables turns the evaluation by means analytical/mathematical approaches even more complex. An alternative to overcome this complexity is to use system-level simulations, which are effective and widely used in the literature.

Therefore, in this thesis, we carry out the performance evaluation using a system-level simulator that models the most important aspects of the downlink of 4G LTE cellular networks based on OFDMA. The simulator is fully written in the C++ programming language using the

object oriented programming paradigm and was developed in the context of research projects, with participation of the author of this thesis, at the Wireless Telecommunications Research Group (GTEL). More details about the modeling of this simulator are given in Chapter 2.

The proposed RRA algorithm and some benchmark algorithms were implemented in the system-level simulator and a performance evaluation was conducted based on the statistical analysis of the simulations results. The studied scenarios were simulated several times based on the Monte Carlo approach so that a reliable statistical confidence was achieved. The duration of the simulations was defined so that the main performance metrics were stable after 10% of the beginning of the simulations.

## 1.4 Contributions

As far as we know, none of the RRA algorithms found in the literature have proposed to maximize the user satisfaction in multi-service scenarios or to leave the choice for the network operator to decide which service has more priority during the resource allocation in the downlink of OFDMA-based cellular networks. Therefore, to the best of our knowledge, the present work is the first one to propose and evaluate the performance of a cross-layer utility-based RRA framework which has the specific objective of simultaneously maximizing the satisfaction of multiple services in the downlink of OFDMA-based systems. The proposed technique is suitable for scenarios composed of video services, web browsing, VoIP, among others.

The main contribution of this thesis are summarized as follows:

- We formulate an utility-based optimization problem with the objective of simultaneously maximizing the satisfaction of users from distinct service classes.
- The problem is mathematically manipulated and a sub-optimal solution is derived, from which we design a low complexity RRA algorithm for maximizing the user satisfaction in recent multi-service scenario composed of video services, web browsing, VoIP, etc. The main features of the proposed algorithm are:
  - **Unified**: the same formulation and policy is applied for all services, regardless of their main QoS metric. Furthermore, only one parametrized utility function is applied for all services, thus providing a fair comparison between the services;
  - **Normalized**: before calculating the utility-based weights for all services, we normalized their main QoS metric by their QoS requirement, so that our framework becomes independent of the QoS metrics;
  - **Adaptive**: where the adaptation is performed based on non-linear steps of one parameter of a utility function. A look-up table was calculated to be employed in the algorithm. The proposed algorithm is dynamically adapted to protected the satisfac-

tion level of a most prioritized service. This type of adaptation is the main contribution of this work;

  - **Flexible**: The proposed framework allows the network operator to decide which service is more protected depending on their strategy, so that its satisfaction level is sustained approximately in a plateau for different traffic loads. As far as we know, this specific characteristic has never been proposed in the literature for the downlink of OFDMA-based cellular networks

- Extensive system-level simulations are conducted and a performance evaluation is carried out using the joint capacity plane, which is a complete form of evaluation since it simultaneously illustrates the algorithms performance for single and multi-service scenarios.

In this thesis, we particularize this general framework and develop a specific algorithm for scenarios composed of mixes between video and throughput-based services, as well as for scenarios comprised by VoIP and throughput-based services.

## 1.5 Thesis Organization

Chapter 2 firstly provides an overview of the LTE/LTE-Advanced (LTE-A) architecture, detailing the main network entities and their functionalities. Then, we present the main assumptions taken into account for the system model used in this thesis, which are based on the 3rd Generation Partnership Project (3GPP) LTE standard.

Chapter 3 describes a broad and general survey about works found in the literature related to the RRA research topic. Also, this chapter presents a detailed description of the benchmark algorithms used for performance comparison in this thesis.

Chapter 4 describes the mathematical formulation of the utility-based optimization problem that targets the user satisfaction maximization in multi-service scenarios. A suboptimal solution is mathematically derived, from which a low complexity and adaptive RRA algorithm is derived, which is called Joint Satisfaction Maximization (JSM).

In Chapter 5, a performance evaluation and comparison of the proposed RRA algorithms and four benchmarking algorithms is conducted. The scenarios analyzed include two multi-service scenarios composed of different mixes of distinct service classes. Furthermore, the impact of CSI imperfection at the transmitter is studied.

The main conclusions of this Master's thesis are summarized in Chapter 6.

## 1.6 Scientific Production

The content and contributions presented in this Master's thesis were submitted with the following information:

- **Roberto P. Antonioli**, Emanuel B. Rodrigues, Tarcisio F. Maciel, Diego A. Sousa and Fco. Rodrigo P. Cavalcanti, "Adaptive Resource Allocation Framework for User Satisfaction Maximization in Multi-Service Wireless Networks". Telecommunication Systems (first round review).

- **Roberto P. Antonioli**, Emanuel B. Rodrigues, Tarcisio F. Maciel, Diego A. Sousa and Fco. Rodrigo P. Cavalcanti, "Alocação de Recursos Adaptativa para Maximização da Satisfação dos Usuários em Redes Celulares". XXXV Brazilian Telecommunications Symposium (SBrT), 2017.

In parallel to the work developed in the Master's program that was initiated on the second semester of 2016, I have been working on other research projects related to analysis and control of trade-offs involving QoS provision. In the context of these projects, I have participated on the following papers and technical reports:

- **Roberto P. Antonioli**, Gabriela C. Parente, Carlos F. M. e Silva, Emanuel B. Rodrigues, Tarcisio F. Maciel and Fco. Rodrigo P. Cavalcanti, "Dual Connectivity for LTE-NR Cellular Networks". XXXV Brazilian Telecommunications Symposium (SBrT), 2017.

- **Roberto P. Antonioli**, Emanuel B. Rodrigues, Tarcisio F. Maciel, Diego A. Sousa and Fco. Rodrigo P. Cavalcanti, "Joint Resource Allocation and Spatial Multiplexing Techniques for Satisfaction Maximization in Multi-Service Cellular Networks", GTEL-UFC-Ericsson UFC.40, Tech. Rep., September 2016, Fourth Technical Report.

- Bruno R. S. Silva, Pedro L. F. Lima, Emanuel B. Rodrigues, **Roberto P. Antonioli**, Diego A. Sousa, F. Hugo C. Neto, Tarcisio F. Maciel and Fco. Rodrigo P. Cavalcanti, "Utility-Based Resource Allocation Framework for QoE/QoS Maximization in OFDMA Networks", GTEL-UFC-Ericsson UFC.40, Tech. Rep., September 2016, Fourth Technical Report.

# 2 SYSTEM MODELING

## 2.1 Introduction

The simulations in this thesis were conducted using a system architecture that follows the 3GPP specifications for the LTE standard. In section 2.2, an overview of the LTE architecture is presented, including a description of the main entities that compose this architecture and the protocol layer functionalities performed by such entities. Then, in section 2.3, we present the main assumptions adopted during the simulations conducted in this master thesis.

## 2.2 General Description of LTE Architecture

The LTE standards have been standardized by 3GPP in the Release 8 with the objective of meeting the increasing performance requirements of mobile broadband services. The efforts in that standardization process resulted in the Evolved Packet System (EPS), which is composed of the core network part, the EPC, and the radio network evolution part, the E-UTRAN. A simplified LTE architecture is illustrated in figure 2.1 showing that the EPC consists of one control-plane node, named MME, and two user-plane nodes, called Packet Data Network Gateway (P-GW) and S-GW. Additionally, it is also illustrated that the LTE radio access network consists of the Base Stations (BSs), also denoted as Evolved Node Bs (eNBs), that are interconnected via the X2 interface[1] and connected to the EPC through S1 interface[2], which is split into S1-MME (for the control channel to the MME) and S1-U (for the data channel to the S-GW).

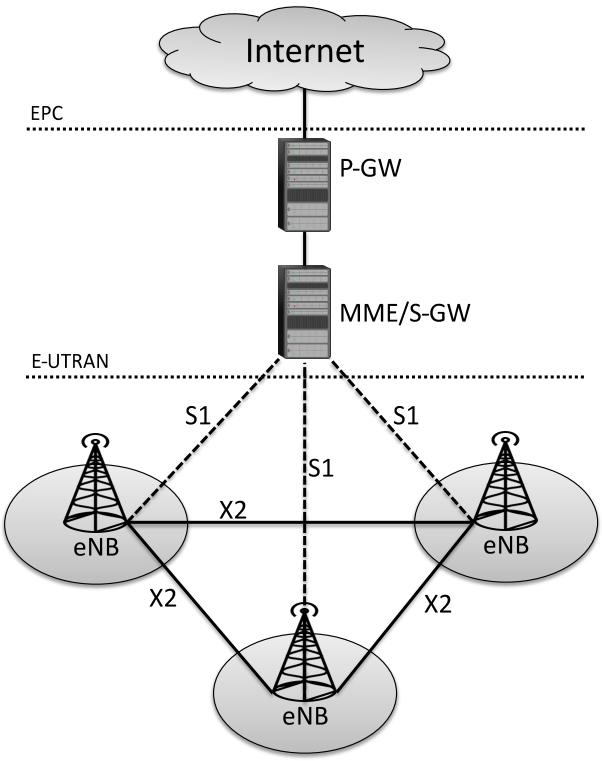
The LTE standard is often referred to as the 4G of mobile networks. However, many people claim that the LTE specified in the 3GPP Release 10, also known as LTE-A, is the true 4G evolution, and the LTE Release 8 is then labeled as ”3.9G”. This comes from the fact that some specifications of the LTE Release 8, such as data rates up 300 Mbits/s in the downlink and 75 Mbits/s in the uplink for user in favorable radio conditions, do not meet the 4G requirements defined by the International Telecommunication Union (ITU), which include data rates up to 1 Gbits/s. Nevertheless, it is worth noting that LTE and LTE-A is the same technology, and that LTE-A is not in any way the final evolution of the LTE standards.

LTE brought advantages for subscribers due to the new applications, such as interactive Television (TV) and user-generated videos, that can be offered by the LTE networks. Furthermore, considering the network operators’ perspective, backward compatibility with legacy networks and simpler architecture are other advantages allowed by the LTE standards. Other benefits provided by this standard are: lower communication latency, higher bandwidth that can

[1] The X2 interface connects two neighboring eNBs in a peer to peer fashion to assist handover and provide a means for rapid co-ordination of radio resources.

[2] The S1 interface is a link between an eNB and an EPC, providing an interconnection point between the E-UTRAN and the EPC. It is also considered as a reference point.

Figure 2.1 – Simplified LTE architecture illustrating the EPC (composed of P-GW, MME and S-GW) and E-UTRAN (comprised by eNBs), which together form the EPS.




Source: Created by the author.

be obtained by means of multi-carrier aggregation, spatial multiplexing on the downlink and uplink allowed by the Multiple Input Multiple Output (MIMO) technology and peak data rates up to 1 Gbps and 500 Mbps in the downlink and uplink, respectively [18].

LTE is a pure packet-based all-Internet Protocol (IP) architecture, which means that the core network is completely packet-switched and based on IP. The main advantage of the all-IP network technology is that it allows operators to offer efficient support to IP-based services with reduced deployment, operational costs and complexity. The LTE networks use the concept of bearer, which is an IP packet flow or logical channel with a defined QoS, to route IP traffic from the Packet Data Network (PDN) (i.e., external IP networks, such as the Internet) to the UE.

As can be seen in figure 2.1, the entry point for the LTE network is the EPC, more specifically the P-GW, which has a direct link to the S-GW/MME. In general, these entities are primarily responsible for the overall control of UEs and, along with the E-UTRAN, for the set up and release of bearers according to the requisitions sent by applications. Specific functions of these EPC entities are [19]:

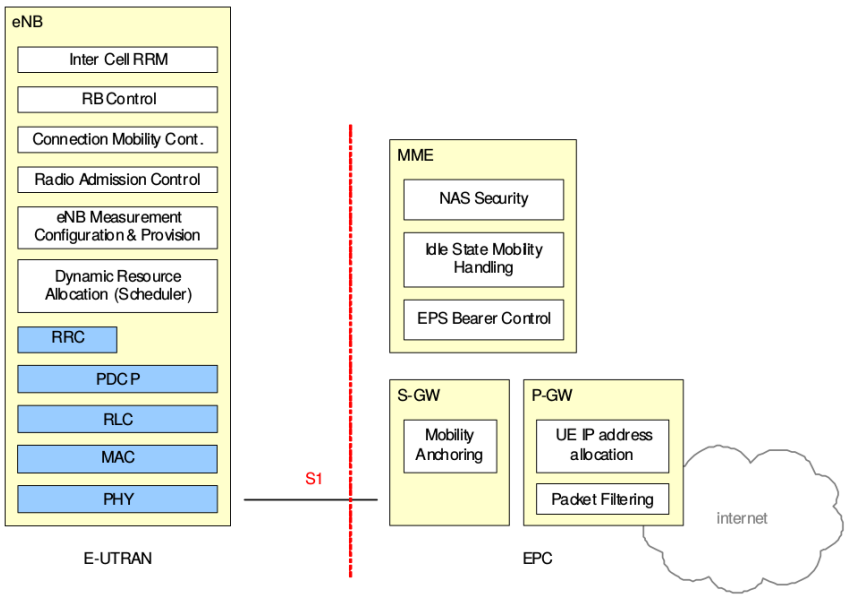
- P-GW: this a user plane node and serves as a default router for UEs connecting to IP networks. Also, it assigns IP addresses to UEs and serves as mobility anchor for non-3GPP access technologies, such as Wireless Fidelity (Wi-Fi).

- MME: this is the control plane node and is responsible for authentication of UEs, radio bearer and user mobility management, and interworking with other 3GPP access systems, such as Global System for Mobile Communications (GSM) and Universal Mobile Telecommunications System (UMTS).
- S-GW: this is the other user plane node and is responsible for managing user data tunnels (i.e., IP packet transfer) between eNBs and the P-GW, as well as acting as local mobility anchor for inter-eNB handover and mobility to other 3GPP access technologies.

The LTE radio access network, known as E-UTRAN, consists of eNBs providing user and control plane protocol terminations towards the UE. In the LTE architecture, the eNBs are responsible for all radio interface related functions, such as: functions for Radio Resource Management (RRM) related to radio admission and connection mobility control, dynamic allocation of radio resources to UEs in both uplink and downlink (scheduling), radio mobility control and radio bearers control, interference management, ciphering, handover management and power control. Also, the eNBs handles the signaling towards MME and S-GW.

In fact, the eNBs provides a user plane termination to the UEs by means of four protocol layers, namely, Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC), Medium Access Control (MAC) and Physical (PHY). The control plane termination is provided by means of the Radio Resource Control (RRC) layer. Figure 2.2 illustrates a summary of all function performed by the entities located at the E-UTRAN and EPC. The protocol layers are also shown in the eNB. Let us now describe the main functionalities performed in each of these protocol layers.

Figure 2.2 – System architecture showing the function split between EPC and E-UTRAN.




Source: [20].

### *2.2.1 Protocol Layer Design*

The RRC protocol layer is responsible for: handling the broadcasting of system information; controlling all procedures related to the establishment, modification and release of RRC connection, including the establishment of control and data radio bearers, handover within LTE and configuration of the lower protocol layers (PDCP, RLC and MAC); controlling the inter-Radio Access Technology (RAT) mobility; measurement configuration and reporting; among others functionalities [21].

The PDCP protocol layer performs ciphering and integrity check of RRC messages in the control plane. Considering the user plane, this layer is responsible for header and payload compression of IP packets. Additionally, this layer supports lossless mobility in case of inter-eNB handovers and provides integrity protection to higher layer-control protocols.

The RLC protocol layer performs the segmentation and reassembly of upper layer packets in order to adapt them to the size that can actually be transmitted over the radio link interface. Another focus of the RLC is on providing lossless transmission of data for radio bearers that require error-free transmission. Additionally, the RLC layer performs reordering to compensate for out-of-order reception due to Hybrid Automatic Repeat Request (HARQ) operation in the layer below [22].

The MAC layer is mainly responsible for handling uplink and downlink scheduling as well as HARQ signaling. This layer performs multiplexing of data from different radio bearers. The MAC layer aims at achieving the negotiated QoS for each radio bearer by deciding the amount of data that can be transmitted from each radio bearer and instruction the RLC to provide packets of a given size. Regarding the uplink, this process involves reporting to the eNB the amount of data at the transmiter buffer [21].

The properties presented by the PHY layer determine the characteristics of a cellular network with respect to peak data rates, latencies and coverage. The LTE PHY supports both Frequency Division Duplex (FDD) and Time Division Duplex (TDD) duplexing schemes. For the downlink, LTE uses the conventional Orthogonal Frequency Division Multiplexing (OFDM) due to robustness to time dispersion of the radio channel. For the uplink, LTE employs a Discrete Fourier Transform (DFT)-spread OFDM (also denoted as Single Carrier - Frequency Division Multiple Access (SC-FDMA)), which provides improved peak-to-average power ratio that enables more power efficient UEs. The PHY layer resources are utilized by physical channels and signals for transmission of data and/or control information from the MAC layer and for supporting physical-layer functionalities, respectively.

The simulation environment adopted in this thesis models the MAC and PHY protocol layers of the LTE standard. Regarding the MAC layer, only the downlink resource allocation is modeled, which is the main focus of this thesis. For the PHY layer, the FDD was adopted as the duplexing scheme and, as the interest here is on the downlink, the OFDMA is employed

as the multiple access technology. In the following, more details are given about the resource allocation architecture and the OFDMA technology.

### *OFDMA*

The 4G cellular networks that follow the LTE standard employ OFDMA in the downlink. The time domain structure of the LTE physical layer is composed of radio frames of 10 ms, where each radio frame is subdivided into 10 subframes of length 1 ms. Looking at the frequency domain, the default subcarrier spacing is 15 kHz and all subcarriers are grouped in sets of 12 subcarriers. These definitions in the time and frequency domains lead us to definition of RB, which is the minimum scheduling unit of the considered simulator, in accordance with the LTE standard. One RB consists of 12 subcarries (e.g., 180 KHz) in the frequency domain and one sub-frame (e.g., 1 ms) in the time domain.

OFDMA is a transmission technology that extends the OFDM technology in order to provide a more flexible access to the RBs [23]. The main benefit brought by OFDMA over OFDM is that the former distributes subcarries to different users at a time, while the latter splits the frequency bandwidth into orthogonal subcarries to transmit data to a single user. Another benefit of OFDMA is the opportunity to take advantage of the frequency, multi-user, time and space diversities. There is a small probability that all frequency resources in a link have the same channel quality because of the frequency diversity. The multi-user diversity relies on the fact that users in different positions within the eNB coverage region experience channel almost independently [24]. The time diversity exists due to the time varying characteristics of the mobile communications channel, where the user speed might be used for estimating the speed the channel state changes. The space diversity, also known as spatial or antenna diversity, is related to the use of two or more antennas to enhance the reliability and quality of the wireless link [15].

A means for taking advantage of all these diversities is the employment of RRA algorithms, also referred as to scheduling algorithms. RRA algorithms are responsible for performing a selection to determine which UEs have access to the system resources and with which configuration. Therefore, RRA algorithms have a significant impact on system performance. Next, more details are given about RRA algorithms.

### *LTE-A MAC resource allocation*

Considering LTE networks, the radio resources are assigned to the UEs by the units of RBs, which is only possible because OFDMA is employed. The resource allocation process takes place in a subframe basis, i.e., it happens every 1 ms and might dynamically change from one millisecond to another. The time duration of 1 ms where the RRA takes is known as Transmission Time Interval (TTI) and has the same duration of one RB. Such a short TTI allows network operators to exploit the channel variations by scheduling UEs depending on their current channel quality.

When a UE wants to receive data from a given application, a connection (or bearer) is established between the UE and the LTE core network (i.e., the EPC). Upon establishing the bearer, a QoS Class Identifier (QCI) is assigned, which specifies whether the bearer is guaranteed bit-rate or not, target delay and loss requirements, for instance. Then, the eNB is able to translate the QCI attributes into requirements for the air interface [24]. The RRA algorithms should take into account these requirements when performing the resource assignment in order to guarantee high UE satisfaction levels.

The MAC scheduler located in the eNB is in charge of assigning both uplink and downlink radio resources. In this thesis, the interest is only on the downlink process. The scheduling decision covers not only the RB assignment but also which modulation and coding scheme are used [22]. Considering the LTE downlink, the available modulation schemes are Quadrature Phase Shift Keying (QPSK), 16-Quadrature Amplitude Modulation (QAM) and 64-QAM.

Let us now explain the scheduling process that happens every TTI in the downlink of LTE networks, which is illustrated in figure 2.3. A packet classifier is responsible for classifying the incoming packets of connected UEs according to their types or QoS attributes, which are used by RRA algorithms for deciding the priority of UEs in a given resource assignment decision. Before the RB assignment process, the eNB's MAC layer first decides which Modulation and Coding Scheme (MCS) the UEs can use according to the Adaptive Modulation and Coding (AMC) strategy implemented by the network operator. The AMC algorithm chooses the MCS based on Channel Quality Indicator (CQI) values reported by the UEs, which is exemplified in figure 2.3 by the Signal to Interference-plus-Noise Ratio (SINR) values. From the chosen MCS, the amount of data (in bits) that can be transmitted on a RB to the UE is decided. Once this decision is finished, the RRA algorithm located at the eNB's MAC layer allocates the appropriate numbers of RB to the UE, according to their QoS metrics.

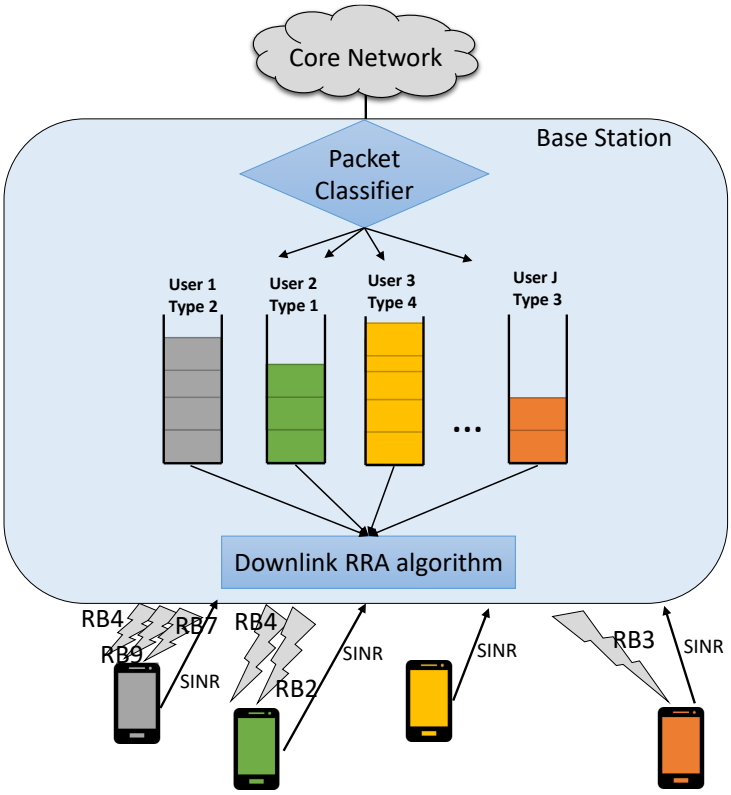
## 2.3 System Layout

We consider the downlink transmission of a single-cell LTE cellular system based on OFDMA, so that we do not need to deal with interference management and only focus on the resource assignment. An eNB located at the center of a three sectored cell serves a set of UEs, represented by $\mathcal{J} = \{1, 2, \ldots, J\}$, which are distributed within its coverage area.

The downlink transmission multiple access scheme is based on OFDMA using a normal cyclic prefix length and considering 14 OFDM symbols per TTI. The minimum allocable radio resource considered herein is referred as a RB, which is a time-frequency chunk comprised of a time slot of 1 ms (TTI) and 12 subcarriers. The total subcarrier bandwidth is 15 kHz, which accounts for both data and pilot symbols. The system disposes of a set of RBs, represented by $\mathcal{K} = \{1, 2, \ldots, K\}$, to be allocated to the UEs.

In this thesis, we do not deal with dynamic power allocations because previous studies

Figure 2.3 – Structure of radio resource allocation algorithm in the downlink of 3GPP LTE systems, which has been implemented in the simulation environment used in this thesis.




Source: Created by the author.

have demonstrated that such techniques do not impact the performance in the considered scenarios, which are comprised of high numbers of users and diversified QoS requirements [25, 26]. Therefore, it is assumed that the total available power $P_t$ of each BS is equally distributed among all RBs during the transmission. Consequently, the power $p_k$ allocated to RB $k$ is $p_k = \frac{P_t}{K}$.

A downlink Single Input Single Output (SISO) channel is considered, i.e., both eNB and UE are equipped with single antennas. The channel transfer function between the user $j$ on RB $k$ and the eNB is represented by $h_{j,k}$, which is considered to be the transfer function of the mid sub-carrier that composes the resource block. The channel transfer function is calculated taking into account the main propagation characteristics of the wireless channel, namely path loss, shadowing (slow fading) and small-scale fading (fast fading).

The channel gains are constant over a TTI, but might vary from one TTI to another. The UEs are uniformly deployed within the eNB coverage area and have no mobility; therefore, in order to capture the system performance in different coverage situations, several independent snapshots considering different user distributions are taken into account during the simulations.

It is worth to note that although the study performed in this thesis considered an OFDMA-based system, the same analysis could be conducted for any multiple access scheme that guar-

antees orthogonality among the resources.

### *Propagation Modeling*

The radio channel model takes into account the effects that traditionally have a significant impact over the signal power received by the mobile station, such as path-loss, shadowing and fast fading. Furthermore, the eNB antenna radiation pattern plays an important role on the calculation of the total propagation gain.

#### *Path Loss*

In this thesis, we analyze two path loss models for a carrier frequency of 2 GHz. The first one is a less severe macro cell path loss model based on the propagation model presented in [27], and is given (in dB) by

$$PL(d) = 15.3 + 37.6 \cdot \log_{10}(d), \tag{2.1}$$

where $d$ is the distance between UE and eNB, and is given in meters.

The second model is more severe (worse coverage) if compared to the first model. It is a modified COST231 Hata urban macro propagation model presented in [28] and is given (in dB) by

$$PL(d) = 34.5 + 35 \cdot \log_{10}(d), \tag{2.2}$$

where $d$ is also expressed in meters.

#### *Shadowing*

The shadowing (slow fading) is modeled as a log-normal random variable, with mean equal to zero and standard deviation $\sigma_{\text{sh}} = 8$ dB [29]. In this thesis, no spatial correlation for shadowing is considered.

#### *Fast Fading*

The fast fading follows a time-and-frequency-correlated Rayleigh distribution taking into account the power delay profile of the Typical Urban (TU) channel from [30]. A fading map is generated for the simulator using the well-known Jakes' model, which is a model for Rayleigh fading based on summing sinusoids, and the Tapped Delay Line (TDL) model. The samples are taken from the map by choosing a random initial position (in time and frequency axes) that is unique for a UE-eNB pair. Then, at each TTI, an offset is chosen according to the current time. This guarantees a degree of decorrelation between the fading samples of different links. The attributed offset depends on the relation between the Doppler spread used in the generation of the map and the one that corresponds to the UE speed. It becomes an integer value that is used to jump in samples inside the map.

*Antenna Gain*

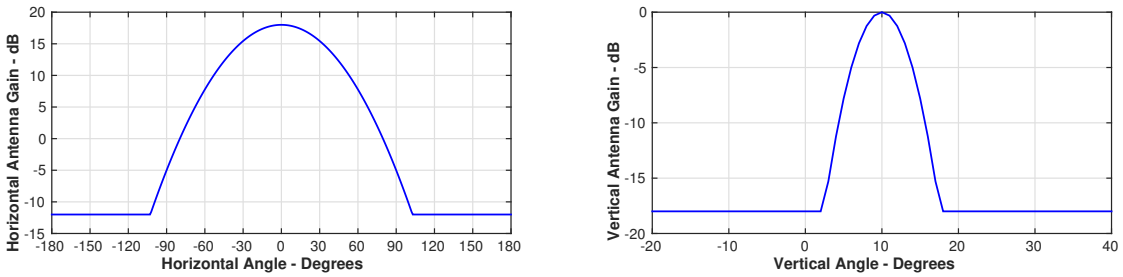
The eNB antenna radiation pattern adopted in this thesis is an extension of the model proposed in [31], where only the horizontal angle was considered. Here, we take into account a model with vertical and horizontal antenna patterns, as proposed in [32]. In equation (2.3), the horizontal and vertical components of this radiation pattern are presented

$$G(\theta,\phi) = G_h(\theta) + G_v(\phi) = -\min\left[12\left(\frac{\theta}{65}\right)^2, 30\right] + 18 + \max\left[-12\left(\frac{\phi-\phi^{\text{tilt}}}{6.2}\right)^2, -18\right], \quad (2.3)$$

where $\theta$ is the horizontal angle relative to the main beam positioning direction, $\phi$ is the negative elevation angle relative to the horizontal plane and $\phi^{\text{tilt}}$ is the downtilt angle, $G_h(\theta)$ is the horizontal gain component, $G_v(\phi)$ is the vertical gain component and $G(\theta,\phi)$ is the final antenna gain. It is shown in figure 2.4 the horizontal and vertical gain for $\phi^{\text{tilt}} = 10°$.

Figure 2.4 – Horizontal and vertical gains for antenna radiation pattern.

(a) Horizontal Gain. (b) Vertical Gain.




Source: Created by the author, adapted from [32].

***Link Adaptation***

Depending on the current channel conditions, an appropriate number of bits might be transmitted on each RB. This is accomplished by the AMC or link adaptation procedure, which adjust the transmission parameters according to the current users' channel conditions.

Considering the downlink of LTE mobile networks, the UEs transmit a CQI to the eNB, which in response chooses the best MCS to be employed in the downlink transmission. The table 2.1 shows the mapping between CQI and MCS in the LTE standard. Notice that the higher the CQI (better channel conditions), the higher is the amount of bits per symbol that can be transmitted.

Different values of MCS causes differences in the BLock Error Rate (BLER) performance, which are shown in figure 2.5. In this figure, the relationship between the Signal to Noise Ratio (SNR), BLER and MCS is presented. Considering the same SNR value, one can see that a

Table 2.1 – Mapping between CQI and MCS in the LTE standard.

| **CQI** | **Modulation** | **Code Rate** [$\times 1024$] | **Rate** [Bits/symbol] |
|---|---|---|---|
| 0 | Out of range | | |
| 1 | QPSK | 78 | 0.1523 |
| 2 | QPSK | 120 | 0.2344 |
| 3 | QPSK | 193 | 0.3770 |
| 4 | QPSK | 308 | 0.6016 |
| 5 | QPSK | 449 | 0.8770 |
| 6 | QPSK | 602 | 1.1758 |
| 7 | 16-QAM | 378 | 1.4766 |
| 8 | 16-QAM | 490 | 1.9141 |
| 9 | 16-QAM | 616 | 2.4062 |
| 10 | 64-QAM | 466 | 2.7305 |
| 11 | 64-QAM | 567 | 3.3223 |
| 12 | 64-QAM | 666 | 3.9023 |
| 13 | 64-QAM | 772 | 4.5234 |
| 14 | 64-QAM | 873 | 5.1152 |
| 15 | 64-QAM | 948 | 5.5547 |

Source: Created by the author, adapted from [33].

higher MCS index results in a higher BLER, meaning that for operating with an acceptable low BLER, a given MCS requires a minimum SNR value [31].

The SNR value for a UE $j$ in the RB $k$ is given by

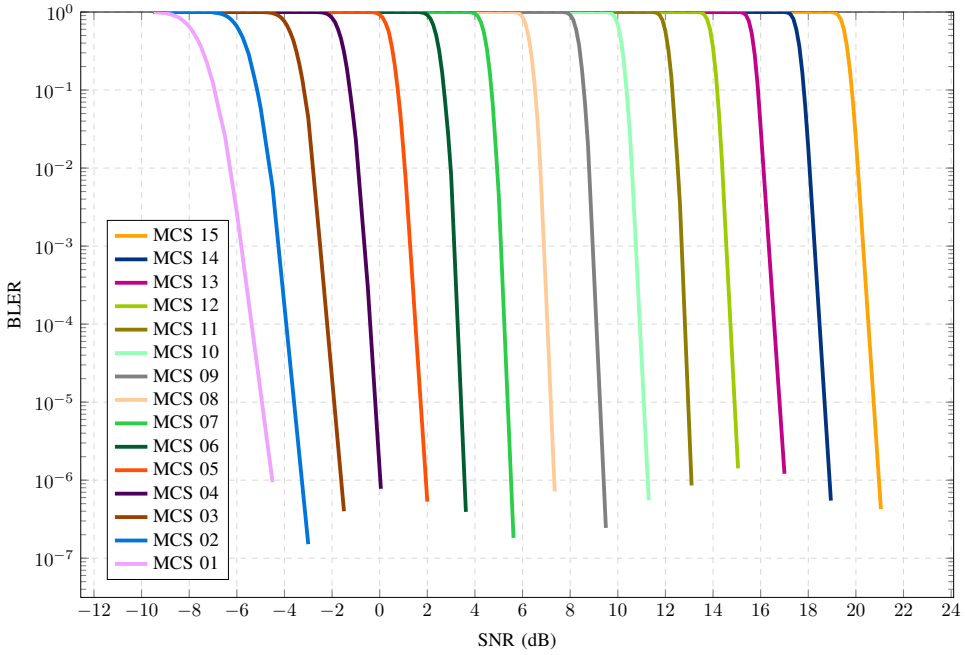
$$\gamma_{j,k} = \frac{p_k \left|h_{j,k}\right|^2}{\sigma^2}, \tag{2.4}$$

where $\sigma^2$ denotes the average Additive White Gaussian Noise (AWGN) power in the frequency band of a RB.

We consider that the eNB employs a link adaptation mechanism that allows different transmission rates depending on the $\gamma_{j,k}$ of the UE $j$ on RB $k$. Given the $\gamma_{j,k}$ value, the eNB selects from a set of 15 MCSs (shown in table 2.1), the one that provides the highest transmit data rate and has an estimated BLER lower than a given threshold. Therefore, the rate allocated by the eNB to the user $j$ on RB $k$ is

$$r_{j,k} = f(\gamma_{j,k}), \tag{2.5}$$

Figure 2.5 – Relationship between SNR, BLER and MCS in the LTE standard.




Source: Created by the author, adapted from [34].

where $f(\cdot)$ represents a link adaptation function. The total rate allocated to UE $j$ is given by

$$R_j = \sum_{k \in \mathcal{K}_j} r_{j,k}, \tag{2.6}$$

where $\mathcal{K}_j \subset \mathcal{K}$ is the subset of RBs assigned to UE $j$.

### *CSI Imperfection*

A noise term and a delay component in the CSI estimation are introduced to characterize imperfections in the estimation performed at the transmitter, i.e., at the BS. The channel is estimated by the UE by means of pilot symbols transmitted by the BS. Imperfect estimated channels can be modeled as described in [35] by

$$\hat{h}_{j,k}[n] = \sqrt{(1-\psi)} h_{j,k}[n] + \sqrt{\psi}\eta[n], \tag{2.7}$$

where: $\psi \in (0, 1)$ represents the degradation of the channel estimation; and $\eta[n] \in \mathbb{C}$ represents a channel estimation error, which is modeled as a Zero Mean Circularly Symmetric Complex Gaussian (ZMCSCG) random variable, with $\mathbb{E}\{|\eta[n]|^2\} = \mathbb{E}\{|h_{j,k}[n]|^2\}$. Additionally, the channel estimations reported by UEs to their BS in order to be used by the RRA algorithms are outdated by $\Delta n$ TTIs, as expressed in the following:

$$\hat{h}_{u,k}[n] = \hat{h}_{u,k}[n - \Delta n]. \tag{2.8}$$

## 2.4 Traffic Models

The multi-service scenarios examined in this thesis are comprised of mixes between two of the following services: CBR services, which are throughput-based; VoIP services, which are delay-based; and video services, which are queue-based. In this section, an explanation is presented describing how these services are modeled.

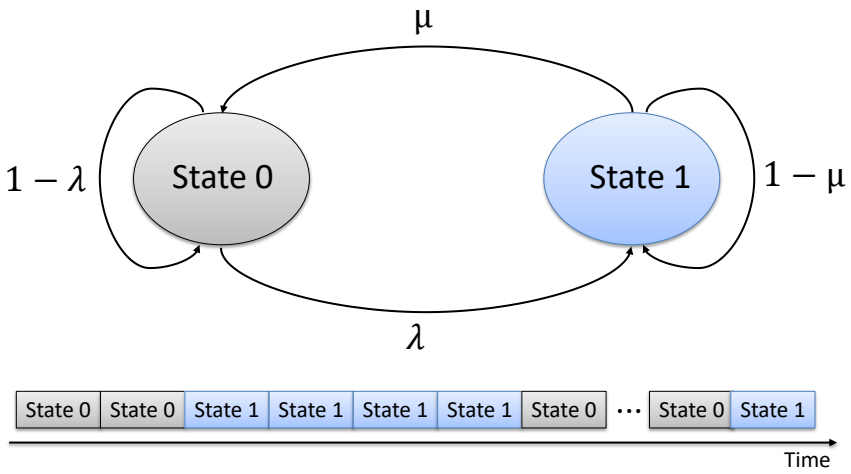
### *2.4.1 CBR*

The CBR traffic model is used for emulating throughput-based services. Some examples of real applications that follow this model are audio and video transmission (Television, pay-per-view) that have a fixed rate of data transmission.

Since we assume that this service is throughput-based, it is delay tolerant. This means that, although this traffic has a target delay of 200 ms, delayed packets are not discarded, which increases the buffer size [23].

The CBR flows have been modeled using two-state (ON/OFF) Markov chains, where the ON and OFF periods are calculated by using an exponential distribution with mean of 1 second. A two-state Markov chain is illustrated in figure 2.6, where $\mu$ and $\lambda$ are the state transitions probabilities. When the flow is active (ON state), fixed-size packets are generated at a fixed inter-arrival time.

Figure 2.6 – Two-state (ON/OFF) Markov chain used for traffic modeling.




Source: Created by the author.

Since the CBR service is a throughput-based service, CBR users are considered satisfied if their individual session throughput is equal or higher than a threshold ($T_j \geq \Phi_{\text{req}}^{\text{thr}}$), where the session time is related to the duration of each independent simulation. The simulation parameters used for the CBR traffic model during the simulations are summarized in table 2.2.

Table 2.2 – Parameters used for CBR traffic model.

| **Parameter** | **Value** |
|---|---|
| Packet generation rate | 512 kbps |
| Packet size | 2048 bits |
| Packet inter-arrival time | 4 ms |
| Session duration | Simulation time |
| Average duration of each Markov chain state | 1 second |
| Activity factor | 50% |
| Target delay | 200 ms |
| Throughput Requirement ($\Phi_{\text{req}}^{\text{thr}}$) | 512 kbps |

Source: Created by the author.

### *2.4.2 VoIP*

During the 1st Generation (1G) of cellular networks, the circuit-switched voice services were the dominant traffic source. In this thesis, we consider an evolution of this service, which is offered by the 4G cellular networks and is known as VoIP.

The VoIP service is characterized by the routing of packets over the Internet or any other IP-based computer network, which transforms the transmission of human voice in a service supported by data networks. The VoIP service is delay sensitive, that is why we consider it as a delay-based service.

The VoIP service has also been modeled using a two-state (ON/OFF) Markov chains (as illustrated in figure 2.6), where the ON and OFF periods are also calculated using an exponential distribution with mean of 1 second. For this traffic model, the active state emulates a talk spurt, when packets are generated at a constant rate and arrive at a fixed inter-arrival time.

Since the VoIP service is delay sensitive, a given packet that is not transmitted within a predetermined limit is discarded. More precisely, if the Head Of Line (HOL) packet at the eNB transmitter buffer is not transmitted within the HOL packet delay requirement, it is discarded. Besides that, packets might be discarded due to channel errors [9]. The amount of discarded packets is used for calculating the FER for a given user $j$ as follows

$$\text{FER}_j[\text{n}] = \frac{\kappa_j^{\text{disc}}[\text{n}]}{\kappa_j^{\text{disc}}[\text{n}] + \kappa_j^{\text{success}}[\text{n}]}, \tag{2.9}$$

where $\kappa_j^{\text{disc}}[\text{n}]$ and $\kappa_j^{\text{success}}[\text{n}]$ represent the amount of discarded and successfully transmitted packets, respectively, of a given user $j$ from the session beginning up to TTI $n$. Then, a user $j$ from a delay-based service is considered to be satisfied when its FER is equal or lower than a given requirement ($\text{FER}_j \leq \text{FER}_{\text{req}}$). In table 2.3, the simulation parameters used for the VoIP traffic model are summarized, which are based on the VoIP G.729 encoder.

Table 2.3 – Parameters used for VoIP traffic model.

| **Parameters** | **Value** |
|---|---|
| Packet generation rate | 16 kbps |
| Packet size | 320 bits |
| Packet inter-arrival time | 20 ms |
| Call duration | Simulation time |
| Average duration of each Markov chain state | 1 second |
| Activity factor | 50% |
| FER requirement ($FER_{req}$) | 1% |
| HOL packet delay requirement ($\Phi_{req}^{delay}$) | 20 ms |

Source: Created by the author.

### *2.4.3 Video*

The video traffic model considered in this work is a queue-based video streaming service that generates packets of variable size. This variability in the packet size attempts to emulate different scenes in the video since fast scenes produces bigger frames and slow scenes generate smaller frames. We use a streaming video traffic model proposed by 3rd Generation Partnership Project 2 (3GPP2) in [36], which has been used for other works in the literature as in [37], and more recently in [38].

The source data rate proposed originally by 3GPP2, which was of 32 Kbps, is too low for present-day videos watched in recent cellular networks. Therefore, it has been adopted that the video flow is encoded at the rate of 242 Kbps, which is based on realistic video trace files that use the H.264 encoder [23]. The video traffic parameters to generate packets at 242 Kbps are detailed in table 2.4, where the distributions that control each information type are described. In figure 2.7, the packet generation of a video flow based on 30 FPS is illustrated.

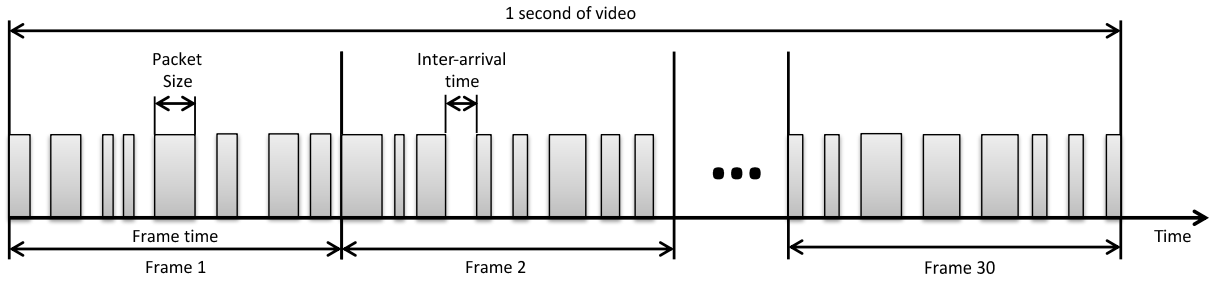
The threshold for the HOL packet delay for the video service is set to be 50 ms [39]. The FER threshold for this service is 2%. The transmitter buffer at the BS is assumed to be

Table 2.4 – Summary of parameters for packet generation of video traffic model.

| **Information Types** | **Distribution** | **Distribution Parameters** |
|---|---|---|
| Inter-arrival time between frames | Deterministic | 33.30 ms (based on 30 FPS) |
| Number of packets per frame | Deterministic | 8 |
| Packet Size | Truncated Pareto (Mean = 126 bytes, Max = 200 bytes) | K = 85 bytes, $\alpha$=1.2 |
| Inter-arrival time between packets | Truncated Pareto (Mean = 2 ms, Max = 4 ms) | K = 1.20 ms, $\alpha$=1.2 |

Source: Created by the author.

Figure 2.7 – Illustration of packet generation for video traffic model based on 30 FPS.




Source: Created by the author.

infinite, resulting in packets being discarded only when they exceed the delay threshold, not due to buffer overflow. For simulation purposes, it is assumed that the de-jitter buffer at the end user is initially full of window size of video streaming service of 2 seconds; this buffer makes the application resilient against latency and jitter [40]. The video play-out rate is equal to source data rate, i.e., packets are consumed from the user-end buffer at 242 Kbps.

Since the video service is a queue-based service, video users are considered satisfied if their individual session throughput is equal or higher than a threshold ($T_j \geq \Phi_{\text{req}}^{\text{thr}}$) and their FER is equal or lower the FER threshold ($\text{FER}_j \leq \text{FER}_{\text{req}}$). Table 2.5 summarizes the video traffic parameters.

Table 2.5 – Parameters used for video traffic model.

| **Parameters** | **Value** |
|---|---|
| Packet generation rate | 242 kbps |
| Packet size | Variable (see table 2.4) |
| Packet inter-arrival time | Variable (see table 2.4) |
| Video duration | Simulation time |
| FER requirement ($\text{FER}_{\text{req}}$) | 2% |
| HOL packet delay requirement ($\Phi_{\text{req}}^{\text{delay}}$) | 50 ms |
| Throughput Requirement ($\Phi_{\text{req}}^{\text{thr}}$) | 242 kbps |

Source: Created by the author.

## 2.5 Performance Metrics

This section presents the performance metrics used for evaluating and comparing the proposed resource allocation algorithm and others found in the literature.

### 2.5.1 User Satisfaction

In terms of satisfaction, the algorithms are evaluated considering the percentage of satisfied users given by

$$\Upsilon[n] = \frac{J^{\text{sat}}[n]}{J}, \tag{2.10}$$

where $J^{\text{sat}}[n]$ is the number of satisfied users in TTI $n$ and $J$ is the total number of users in the system. The definition of satisfaction depends on the user service class, which have been described in the previous section. Since we are analyzing multi-service scenarios, there is one satisfaction value for each service. Thus, the total set of users is split according to their service and one satisfaction index is calculated for each subset.

### 2.5.2 Fairness

In terms of fairness in the resource allocation process, the algorithms are evaluated by means of the well-known Jain's index [41]. For a generic QoS metric $\mathbf{x} = [x_1, \cdots, x_j, \cdots, x_J]$, the Jain's fairness index can be calculated using the expression

$$F(\mathbf{x}) = \frac{\left(\sum_{j=1}^{J} x_j\right)^2}{J \cdot \sum_{j=1}^{J} x_j^2}. \tag{2.11}$$

The Jain's fairness index is independent of the QoS metric being considered and is bounded between 0 and 1 (0% and 100%). When all $x_j$'s are equal, it means that a totally fair allocation was achieved and the fairness index is equal to 1. On the other hand, when all resources are given to only one user, a totally unfair allocation happens and the fairness index is equal to $1/J$, which is 0 in the limit as $J \to \infty$. For throughput- and queue-based services, $x_j$ is given by the throughput $T_j[n]$ normalized by the throughput requirement, while for delay-based services, $x_j$ corresponds to the (1-$\text{FER}_j$), which is a consequence of exceeding the HOL packet requirement. Also for the fairness index, one value is calculated for each service class in the systems.

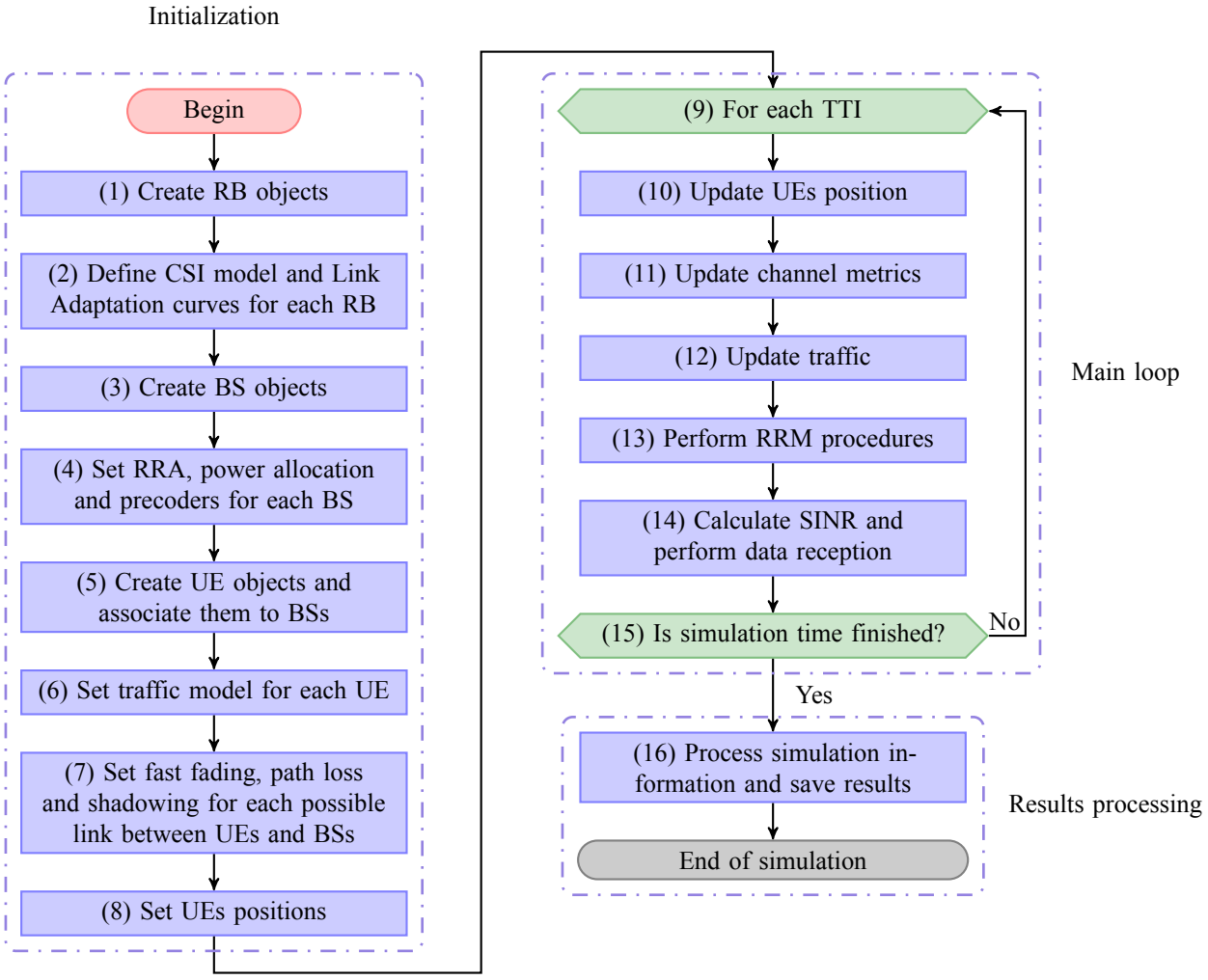
### 2.5.3 Total Cell Throughput

The total cell throughput accounts for the summation of the total individual users' throughput at the simulation end. This performance indicator along with the satisfaction index allows us to analyze the trade-off between the utilization of radio resources in terms of data rate and how well the users' QoS requirements were satisfied.

## 2.6 Simulator Flowchart

In figure 2.8, we present a general simulation flowchart where we can clearly see the sequence of functions that are executed during simulation. Some steps of this general flowchart might change depending on the scenario we desire to simulate, as will be explained in the following.

Figure 2.8 – General simulation flowchart.




Source: Created by the author.

Basically, the simulator is composed of three main parts: **initialization**, **main loop** and **results processing**. In the **initialization** part, blocks 1 to 8, all the simulation objects and data structures that are needed by the simulation are created. Let us now describe what is performed in each of these steps:

- **Create RBs objects and define their CSI model and link adaptation curves (blocks 1 and 2)**: The first step is to create a list of RBs and set their channel characteristics and link adaptation models.
- **Create BS objects (block 3)**: The LTE BSs are created and their main characteristics are defined, such as: height, transmission power, antenna model.
- **Set RRA, power allocation and precoders for each BS (block 4)**: After positioning the BSs, we define the RRA and power allocation techniques that will be performed by the BSs, as well as the precoders used by the BSs. In this thesis, the power allocation technique used is always the Equal Power Allocation (EPA) technique.

- **Create UE objects and associate them to BSs (block 5)**: Each BS is initialized with a pre-defined number of UEs in its coverage area. As the UEs are static, the number of UEs connected to each BSs does not change during the simulation.
- **Set traffic model for each UE (block 6)**: As the simulator is equipped with a variety of traffic models, we need to define which service type the UEs are consuming. This choice directly impacts the performance of scheduling algorithms since different traffics have distinct QoS requirements.
- **Set fast fading, path loss and shadowing (block 7)**: The path-loss model and shadowing are defined in this step. The channel fading maps are built based on the selected model.
- **Set UEs positions (block 8)**: The final step in the initialization is to effectively position the UEs in the system. The UEs are uniformly distributed in the coverage area of the BSs to which they were initially associated.

The core of the simulator is within the **main loop**, blocks 9 to 15. Basically, in this part the system evolves according to a predefined time step (the duration of a TTI of the LTE standard). The simulator runs based on the shortest time step, which in our case will be the time step from the NR RAT. In each iteration, some tasks are performed:

- **Update UEs position (block 10)**: As mentioned before, the user are static, meaning that they have no mobility. Thus, for the studies of this thesis, the user mobility option was deactivated setting the UE speed to zero.
- **Update channel metrics (block 11)**: The metrics related to the link between BSs and UEs are updated.
- **Update traffic (block 12)**: There are different service types with distinct characteristics regarding packet size and packet generation frequency. Therefore, at each iteration the state of traffic objects are checked in order to evaluate if new packets should be generated.
- **RRM procedure (block 13)**: Once channel state and traffic are updated, RRM procedure are performed. RRM procedures could exploit the available information about channel and traffic states as well as QoS/QoE and satisfaction requirements.
- **Calculation of SINR and data reception (block 14)**: When the system resources are assigned to the selected users, data reception should be performed in order to evaluate whether data packets were successfully received or not. To do that, the receive SINR should be calculated and passed to a link-to-system mapping function in order to estimate packet error probabilities, where the link adaptation curve was defined in block (2). During the reception, the transmitter buffer of each BS associated with each connected user should be updated according to the amount of data that was correctly received during the reception.

Finally, in the **result processing**, block 16, the simulation objects are processed in order to obtain statistics that will be useful to analyze the system performance. Through the analysis of these statistics, the system performance can be optimized by the proposal of new RRM strategies.

# 3 RELATED WORK

## 3.1 Introduction

This chapter presents a survey of works found in the literature related to the study area of RRA in the downlink of OFDMA cellular networks. Specifically, in section 3.2, we present a broad and general survey of related works. Then, in section 3.3, we describe in more details the benchmark algorithms used for performance comparison in this thesis.

## 3.2 Related Work

There have been a variety of research works focusing on RRA in single and multi-service scenarios for the downlink of OFDMA-based cellular networks.

Opportunistic algorithms have been widely studied in the literature. In [26], the well-known Rate Maximization (RM) algorithm was analyzed, which has the main objective of maximizing the system total throughput by assigning RBs to UEs experiencing the best channel quality on each RB. Considering this approach, the RM algorithm assigns UEs that are able to transmit the highest amount of data on each RB, thus maximizing the system capacity. The Proportional Fair (PF) algorithm is another well-known opportunistic algorithm, which is appropriate for scheduling throughput-based services [42] and guarantees the fairness between users by allocating the RBs according to a ratio between their current throughput and achievable data rate on a given RB. However, the PF algorithm is not efficient in scenarios with delay sensitive applications because it does not consider the packet delay during the resource allocation [43, 8].

Opportunistic algorithms that can be applied in scenarios with delay sensitive applications have also been proposed in the literature. The Modified Largest Weighted Delay First (MLWDF) scheme considers both channel and queue state aiming to keep the HOL packet delay of most users below a given requirement [44]. This algorithm supports flows with different QoS requirements by prioritizing one service over the other according to the maximum tolerable packet delay. The MLWDF algorithm was extended in [23] and was named Queue-HOL-MLWDF (QHMLWDF), which allocates the resources taking into account the queue size (in bits) and the HOL packet delay. The QHMLWDF algorithm is targeted for multi-service scenarios and has been analyzed in [23] for scenarios composed of video, VoIP and CBR flows. Another example is the Exponential/PF (EXP/PF) algorithm. The authors in [37] extended the classical Exponential (EXP) rule [45] and PF algorithms, proposing the EXP/PF algorithm, which is a resource allocation scheme that schedules delay sensitive services according to the EXP rule and throughput-based users based on the PF scheduler.

The opportunistic RRA algorithms are usually more interested in maximizing the overall

system throughput then in satisfying the individual users' QoS requirements. On the other hand, the QoS-based algorithm consider users' QoS metrics during the resource allocation aiming at meeting the users' QoS requirements. Notice that some opportunistic algorithms take into account QoS metrics for performing the resource allocation, such as the MLWDF e EXP rule. However, their main focus is not on meeting QoS requirements.

Considering single service scenarios, [46] and [47] propose, respectively, the Satisfaction-Oriented Resource Allocation for Real Time Services (SORA-RT) and Satisfaction-Oriented Resource Allocation for Non-Real Time Services (SORA-NRT), which are heuristic resource allocation algorithms that have the objective of maximizing the user satisfaction in scenarios composed of only delay sensitive (RT) or throughput-based (NRT) services, respectively. The single service cases analysis described in [46] and [47] were further evaluated in [24], when the Capacity-driven Resource Allocation (CRA) was proposed, which dynamically controls the resource sharing among flows of different services (delay sensitive and throughput-based) and exploits channel-quality knowledge in order to provide gains in the joint system capacity in multi-service scenarios. The authors in [48] propose a scheduling algorithm that has the objective of minimizing the Packet Loss Ratio (PLR) of delay sensitive services while guaranteeing the QoS requirements of Guaranteed Bit Rate (GBR) services, where the authors show that their algorithm can achieve a good trade-off between satisfaction of QoS requirement and fairness while decreasing the PLR. In [5], a well-known bipartite matching algorithm and a standard gradient scheduling algorithm are employed in order to satisfy the QoS requirements of delay sensitive services and for maximizing the utility of throughput-based services, respectively.

The Utility Theory has also been widely applied for designing RRA algorithms. A utility-based algorithm for scenarios composed of only delay sensitive application was proposed in [6], where the authors use some utility functions that were specially conceived for the scenarios studied therein. In [7], the authors propose a utility-based algorithm that aims at guaranteeing PLR and the play-out outage rate of video services, where a barrier function was applied as a utility function. Another utility-based RRA algorithm is the Urgency and Efficiency-based Packet Scheduling (UEPS) [8], which uses a delay-based utility function as an urgency factor and the user channel state as an efficiency indicator of radio resource usage during the resource allocation. In [9], the authors propose a unified utility-based framework, from which they derive two resource allocation policies named Delay-based Satisfaction Maximization (DSM) and Throughput-based Satisfaction Maximization (TSM) to be employed in single service scenarios composed of only delay sensitive or throughput based services, respectively. These two policies use sigmoid function to maximize the user satisfaction in their respective scenarios. The work proposed in [9] is the basis of this thesis proposal; herein, we propose to extent the single-service cases analyzed in [9] for multi-service scenarios. A resource allocation scheme based on game theory that uses a sigmoid utility function and a delay based scheduler for mixed traffic is proposed in [49]. In [10, 11], a utility-based scheduling scheme, named Max-Delay-Utility (MDU), was proposed using two different step-like utility functions to define the users' priority during

allocation. A packet scheduling algorithm based on CSI and utility functions of HOL packet delay is proposed in [12], which uses a z-shaped function for scheduling RT services and an exponential function for allocating resources to NRT services.

The authors in [50] propose a low-complexity resource and power allocation algorithm that targets the maximization of total capacity of delay-tolerant and delay-sensitive users in two-tier networks comprised of frequency-sharing femtocells and macrocells. In [51], the authors propose a fairness-based resource allocation algorithm that aims at guaranteeing a fairly resource allocation among all services in the system. In [52], a resource allocation scheme for cognitive femtocells is proposed, aiming at maximizing the total capacity of all femtocells and achieving fairness among them. The authors in [53] propose two resource allocation algorithms to maximize the total rate while guaranteeing fairness among the services in OFDMA based systems. In [54], the authors derive a near optimal bargaining resource allocation strategy based on a cooperative bargaining game theory and considering fairness in the resource allocation as well as interference mitigation.

In the light of this survey of related work form the literature, this thesis proposes a RRA algorithm that targets the user satisfaction maximization in multi-service cellular networks. Besides being normalized and unified across all service classes, the proposed algorithm employs an innovative service prioritization that is adapted to meet the satisfaction target of the most prioritized service. This specific feature allows network operators to flexibly define their strategy.

## 3.3 Benchmarking Algorithms

This section presents the specific RRA algorithms that have been used for performance comparison against the proposal from this thesis. Together with [9], the works in [10, 11, 12] were also used for designing this thesis proposal. Let us now describe in more details the benchmark algorithms.

### *Max-Delay-Utility (MDU)*

The MDU algorithm proposed in [10] and analyzed in details in [11]. The MDU algorithm is based on the Utility Theory and uses the users' channel information, transmit buffer size of each user at the BS and the packet generation rate to determine the resource allocation.

In [10], the author conducts a rigorous mathematical demonstration based on a utility-based optimization problem and derives the resource allocation policy used by the MDU algorithm. The policy employed by this algorithm selects the user $j^{\star}$ to transmit on RB $k$ in TTI $n$ according to

$$j^{\star} = \arg\max_{j} \left\{ \frac{|U^{'}(w_j[n])|}{\lambda_j} \cdot r_{j,k}[n] \right\}, \tag{3.1}$$

where $r_{j,k}[n]$ denotes the instantaneous achievable transmission rate of user $j$ with respect to RB $k$ at TTI $n$ and $U^{'}(\mathrm{w}_j[n])$ denotes the first derivative of a utility function $U(\mathrm{w}_j[n])$ with

respect to $\mathrm{w}_j[n]$. The term $\mathrm{w}_j[n]$ is defined as the average waiting time of user $j$ at TTI $n$ and is calculated using the following equation [10]

$$w_j[n] = \frac{\overline{Q}_j\,[n]}{\lambda_j}, \tag{3.2}$$

where $\overline{Q}_j\,[n]$ is the transmit buffer average size of user $j$ and $\lambda_j$ is the packet generation rate of user $j$.

In [11], the authors present the utility function based on the QoS requirements of VoIP, streaming and best effort services. Considering the VoIP service, the utility function adopted was:

$$\left|U_V^{'}(\mathrm{w})\right| = \begin{cases} \mathrm{w}[n], & \text{if } \mathrm{w}[n] \leq 25 \text{ ms} \\ \mathrm{w}[n]^{1,5} - 25^{1,5} + 25, & \text{if } \mathrm{w}[n] > 25 \text{ ms.} \end{cases} \tag{3.3}$$

The value 25 ms refers to one quarter of the considered end-to-end delay, which was 100 ms. For the streaming service, the utility function was defined as:

$$\left|U_S^{'}(\mathrm{w})\right| = \begin{cases} \mathrm{w}[n]^{0,6}, & \text{if } \mathrm{w}[n] \leq 100 \text{ ms} \\ \mathrm{w}[n] - 100 + 100^{0,6}, & \text{if } \mathrm{w}[n] > 100 \text{ ms.} \end{cases} \tag{3.4}$$

The value 100 ms refers to one quarter of the considered end-to-end delay for streaming services, which is between 150 ms and 400 ms. Finally, for the best effort service (such as the CBR service considered in this thesis), the chosen utility function was:

$$\left|U_B^{'}(\mathrm{w})\right| = \begin{cases} \mathrm{w}[n]^{0,5}, & \text{if } \mathrm{w}[n] \leq 100 \text{ ms} \\ 100^{0,5}, & \text{if } \mathrm{w}[n] > 100 \text{ ms.} \end{cases} \tag{3.5}$$

In [11], the authors did not explain the reason behind the choice of the 100 ms value. However, they mentioned that considering (3.5) and 100 ms, the MDU algorithm behaves as the PF algorithm.

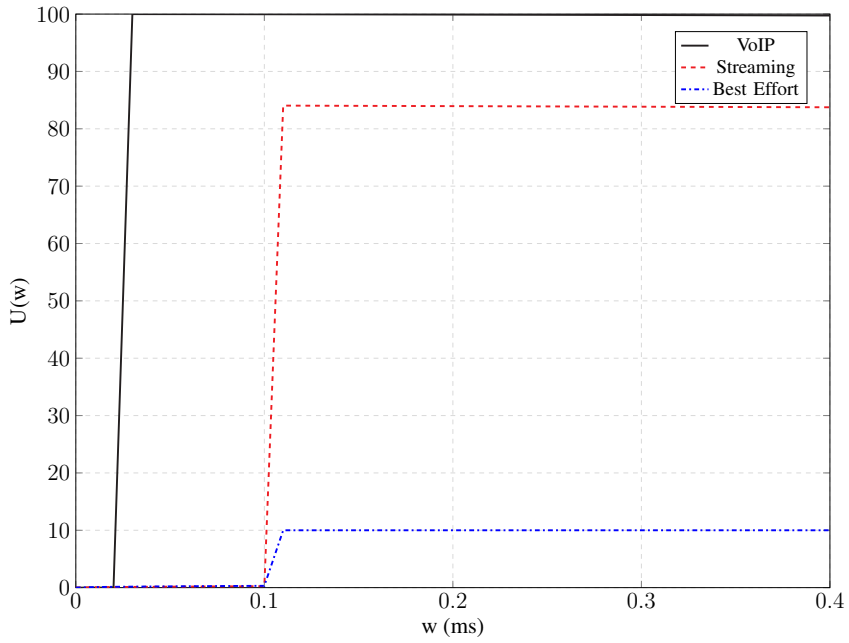
In figure 3.1, the shape of the utility function defined in (3.3), (3.4) and (3.5). Notice that the VoIP service has always the highest priority during the resource allocation process, followed by the streaming service and at last the best effort service. This approach is commonly adopted in the literature, where the VoIP service has the highest priority among all the services and the best effort has the lowest priority, receiving resources only when no other service is in need.

The authors of MDU compared its performance with the classical PF, EXP and MLWDF algorithms, showing that the MDU algorithm presents a better performance in terms of users' throughput and packet delay.

***Algorithm proposed by Lei et al. [12]***

Another utility-based algorithm was proposed in [12]. Since only the proposed algorithm in [12] was based on the utility theory, the authors just referred to the algorithm as Utility in the

Figure 3.1 – Utility functions employed by the MDU algorithm for VoIP, streaming (such as video) and best effort (such as the CBR considered in this thesis) services.




Source: Created by the author, adapted from [11].

paper. As in this thesis more algorithms are based on the utility theory, we refer to this algorithm as ”Lei algorithm”. This Lei algorithm was proposed for multi-service scenarios composed on RT and NRT services, and considers the users' channel information and packet delay during the resource allocation.

Similarly to [11], the authors in [12] formulated a utility-based optimization problem and after some mathematical assumptions, a resource allocation policy was derived, which selects the user $j^{\star}$ to transmit on RB $k$ in TTI $n$ according to

$$j^{\star} = \arg\max_{j} \left\{ |U^{'}(d_j^{\text{hol}}[n])| \cdot \frac{r_{j,k}[n]}{T_j\,[n]} \right\}, \tag{3.6}$$

where $T_j\,[n]$ denotes the throughput of user $j$ at TTI $n$ and $U^{'}(d_j^{\text{hol}}[n])$ denotes the first derivative of $U(d_j^{\text{hol}}[n])$ with respect to $d_j^{\text{hol}}[n]$. Notice that differently from what was adopted in [11], the authors of the Lei algorithm use the HOL packet delay as the utility function parameter. The utility functions chosen for defining the users priority during the resource allocation were

$$U_{\text{RT}}(d_j^{\text{hol}}[n]) = 1 - (1 + e^{-\beta(d_j^{\text{hol}}[n] - \Phi_{\text{req}}^{\text{delay}})}), \tag{3.7}$$

for the RT service (where $\beta > 0$ determines the function slope and $\Phi_{\text{req}}^{\text{delay}} > 0$ defines the location of the inflection point of the function, which is defined as the HOL packet delay requirement) and for the NRT service, the following function was employed

$$U_{\text{NRT}}(d_j^{\text{hol}}[n]) = 1 - b \cdot e^{a \cdot (d_j^{\text{hol}}[n] - c)}, \tag{3.8}$$

where $a > 0$, $b > 0$, $c > 0$ determine the slope and amplitude of the function, and the QoS requirement, respectively.

In order to calculate the priority during the resource allocation, it is necessary to calculate the first derivative of (3.7) and (3.8) with respect to the HOL delay. After the derivative, the following equation are obtained:

$$U'_{\text{RT}}(d_j^{\text{hol}}[n]) = \frac{\beta e^{-\beta(d_j^{\text{hol}}[n]-\Phi_{\text{req}}^{\text{delay}})}}{\left(1+e^{\beta(d_j^{\text{hol}}[n]-\Phi_{\text{req}}^{\text{delay}})}\right)^2}, \tag{3.9}$$

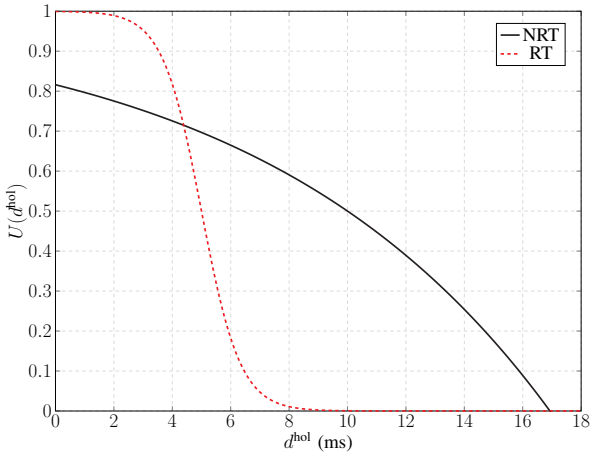
and

$$U'_{\text{NRT}}(d_j^{\text{hol}}[n]) = \text{a}\cdot\text{b}\cdot e^{\text{a}(d_j^{\text{hol}}[n]-c)}. \tag{3.10}$$

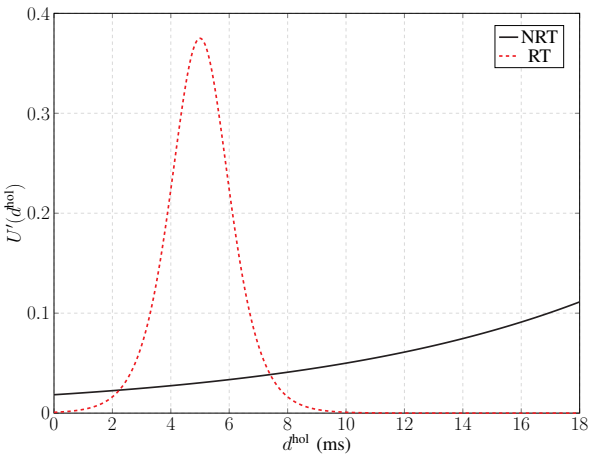
The curves obtained from (3.7) and (3.8) are shown in figure 3.2a for the following parameters values, proposed in [12]: $\beta = 1.5$, $\Phi_{\text{req}}^{\text{delay}} = 5$, $a = 0.1$, $b = 0.5$ e $c = 10$. The curves for (3.9) and (3.10) are depicted in figure 3.2b. Notice that for values of $d_j^{\text{hol}}[n]$ close to $\Phi_{\text{req}}^{\text{delay}}$, the priority for RT services is higher. As $d_j^{\text{hol}}[n]$ increases, the priority for NRT services becomes higher.

Figure 3.2 – Utility functions employed by the Lei algorithm for RT (such as the VoIP and video services) and NRT (such as the CBR considered in this thesis) services.

(a) Utility functions for RT and NRT services.




(b) First derivative of utility functions for RT and NRT services.




Source: Created by the author, adapted from [12].

The performance of the Lei algorithm was compared with the classical PF, EXP and MLWDF algorithms, showing that the Lei algorithm presents a better performance in terms of users' throughput and PLR. However, the performance of the Lei in algorithm terms of fairness was worse than the classical algorithms.

### *Queue-HOL-MLWDF (QHMLWDF)*

The third benchmark algorithm was proposed in [23] and is named QHMLWDF. This algorithm aims at guaranteeing a fair resource allocation in the system. Differently from the MDU and algorithm Lei, the QHMLWDF does come from an optimization problem. This algorithm is presented by the authors in [23] as an enhanced version of the MLWDF [44] and Virtual Token MLWDF (VTMLWDF) [55] algorithms.

The QHMLWDF combines, according to [23], the main QoS metrics used by the MLWDF and VTMLWDF: 1) the HOL packet delay, that increases the priority of users with higher HOL packet delay and close to the requirement; 2) the transmit queue size, that quantifies the service priority according to the queue size in bits. Besides that, the QHMLWDF uses only one criterion for scheduling the users, regardless of service. In this algorithm, the user $j^\star$ is selected to transmit on RB $k$ in TTI $n$ according to

$$j^\star = \arg\max_j \left\{ \frac{\alpha[n] \cdot d_j^{\text{hol}}[n] \cdot \text{Q}_j[n]}{T_j[n]} \cdot r_{j,k}[n] \right\}, \tag{3.11}$$

where $\alpha[n]$ represents the maximum allowed probability for packets exceeding delay requirement at TTI $n$.

Since the QHMLWDF was an enhancement of the MLWDF and VTMLWDF algorithms, the authors in [23] compare its performance with the predecessors algorithms. The QHMLWDF present some performance improvement in terms of PLR, throughput, fairness and spectral efficiency considering scenarios composed of video, VoIP and best effort services.

### *Exponential/Proportional Fair (EXP/PF)*

The resource allocation algorithm known as Exponential/Proportional Fair (EXP/PF) is a combination of the classical single-service EXP and PF algorithms, which are used for RT and throughput-based services, respectively. Therefore, in the EXP/PF algorithm, the resource allocation for RT services is performed by the EXP algorithm, which was initially proposed for Code-Division Multiple Access (CDMA)-based systems composed of a single carrier and with a shared downlink spectrum [45]. The EXP algorithm selects the user $j^\star$ to transmit on RB $k$ in TTI $n$ according to

$$j^\star = \arg\max_j \left\{ \exp\left( \frac{d_j^{\text{hol}}[n]}{1 + \sqrt{\overline{d_j^{\text{hol}}}[n]}} \right) \cdot \frac{r_{j,k}[n]}{T_j[n]} \right\}, \tag{3.12}$$

where $\overline{d_j^{\text{hol}}}[n]$ is the mean HOL packet delay of all active RT users at TTI $n$. Notice that according to this equation, the EXP/PF algorithm attempts to guarantee that RT users transmit as soon as they become active. As the HOL packet delay of user $j$ increases approaching the $\overline{d_j^{\text{hol}}}[n]$, the priority of user $j$ also increases.

For the throughput-based services, the EXP/PF algorithm schedules users according the PF policy, given by

$$j^{\star} = \arg\max_{j}\left\{\frac{r_{j,k}[n]}{T_j\,[n]}\right\}. \tag{3.13}$$

The PF algorithm schedules users guaranteeing the fairness between users by allocating the RBs according to a ratio between their current throughput and achievable data rate on a given RB.

The EXP/PF algorithm was originally conceived in [56] for supporting multimedia applications in Time Division Multiple Access (TDMA)-based systems. The algorithm EXP/PF attempts to allocate users from RT services within a given time limit and to maximize the system throughput. In [37], the authors adapted the original proposal for OFDMA-based systems, where it was shown that the EXP/PF presented better performance than the MLWDF algorithm.

# 4 SCHEDULING FRAMEWORK FOR JOINT SATISFACTION MAXIMIZATION

## 4.1 Introduction

The Utility Theory has been widely used in the literature for designing RRA algorithms for cellular networks. For instance, this theory has been used for single service scenarios in [6, 7, 8, 9], and for multi-service scenarios in [11, 12]. The Utility Theory was initially conceived for applications in the economics area, where it was applied to explain the consumers' behavior and help in the decision-taking process [13, 14]. However, this theory has also received some attention of the wireless communications research community over the last years [15].

The RRA algorithms for cellular networks aim at guaranteeing a trade-off between QoS, spectral efficiency and fairness in the RBs allocation. In the economics field, the utility theory has been used to study the problem of providing a fair and efficient resource allocation, where utility functions have been applied for quantifying the advantage of using particular resources [6]. A similar approach can be employed in the area of cellular networks, where an evaluation of how well the network is satisfying the users' applications requirements could be conducted by using metrics such as throughput, FER or outage probability [16].

Therefore, the utility theory emerges as a powerful tool for the conception of RRA algorithms since it allows us to quantify the user satisfaction levels for a given resource allocation. Thus, it is possible to design RRA algorithm capable of achieving different levels of fairness and user satisfaction in the resource allocation process [15, 11].

This chapter firstly presents the general formulation of the proposed utility-based optimization problem that is generalized for distinct services classes. Then, this general formulation is particularized for multi-service scenarios composed of throughput-based, delay-based and queue-based services, which are defined here as the services that require a minimum throughput and maximum FER. Finally, the resource allocation scheme derived from this formulation is presented.

## 4.2 General Formulation

As aforementioned, the utility theory allows us to capture the satisfaction level of users for a certain resource assignment. Furthermore, this theory allows the combination of metrics from the PHY and MAC layers to achieve cross-layer optimization [25].

In order to address the problem of simultaneously maximizing the satisfaction of users from distinct service classes, we have designed a general utility-based optimization problem with the objective of maximizing the users' utility derived from the network (satisfaction) for all services simultaneously. Considering the total set of users $\mathcal{J} = \{1, 2, \ldots, J\}$ and the total

set of RBs $\mathcal{K} = \{1, 2, \ldots, K\}$ in the system, the proposed problem is formulated as

$$\max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} V\left[U\left(x_j\right)\right] \tag{4.1a}$$

$$\text{subject to } \rho_{j,k} \in \{0, 1\},\ \forall j \in \mathcal{J} \text{ and } \forall k \in \mathcal{K}, \tag{4.1b}$$

$$\sum_{j \in \mathcal{J}} \rho_{j,k} = 1,\ \forall k \in \mathcal{K}, \tag{4.1c}$$

$$\sum_{k \in \mathcal{K}} p_k = P_{\mathrm{t}}, \tag{4.1d}$$

$$p_k \geq 0,\ \forall k \in \mathcal{K}, \tag{4.1e}$$

where $\mathcal{J}$ is the set of all users in the system; $J$ is the total number of users in a cell served by the BS; $\mathcal{K}$ is the set of all RBs in the system; $K$ is the total number of RBs in the system to be assigned to the users; $\rho_{j,k}$ is an assignment variable that assumes the value 1 if the RB $k$ of the BS is assigned to the user $j$ and 0 otherwise; $p_k$ is the transmit power allocated on each RB $k$ of the BS; $P_{\mathrm{t}}$ is total transmit power of the BS; $U\left(x_j\right)$ is a user utility function based on a generic variable $x_j$ that can represent a resource usage or QoS metric associated to user $j$; and $V\left(\cdot\right)$ is a service utility function that depends on the user utility functions $U(x_j)$'s and differentiates the service classes in the system. Constraints (4.1b) and (4.1c) state that the RBs of the BS are discrete and that the same RB cannot be shared by more than two users in the same TTI. Constraints (4.1d) and (4.1e) require that the total sum of the powers over all RBs must not surpass the total transmit power of the BS, and that these powers must be non-negative.

In general, it is very hard to find the optimum solution for the proposed optimization problem. Therefore, a problem-splitting approach has been used, similar to the ones commonly used in the literature, in order to derive a sub-optimum solution. Firstly, a Dynamic Resource Assignment (DRA) is performed with fixed power allocation; then, there is a stage of EPA with fixed resource assignment [57].

## 4.3 General Multi-Service Formulation

This work is an extension of the single service cases evaluated in [9], where the authors particularized the general utility-based optimization problem for scenarios with only throughput-based or delay-based services. The authors in [9] have demonstrated that it is possible to derive simplified optimization problems equivalent to the original optimization problems for both throughput-based and delay-based services separately, where the simplified problems have objective functions in terms of the users' instantaneous data rate.

As a generalization of the scenarios presented in [9], in this work we consider a scenario composed of $S$ distinct services. The optimization problem is the maximization of the total utility with respect to the users' QoS, which can be throughput, HOL packet delay or queue size for throughput-based, delay-based or queue-based services, respectively. For example, let us assume that the set $\mathcal{J}$ of the users in the system is separated in $S$ subsets: $\mathcal{J}_1, \mathcal{J}_2, \ldots, \mathcal{J}_s, \ldots, \mathcal{J}_S$

for users from service $1, 2, \ldots, s, \ldots, S$, respectively. Based on this assumption, the objective function (4.1a) can be re-written as:

$$\max_{\rho_{j,k},\, p_k} \left\{ \sum_{j \in \mathcal{J}_1} V\left[U_1\left(x_j^1\right)\right] + \ldots + \sum_{j \in \mathcal{J}_s} V\left[U_s\left(x_j^s\right)\right] + \ldots + \sum_{j \in \mathcal{J}_S} V\left[U_S\left(x_j^S\right)\right] \right\}, \tag{4.2}$$

where $U_s\left(x_j^s\right)$ is a user utility function based on a generic variable $x_j^s$ that represents a QoS metric associated to user $j$ from service $s$.

The service utility function is used for multi-service scenarios in order to give specific weights for different services depending on the desired objective to be achieved. This utility function may be designed to provide equal priority among services, a fixed priority to one service over the other, or even to be an adaptive function that can be changed dynamically in order to meet some QoS requirement of the most prioritized service.

In this work, we propose an adaptive service utility function that is used for prioritizing the services in the system. This feature is the main novelty proposed here; as far as we know, there is no other work in the literature that has proposed a similar approach.

## 4.4 Scenario Particularization

In this section, we particularize the general utility-based optimization problem presented in the previous subsection for scenarios with throughput-based services, delay-based services, queue-based services, and a traffic mix between these types of services.

### *4.4.1 Throughput-Based Single Service Scenario*

The problem regarding a scenario with only throughput-based services is similar to the problem presented in [9] for the NRT single service scenario. We refer to the utility function for the throughput-based service as $U_{\text{thr}}(\cdot)$. The optimization problem is given by the maximization of the total utility with respect to the users' throughput, where the objective function is

$$\max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} U_{\text{thr}}\left(T_j\left[n\right]\right), \tag{4.3}$$

while the constraints of the general optimization problem remain unaltered. $U_{\text{thr}}\left(T_j\left[n\right]\right)$ is an increasing utility function based on the current throughput $T_j\left[n\right]$ of user $j$ at TTI $n$. Since we are dealing with a single service in this subsection (no need for service differentiation), the service utility function $V\left(\cdot\right)$ is not used.

As demonstrated in appendix A, it is possible to derive a simplified optimization problem that is equivalent to our original problem regarding throughput-based services. From appendix A, the objective function of the simplified problem is linear in terms of the instantaneous user's data rate and given by

$$\max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} U'_{\text{thr}}\left(T_j\left[n-1\right]\right) \cdot R_j\left[n\right] = \max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} w_j^{\text{thr}} \cdot R_j\left[n\right], \tag{4.4}$$

where $U'_{\text{thr}}\left(T_j\left[n-1\right]\right) = \left.\dfrac{\partial U_{\text{thr}}}{\partial T_j}\right|_{T_j=T_j[n-1]}$ is the marginal utility of user $j$ with respect to its throughput in the previous TTI $n-1$, and $R_j\left[n\right]$ is the instantaneous data rate of user $j$ at TTI $n$.

The assumptions and mathematical simplifications described in appendix A allows us to affirm that the instantaneous optimization that maximizes (4.4) leads to a long-term optimization that maximizes (4.3). The simplified objective function (4.4) characterizes a weighted sum rate maximization problem [58], whose weights are adaptively controlled by the user marginal utilities $w_j^{\text{thr}}$.

For simplicity of notation, we represent the user marginal utility corresponding to the throughput-based user $j$ as the weight

$$w_j^{\text{thr}} = U'_{\text{thr}}\left(T_j\left[n-1\right]\right). \tag{4.5}$$

### *4.4.2 Delay-Based Single Service Scenario*

For the case of delay-based services, the optimization problem is given by the maximization of the total utility with respect to the users' HOL packet delays. We refer to the utility function for the delay-based service as $U_{\text{delay}}(\cdot)$. The objective function (4.1a) of problem (4.1) becomes

$$\max_{\rho_{j,k},\, p_k} \sum_{j\in\mathcal{J}} U_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right), \tag{4.6}$$

while constraints (4.1b)-(4.1e) remain valid. $U_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right)$ is a decreasing utility function based on the current HOL delay $d_j^{\text{hol}}\left[n\right]$ of user $j$ at TTI $n$. The HOL delay is the time that the oldest packet in the user's buffer has to wait before being transmitted over the access network. Notice again that the service utility function $V\left(\cdot\right)$ is not used in this single service scenario because we are dealing with only one service.

It is also possible to derive a simplified optimization problem that is equivalent to our original problem regarding delay-based services. This configuration is similar to the problem presented in [9] for the RT single service scenario. According to appendix B, the objective function of our simplified optimization problem is also linear in terms of the instantaneous user's data rate and given by

$$\max_{\rho_{j,k},\, p_k} \sum_{j\in\mathcal{J}} \left|U'_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right)\right| \cdot R_j\left[n\right] = \max_{\rho_{j,k},\, p_k} \sum_{j\in\mathcal{J}} w_j^{\text{delay}} \cdot R_j\left[n\right], \tag{4.7}$$

where $U'_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right) = \left.\dfrac{\partial U_{\text{delay}}\left(d_j^{\text{hol}}\right)}{\partial d_j^{\text{hol}}}\right|_{d_j^{\text{hol}}=d_j^{\text{hol}}[n]}$ is the user marginal utility of user $j$ with respect to its current HOL delay.

The assumptions and mathematical simplifications described in appendix B allows us to claim that the instantaneous optimization that maximizes (4.7) and also leads to a long-term optimization that maximizes (4.6). The objective function (4.7) also characterizes a weighted sum

rate maximization [58], where the weights are given by the absolute value of the user marginal utility with respect to the current HOL delay.

For the sake of simplicity, let us also define a user-specific weight $w_j^{\text{delay}}$ given by

$$w_j^{\text{delay}} = \left| U'_{\text{delay}} \left( d_j^{\text{hol}} [n] \right) \right|, \tag{4.8}$$

for the case when user $j$ uses a delay-based service.

### *4.4.3 Queue-Based Single Service Scenario*

For the case of services with minimum throughput and maximum FER requirements, we have chosen to use the average queue size as the metric to allocate resources for this type of traffic. The reason for this is that by keeping this metric in low values, we can maximize the throughput and minimize the FER for a certain user $j$. Therefore, for this service class, which are denoted as queue-based, the considered optimization problem is the maximization of the total utility with respect to the predicted average queue size (in bits) over a time window of a user $j$, where this queue is located at the BS. This specific formulation is one of the contributions of this work.

For simplicity, we will refer to the utility function for the queue-based services as $U_{\text{queue}}(\cdot)$. The considered optimization problem is the maximization of the total utility with respect to the average queue size of user $j$, where the objective function is

$$\max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} U_{\text{queue}} \left( \overline{Q}_j [n+1] \right), \tag{4.9}$$

while the constraints of the general optimization problem remain valid. $U_{\text{queue}} \left( \overline{Q}_j [n+1] \right)$ is a decreasing utility function based on the predicted average queue size $\overline{Q}_j [n+1]$ of user $j$ at TTI $n+1$. Notice again that the service utility function $V(\cdot)$ is not used in this single service scenario.

It is also possible to derive a simplified optimization problem that is equivalent to our original problem regarding services with minimum throughput and maximum FER requirements. According to appendix C, the objective function of our simplified optimization problem is also linear in terms of the instantaneous user's data rate and given by

$$\max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} \left| U'_{\text{queue}} \left( \overline{Q}_j [n] \right) \right| \cdot R_j [n] = \max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} w_j^{\text{queue}} \cdot R_j [n], \tag{4.10}$$

where $U'_{\text{queue}} \left( \overline{Q}_j [n] \right) = \left. \dfrac{\partial U_{\text{queue}} \left( \overline{Q}_j \right)}{\partial \overline{Q}_j} \right|_{\overline{Q}_j = \overline{Q}_j [n]}$ is the marginal utility of user $j$ with respect to its current average queue size.

The appendix C describes some mathematical simplifications that allows us to assume that the instantaneous optimization that maximizes (4.10) leads to a long-term optimization that

maximizes (4.9). The objective function (4.10) also characterizes a weighted sum rate maximization [58], where the weights are given by the absolute value of the user marginal utility with respect to the current average queue size, which is a metric available at the transmitter buffer at the BS.

For the sake of simplicity, let us also define a user-specific weight $w_j^{\text{queue}}$ given by

$$w_j^{\text{queue}} = \left| U'_{\text{queue}} \left( \overline{Q}_j \left[ n \right] \right) \right|, \tag{4.11}$$

for the case when user $j$ uses queue-based services.

### *4.4.4 Particularized Multiple Services Scenario*

Considering a scenario with throughput-based, delay-based and queue-based services, the optimization problem is done by the maximization of the total utility with respect to the users' QoS, namely throughput, HOL packet delay and average queue size for throughput-based, delay-based and queue-based services, respectively. Let us assume that the set $\mathcal{J}$ of the users in the system is separated in three subsets: $\mathcal{J}_{\text{thr}}$, $\mathcal{J}_{\text{delay}}$ and $\mathcal{J}_{\text{queue}}$ for throughput-based, delay-based and queue-based users, respectively. Therefore, the objective function of the general optimization problem can be re-written as

$$\max_{\rho_{j,k},\, p_k} \left\{ \sum_{j \in \mathcal{J}_{\text{thr}}} V \left[ U_{\text{thr}} \left( T_j \left[ n \right] \right) \right] + \sum_{j \in \mathcal{J}_{\text{delay}}} V \left[ U_{\text{delay}} \left( d_j^{\text{hol}} \left[ n \right] \right) \right] + \sum_{j \in \mathcal{J}_{\text{queue}}} V \left[ U_{\text{queue}} \left( \overline{Q}_j \left[ n+1 \right] \right) \right] \right\}. \tag{4.12}$$

Notice that now the service utility function $V\left(\cdot\right)$ is used in this multi-service scenario, which is used for differentiating the priority of different services. This utility function could be designed to provide equal priority among services, a fixed priority to one service over the other, or even an adaptive function could be changed dynamically in order to meet some QoS requirement of the most prioritized service.

According to appendix D, it is also possible to derive a simplified optimization problem that is equivalent to our original problem regarding multiples services. The objective function of our simplified optimization problem is also linear in terms of the instantaneous user's data rate and given by

$$\begin{aligned} \max_{\rho_{j,k},\, p_k} \Bigg\{ & \sum_{j \in \mathcal{J}_{\text{thr}}} V' \left( U_{\text{thr}} \left( T_j \left[ n-1 \right] \right) \right) \cdot U'_{\text{thr}} \left( T_j \left[ n-1 \right] \right) \cdot R_j \left[ n \right] \\ & + \sum_{j \in \mathcal{J}_{\text{delay}}} V' \left( U_{\text{delay}} \left( d_j^{\text{hol}} \left[ n \right] \right) \right) \cdot \left| U'_{\text{delay}} \left( d_j^{\text{hol}} \left[ n \right] \right) \right| \cdot R_j \left[ n \right] \\ & + \sum_{j \in \mathcal{J}_{\text{queue}}} V' \left( U_{\text{queue}} \left( \overline{Q}_j \left[ n \right] \right) \right) \cdot \left| U'_{\text{queue}} \left( \overline{Q}_j \left[ n \right] \right) \right| \cdot R_j \left[ n \right] \Bigg\}. \end{aligned} \tag{4.13}$$

Generalizing the notation, we can re-write (4.13) as

$$\max_{\rho_{j,k},\, p_k} \left\{ \sum_{j \in \mathcal{J}_{\text{thr}}} w_j^s \cdot w_j^{\text{thr}} \cdot R_j\left[n\right] + \sum_{j \in \mathcal{J}_{\text{delay}}} w_j^s \cdot w_j^{\text{delay}} \cdot R_j\left[n\right] + \sum_{j \in \mathcal{J}_{\text{queue}}} w_j^s \cdot w_j^{\text{queue}} \cdot R_j\left[n\right] \right\} = \max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} w_j^s \cdot w_j \cdot R_j\left[n\right]. \tag{4.14}$$

where $w_j^s$ is the utility-based service weight associated to user $j$ and service $s$, and $w_j$ is the utility-based user weight associated to user $j$. According to (4.13) and (4.14), we have that

$$w_j^s = \begin{cases} V'\left(U_{\text{thr}}\left(T_j\left[n-1\right]\right)\right), & \text{if } j \text{ is a user from a throughput-based service} \\ V'\left(U_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right)\right), & \text{if } j \text{ is a user from a delay-based service} \\ V'\left(U_{\text{queue}}\left(\overline{Q}_j\left[n\right]\right)\right), & \text{if } j \text{ is a user from a queue-based service} \end{cases} \tag{4.15}$$

and

$$w_j = \begin{cases} U'_{\text{thr}}\left(T_j\left[n-1\right]\right), & \text{if } j \text{ is a user from a throughput-based service} \\ \left|U'_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right)\right|, & \text{if } j \text{ is a user from a delay-based service} \\ \left|U'_{\text{queue}}\left(\overline{Q}_j\left[n\right]\right)\right|, & \text{if } j \text{ is a user from a queue-based service.} \end{cases} \tag{4.16}$$

Notice that the equations in (4.16) are equal to (4.5), (4.8) and (4.11). In eq. (4.15), the innovative service utility weights are presented, feature that has not been found in any other work in the literature. These utility-based weights play an important role in the decision making of the RRA framework proposed in the sequel.

## 4.5 Proposed Resource Allocation Algorithm

In the last subsection, we have described an optimization problem that can be employed in any current or future cellular system. We propose herein a resource allocation framework suitable for air interfaces based on OFDMA. This framework might be employed in current and future air interfaces that use this multiple access scheme or as long as the orthogonal resource allocation is guaranteed. Additionally, the proposed RRA framework is able to adaptively exploit the diversities in wireless systems, such as: time, frequency, space, multi-user and traffic diversities.

The simplified objective function in eq. (4.14) characterizes a weighted sum rate maximization problem [58], whose weights are $w_j^s$ and $w_j$, and the objective function is linear with respect to $R_j\left[n\right]$. Because of this linearity, when the objective function is given by eq. (4.14), the DRA problem described previously has a closed form solution [59]. Considering the more realistic and complex scenario composed of multiple services, the user with index $j^\star$ is chosen to transmit on RB $k$ in TTI $n$ if it satisfies the condition given by

$$j^\star = \arg\max_j \left\{ w_j^s \cdot w_j \cdot r_{j,k}\left[n\right] \right\}, \tag{4.17}$$

where $r_{j,k}[n]$ denotes the instantaneous achievable transmission rate of user $j$ with respect to RB $k$ at TTI $n$. If more than one user has the same priority for the same RB $k$, a tiebreaker process selects the user with the highest SNR.

Equation (4.17) describes the first step of the problem-splitting approach mentioned at the end of section 4.2, which is the DRA procedure. Then we perform an EPA considering the resource assignment obtained from the first step. At this point, the total available power $P_t$ of each BS is equally distributed among all RBs. Consequently, the power $p_k$ allocated to each RB $k$ is $p_k = \frac{P_t}{K}$.

It is worth noting here the simplicity of our proposed algorithm. In the beginning, there was a utility-based optimization problem whose solution is hard to be found. After some mathematical assumptions and simplifications, we derived an RRA formula (eq. (4.17)) that can be easily comprehended since it only involves the multiplication of two weights ($w_j^s$ and $w_j$) by $r_{j,k}[n]$.

## 4.6 Utility Functions for Maximization of User Satisfaction

This thesis proposes an RRA policy that is capable of maximizing the number of satisfied users in a cellular system where multiple service classes are present. The policy proposed in this work is called JSM and is an extension of the RRA policies introduced in [9], namely TSM and DSM. Besides the JSM policy, we also propose the Queue-based Satisfaction Maximization (QSM) policy, which is a contribution of this work, that is suitable for maximizing the user satisfaction in scenarios composed of queue-based services, such as the video service.

### *4.6.1 Utility Function for User Prioritization*

The authors in [9] have originally proposed the use of a sigmoidal utility function based on a generic QoS metric $x_j[n]$ of the user $j$ for the TSM and DSM policies and demonstrated that the use of this particular function provides higher levels of user satisfaction when compared with classical algorithms. As mentioned before, besides extending the policies TSM and DSM, we propose a new policy called QSM. In this work, we propose the use of the logistic function for the three policies aforementioned, which is an equivalent form of the sigmoidal function as indicated below

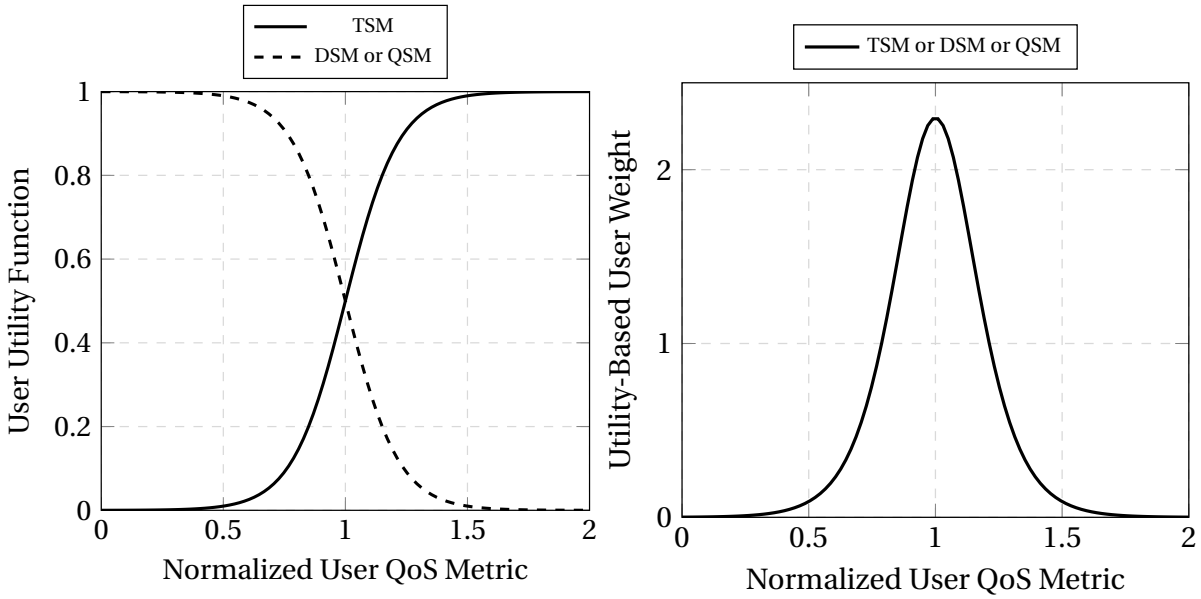
$$U(x_j[n]) = \frac{1}{1 + e^{\mu(x_j[n] - x_j^{\text{req}})/\sigma}}. \tag{4.18}$$

This function was chosen to be employed in our framework due to the properties allowed by its parametrization and because it is a continuous and differentiable function. One of the parameters ($\sigma$) has a non-negative value that allows the regulation of the logistic function shape; the second parameter ($\mu$) is a constant that defines whether the function is ascending ($\mu = -1$) or descending ($\mu = 1$); and the last one ($x_j^{\text{req}}$) is the QoS requirement of a given service and determines the abscissa shift of the function.

Figure 4.1 – Functions for user prioritization.

(a) Logistic utility function.

(b) Bell-shaped marginal utility.




Source: Created by the author.

The input variable ($x_j[n]$) of the logistic function is the QoS metric of each service. Once distinct service classes have different metrics, we normalize both the input QoS metric ($x_j[n]$) and the QoS requirement ($x_j^{\text{req}}$) by the QoS requirement. Therefore, after the normalization, our framework is independent of the QoS metric being considered.

A function of $x_j^{\text{req}}$ was developed in order to obtain a value of $\sigma$ such that a desired step-shaped logistic function can be achieved. The expression

$$\sigma = \frac{\mu \cdot (\rho - 1) \cdot x_j^{\text{req}}}{\log\left(\frac{1}{\delta} - 1\right)} \tag{4.19}$$

states that the logistic function is equal to a given value $\delta$ when the QoS metric $x_j$ achieves a proportion $\rho$ of the QoS requirement $x_j^{\text{req}}$.

The TSM policy uses an increasing utility function based on the users' throughput and centered at $x^{\text{req}}$. The throughput-based utility function employed by the TSM algorithm is shown in figure 4.1a. As a result of the curve shape, a certain user becomes satisfied rapidly when its throughput approaches and exceeds the throughput requirement. Satisfactory results were obtained in [9] for the NRT utility function when $\delta = 0.01$, $\rho = 0.50$ and $x_j^{\text{req}} = 1$, which gives the value $\sigma_{\text{thr}} = 0.1088$.

The DSM policy uses a decreasing utility function based on the users' HOL packet delay and also centered at $x^{\text{req}}$, as shown in figure 4.1a. A decreasing delay-based utility function

means that the higher the delay users are experiencing, the lower the users' utility derived from the network. Consequently, users become satisfied when the user HOL packet delay decreases to values below the requirement. Considering this service, satisfactory results were obtained when the parameters to calculate $\sigma$ using the proposed formula are: $\delta = 0.99$, $\rho = 0.50$ and $x_j^{\text{req}} = 1$, which also returns $\sigma_{\text{delay}} = 0.1088$.

The QSM policy employs the same utility function used by the DSM policy, as shown in figure 4.1a. However, instead of using the users' HOL packet delay as the QoS metric $x_j$, the QSM policy uses the average queue size. The reason behind this is that this policy is applied for queue-based services that have simultaneously throughput and FER requirements, and a decreasing utility function means that the higher the average queue size the users are experiencing, the lower the users' utility derived from the network. As a consequence, users become satisfied when the user average queue size decreases to values below the requirement.

Assessing the DRA algorithm described in the previous section, it is possible to conclude that the higher the utility-based weights values for a given user, the higher the priority of that user to get a resource. The utility-based weight $w_j$ in eq. (4.17) is obtained by a derivative operation, as explained in section 4.4.4, given by

$$w_j = \frac{\partial U(x_j[n])}{\partial x_j[n]} = \frac{-\mu e^{\mu(x_j[n]-x_j^{\text{req}})/\sigma}}{\sigma(1+e^{\mu(x_j[n]-x_j^{\text{req}})/\sigma})^2}. \tag{4.20}$$

As can be seen in figure 4.1b, the marginal utility is a bell-shaped function centered at $x_j^{\text{req}} = 1$ for the TSM, DSM or QSM.

It is worth highlighting that due to the independence of the QoS metric being considered in our normalized framework, no adjustment is needed to compare the utility-based weight for throughput-based, delay-based or queue-based services. Therefore, the utility-based weight has the same behavior for all services.

### *4.6.2 Utility Function for Service Prioritization*

So far, in this section, we have described RRA algorithms that deal with throughput-based, delay-based or queue-based services separately, i.e., only the utility-based user weights have been mentioned. However, the main novelty of our proposed technique is that it addresses the maximization of user satisfaction in more realistic scenarios, where multiple service classes are considered at the same time. Recall that the JSM technique proposes an adaptability in this service utility function to enhance the differentiation in the service priority.

The service utility function $V(\cdot)$ employed in this work is given by

$$V\left(z_j[n]\right) = \log(1 + e^{\mu(z_j[n]-1)/\lambda}). \tag{4.21}$$

This function was chosen due to the properties of its first derivative, which is a scaled version of the hyperbolic tangent, that allows us to define regions for the service differentiation;

therefore, each service has a specific priority related to the magnitude of the function in the corresponding region. This function is depicted in figure 4.2a for $\mu = 1$ and 5 values of sigma: $\lambda = 0.1088$, $\lambda = 0.4683$, $\lambda = 5000$, $\lambda = -0.4683$, and $\lambda = -0.1088$. The reason behind these values of $\lambda$ are explained in the sequel.

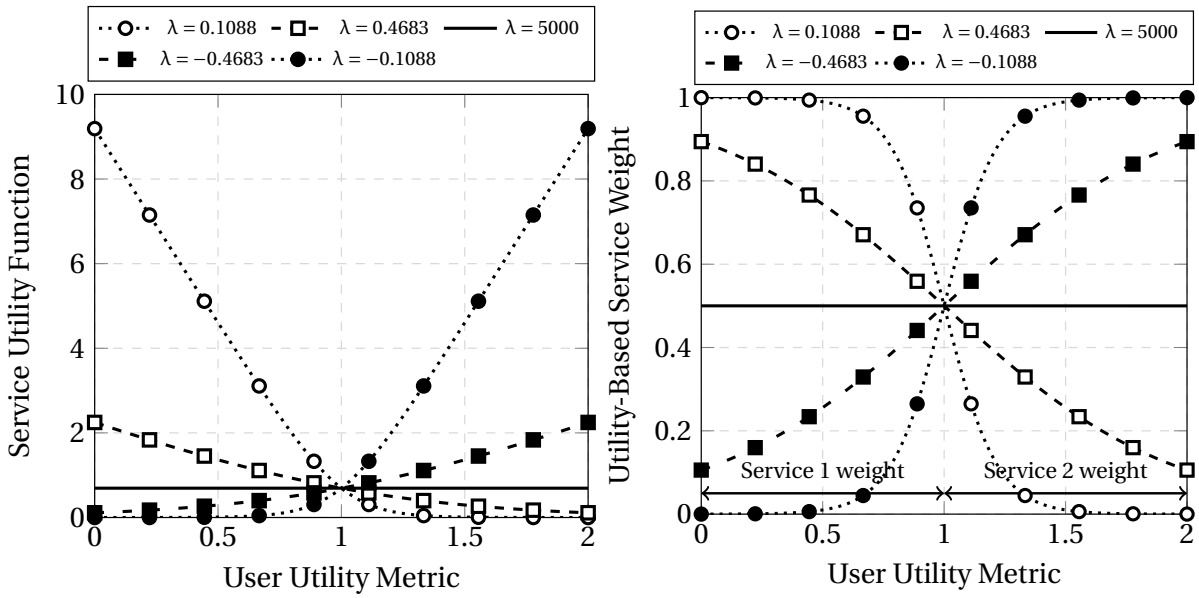
The marginal utility function $V^{'}(z_j[n])$ is given by the first derivative of $V(z_j[n])$. Therefore, $V^{'}(z_j[n])$, also known as the utility-based service weight $w_j^s$, is a logistic marginal utility given by

$$w_j^s = \frac{\partial V(z_j[n])}{\partial z_j[n]} = \frac{1}{1 + e^{\mu(z_j[n]-1)/\lambda}}. \tag{4.22}$$

The parameters of this function are set as: $\mu = 1$, $\lambda$ is the parameter that controls the function shape, and $z_j[n]$ can be $U_{\text{thr}}\left(T_j\left[n-1\right]\right)$ if $j$ is a throughput-based user, $U_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right)$ if $j$ is a delay-based user or $U_{\text{queue}}\left(\overline{Q}_j\left[n\right]\right)$ if $j$ is a queue-based user, as explained in section 4.4.4. $V^{'}(z_j[n])$ is depicted in figure 4.2b for five values of $\lambda$. It is important to notice here that in this thesis, we particularize the multiple services scenario for two services due to the chosen utility function. However, other scenarios composed of more than two services can be exploited by selecting another service utility function with three separate regions for service prioritization.

Figure 4.2 – Functions for Service Prioritization.

(a) Service utility Function. (b) Logistic marginal utility function.




Source: Created by the author.

Notice that $U_s\left(x_j^s\right)$ utility function vary in amplitude from 0 to 1 (see figure 4.1a), and the abscissa axis of $w_j^s$ varies from 0 to 2 (see figure 4.2b). In order to match the mapping between the $U_s\left(x_j^s\right)$ utility functions and utility-based service weight, an adjustment is employed by

defining

$$z_j[n] = \begin{cases} U_1\left(x_j^1\right), & \text{if } j \text{ is a user from service 1} \\ U_2\left(x_j^2\right) + 1. & \text{if } j \text{ is a user from service 2.} \end{cases} \tag{4.23}$$

Therefore, after this adjustment, the values of $w_j^s$ for the abscissa axis varying from $0$ to $1$ represent the priority weights for users from service 1, and the values of $w_j^s$ for the abscissa axis varying from $1$ to $2$ (due to the unitary shift introduced in $z_j[n]$) determine the priority weights for users from service 2, as shown in figure 4.2b.

The JSM technique considers the service utility function $V(z_j[n])$ as an adaptive function, changing its shape dynamically in order to protect the most prioritized service. From now on, let us assume that the service 1 is the most prioritized service. In our approach, the percentage of satisfied users from service 1 ($\Upsilon^1[n]$) must remain above a threshold of 90% by adapting the $\lambda$ parameter in the utility-based service weight. The approach generally followed in the literature is to protect service with RT QoS requirements. Considering our definitions, queue-based services (such as the video service) or delay-based services (such as VoIP) have RT QoS requirements. Therefore, as examples of this approach, we could have queue-based services or delay-based services protected over the throughput-based service.

The technique used for performing the adaptation of the $\lambda$ parameter is explained in the following. In order to adapt the value of $\lambda$, a look-up table was created by using a curve fitting tool. The look-up table is calculated off-line and only once for all simulations. Equal priority for both services is achieved when $\lambda = 5000$, which was calculated to obtain a horizontal line at $V^{'}(z_j[n]) = 0.5$. Looking at figure 4.2b, one can notice that when $\lambda = 0.1088$, the service 1 has the highest possible service priority weight, while the lowest achievable service priority weight is given for the service 2. On the other hand, the service 2 is highly prioritized over the service 1 when $\lambda = -0.1088$. Aiming to gather curves in between these two extremes, a cubic interpolant function was employed in a curve fitting tool to obtain 41 equally spaced curves in $z_j[n] = 0.5$ (which is equivalent to $z_j[n] = 1.5$ because of the function symmetry), which is the middle point of the priority weight region for each service. The result of this procedure is a look-up table comprised of 41 non-linear spaced values of $\lambda$. Even though our look-up table is composed of 41 values of $\lambda$, a look-up table with more values of $\lambda$ could be calculated so that a thinner adaptation is performed. Some values of $\lambda$ taken from this table and their corresponding utility-based service weights are depicted in figure 4.2b.

It is worth mentioning that the choice of which service is given higher priority can be changed according to the users' behavior and/or network operator's interest. For instance, if the main objective of the network operator is to increase the system capacity, the throughput-based service can be prioritized over the delay-based service.

### *4.6.3 Flow Chart of Proposed Algorithm*

A summary of all steps involved in the resource allocation performed every TTI by the JSM technique is depicted in figure 4.3, showing the simplicity of the proposed framework.

The steps 1 and 2 involves initialize the set of active users in the system (i.e., users with some data to be received) and separate this set in two subsets according to the users' services. In step 3, the percentage of satisfied users from the protected service is estimated. Then, step 4 compares the estimated satisfaction value to the target to be achieved and step 5 adapts the service weight according to the relationship between estimated and target satisfaction value. In step 6, the QoS metrics are estimated for all users and step 7 calculates the service and user weights from the estimated QoS metrics. Finally, step 8 decides which user is assigned to each RB and step 9 performs an EPA among the all RBs.

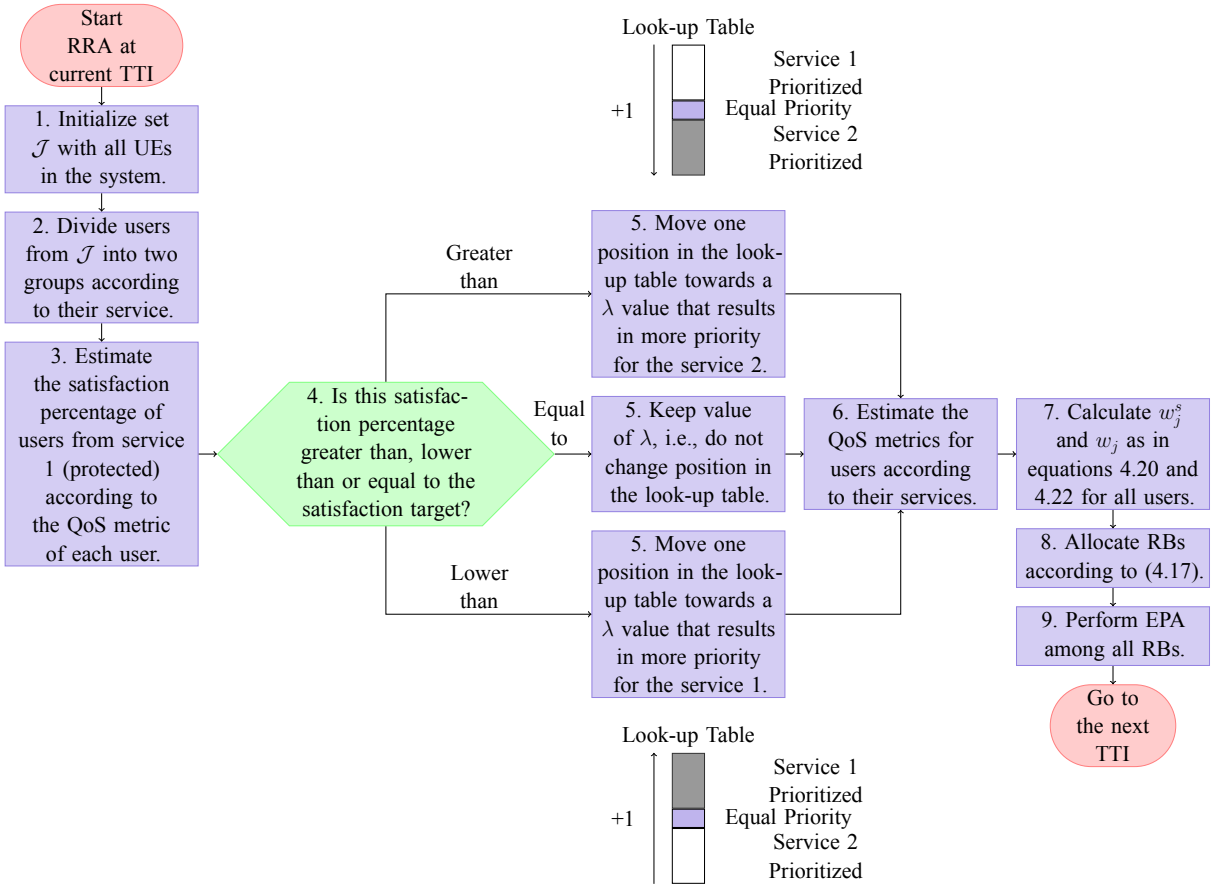



Figure 4.3 – Flow chart explaining all steps involved in the proposed framework.

Look-up tables are shown above or below their corresponding step in the algorithm, where the gray region is the target region of $\lambda$ values for that particular condition. For instance, when $\Upsilon^1[n] < 90\%$, one position is moved in the look-up table towards a value of $\lambda$ that results in giving more priority to the users from service 1, as shown at the bottom of figure 4.3. The opposite occurs when $\Upsilon^1[n] > 90\%$ (upper region of figure 4.3). All the other steps presented in this flow chart have been discussed throughout this chapter.

### *4.6.4 Pseudocode and complexity of JSM Algorithm*

The pseudocode of the JSM algorithm is presented in Alg. 1, showing again the simplicity of the proposed framework. The steps described in Alg. 1 are equal to the steps in figure 4.3, but in a pseudocode format. The first step involves splitting the set of UEs $\mathcal{J}$ into two service groups, namely $\mathcal{J}_1$ and $\mathcal{J}_2$. Then, the QoS metrics of the UEs are estimated according to their service. For example, considering throughput-based services, estimate the current and required throughput. Then, estimate the percentage of satisfied UEs from the protected service ($\Upsilon^1$) and adapts the $\lambda$ value. Finally, the RBs are allocated following eq. (4.17) and an EPA is performed.

The complexity of the algorithm is mainly driven by lines 14 to 17 of Alg. 1. These steps calculate $J$ priority values and find the maximum between them, which is repeated $K$ times. Therefore, the complexity of the proposed algorithm is $\mathcal{O}(JK)$.

**Algorithm 1** Pseudo code of proposed RRA algorithm.

1: Initialize set of UEs $\mathcal{J} = \{1, 2, \ldots, J\}$ and set of RBs $\mathcal{K} = \{1, 2, \ldots, K\}$
2: Split $\mathcal{J}$ into $\mathcal{J}_1$ and $\mathcal{J}_2$
3: Estimate QoS metrics for UEs according to their services.
4: Calculate $\Upsilon^1$ (percentage of satisfied UEs from $\mathcal{J}_1$)
5: **if** $\Upsilon^1 = 90\%$ **then** ▷ Satisfaction equals to target
6: Keep current $\lambda$ value in look-up table
7: **else**
8: **if** $\Upsilon^1 > 90\%$ **then** ▷ Satisfaction higher than target
9: Move one position in look-up table to a $\lambda$ value that gives more priority to $\mathcal{J}_2$
10: **else** ▷ Satisfaction lower than target
11: Move one position in look-up table to a $\lambda$ value that gives more priority to $\mathcal{J}_1$
12: **end if**
13: **end if**
14: Calculate $w_j^s$ and $w_j$ as in eq. (4.13) and eq. (4.14) $\forall j \in \mathcal{J}$
15: **for** $k = 1$ to $K$ **do**
16: Allocate RB according to eq. (4.17) ▷ RB allocation
17: **end for**
18: **for** $k = 1$ to $K$ **do**
19: $p_k = P_t/K$ ▷ Perform EPA
20: **end for**

# 5 PERFORMANCE EVALUATION

## 5.1 Introduction

In this chapter, the performance of the proposed JSM algorithm is evaluated and compared against four benchmark algorithms (which have been described in details in Chapter 3) by means of system-level simulations. Section 5.2 describes the simulation parameters along with the evaluated scenarios and metrics and section 5.3 presents the performance evaluation and discussions.

## 5.2 Simulation Assumptions

The system modeling presented in Chapter 2 was adopted for all simulations conducted in this thesis. Table 5.1 presents the main simulation parameters adopted during the simulations, which are aligned with the 3GPP LTE architecture [60, 30].

| Parameter | Value | Ref.'s |
|---|---|---|
| Maximum eNB transmit power ($P_t$) | 20 W and 12 W | [60] |
| eNB antenna radiation pattern | Three-sectored | [60] |
| Cell radius | 1 km | |
| UE speed | 3 km/h | [60] |
| Carrier frequency | 2 GHz | [60] |
| System bandwidth | 5 MHz and 3 MHz | [61] |
| Number of RBs ($K$) | 25 and 15 | [61] |
| Path loss model 1[a] | $15.3 + 37.6 \log_{10}(d)$ | [60] |
| Path loss model 2[a] | $34.5 + 35 \log_{10}(d)$ | [28] |
| Antenna Gain [b] | $G_h(\theta_h) + G_v(\theta_v)$ | [32] |
| Downtilt angle ($\phi^{\text{tilt}}$) | 8° | |
| Log-normal shadowing standard deviation | 8 dB | [60] |
| Small-scale fading | 3GPP Typical Urban | [30] |
| AWGN power per sub-carrier | -123.24 dBm | |
| Noise figure | 9 dB | |
| Link adaptation | Link level curves of LTE | [34] |
| CSI delay ($\Delta n$) | 0, 10, 20, 40 TTIs | |
| TTI duration | 1 ms | |
| Simulation duration | 30 seconds (30,000 TTIs) | |
| Number of repetitions | 150 | |
| Confidence interval | 95% | |

[a] $d$ is the distance between eNB and UE in meters.
[b] $\theta_h$ and $\theta_v$ are the horizontal and vertical angle with respect to the eNB, respectively.

Table 5.1 – General simulation parameters.

Notice that in table 5.1, there are two values for maximum eNB transmit power, system bandwidth and number of RBs. The first values are used in scenarios composed of CBR and video services, and the second values for the scenarios with CBR and VoIP services.

### *5.2.1 Scenarios and Evaluation Method*

The performance evaluation of the JSM and the four benchmark algorithms was carried out considering four different scenarios. Each scenario is described in the following:

(S-I). The first scenario considered is composed of CBR and VoIP services. For this scenario, the path loss model 2 was adopted, which is more severe than path loss model 1. In order to perform a complete evaluation, we simulate the two single service cases (i.e., only CBR (100% CBR + 0% VoIP) or only VoIP (0% CBR + 100% VoIP) services) and three multi-service cases, where different proportion between the services are considered. The proportions are explained in section 5.3.1.

(S-II). The second analysis was conducted in scenarios composed of CBR and video services. The path loss model 1 was adopted for this analysis only, which is less severe when compared to the other path loss model. The same approach for simulating single service and multi-service scenarios was employed here, where we simulate the two single service cases (i.e., only CBR (100% CBR + 0% video) or only video (0% CBR + 100% video) services) and different multi-service cases, namely: 75% CBR + 25% video, 50% CBR + 50% video, and 25% CBR + 75% video.

(S-III). The third scenario is similar to the (S-II), but considering the path loss model 2.

(S-IV). In the last one, we analyzed the impact of different values of CSI reporting delay. The mix of services was composed of CBR and video services. The same proportions of the (S-II) were simulated. For this last scenario, only the JSM and the two best benchmark algorithms from the previous scenarios were analyzed.

The metrics adopted to evaluate the different algorithms have been described in section 2.5, which are user satisfaction, fairness and total cell throughput.

In this thesis, after simulating all single and multi-service scenarios, the performance evaluation is conducted using the joint capacity plane, which is a complete form of evaluation since it simultaneously illustrates the algorithms' performance for single and multi-service scenarios. The joint capacity plane is a powerful tool to evaluate how well resource allocation algorithms perform in multiple services scenarios. This plane shows the system capacity regions that can be defined as the number of users for which predefined QoS levels are sustained for all service classes simultaneously.

The abscissa and ordinate points in the capacity plane represent the last load such that the algorithms are able of keeping the user satisfaction percentage above a threshold of 90% for

the single service cases. The joint system capacity (interior points of the capacity curves) are obtained from the multiple services scenarios. The joint system capacity is defined as the last load of users such that acceptable system-level quality (user satisfaction threshold of 90%) is sustained for both service classes simultaneously. Thus, in order to obtain all the points that compose the joint capacity plane, several single and multi-service simulations need to be conducted aiming to find the last load of users at which the satisfaction level of 90% is maintained.

Since we consider 90% as the satisfaction level to be achieved, in all figures showing the percentage of satisfied users, a horizontal line at the target satisfaction level is also depicted. Also, all the results in this thesis are presented with the 95% bootstrap confidence interval of the mean of the samples.

### *5.2.2 Settings for JSM and Benchmarking Algorithms*

This subsection describes the parameters values for the JSM and the benchmark algorithms. First of all, it is worthy mentioning that the complexity of the proposed algorithm is $\mathcal{O}(JK)$, which is same complexity of the benchmark algorithms, so the comparison conducted here compares algorithms with the same complexity. The EXP/PF [37] algorithm does not have any parameters to be set. For the QHMLWDF [23] algorithm, the $\alpha[n]$ parameter (maximum allowed probability of delayed packets) is set to $\alpha[n] = 0.1$, which reflects the fact that the QoS level to be maintained is 90%. The parameters of the MDU [11] and Lei [12] algorithms were adapted to our simulation environment keeping the same proportion of the original proposals. More specifically, the adjustments in the MDU and Lei utility functions were employed because the QoS requirements in this thesis are different from the values in [11] and [12].

The utility functions for the MDU algorithm were

$$\left|U_V^{'}(\mathrm{w})\right| = \begin{cases} \mathrm{w}[n], & \text{if } \mathrm{w}[n] \leq 5 \text{ ms} \\ \mathrm{w}[n]^{2.9} - 5^{2.9} + 5, & \text{if } \mathrm{w}[n] > 5 \text{ ms}, \end{cases} \tag{5.1}$$

the VoIP service,

$$\left|U_S^{'}(\mathrm{w})\right| = \begin{cases} \mathrm{w}[n]^2, & \text{if } \mathrm{w}[n] \leq 12.5 \text{ ms} \\ \mathrm{w}[n]^{1.9} - 13^{1.9} + 13, & \text{if } \mathrm{w}[n] > 12.5 \text{ ms}, \end{cases} \tag{5.2}$$

for the streaming (video) service and

$$\left|U_B^{'}(\mathrm{w})\right| = \begin{cases} \mathrm{w}[n]^{1.25}, & \text{if } \mathrm{w}[n] \leq 50 \text{ ms} \\ \mathrm{w}[n] - 50 + 50^{1.25}, & \text{if } \mathrm{w}[n] > 50 \text{ ms}. \end{cases} \tag{5.3}$$

best effort service (CBR service).

For the Lei algorithm, the utility function used in our simulations were

$$U_{\text{VoIP}}^{'}(d_j^{\text{hol}}[n]) = \frac{450 \cdot e^{-450 \cdot (d_j^{\text{hol}}[n] - 20)}}{\left(1 + e^{450 \cdot (d_j^{\text{hol}}[n] - 20)}\right)^2}, \tag{5.4}$$

for the VoIP service,

$$U'_{\text{Video}}(d_j^{\text{hol}}[n]) = \frac{250 \cdot e^{-250\cdot(d_j^{\text{hol}}[n]-50)}}{\left(1+e^{250\cdot(d_j^{\text{hol}}[n]-50)}\right)^2}, \tag{5.5}$$

for the streaming (video) service and

$$U'_{\text{CBR}}(d_j^{\text{hol}}[n]) = 5 \cdot 0.5 \cdot e^{5\cdot(d_j^{\text{hol}}[n]-200)}, \tag{5.6}$$

best effort service (CBR service).

Most of the parameters used for JSM have been discussed in the Chapter 4. The only parameter that has not been yet discussed is the $x_j^{\text{req}}$ value before the normalization for each single service policy. The function of the throughput-based user weight $w_j^{\text{thr}}$ is centered at $x_j^{\text{req}} = \text{TSM}_{\text{req}} = \Phi_{\text{req}}^{\text{thr}} = 512\,\text{kbps}$, which is the throughput requirement for the CBR service. The function of the delay-based user weight $w_j^{\text{delay}}$ is centered at $x_j^{\text{req}} = \text{DSM}_{\text{req}} = \Phi_{\text{req}}^{\text{delay}} = 20\,\text{ms}$, which is the delay requirement for the VoIP service. For the function of the queue-based user weight $w_j^{\text{queue}}$, we have chosen the center to be the queue size (in bits) related to 50% of the window size of video streaming, which we assumed as 2 seconds of video. Therefore, $x_j^{\text{req}} = \text{QSM}_{\text{req}} = 0.5 \cdot 2 \cdot 242\,\text{Kbps}$, which means that the proposed algorithm attempts to keep the transmitter buffer size with at most 1 second of video. Notice that the choice of the function center does not change the function shape presented in figure 4.1b because of the normalized framework.

## 5.3 Performance Evaluation

### *5.3.1 Case Study I*

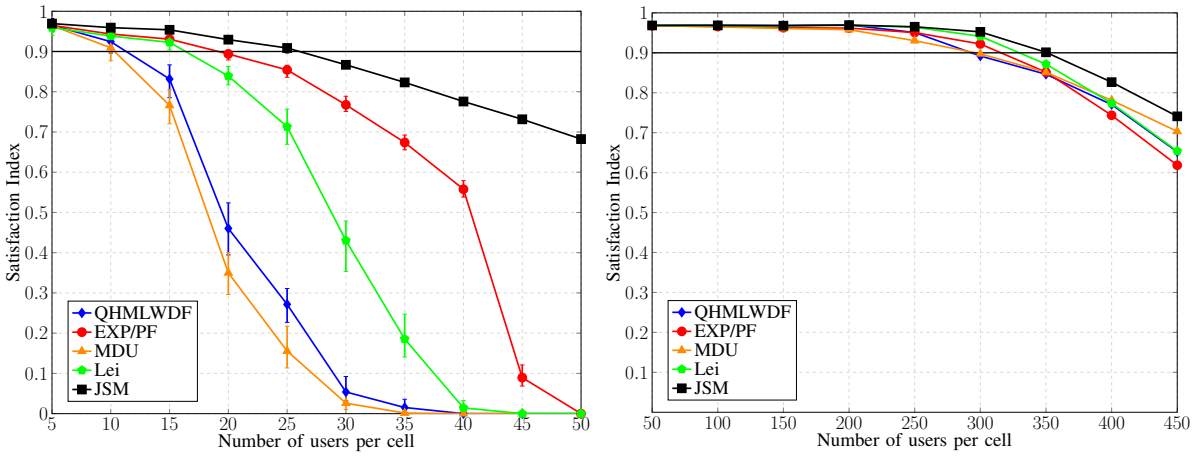
The first study case analyses a scenario composed of CBR and VoIP service mixes, as well as path loss model 2, which is more severe when compared to path loss model 1. Also, recall that for this scenario only, the eNB disposes of only 15 RBs and maximum transmit power of 12 W. This scenario was constructed because the traffic load generated by the VoIP traffic is very low (see table 2.3), thus too many VoIP users would be needed to see a performance decrease with the 25 RBs configuration.

Figure 5.1a presents the CBR single service case. Notice that since the light system loads, the JSM provides higher levels of user satisfaction. This happens because the proposed algorithm provides a more efficient distribution of the limited available resources (only 15 RBs available for transmission). The two worst performances were presented by the Lei and MDU algorithms, which were the first algorithms to have the satisfaction level dropping below 90%. The performances of the QHMLWDF, EXP/PF and JSM algorithms were similar up to 15 users, but from this point on the JSM presented a higher performance. Notice that when the system load increased, the JSM algorithm once more showed its stability and well designed utility functions by keeping the satisfaction level even higher than the other algorithms. The Lei, EXP/PF and

JSM algorithms presented the best performance because they consider directly in their priority the UEs' throughput, which is the main QoS metric that defines the satisfaction for CBR videos.

Figure 5.1 – Satisfaction index for the single services scenarios.

(a) Single service scenario with only CBR service. (b) Single service scenario with only VoIP service.



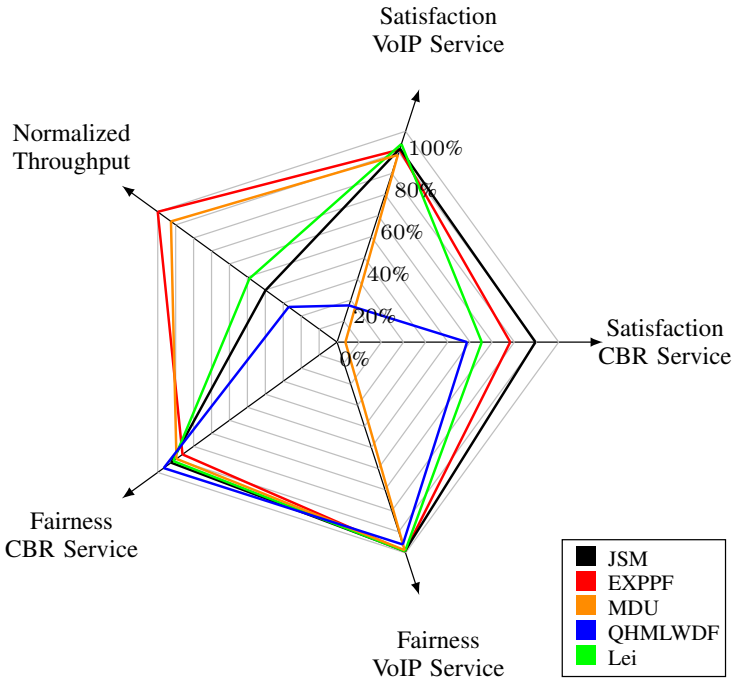

Source: Created by the author.

Figure 5.1b depicts the results for the VoIP single service case. Notice that the performance of all algorithms is very similar up to 200 VoIP users. The QHMLWDF and MDU algorithms presented the worst performances, which is explained by the fact that these algorithms do not primarily take into account the HOL packet delay during the resource allocation. The third best performance is presented by the EXP/PF algorithm, which considers the HOL packet delay of each UE and the mean delay value considering all active VoIP users. Finally, the two best performances are presented by the Lei and JSM algorithms, which both consider the HOL packet delay of users and utility functions. However, the more well-designed and normalized user utility functions employed by the JSM algorithm allows higher satisfaction levels than the Lei algorithm. Recall that the user utility functions employed by the JSM framework are designed to be unified across all policies, where the shape parameter has the same value for all policies. Thus, according to the results presented, the best approach to allocate delay-based services is to consider the HOL packet delay and well-designed utility functions during the resource allocation.

Figure 5.2 depicts a spider chart obtained after simulations of a scenario composed of 200 VoIP users and 14 CBR users. Each axis presents a different metric for comparison: fairness (based on the well-known Jain's fairness index [41]) and satisfaction index for each service, and the normalized total cell throughput. The JSM algorithm presented the most stable performance since it can balance the satisfaction of both services without penalizing one over the other. Notice that this behavior is achieved without losing performance with respect to fairness during the re-

source allocation. However, the normalized total cell throughput achieved by the JSM algorithm was the second worse among all algorithms. This happened due to the fact that JSM prioritizes the VoIP service over the CBR service, and the VoIP service generates much less traffic than the CBR service (as can be seen by comparing tables 2.2 and 2.3). Therefore, by prioritizing the VoIP users during transmission, less bits are transmitted and the total cell throughput diminishes at a cost of preserving the VoIP satisfaction. The results showed for this spider chart are accomplished not only for the scenario presented, but also in all other traffic mixes and loads, which will be presented next. The Lei and EXP/PF algorithms keep the satisfaction of the VoIP service above 90%, but the satisfaction for the CBR service is highly penalized, while JSM also keeps the CBR satisfaction above 90%. Thus, only JSM presents a satisfactory commitment between the satisfaction of both services.

Figure 5.2 – Spider chart of mix composed of 200 VoIP users and 14 CBR users.




Source: Created by the author.

The multi-service scenarios for this case study are constructed based on the results of the single service cases. As can be seen in figure 5.1, the number of VoIP and CBR users supported is highly different, thus making it unfeasible to consider proportions such as 75% CBR + 25% VoIP, 50% CBR + 50% VoIP, and 25% CBR + 75% VoIP in scenarios with 400 users, for example. Therefore, the approach adopted for this scenario is to fix the amount of VoIP users and increase the load of CBR users in the system to identify the number of VoIP and CBR users supported considering the satisfaction of both services above 90%. The fixed amount of VoIP users chosen were 150, 200 and 250 users, which are loads for which all algorithms are able to keep the satisfaction above 90% for the VoIP single service scenario (see figure 5.1b). Then, the amount of CBR users is increased for each fixed amount of VoIP users, so that the interior

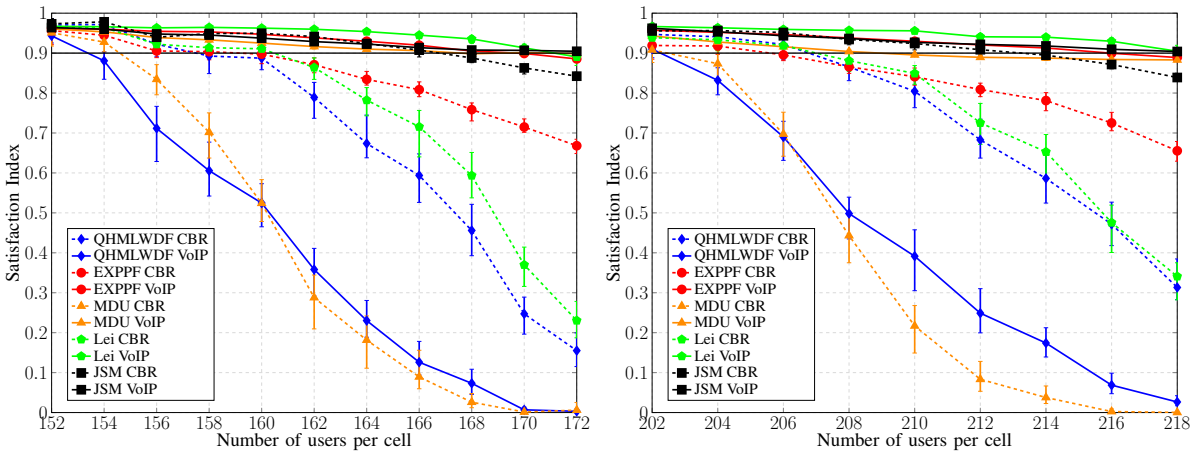
points of the capacity plane are obtained.

In figure 5.3a, the satisfaction index as a function of the total number of users (VoIP plus CBR users) in the system is depicted for the mix composed of 150 VoIP users. The best performance in this traffic mix is achieved by the JSM, supporting 167 users (150 VoIP users plus 17 CBR users) above the 90% satisfaction threshold. Notice that the performance of JSM is limited by the CBR service, which is the service with lower priority. It is worth mentioning that the satisfaction level of the VoIP service, which is the protected (with higher priority) service, is maintained above 90% for all system loads by the JSM algorithm, which is accomplished by the adaptation of the shape parameter ($\lambda$ value) performed by the proposed algorithm. This

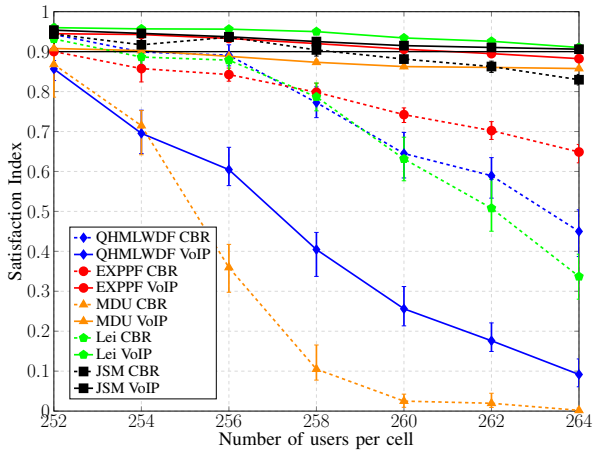
Figure 5.3 – Satisfaction index for different traffic mixes composed of CBR and VoIP services.

(a) 150 VoIP users plus increasing number of CBR users.

(b) 200 VoIP users plus increasing number of CBR users.




(c) 250 VoIP users plus increasing number of CBR users.




Source: Created by the author.

behavior is also presented for the other two traffic mixes, as can be seen in figures 5.3b and 5.3c, where the JSM algorithm always keeps the satisfaction of the video service above the threshold of 90%. The Lei algorithm presents the second best performance in terms of the load to which the satisfaction of one of the services drops below 90%, having performance very close to the EXP/PF algorithm. The performance of the MDU and QHMLWDF algorithms were the worst, which reflects their worst performance obtained in the single service cases.

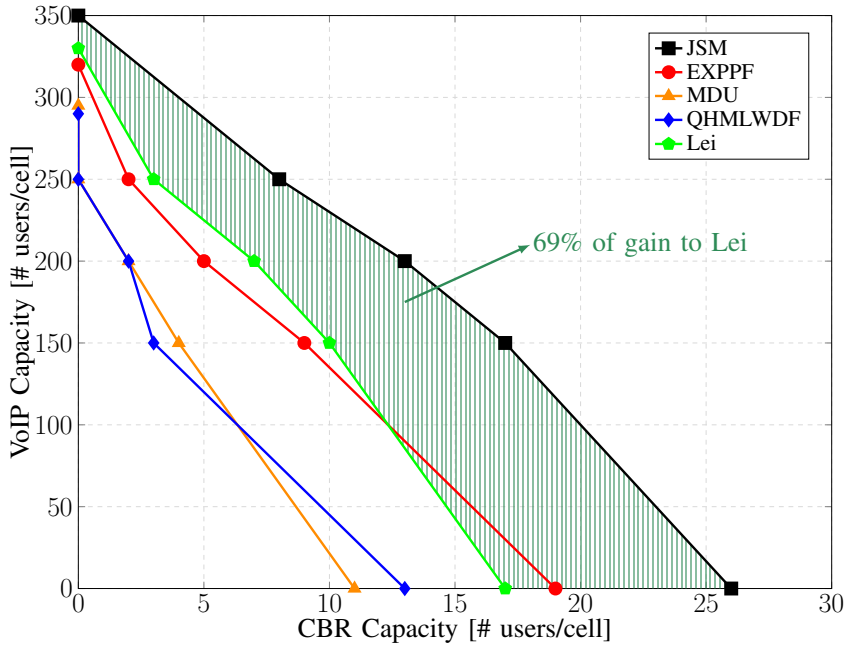
An important point to be highlighted is that all the benchmark algorithms have a static service prioritization, i.e., there is no adaptation in the service priority as in JSM. This is reflected by the fact that the satisfaction level of the most prioritized service is always kept as high as possible, but the satisfaction percentage of the other service dramatically decreases. For example, look at the results presented by the Lei and EXP/PF algorithms in figures 5.3b and 5.3c, where in the final system load the satisfaction of the VoIP service dropped below 90%. Also, the performance of the QHMLWDF algorithm is quite different from the others. For this specific mix (CBR plus VoIP), the satisfaction of the CBR service is always higher than the satisfaction of the VoIP service. This is not common since the delay sensitive services (such as the VoIP service) are usually prioritized over the others. The reason for this behavior is that the QHMLWDF algorithm considers the transmit buffer queue size during the allocation (see eq. 3.11). Since the CBR service generates more traffic than the VoIP service, the transmit buffer size for this service is always higher, thus yielding higher priority for the CBR service, which causes this different behavior.

Now, we present the joint capacity plane, which shows the system capacity regions that are defined as the number of users for which the satisfaction is kept above 90% for all service classes simultaneously. The larger the area below the capacity curve, the higher the number of satisfied users respecting the minimum satisfaction limit of 90% for both services. Therefore, the main result seen in figure 5.4 is that the proposed JSM algorithm significantly enhances the overall system capacity when compared with the benchmark algorithms. In order to calculate the total system capacity gain, the area under each capacity curve was calculated for each algorithm. The gains in terms of area below the capacity curve were of 69.37% with respect to Lei algorithm and 75.7% with respect to the EXP/PF algorithm. The shaded region in figure 5.4 highlights the capacity gain with respect to the Lei algorithm, which achieved the second best performance. The gains obtained by JSM occur mainly due to the fact the proposed algorithm is able to more efficiently exploit the frequency diversity provided by the path loss model 2, besides the balancing between the satisfaction of both services in the system and protection of one most prioritized service by adapting its service priority function.

### *5.3.2 Case Study II*

The second scenario considers CBR and video service mixes, as well as path loss model 1. In figure 5.5, we present the satisfaction index (percentage of satisfied users) as a function of

Figure 5.4 – Joint capacity plane showing the system capacity regions for different traffic mixes for JSM and benchmark algorithms.




Source: Created by the author.

the number of users in the system for the single service cases.

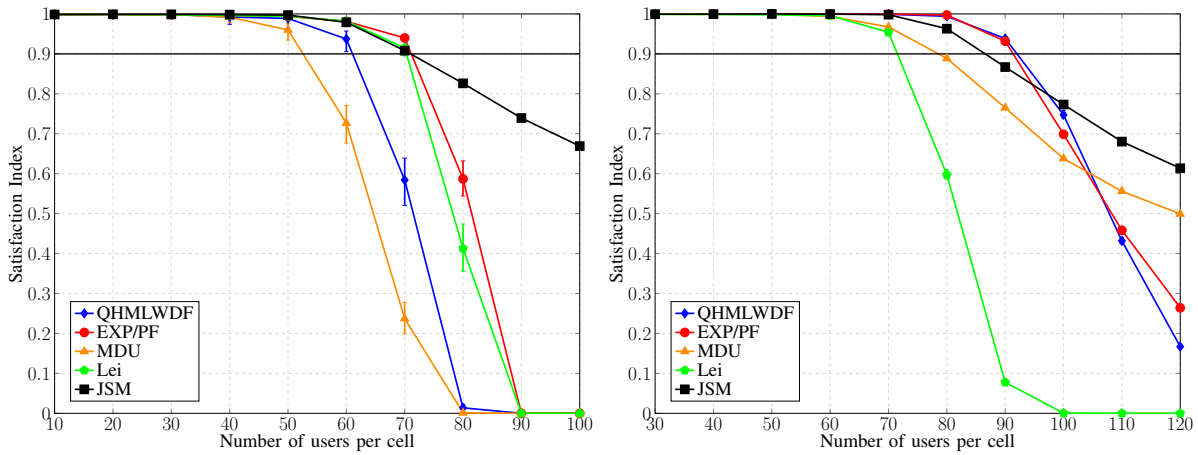
In figure 5.5a, the CBR single service case is presented. Notice that for light system loads, all algorithms achieved similar performances. However, as the system load becomes heavier, the performance of some algorithms started to deteriorate more rapidly. Again, the two worst performances were presented by the MDU and QHMLWDF algorithms. Then, the performance of the Lei, EXP/PF and JSM algorithms were similar up to the point where their satisfaction levels dropped below 90%, which is around 70 users. However, one can see that as the system load becomes even heavier, the JSM algorithm presented higher satisfaction levels due to its stability and well-designed utility functions.

Figure 5.5b depicts the results for the video single service case. Once again, for light system loads, all algorithms achieved similar performances. The two worst performances were presented by the Lei and MDU algorithms. The performances of the QHMLWDF, EXP/PF and JSM algorithms were very similar, but the system load where the satisfaction level dropped below 90% for the JSM algorithm was lighter. However, when the system load increased, the JSM showed again its stability and well designed utility functions by keeping the satisfaction level even higher than the other algorithms. The JSM and QHMLWDF algorithms presented satisfactory results because they consider the transmit buffer queue size when allocating the resources. The EXP/PF algorithm keeps the fairness between the users, which yields good satisfaction levels up to a given point; however due to its opportunistic behavior, when the system load increases, the satisfaction level decreases rapidly. The MDU algorithm also considers the

queue size during resource allocation, but this algorithm does not change the users' priority as the queue size increases (see figure 3.1), which deteriorates the satisfaction of users with high values of queue size.

Figure 5.5 – Satisfaction index for the single services scenarios.

(a) Single service scenario with only CBR service. (b) Single service scenario with only video service.

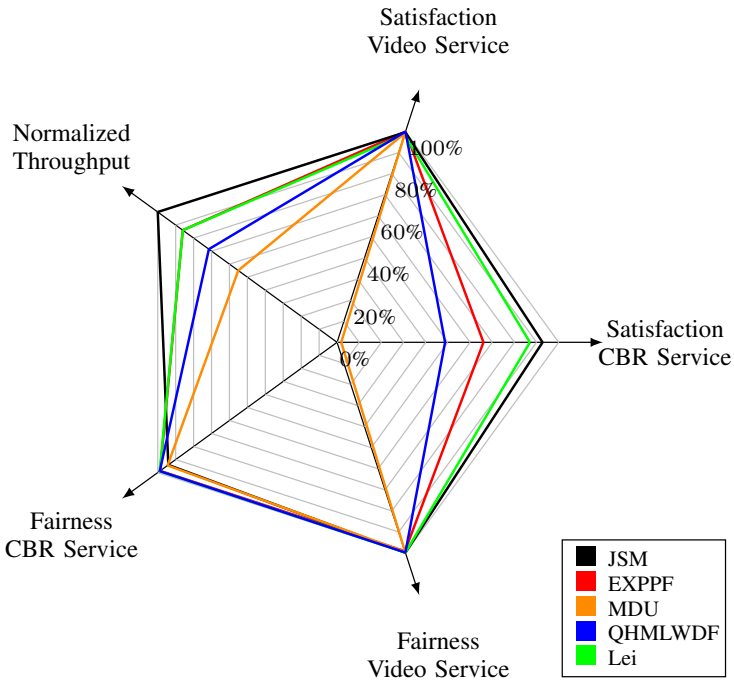



Source: Created by the author.

Figure 5.6 depicts a spider chart obtained after simulations of a scenario with the mix 50%CBR and 50%video composed of 70 users in total, thus 35 users of each service. The approach adopted in (S-I) was applied here, where each axis presents a different metric for comparison: fairness (based on the well-known Jain's fairness index) and satisfaction index for each service, and the normalized total cell throughput. The JSM algorithm presented the most stable performance since it can balance the satisfaction of both services without penalizing one over the other. Notice that this behavior is achieved without losing performance with respect to fairness during the resource allocation and total cell throughput. This is accomplished not only for the scenario presented in this spider chart, but also in all other traffic mixes and loads, which will be presented next. The benchmark algorithms were able to keep the satisfaction of the video service above 90%, but the satisfaction for the CBR service is highly penalized, while JSM also keeps the satisfaction of the CBR service above 90%. Therefore, only JSM presents a fair commitment between the satisfaction of both services.

Widening our analysis, let us now present other mixes of services and system loads. For multi-service scenarios, we need to check the last system load to which the satisfaction level of both services are above 90%. In figure 5.7a, the satisfaction index for the mix 25%CBR/75%video is depicted. One can see that the best performance in this mix is presented by the JSM, supporting 72 users above the 90% satisfaction threshold. Notice that the performance of JSM is limited by the CBR service, which is the service with lower priority. It is important to highlight that the

Figure 5.6 – Spider chart of mix 50%CBR and 50%video composed of 70 users in total.




Source: Created by the author.

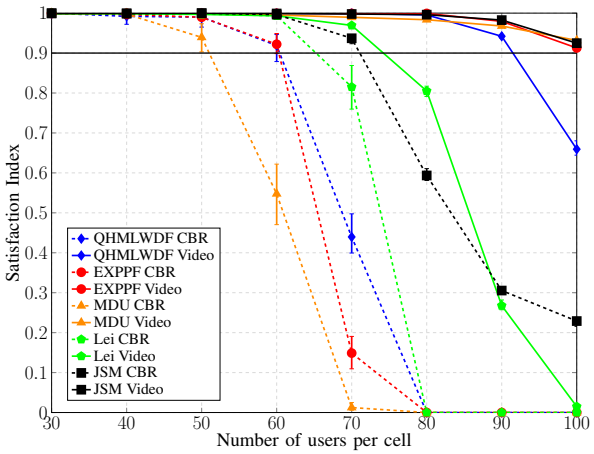
satisfaction level of the video service, which is the protected (with higher priority) service, is maintained above 90% for all system loads. This is achieved by the adaptation of the $\lambda$ parameter performed by the proposed algorithm. This behavior is also presented for the other two traffic mixes, as can be seen in figures 5.7b and 5.7c, where JSM algorithm always keeps the satisfaction of the video service above the threshold of 90%. Notice in figure 5.7c that the last load with satisfaction above 90% is similar for the JSM, EXP/PF and Lei algorithms, but when the system load increases, only JSM is able to control and keep the satisfaction level of the CBR service still high. The Lei algorithm presents the second best performance in terms of the load to which the satisfaction of one of the services drops below 90%. However, notice that when the system load increases, the satisfaction for both services dramatically decreases, which does not happen with the JSM algorithm due to the video service protection. The performance of the MDU algorithm was the worst, which reflects its worst performances obtained in the single service cases.

The same point mentioned for the previous scenario is applied here, where all the benchmark algorithms have a static service prioritization, i.e., there is no adaptation in the service priority as in JSM. This is reflected by the fact that the satisfaction level of the statically most prioritized service is always kept as high as possible, but the satisfaction percentage of the other service dramatically decreases. For example, look at the results presented by the Lei and EXP/PF algorithms in figure 5.7a. The satisfaction of the video service for light traffic loads is maintained always above 90%. However, the satisfaction level of the CBR users for loads above 80 users is equal to 0%, i.e., the satisfaction of this service was highly degraded. A similar behavior is also presented by the Lei and EXP/PF algorithms in figures 5.7b and 5.7c, and by the other bench-

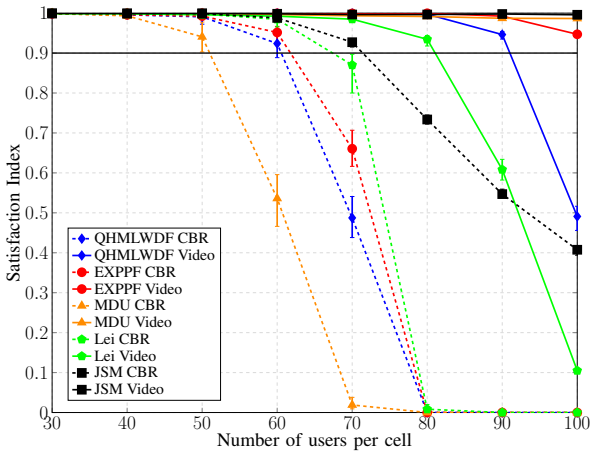
mark algorithms in all traffic mixes. This shows that the benchmark algorithms do not present a satisfactory commitment between the satisfaction of all services present in the system, as done by JSM.

Figure 5.7 – Satisfaction index for different traffic mixes composed of CBR and video services.

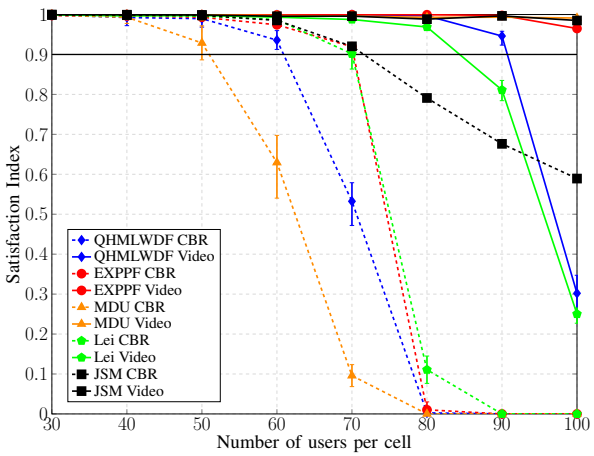
(a) Mix 25%CBR and 75%video (c25v75).




(b) Mix 50%CBR and 50%video (c50v50).




(c) Mix 75%CBR and 25%video (c75v25).

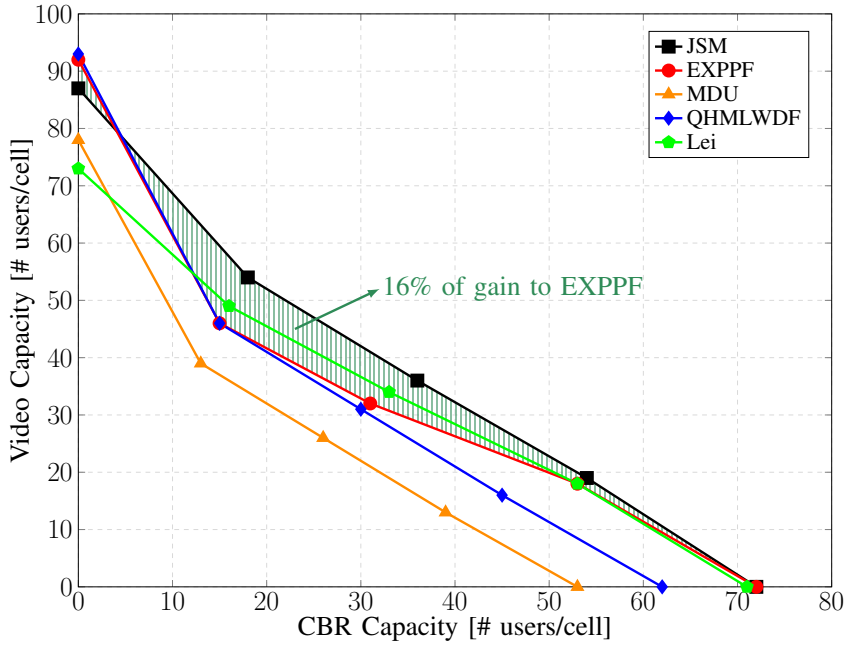



Source: Created by the author.

Figure 5.8 presents a joint capacity gain similar to the one presented in the (S-I). The larger the area below the capacity curve, the higher the number of satisfied users respecting the minimum satisfaction limit of 90% for both services. Therefore, the main result seen in the joint capacity plane shown in figure 5.8 is that the JSM is able to increase the overall system capacity when compared with the benchmark algorithms. Notice that the gain presented in this scenario is not very high, which is due to the fact that path loss model considered yields good channels condition for most users, thus the Lei and EXP/PF algorithms presented similar performances to JSM. The reason for the best performance of JSM is that this algorithm balances the satisfaction

of both services in the system and protects the satisfaction level of one service by adapting its service priority function. It is worth noting that the highest gains were obtained when video was the dominant service (the mix c25v75), which can be considered a commonly found realistic scenario because of the increasing number of video users present in recent networks. The second and third best performances were achieved by the EXP/PF and Lei algorithms, respectively. The gains in terms of area below the capacity curve were of 16.09% with respect to EXP/PF and 16.9% with respect to the Lei algorithm. The shaded region in figure 5.8 highlights the gain with respect to the EXP/PF algorithm, which achieved the second best performance. However, recall that in figures 5.7a, 5.7b and 5.7c, the JSM algorithm was able to keep the satisfaction index of the video service always above 90% by protecting its satisfaction. On the other hand, even though the the EXP/PF and Lei algorithms presented the second and third best performances in terms of the load to which the satisfaction of one of the services drops below 90%, the satisfaction index for both services significantly decreased for high system loads for these algorithms. Therefore, the gains obtained by the JSM algorithm in the capacity plane are even higher if we consider this fact. Furthermore, this shows that only JSM presents a satisfactory commitment between the satisfaction of both services, which is more desirable for network operators.

Figure 5.8 – Joint capacity plane showing the system capacity regions for different traffic mixes for JSM and benchmark algorithms.




Source: Created by the author.

### *5.3.3 Case Study III*

The third scenario analyses the same service mix of (S-II), but now considering the path loss model 2, which is more severe compared to path loss model 1. This severity can be seen by

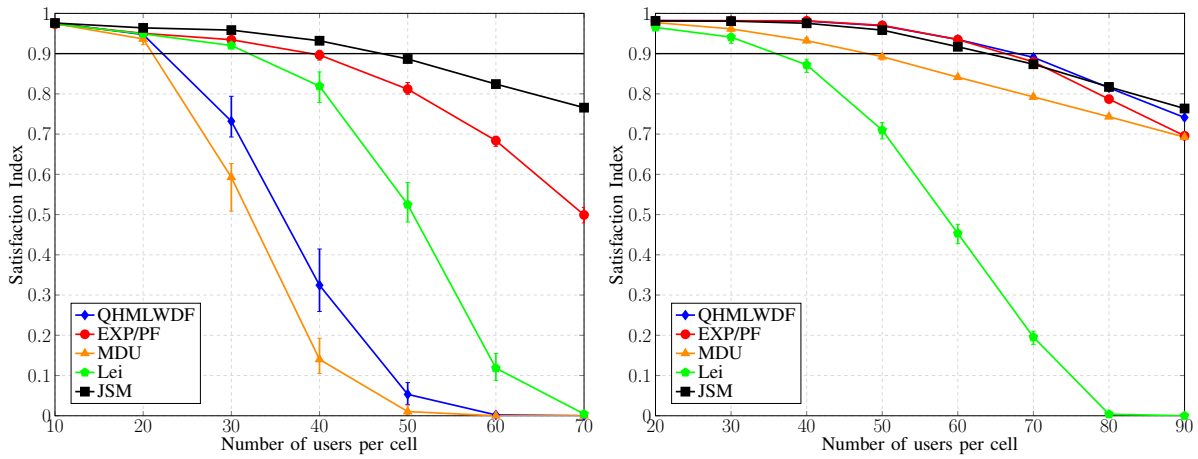
the number of satisfied users in the abscissa axises, which is always lower for the scenarios in this section.

In figure 5.9a, the CBR single service case is presented. Firstly, one can notice by the abscissa axis, that the system load to which the algorithms keep the satisfaction above 90% is lower for this study case, which reflects the severity of the considered path loss model. This behavior can be seen in all results for this study case. Differently from what was obtained in (S-II), in the CBR single service scenario the performance of the JSM algorithm is clearly better than the benchmark algorithms. Recall that in figure 5.5a, the JSM, EXP/PF and QHMLWDF algorithms presented very similar performances. However, with the more severe path loss, the JSM performs better even for low system loads, and as the system load increases, the performance gap increases. This shows that the resource allocation decision performed by the proposed algorithm is able to explore more efficiently the frequency diversity due to its utility function shape and considered QoS metric.

Figure 5.9b presents the results for the video single service case. Once again, the performance of the JSM algorithm presented better relative improvements compared to the benchmark algorithms. In figure 5.5b, the point where the satisfaction dropped below 90% for the JSM algorithm was considerable lower than the QHMLWDF and EXP/PF algorithms. In figure 5.9b, there is a technical draw considering this point of interest. However, we can see by the load of 90 users that, as the load increases, the performance of EXP/PF and QHMLWDF decreases more rapidly than the performance of JSM.

Figure 5.9 – Satisfaction index for the single services scenarios.

(a) Single service scenario with only CBR service. (b) Single service scenario with only video service.

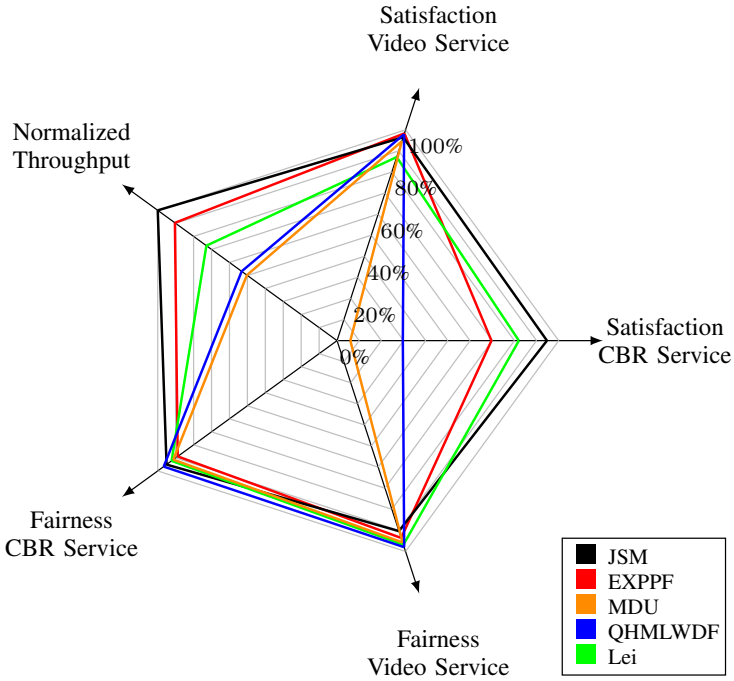



Source: Created by the author.

Figure 5.10 depicts a spider chart obtained after simulations of a scenario with the mix 25%CBR and 75%video composed of 40 users in total, thus 10 CBR users and 30 video users.

Again, the JSM algorithm presents the most stable performance since it can balance the satisfaction of both services without penalizing one over the other. Notice that this behavior is achieved without losing performance with respect to fairness during the resource allocation and total cell throughput. The same behavior seen for (S-II) can be observed here, where the benchmark algorithms are able to keep the satisfaction of the video service above 90%, but the satisfaction for the CBR service is highly penalized, while JSM also keeps the satisfaction of the CBR service above 90%. Therefore, only JSM presents a satisfactory commitment between the satisfaction of both services.

Figure 5.10 – Spider chart of mix 25%CBR and 75%video composed of 40 users in total.




Source: Created by the author.

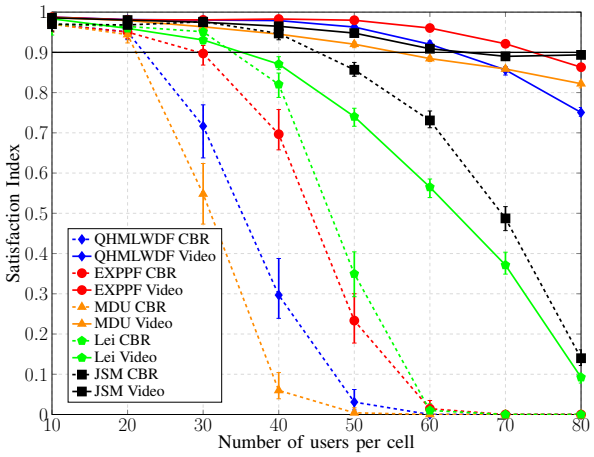
In figure 5.11a, the satisfaction index for the mix 25%CBR/75%video is depicted. Notice again that the best performance in this mix is presented by the JSM, supporting 46 users above the 90% satisfaction threshold. Once again, the performance of JSM is limited by the CBR service, which is the service with lower priority. The satisfaction level of the video service, which is the protected (with higher priority) service, is maintained above 90% for all system loads, which is achieved by the adaptation of the $\lambda$ parameter. This behavior is also presented for the other two traffic mixes, as can be seen in figures 5.7b and 5.7c, where the JSM algorithm always keeps the satisfaction of the video service above the threshold of 90%. The Lei algorithm presents the second best performance in terms of the load to which the satisfaction of one of the services drops below 90%. However, notice that when the system load increases, the satisfaction for both services dramatically decreases, which does not happen with the JSM algorithm due to the protection of the video service. Considering the more severe path loss, we can see that the gains obtained by the JSM algorithm were considerable higher, showing also for the multi-service

scenario that the proposed solutions better exploits the frequency diversity of the system.

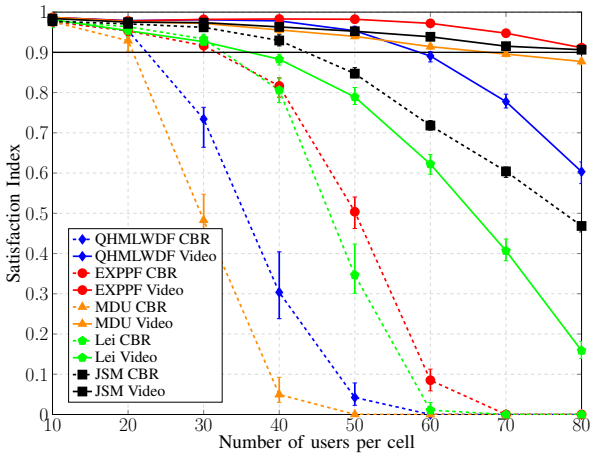
Looking again at a specific case in figure 5.11a, for the load of 50 users, the satisfaction for the CBR service obtained by the JSM algorithm is higher than 80%. However, the CBR satisfaction levels for the EXP/PF and Lei algorithms are below 40%, showing a performance gain of JSM higher than 100%. This shows that the benchmark algorithms do not present a satisfactory commitment between the satisfaction of all services present in the system, as performed by JSM.

Figure 5.11 – Satisfaction index for different traffic mixes composed of CBR and video services.

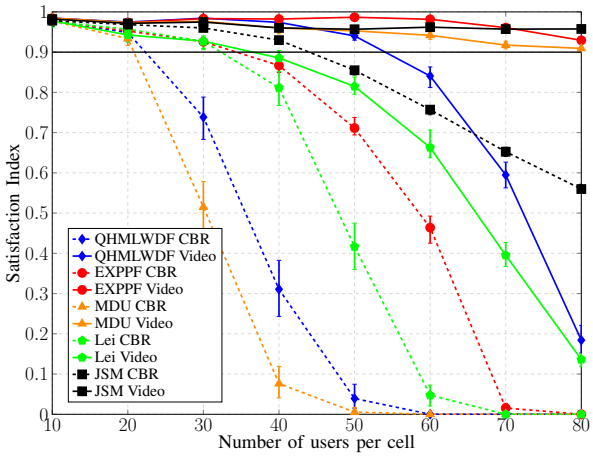
(a) Mix 25%CBR and 75%video (c25v75).




(b) Mix 50%CBR and 50%video (c50v50).




(c) Mix 75%CBR and 25%video (c75v25).

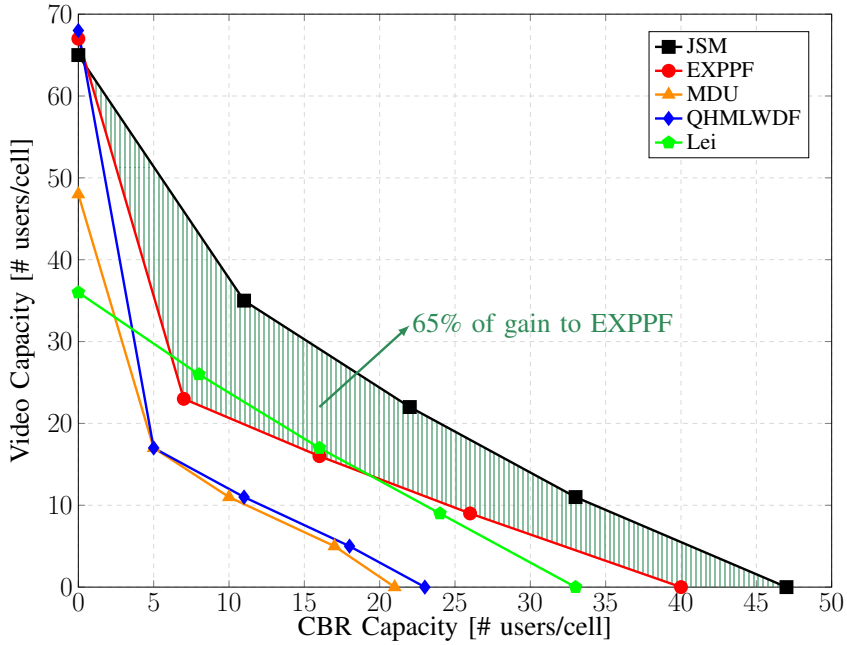



Source: Created by the author.

Figure 5.12 presents a joint capacity gain similar to the one presented for (S-II). Notice again that JSM is able to increase the overall system capacity when compared with the benchmark algorithms. However, this time, the gain is considerable higher if compared to the (S-II).

Once again, the area under each capacity curve was calculated for each algorithm in order to calculate the total system capacity gain. The gains in terms of area below the capacity curve were of 65.36% with respect to EXP/PF and 98.76% with respect to the Lei algorithm. The shaded region in figure 5.12 highlights the gain with respect to the EXP/PF algorithm, which achieved the second best performance. This is mainly due to the fact the proposed algorithm is able to more efficiently exploit the frequency diversity, besides the balancing between the satisfaction of both services in the system and protection of one most prioritized service by adapting its service priority function. It is worth noting that the highest gains were obtained when video was the dominant service (the mix c25v75), which can be considered a commonly found realistic scenario because of the increasing number of video users present in recent networks. The second and third best performances are achieved again by the EXP/PF and Lei algorithms. However, recall that in figures 5.11a, 5.11b and 5.11c, the JSM algorithm was able to keep the satisfaction index of the video service always above 90% by protecting the satisfaction of this service. On the other hand, even though the Lei and EXP/PF algorithms present similar performances in terms of the load to which the satisfaction of one of the services drops below 90%, the satisfaction index for both service significantly decreased for high system loads for these algorithms. Therefore, the gains obtained by the JSM algorithm in the capacity plane are even higher if we consider this fact.

Figure 5.12 – Joint capacity plane showing the system capacity regions for different traffic mixes for JSM and benchmark algorithms.




Source: Created by the author.

### *5.3.4 Case Study IV*

The performance of the three best algorithms from the previous scenarios are analyzed in this case study for imperfect CSI estimation at the transmitter. The path loss model 2 is adopted and the service mix is composed of CBR and video services. The imperfection analyzed here relates only to eq. 2.8, which means that the CSIs used by the resource allocation algorithms are outdated by $\Delta n$ TTIs. The users are able to estimates their CSIs perfectly (i.e., no error is inserted in the CSI), thus the algorithms use the exact CSI value, but outdated by $\Delta n$ TTIs. The values of $\Delta n$ are 0, 10, 20 and 40 TTIs. Notice that the results from section 5.3.3 were obtained for $\Delta n = 0$, thus these results are also used here.

In all figures in this section, the performance of each algorithm is represented by a different color and line type (solid, dashed or dotted), while the different values of CSI reporting delay are differentiated by distinct markers on the performance curves.

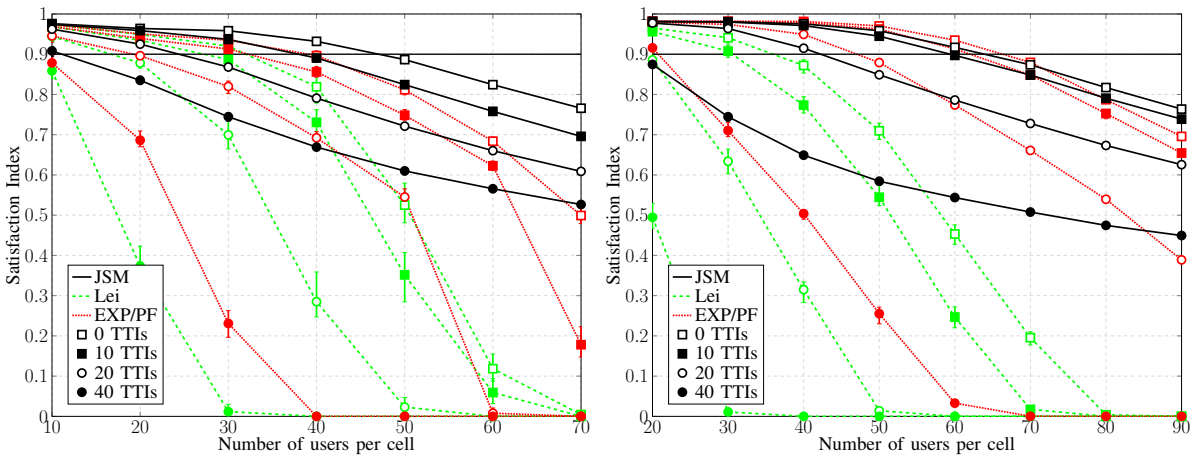
In figures 5.13a and 5.13b, the satisfaction indexes for the CBR and video single service cases are presented, respectively. It can be noticed that the performances of all algorithms are negatively impacted by the CSI imperfections, which is expected because transmission errors occur when the eNB assigns a certain data rate to a user based on a given CSI that cannot be supported by the true channel state conditions. However, one can see that the performance of the JSM algorithm is less sensitive to the increase in the CSI reporting delay. As an example, when the CSI delay was of 40 TTIs in figure 5.13a, only the JSM was able to keep the satisfaction above 90% for 10 users. Also, for 40 TTIs of CSI delay in figure 5.13a, the satisfaction index for more than 40 users was 0% for the EXP/PF and Lei algorithms, while JSM maintained the satisfaction above 50% even for this very severe CSI imperfection. A similar behavior is also presented in figure 5.13b. The performance of the EXP/PF is highly impacted by the CSI imperfection due to its opportunistic characteristics, which explains the fact that the performance degradation gap presented by EXP/PF increases as the CSI reporting delay increases. The Lei algorithm has already presented the worse performance among these algorithms in the scenario with perfect CSI. Notice that the performance degradation gap presented by Lei algorithm is similar to the one presented by EXP/PF, while JSM presented a more stable degradation gap as the CSI delay increased.

The same service mixes presented in section 5.3.3 were simulated for analyzing the CSI imperfections. However, the satisfaction of the CBR and video services are presented in separate graphs due to the high amounts of curves.

In figures 5.14a and 5.14b, the satisfaction index for the CBR and video services are presented considering the service mixes composed of 25%CBR and 75%video. For the scenario with perfect CSI information, this service mix was the one where JSM presented the highest gain. The same relative gains were maintained even under CSI imperfection since the JSM algorithm balances the satisfaction of both service by performing a service priority adaptation. Notice in

Figure 5.13 – Satisfaction index for the single services scenarios for different values of CSI delay.

(a) Single service scenario with only CBR service. (b) Single service scenario with only video service.



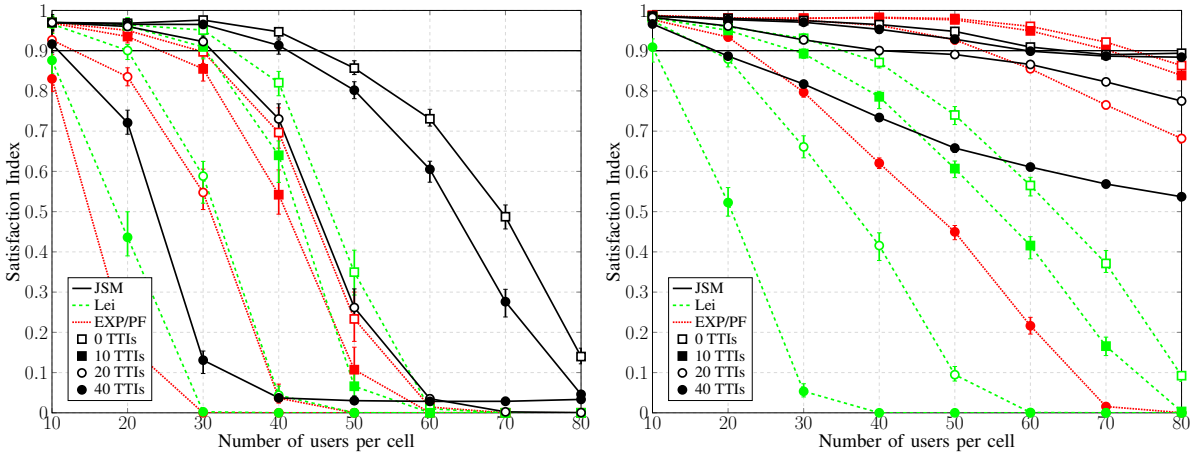

Source: Created by the author.

figure 5.14b that even with 10 TTIs of CSI delay reporting delay, the JSM algorithm was able to keep the satisfaction of the video service very close to 90% for high system loads. As seen for the perfect CSI scenario, the satisfaction performance of the Lei algorithm severely drops for both services, while for the EXP/PF the performance of the video service is maintained as high as possible at a cost of severely penalizing the CBR satisfaction.

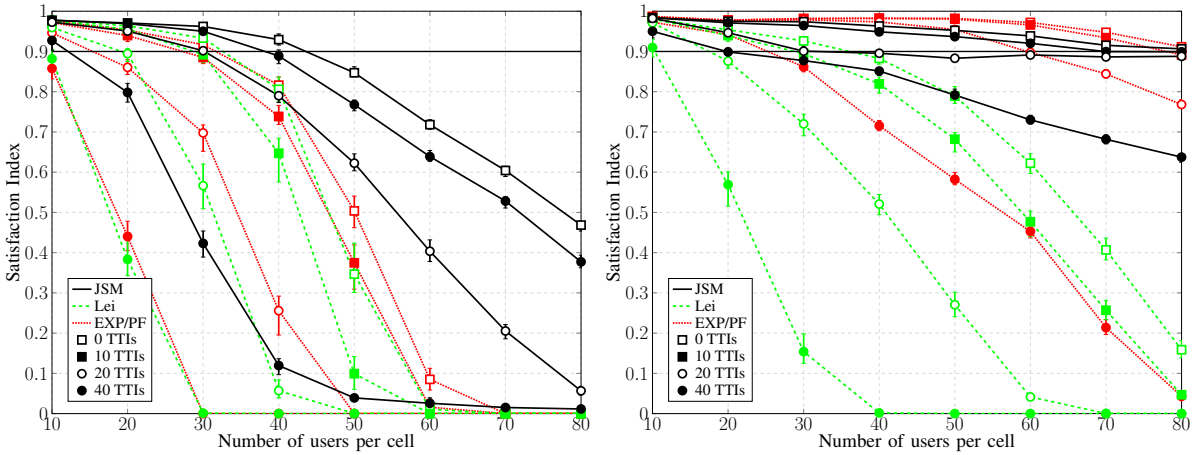
In figures 5.14c and 5.14d, the satisfaction index for the CBR and video services are presented considering the service mixes composed of 50%CBR and 50%video. The same behavior presented in the service mix 25%CBR-75%video can be seen for the mix 50%CBR and 50%video. Furthermore, notice that since there are less video users in this mix, the JSM algorithm is able to maintain the video satisfaction very close to 90% for higher CSI reporting delays. Also, in figures 5.14e and 5.14f, where the mix is composed by 75%CBR-25%video, the JSM algorithm maintains the video satisfaction very close to 90% for all analyzed CSI reporting delays, which not achieved by any of the benchmark algorithms.

Figure 5.14 – Satisfaction index for different traffic mixes and CSI delays.

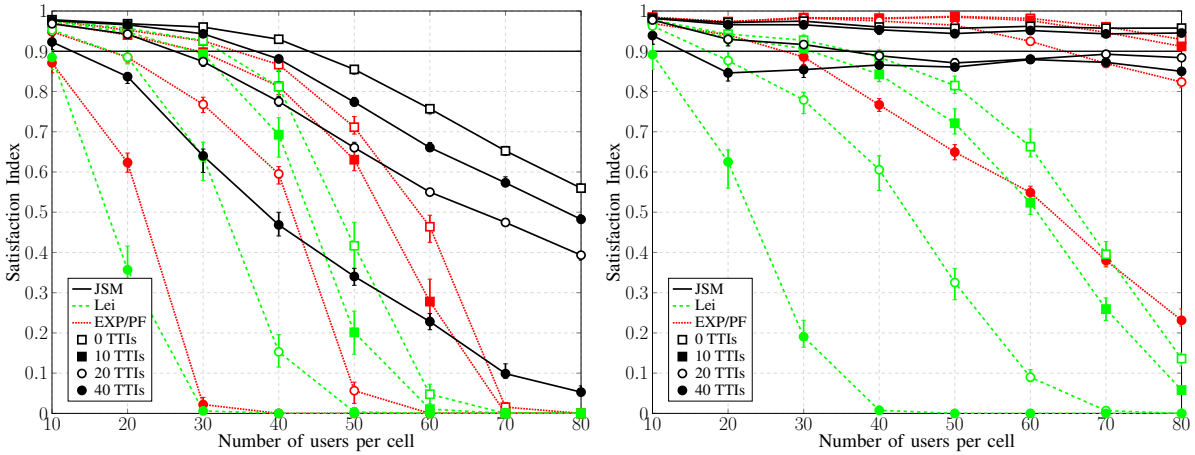
(a) CBR satisfaction for mix 25%CBR-75%video. (b) Video satisfaction for mix 25%CBR-75%video.




(c) CBR satisfaction for mix 50%CBR-50%video. (d) Video satisfaction for mix 50%CBR-50%video.




(e) CBR satisfaction for mix 75%CBR-25%video. (f) Video satisfaction for mix 75%CBR-25%video.




Source: Created by the author.

# 6 CONCLUSIONS AND FUTURE WORK

In this master's thesis, we have studied a utility-based optimization problem that targets the maximization of the total utility (user satisfaction) in wireless networks composed of multiple services. The optimization problem has constraints assuming that the available resources are discrete and cannot be shared by two or more user at a given transmission time. The studied problem belongs to the research fields of QoS provisioning and RRA.

After formulating the optimization problem, it has been found that its optimum solution is very hard to be found. To overcome this problem, we have adopted a commonly applied problem-splitting technique where firstly a DRA is performed with fixed power allocation; then, there is a stage of EPA with fixed resource assignment. The problem was then reformulated for dealing with delay-, throughput- and queue-based services. After the reformulation, the problem was mathematically manipulated and reduced to a simplified optimization problem which is linear in terms of the instantaneous users' data rate, from which a unified low complexity sub-optimum was derived. The unification of the solution found relates to the fact that the same formulation is applied for all services, regardless of their main QoS metric.

The analyzed optimization problem can be employed in any current or future cellular system. In this master's thesis, we have proposed a heuristic resource allocation framework for application in air interfaces based on OFDMA. However, it is worth noting that even though the study performed in this thesis considered an OFDMA-based system, the same analysis could be conducted for any current and future multiple access scheme that guarantees orthogonality among the resources.

The resource allocation framework developed herein is composed of user weights and an innovative service weight which has a parameter adapted to meet the satisfaction target of the most prioritized service chosen by the network operator. The user utility functions employed by the proposed framework, called JSM, are designed to be unified across all policies, where the shape parameter has the same value for all considered policies. The only difference in the user utility functions for different services is the derivative (positive or negative), which depends on the intrinsic characteristics of the service. Furthermore, a QoS metric normalization is performed, so that our framework becomes independent of the QoS metric being considered.

The performance of the proposed JSM algorithm was then analyzed and compared to four benchmark algorithms in four different scenarios considering different service mixes and coverage situation. In the first case study, where a low coverage scenario was considered and the service mix was composed of CBR and VoIP services, the JSM significantly improved the system capacity providing gains of 69%. In the second scenario, where we had a mix of CBR and video services as well as a good coverage scenario, the gains provided by the JSM algorithm

were of 16% in the system capacity plane. Then, in the third case study, which was similar to the second one, but considering a low coverage scenario, the gain went from 16% to 65%, showing that the proposed algorithm is able to more efficiently exploit the frequency diversity. Furthermore, in the fourth scenario we analyzed CSI imperfect estimation at the transmitter. Considering this imperfection, all algorithms have been negatively impacted, however the relative results did not considerably change from the ones seen in the other scenarios, where the JSM significantly outperformed the benchmark algorithms.

The main reason for the significant gains obtained by the JSM algorithm is the adaptation of the service priority. This is performed by adapting the shape parameter of the service utility function. By means of this adaptation, the JSM algorithm provides a stable performance by balancing the satisfaction of both services without penalizing one over the other and guaranteeing that the satisfaction level of the protected (with higher priority) service is maintained above 90% for all system loads. To the best of our knowledge, this specific feature has never been addressed in the literature. Furthermore, the proposed algorithm presented a very stable performance degradation as the system load increased, which is also a feature desired by network operators. Finally, it is worthy mentioning that the complexity of the proposed algorithm is $\mathcal{O}(JK)$ (same complexity of the benchmark algorithms used), so the gains were achieved without increasing the complexity compared to the benchmark algorithms.

As perspective of future works for further developing the work initiated in this master's thesis, we have:

- **Multiple antennas:** the scenarios analyzed in this work considered only the case of SISO. A possible extension would be to consider Multiple Input Single Output (MISO) or even MIMO scenarios in order to exploit space multiplexing and antenna array directivity gains. Also, in the multiple antennas scenarios, we could design dynamic power allocation techniques to enhance the resource usage efficiency. This would allow us to further improve the performance of the proposed technique.

- **Quality of Experience (QoE) provisioning:** this work considered QoS provisioning for different service classes. In the literature, some QoS to QoE mapping functions have been designed. These mapping function could be incorporated in the proposed framework for developing a QoE-aware algorithm focused on QoE provisioning.

- **Scenarios with more services:** this master's thesis considered scenarios composed of two services. A perspective for future work would be to propose an extension for more than two services, where: (1) the services could be divided in two general classes (one with higher priority) and the service utility function used in this master's thesis could still be applied or; (2) a new service utility function could be designed/found which would need to have three or more independent regions for service prioritization, instead of only two as in the service utility function studied in this master's thesis.

- **Multicell scenarios:** the scenarios analyzed in this thesis considered the singlecell case. Thus, another perspective for the continuation of this study would be to consider the multicell case, where the impact of interference could be analyzed for the different service classes.

# BIBLIOGRAPHY

1 CISCO. **Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020**. *Whitepaper*, fev. 2016. Disponível em: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf>.

2 HOSSAIN, E.; HASAN, M. 5G cellular: Key enabling technologies and research challenges. ***IEEE Instrumentation and Measurement Magazine***, v. 18, n. 3, p. 11–21, jun. 2015. ISSN 1094-6969.

3 OSSEIRAN, A. et al. Scenarios for 5G mobile and wireless communications: The vision of the METIS project. ***IEEE Communications Magazine***, v. 52, n. 5, p. 26–35, may 2014. ISSN 0163-6804.

4 ERICSSON. **Ericsson Mobility Report: On the Pulse of the Networked Society**. *Whitepaper*, jun. 2016. Disponível em: <https://www.ericsson.com/res/docs/2016/ericsson-mobility-report-2016.pdf>.

5 KIM, Y.; SON, K.; CHONG, S. **QoS Scheduling for Heterogeneous Traffic in OFDMA-Based Wireless Systems**. In: *IEEE Global Telecommun. Conf.* [S.l.: s.n.], 2009. p. 1–6. ISSN 1930-529X.

6 SONG, G.; LI, Y. G. Utility-based resource allocation and scheduling in OFDM-based wireless broadband networks. ***IEEE Communications Magazine***, v. 43, n. 12, p. 127–134, dez. 2005. ISSN 0163-6804.

7 LEI, H. et al. **QoS aware packet scheduling algorithm for OFDMA systems**. In: *IEEE Vehic. Tech. Conf. (VTC)*. [S.l.: s.n.], 2007. p. 1877–1881. ISSN 1090-3038.

8 RYU, S. et al. **Urgency and efficiency based wireless downlink packet scheduling algorithm in OFDMA system**. In: *IEEE Vehic. Tech. Conf. (VTC)*. [S.l.: s.n.], 2005. v. 3, p. 1456–1462. ISSN 1550-2252.

9 RODRIGUES, E. B. et al. Maximization of user satisfaction in OFDMA systems using utility-based resource allocation. ***Wireless Commun. and Mob. Computing***, p. 1–17, set. 2014.

10 SONG, G. et al. **Joint channel-aware and queue-aware data scheduling in multiple shared wireless channels**. In: *IEEE Wireless Commun. and Networking Conf. (WCNC)*. [S.l.: s.n.], 2004. v. 3, p. 1939–1944. ISSN 1525-3511.

11 SONG, G. ***Cross-Layer Resource Allocation and Scheduling in Wireless Multicarrier Networks***. Tese (Doutorado) — Georgia Institute of Technology, Georgia, USA, 2005.

12 LEI, H. et al. **A Packet Scheduling Algorithm Using Utility Function for Mixed Services in the Downlink of OFDMA Systems**. In: *IEEE Vehic. Tech. Conf. (VTC)*. [S.l.: s.n.], 2007. p. 1664–1668. ISSN 1090-3038.

13 FISHBURN, P. C. **Utility Theory**. *Management Science - Theory Series*, v. 14, n. 5, 1968. Disponível em: <http://dx.doi.org/10.1287/mnsc.14.5.335>.

14 NORTH, D. W. A tutorial introduction to decision theory. ***IEEE Transactions on Systems Science and Cybernetics***, v. 4, n. 3, p. 200–210, fev. 1968. ISSN 0536-1567.

15 RODRIGUES, E. B. ***Adaptive Radio Resource Management for OFDMA-Based Macro- and Femtocell Networks***. Tese (Doutorado) — Universitat Politècnica de Catalunya, Barcelona, Spain, 2011. Disponível em: <http://www.grcm.tsc.upc.edu/sites/default/files/thesis_emanuel_bezerra_final.pdf>.

16 SHENKER, S. Fundamental design issues for the future Internet. ***IEEE Journal on Selected Areas in Communications***, v. 13, n. 7, p. 1176–1188, set. 1995.

17 JERUCHIM, M. C.; BALABAN, P.; SHANMUGAN, K. S. ***Simulation of communication systems: modeling, methodology and techniques***. [S.l.]: Springer Science & Business Media, 2006.

18 DAHLMAN, E.; PARKVALL, S.; SKOLD, J. ***4G: LTE/LTE-Advanced for mobile broadband***. [S.l.: s.n.], 2013.

19 POLESE, M. ***Performance Comparison of Dual Connectivity and Hard Handover for LTE-4G Tight Integration in mmWave Cellular Networks***. Dissertação (Mestrado) — Department of Information Engineering, University of Padova, Padova, Italy, jul. 2016.

20 3GPP. ***Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description***. [S.l.], 2017. Disponível em: <http://www.3gpp.org/ftp/Specs/html-info/36300.htm>.

21 SESIA, S.; BAKER, M.; TOUFIK, I. ***LTE-The UMTS Long Term Evolution: from theory to practice***. [S.l.]: John Wiley & Sons, 2011.

22 LARMO, A. et al. The LTE link-layer design. ***IEEE Communications Magazine***, v. 47, n. 4, p. 52–59, abr. 2009. ISSN 0163-6804.

23 NASRALLA, M. M.; MARTINI, M. G. **A downlink scheduling approach for balancing QoS in LTE wireless networks**. In: *IEEE Personal, Indoor and Mob. Radio Commun. (PIMRC)*. [S.l.: s.n.], 2013. p. 1571–1575. ISSN 2166-9570.

24 LIMA, F. R. M. et al. Scheduling for improving system capacity in multiservice 3GPP LTE. ***J. of Electrical and Computer Engineering***, Hindawi Publishing Corp., v. 2010, p. 16, jun. 2010.

25 SONG, G.; LI, Y. Cross-layer optimization for OFDM wireless networks - part I: Theoretical framework. ***IEEE Transactions on Wireless Communications***, v. 4, n. 2, p. 614–624, mar. 2005. ISSN 1536-1276.

26 JANG, J.; LEE, K. B. Transmit power adaptation for multiuser OFDM systems. ***IEEE Journal on Selected Areas in Communications***, v. 21, n. 2, p. 171–178, jan. 2003. ISSN 0733-8716.

27 3GPP. ***Physical layer aspect for evolved Universal Terrestrial Radio Access (UTRA)***. [S.l.], 2006. Disponível em: <http://www.3gpp.org/ftp/Specs/html-info/25814.htm>.

28 3GPP. ***Spatial channel model for Multiple Input Multiple Output (MIMO) simulations***. [S.l.], 2009. Disponível em: <http://www.3gpp.org/ftp/Specs/html-info/25996.htm>.

29 RAPPAPORT, T. S. ***Wireless Communications: Principles and Practice***. 2nd. ed. [S.l.]: Prentice Hall, 2002.

30 3GPP. ***Deployment aspects***. [S.l.], 2009. Disponível em: <http://www.3gpp.org/ftp/Specs/html-info/25943.htm>.

31 3GPP. ***Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios***. [S.l.], 2012. Disponível em: <http://www.3gpp.org/ftp/Specs/html-info/36942.htm>.

32 GUNNARSSON, F. et al. **Downtilted Base Station Antennas - A Simulation Model Proposal and Impact on HSPA and LTE Performance**. In: *IEEE Vehic. Tech. Conf. (VTC)*. [S.l.: s.n.], 2008. p. 1–5. ISSN 1090-3038.

33 3GPP. ***Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures***. [S.l.], 2010. Disponível em: <http://www.3gpp.org/ftp/Specs/html-info/36213.htm>.

34 MEHLFÜHRER, C. et al. **Simulating the Long Term Evolution Physical Layer**. In: *European Signal Processing Conf.* Glasgow, Scotland: [s.n.], 2009. p. 1471–1478.

35 MACIEL, T. F.; KLEIN, A. On the performance, complexity, and fairness of suboptimal resource allocation for multiuser MIMO-OFDMA systems. ***IEEE Transactions on Vehicular Communications***, v. 59, n. 1, p. 406–419, fev. 2010.

36 3GPP2. ***cdma2000 Evaluation Methodology - Revision B***. [S.l.], 2009. Disponível em: <http://www.3gpp2.org/public_html/specs/C.R1002-B>.

37 BASUKALA, R.; RAMLI, H. A. M.; SANDRASEGARAN, K. **Performance analysis of EXP/PF and M-LWDF in downlink 3GPP LTE system**. In: *First Asian Himalayas Internat. Conf. on Internet*. [S.l.: s.n.], 2009. p. 1–5. ISSN 1089-7801.

38 PALIT, B.; DAS, S. S. Performance evaluation of mixed traffic schedulers in OFDMA networks. ***Wireless Personal Commun.***, v. 83, n. 2, p. 895–924, mar. 2015.

39 HE, L.; LIU, G. Quality-driven cross-layer design for H.264/AVC video transmission over OFDMA system. ***IEEE Transactions on Wireless Communications***, v. 13, n. 12, p. 6768–6782, dez. 2014. ISSN 1536-1276.

40 LUNDEVALL, M. et al. **Streaming applications over HSDPA in mixed service scenarios**. In: *IEEE Vehic. Tech. Conf. (VTC)*. [S.l.: s.n.], 2004. v. 2, p. 841–845. ISSN 1090-3038.

41 JAIN, R.; CHIU, D.-M.; HAWE, W. R. ***A quantitative measure of fairness and discrimination for resource allocation in shared computer system***. [S.l.]: Eastern Research Laboratory, Digital Equipment Corporation Hudson, MA, 1984. v. 38.

42 BENDAOUD, F.; ABDENNEBI, M.; DIDI, F. Survey on scheduling and radio resources allocation in LTE. ***International Journal of Next-Generation Networks***, v. 6, n. 1, p. 17–29, mar. 2014.

43 KWAN, R.; LEUNG, C.; ZHANG, J. Proportional fair multiuser scheduling in LTE. ***IEEE Signal Processing Letters***, v. 16, n. 6, p. 461–464, jun. 2009. ISSN 1070-9908.

44 ANDREWS, M. et al. Providing quality of service over a shared wireless link. ***IEEE Communications Magazine***, v. 39, n. 2, p. 150–154, fev. 2001. ISSN 0163-6804.

45 SHAKKOTTAI, S.; STOLYAR, A. L. Scheduling for multiple flows sharing a time-varying channel: The exponential rule. ***Translations of the American Mathematical Society-Series 2***, v. 207, p. 185–202, 2000.

46 LIMA, F. R. M.; FREITAS, W. C.; CAVALCANTI, F. R. P. **Scheduling Algorithm for Improved System Capacity of Real-Time Services in 3GPP LTE**. In: *Brazilian Telecommun. Symp. (SBrT)*. [S.l.: s.n.], 2009. p. 1–6.

47 SANTOS, R. B. et al. **QoS based Radio Resource Allocation and Scheduling with Different User Data Rate Requirements for OFDMA Systems**. In: *IEEE Personal, Indoor and Mob. Radio Commun. (PIMRC)*. [S.l.: s.n.], 2007. p. 1–5. ISSN 2166-9570.

48 WU, X.; HAN, X.; LIN, X. **QoS oriented heterogeneous traffic scheduling in LTE downlink**. In: *IEEE Internat. Conf. on Commun. (ICC)*. [S.l.: s.n.], 2015. p. 3088–3093. ISSN 1550-3607.

49 ALI, S.; ZEESHAN, M. **A utility based resource allocation scheme with delay scheduler for LTE service-class support**. In: *IEEE Wireless Commun. and Networking Conf. (WCNC)*. [S.l.: s.n.], 2012. p. 1450–1455. ISSN 1525-3511.

50 ZHANG, H. et al. Resource allocation in spectrum-sharing OFDMA femtocells with heterogeneous services. ***IEEE Transactions on Communications***, v. 62, n. 7, p. 2366–2377, jul. 2014. ISSN 0090-6778.

51 MADI, N. K. et al. Two-level QoS-aware frame-based downlink resources allocation for RT/NRT services fairness in lte networks. ***Telecommunication Systems***, Springer, p. 1–19, 2017. ISSN 1572-9451.

52 ZHANG, H. et al. Interference-limited resource optimization in cognitive femtocells with fairness and imperfect spectrum sensing. ***IEEE Transactions on Vehicular Technology***, v. 65, n. 3, p. 1761–1771, mar. 2016. ISSN 0018-9545.

53 ALSAHAG, A. M. et al. Maximum rate resource allocation algorithms with multiuser diversity and QoS support for downlink OFDMA based WiMAX system. ***Telecommunication Systems***, Springer, v. 63, n. 1, p. 1–14, 2016. ISSN 1572-9451.

54 ZHANG, H. et al. Resource allocation for cognitive small cell networks: A cooperative bargaining game theoretic approach. ***IEEE Transactions on Wireless Communications***, v. 14, n. 6, p. 3481–3493, jun. 2015. ISSN 1536-1276.

55 ITURRALDE, M. et al. **Performance study of multimedia services using virtual token mechanism for resource allocation in LTE networks**. In: *IEEE Vehic. Tech. Conf. (VTC)*. [S.l.: s.n.], 2011. p. 1–5. ISSN 1090-3038.

56 RHEE, J.-H.; HOLTZMAN, J. M.; KIM, D.-K. **Scheduling of real/non-real time services: adaptive EXP/PF algorithm**. In: *IEEE Vehic. Tech. Conf. (VTC)*. [S.l.: s.n.], 2003. v. 1, p. 462–466. ISSN 1090-3038.

57 GROSS, J.; BOHGE, M. ***Dynamic Mechanisms in OFDM Wireless Systems: A Survey on Mathematical and System Engineering Contributions***. Technical University Berlin, 2006.

58 HOO, L. M. C. et al. Multiuser transmit optimization for multicarrier broadcast channels: asymptotic FDMA capacity region and algorithms. ***IEEE Transactions on Communications***, v. 52, n. 6, p. 922–930, jun. 2004. ISSN 0090-6778.

59 SONG, G.; LI, Y. Cross-layer optimization for OFDM wireless networks - part II: Algorithm development. ***IEEE Transactions on Wireless Communications***, v. 4, n. 2, p. 625–634, mar. 2005. ISSN 1536-1276.

60 3GPP. ***Evolved Universal Terrestrial Radio Access (E-UTRA): Further advancements for E-UTRA physical layer aspects***. [S.l.], 2010. Disponível em: <http://www.3gpp.org/ftp/Specs/html-info/36814.htm>.

61 PELCAT, M.; ARIDHI, S.; PIAT, J. ***Physical Layer Multi-Core Prototyping - A Dataflow-Based Approach for LTE eNodeB***. 1st. ed. [S.l.]: Springer Science & Business Media, 2013.

# APPENDIX A – OPTIMIZATION FORMULATION FOR THROUGHPUT-BASED SERVICES

As explained in section 4.4.1, the considered optimization problem for throughput-based services is the maximization of the total utility with respect to the users' throughput. Thus, the objective function is

$$\max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} V\left[U_{\text{thr}}\left(T_j\left[n\right]\right)\right], \tag{A.1}$$

where $V\left(\cdot\right)$ is the service utility function and $U_{\text{thr}}\left(\cdot\right)$ is the user utility function that is associated to the UE $j$ that uses a throughput-based service.

The throughput of user $j$ is calculated using an exponential smoothing filtering, as indicated below:

$$T_j\left[n\right] = \left(1 - f_{\text{thru}}\right) \cdot T_j\left[n-1\right] + f_{\text{thru}} \cdot R_j\left[n\right], \tag{A.2}$$

where $R_j\left[n\right]$ is the instantaneous data rate of user $j$ and $f_{\text{thru}}$ is a filtering constant.

Evaluating the objective function in equation (A.1) and the throughput expression in equation (A.2), the derivative of $V\left[U_{\text{thr}}\left(T_j\right)\right]$ with respect to the transmission rate $R_j$ is given by:

$$\begin{aligned}\frac{\partial V\left[U_{\text{thr}}\left(T_j\right)\right]}{\partial R_j} &= \frac{\partial V}{\partial U_{\text{thr}}} \cdot \frac{\partial U_{\text{thr}}}{\partial T_j} \cdot \frac{\partial T_j}{\partial R_j} \\ &= \left.\frac{\partial V}{\partial U_{\text{thr}}}\right|_{U_{\text{thr}}=U_{\text{thr}}(T_j)} \cdot \left.\frac{\partial U_{\text{thr}}}{\partial T_j}\right|_{T_j=(1-f_{\text{thru}})\cdot T_j[n-1]+f_{\text{thru}}\cdot R_j[n]} \cdot f_{\text{thru}}.\end{aligned} \tag{A.3}$$

In the case that $f_{\text{thru}}$ is sufficiently small, the expression above can be simplified as follows [59]:

$$\frac{\partial V\left[U_{\text{thr}}\left(T_j\right)\right]}{\partial R_j} \approx f_{\text{thru}} \cdot \left.\frac{\partial V}{\partial U_{\text{thr}}}\right|_{U_{\text{thr}}=U_{\text{thr}}(T_j)} \cdot \left.\frac{\partial U_{\text{nrt}}}{\partial T_j}\right|_{T_j=T_j[n-1]}, \tag{A.4}$$

where the previous resource allocation totally determines the current values of the marginal utilities. Using the one-order Taylor formula [59, 15] and considering equation (A.4), we have

$$\begin{aligned}\sum_{j \in \mathcal{J}} V\left[U_{\text{thr}}\left(T_j\left[n\right]\right)\right] \approx & \sum_{j \in \mathcal{J}} V\left[U_{\text{thr}}\left(T_j\left[n-1\right]\right)\right] \\ & + \sum_{j \in \mathcal{J}} \left.\frac{\partial V}{\partial U_{\text{thr}}}\right|_{U_{\text{thr}}=U_{\text{thr}}(T_j)} \\ & \cdot \left.\frac{\partial U_{\text{thr}}}{\partial T_j}\right|_{T_j=T_j[n-1]} \\ & \cdot \left(f_{\text{thru}} \cdot R_j\left[n\right] - f_{\text{thru}} \cdot T_j\left[n-1\right]\right).\end{aligned} \tag{A.5}$$

Let us consider the maximization of equation (A.5). Notice that the maximization of the left side of equation (A.5) is our original optimization problem given by equation (A.1). The maximization of the right side of equation (A.5) is the new simplified optimization problem. Since $f_{\mathrm{thru}}$ is a constant and $T_j\left[n-1\right]$ is known and fixed before the resource allocation at the current TTI $n$, the new simplified optimization problem becomes linear in terms of the instantaneous user's data rate, and is given by

$$\max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} V^{'}\left(U_{\mathrm{thr}}\left(T_j\left[n-1\right]\right)\right) \cdot U_{\mathrm{thr}}^{'}\left(T_j\left[n-1\right]\right) \cdot R_j\left[n\right]. \tag{A.6}$$

Notice that we started with an optimization formulation based on throughput given by equation (A.1), made some logical assumptions and mathematical simplifications, and ended up with a linear optimization formulation based on instantaneous rates given by equation (A.6). According to these arguments, we claim that the instantaneous optimization maximizing equation (A.6) leads to a long-term optimization that maximizes equation (A.1).

# APPENDIX B – OPTIMIZATION FORMULATION FOR DELAY-BASED SERVICES

According to section 4.4.2, the considered optimization problem for delay-based services is the maximization of the total utility with respect to the users' HOL packet delays. The objective function is given by

$$\max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} V\left[U_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right)\right], \tag{B.1}$$

where $V\left(\cdot\right)$ is the service utility function and $U_{\text{delay}}\left(\cdot\right)$ is the user utility function that is associated to the user $j$ that makes use of a delay-based service.

In this work, we consider a recursive model for calculating an approximate value of the HOL delay [9]. The recursive equation is

$$d_j^{\text{hol}}\left[n+1\right] = d_j^{\text{hol}}\left[n\right] + t_{\text{tti}} - \frac{1}{L} \cdot \left(\frac{R_j\left[n\right] \cdot t_{\text{tti}}}{S_{\text{p}}}\right), \tag{B.2}$$

where $t_{\text{tti}}$ is the duration of the TTI in seconds, $L$ is the packet arrival rate, $S_{\text{p}}$ is the packet size, and $R_j\left[n\right]$ is the instantaneous achievable transmission rate on TTI $n$.

Assessing the objective function in equation (B.1) and the HOL delay expression in equation (B.2), we can see that the derivative of $V\left[U_{\text{delay}}\left(d_j^{\text{hol}}\right)\right]$ with respect to the transmission rate $R_j$ can be expressed as

$$\begin{aligned}\frac{\partial V\left[U_{\text{delay}}\left(d_j^{\text{hol}}\right)\right]}{\partial R_j} &= \frac{\partial V}{\partial U_{\text{delay}}} \cdot \frac{\partial U_{\text{delay}}}{\partial d_j^{\text{hol}}} \cdot \frac{\partial d_j^{\text{hol}}}{\partial R_j} \\ &= \left.\frac{\partial V}{\partial U_{\text{delay}}}\right|_{U_{\text{delay}}=U_{\text{delay}}\left(d_j^{\text{hol}}\right)} \cdot \left.\frac{\partial U_{\text{delay}}}{\partial d_j^{\text{hol}}}\right|_{d_j^{\text{hol}}=d_j^{\text{hol}}[n]} \cdot \left(-\frac{t_{\text{tti}}}{L \cdot S_{\text{p}}}\right).\end{aligned} \tag{B.3}$$

Using the result above and assuming that the TTI duration is sufficiently small, the La-

grange theorem of the mean can be used [12, 15], which says that

$$\begin{aligned}\sum_{j\in\mathcal{J}} V\left[U_{\text{delay}}\left(d_j^{\text{hol}}[n+1]\right)\right] &\approx \sum_{j\in\mathcal{J}} V\left[U_{\text{delay}}\left(d_j^{\text{hol}}[n]\right)\right] + \sum_{j\in\mathcal{J}} \left.\frac{\partial V}{\partial U_{\text{delay}}}\right|_{U_{\text{delay}}=U_{\text{delay}}\left(d_j^{\text{hol}}\right)} \\ &\quad \cdot \left.\frac{\partial U_{\text{delay}}}{\partial R_j}\right|_{R_j=R_j[n-1]} \cdot \left(R_j[n] - R_j[n-1]\right) \\ &= \sum_{j\in\mathcal{J}} V\left[U_{\text{queue}}\left(d_j^{\text{hol}}[n]\right)\right] + \sum_{j\in\mathcal{J}} \left.\frac{\partial V}{\partial U_{\text{delay}}}\right|_{U_{\text{delay}}=U_{\text{delay}}\left(d_j^{\text{hol}}\right)} \\ &\quad \cdot \left.\frac{\partial U_{\text{delay}}}{\partial d_j^{\text{hol}}}\right|_{d_j^{\text{hol}}=d_j^{\text{hol}}[n]} \cdot \left(-\frac{t_{\text{tti}}}{L\cdot S_{\text{p}}}\right) \cdot \left(R_j[n] - R_j[n-1]\right) \\ &= \sum_{j\in\mathcal{J}} V\left[U_{\text{delay}}\left(d_j^{\text{hol}}[n]\right)\right] + \sum_{j\in\mathcal{J}} \left.\frac{\partial V}{\partial U_{\text{delay}}}\right|_{U_{\text{delay}}=U_{\text{delay}}\left(d_j^{\text{hol}}\right)} \\ &\quad \cdot \left|\left.\frac{\partial U_{\text{delay}}}{\partial d_j^{\text{hol}}}\right|\right|_{d_j^{\text{hol}}=d_j^{\text{hol}}[n]} \cdot \frac{t_{\text{tti}}}{L\cdot S_{\text{p}}} \cdot \left(R_j[n] - R_j[n-1]\right). \end{aligned} \tag{B.4}$$

The absolute value operator is used in equation (B.4) because the utility function is assumed to be decreasing, which yields negative marginal utilities and cancels the negative sign in equation (B.4).

On one hand, the maximization of the left side of equation (B.4) is our original optimization problem given by equation (B.1). On the other hand, the maximization of the right side of equation (B.4) is the new simplified optimization problem. We have that $t_{\text{tti}}$, $L$ and $S_{\text{p}}$ are constants, and that $d_j^{\text{hol}}[n]$ and $R_j[n-1]$ are known and fixed before the resource allocation at TTI $n$. Therefore, the new simplified optimization problem becomes linear in terms of the instantaneous user's data rate, and is given by

$$\max_{\rho_{j,k},\, p_k} \sum_{j\in\mathcal{J}} V'\left(U_{\text{delay}}\left(d_j^{\text{hol}}[n]\right)\right) \cdot \left|U'_{\text{delay}}\left(d_j^{\text{hol}}[n]\right)\right| \cdot R_j[n]. \tag{B.5}$$

Taking into account equation (B.4), we are able to assume that the instantaneous optimization maximizing equation (B.5) leads to a long-term optimization that maximizes equation (B.1).

# APPENDIX C – OPTIMIZATION FORMULATION FOR QUEUE-BASED SERVICES

As explained in section 4.4.3, the considered optimization problem for queue-based services is the maximization of the total utility with respect to the predicted average queue size over a time window (in bits) of a user $j$. Thus, the objective function is

$$\max_{\rho_{j,k},\, p_k} \sum_{j \in \mathcal{J}} V\left[U_{\text{queue}}\left(\overline{Q}_j\left[n+1\right]\right)\right], \tag{C.1}$$

where $V(\cdot)$ is the service utility function and $U_{\text{queue}}(\cdot)$ is the user utility function that is associated to the user $j$ that makes use of a throughout- and delay-based service, which is referred to as queue-based service.

The average queue size over a time window of user $j$ is calculated using an exponential smoothing filtering, as indicated below:

$$\overline{Q}_j\left[n\right] = (1 - f_{\text{queue}}) \cdot \overline{Q}_j\left[n-1\right] + f_{\text{queue}} \cdot Q_j\left[n\right], \tag{C.2}$$

where $Q_j\left[n\right]$ is the instantaneous queue size of user $j$ and $f_{\text{queue}}$ is a filtering constant.

The queue size of user $j$ at TTI $n+1$ can be expressed as [10]

$$Q_j\left[n+1\right] = Q_j\left[n\right] - R_j\left[n\right] \cdot t_{\text{tti}} + \alpha_j[n], \tag{C.3}$$

where $\alpha_j[n]$ is the amount of arrival bits of UE $j$ during TTI $n$, $R_j\left[n\right]$ is the instantaneous data rate of user $j$ at TTI $n$ and $t_{\text{tti}}$ is the duration of a single TTI.

At the beginning of TTI $n$, given the instantaneous data rate $R_j\left[n\right]$, the predicted average queue size at the end of TTI $n$ (beginning of TTI $n+1$) is obtained by $E_{\alpha_j[n]}\left\{\overline{Q}_j\left[n+1\right]\right\}$, which is the expectation with respect to $\alpha_j[n]$ [10]. According to equation (C.2) and equation (C.3), we have

$$E_{\alpha_j[n]}\left\{\overline{Q}_j\left[n+1\right]\right\} = (1 - f_{\text{queue}}) \cdot \overline{Q}_j\left[n\right] + f_{\text{queue}} \cdot \left(Q_j\left[n\right] - R_j\left[n\right] \cdot t_{\text{tti}} + E\left\{\alpha_j[n]\right\}\right), \tag{C.4}$$

where $E\left\{\alpha_j[n]\right\} = \omega_j \cdot t_{\text{tti}}$, and $\omega_j$ is the source data rate of the service consumed by user $j$.

Evaluating the objective function in equation (C.1) and the expected queue size expression in equation (C.4), the derivative of $V\left[U_{\text{queue}}\left(\overline{Q}_j\right)\right]$ with respect to the transmission rate

$R_j$ is given by:

$$\begin{aligned}\frac{\partial V\left[U_{\text{queue}}\left(\overline{Q}_j\right)\right]}{\partial R_j} &= \frac{\partial V}{\partial U_{\text{queue}}} \cdot \frac{\partial U_{\text{queue}}}{\partial \overline{Q}_j} \cdot \frac{\partial \overline{Q}_j}{\partial R_j} \\ &= \left.\frac{\partial V}{\partial U_{\text{queue}}}\right|_{U_{\text{queue}}=U_{\text{queue}}\left(\overline{Q}_j\right)} \\ &\cdot \left.\frac{\partial U_{\text{queue}}}{\partial \overline{Q}_j}\right|_{\overline{Q}_j=(1-f_{\text{queue}})\cdot\overline{Q}_j[n]+f_{\text{queue}}\cdot(Q_j[n]-R_j[n]\cdot t_{\text{tti}}+\omega_j\cdot t_{\text{tti}})} \\ &\cdot f_{\text{queue}} \cdot t_{\text{tti}}. \end{aligned} \tag{C.5}$$

In the case that $f_{\text{queue}}$ is sufficiently small, the expression above can be simplified as follows [59]:

$$\frac{\partial V\left[U_{\text{queue}}\left(\overline{Q}_j\right)\right]}{\partial R_j} \approx f_{\text{queue}} \cdot t_{\text{tti}} \cdot \left.\frac{\partial V}{\partial U_{\text{queue}}}\right|_{U_{\text{queue}}=U_{\text{queue}}\left(\overline{Q}_j\right)} \cdot \left.\frac{\partial U_{\text{queue}}}{\partial \overline{Q}_j}\right|_{\overline{Q}_j=\overline{Q}_j[n]}, \tag{C.6}$$

where the previous resource allocation totally determines the current values of the marginal utilities. Using the one-order Taylor formula [59, 15] and considering equation (C.6), we have

$$\begin{aligned}\sum_{j\in\mathcal{J}} V\left[U_{\text{queue}}\left(\overline{Q}_j\left[n+1\right]\right)\right] \approx \sum_{j\in\mathcal{J}} V\left[U_{\text{queue}}\left(\overline{Q}_j\left[n\right]\right)\right] \\ + \sum_{j\in\mathcal{J}} \left.\frac{\partial V}{\partial U_{\text{queue}}}\right|_{U_{\text{queue}}=U_{\text{queue}}\left(\overline{Q}_j\right)} \\ \cdot \left|\frac{\partial U_{\text{queue}}}{\partial \overline{Q}_j}\right|\Bigg|_{\overline{Q}_j=\overline{Q}_j[n]} \\ \cdot \left\{f_{\text{queue}} \cdot \left[Q_j\left[n\right] + t_{\text{tti}} \cdot \left(\omega_j - R_j\left[n\right]\right)\right] - f_{\text{queue}} \cdot \overline{Q}_j\left[n\right]\right\}. \end{aligned} \tag{C.7}$$

The absolute value operator is used in equation (C.7) because the utility function is assumed to be decreasing, which yields negative marginal utilities and cancels the negative sign in equation (C.7).

Considering the maximization of equation (C.7), one can see that the maximization of the left side of equation (C.7) is our original optimization problem given by equation (C.1), while the maximization of the right side of equation (C.7) is the new simplified optimization problem. Since $f_{\text{queue}}$, $\omega_j$ and $t_{\text{tti}}$ are constant and $\overline{Q}_j\left[n\right]$ and $Q_j\left[n\right]$ are known and fixed before the resource allocation at the current TTI $n$, the new simplified optimization problem becomes linear in terms of the instantaneous user's data rate, and is given by

$$\max_{\rho_{j,k},\, p_k} \sum_{j\in\mathcal{J}} V'\left(U_{\text{queue}}\left(\overline{Q}_j\left[n\right]\right)\right) \cdot \left|U'_{\text{queue}}\left(\overline{Q}_j\left[n\right]\right)\right| \cdot R_j\left[n\right]. \tag{C.8}$$

Notice that we started with an optimization formulation based on the predicted average queue size given by equation (C.1), made some logical assumptions and mathematical simplifications, and ended up with a linear optimization formulation based on instantaneous rates given by equation (C.8). According to these arguments, we claim that the instantaneous optimization maximizing equation (C.8) leads to a long-term optimization that maximizes equation (C.1).

# APPENDIX D – OPTIMIZATION FORMULATION FOR MULTIPLE SERVICES

As described in section 4.4.4 and considering a scenario with throughput-based, delay-based and queue-based services, the optimization problem is the maximization of the total utility with respect to the users' QoS, namely throughput, HOL packet delay and average queue size for throughput-based, delay-based and queue-based services, respectively.

Let us assume that the set $\mathcal{J}$ of the users in the system is separated in three subsets: $\mathcal{J}_{\text{thr}}$, $\mathcal{J}_{\text{delay}}$ and $\mathcal{J}_{\text{queue}}$ for throughput-based, delay-based and queue-based users, respectively. Therefore, the objective function of the general optimization problem can be re-written as

$$\max_{\rho_{j,k},\, p_k} \left\{ \sum_{j \in \mathcal{J}_{\text{thr}}} V\left[U_{\text{thr}}\left(T_j\left[n\right]\right)\right] + \sum_{j \in \mathcal{J}_{\text{delay}}} V\left[U_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right)\right] + \sum_{j \in \mathcal{J}_{\text{queue}}} V\left[U_{\text{queue}}\left(\overline{Q}_j\left[n+1\right]\right)\right] \right\}. \tag{D.1}$$

The summation in equation (D.1) regarding queue-based, delay-based and queue-based services were analyzed in appendices A, B, C, respectively. Replacing the approximate expressions in equations (A.5), (B.4) and (C.7) into equation (D.1), and taking into account that $f_{\text{thru}}$, $L$, $S_{\text{p}}$, $f_{\text{queue}}$, $\omega_j$ and $t_{\text{tti}}$ are constant and $T_j\left[n-1\right]$, $d_k^{\text{hol}}\left[n\right]$, $\overline{Q}_i\left[n\right]$ and $Q_i\left[n\right]$ are known and fixed before the resource allocation at the current TTI $n$, we have that the objective function of the mixed services problem becomes

$$\begin{aligned} \max_{\rho_{j,k},\, p_k} \Bigg\{ & \sum_{j \in \mathcal{J}_{\text{thr}}} V'\left(U_{\text{thr}}\left(T_j\left[n-1\right]\right)\right) \cdot U'_{\text{thr}}\left(T_j\left[n-1\right]\right) \cdot R_j\left[n\right] \\ & + \sum_{j \in \mathcal{J}_{\text{delay}}} V'\left(U_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right)\right) \cdot \left|U'_{\text{delay}}\left(d_j^{\text{hol}}\left[n\right]\right)\right| \cdot R_j\left[n\right] \\ & + \sum_{j \in \mathcal{J}_{\text{queue}}} V'\left(U_{\text{queue}}\left(\overline{Q}_j\left[n\right]\right)\right) \cdot \left|U'_{\text{queue}}\left(\overline{Q}_j\left[n\right]\right)\right| \cdot R_j\left[n\right] \Bigg\}. \end{aligned} \tag{D.2}$$

Notice that the new simplified optimization problem given by equation (D.2) is linear in terms of the instantaneous user's data rate. Based on the arguments and assumptions made in appendices A, B and C, we claim that the instantaneous optimization maximizing equation (D.2) leads to a long-term optimization that maximizes equation (D.1).

## APPENDIX E – LOOK-UP TABLE OF JSM ALGORITHM

As explained in section 4.6.2, a cubic interpolant function was employed in a curve fitting tool to obtain a look-up table comprised of 41 non-linear spaced values of $\lambda$ (shape parameter) of the service utility function. In table E.1, all the $\lambda$ values calculated are illustrated. The value -0.1088 is located in position 1 of the look-up table, the value -5.4529 is in position 20, 5000 is in the position 21, the value 5.4529 is in position 22 and the last value in the look-up table (position 41) is 0.1088.

The algorithm starts with the $\lambda$ value equals to 5000, i.e., both services have the same priority. Then, every TTI, the algorithm checks the satisfaction of service 1. If it is above the target, the position in the look-up table is incremented by one, so that the priority of service 2 increases. Otherwise, the position in the look-up table is decremented by one, so that the priority of service 1 increases.

| **Higher priority for service 1** | **Equal priority for both services** | **Higher priority for service 2** |
|---|---|---|
| -0.1088 | | 5.4529 |
| -0.1502 | | 2.6241 |
| -0.1808 | | 1.7221 |
| -0.2091 | | 1.2771 |
| -0.2371 | | 1.0107 |
| -0.2667 | | 0.8324 |
| -0.2981 | | 0.7041 |
| -0.3329 | | 0.6071 |
| -0.3717 | | 0.5306 |
| -0.4163 | 5000 | 0.4683 |
| -0.4683 | | 0.4163 |
| -0.5306 | | 0.3717 |
| -0.6071 | | 0.3329 |
| -0.7041 | | 0.2981 |
| -0.8324 | | 0.2667 |
| -1.0107 | | 0.2371 |
| -1.2771 | | 0.2091 |
| -1.7221 | | 0.1808 |
| -2.6241 | | 0.1502 |
| -5.4529 | | 0.1088 |

Table E.1 – Look-up table employed in the JSM algorithm.