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

DIEGO AGUIAR SOUSA

RADIO RESOURCE MANAGEMENT FOR RATE MAXIMIZATION WITH QOS/QOE PROVISIONING IN WIRELESS NETWORKS

FORTALEZA
2018

DIEGO AGUIAR SOUSA

RADIO RESOURCE MANAGEMENT FOR RATE MAXIMIZATION WITH QOS/QOE PROVISIONING IN WIRELESS NETWORKS

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

Orientador: Prof. Dr. Tarcisio Ferreira Maciel

Coorientador: Prof. Dr. Francisco Rafael Marques Lima

FORTALEZA
2018

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)

S696r Sousa, Diego Aguiar.
Radio Resource Management for Rate Maximization with QoS/QoE Provisioning in Wireless Networks / Diego Aguiar Sousa. – 2018.
130 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2018.
Orientação: Prof. Dr. Tarcisio Ferreira Maciel.
Coorientação: Prof. Dr. Francisco Rafael Marques Lima.

1. radio resource allocation. 2. rate maximization. 3. multi-service. 4. user satisfaction. I. Título.
CDD 621.38

DIEGO AGUIAR SOUSA

RADIO RESOURCE MANAGEMENT FOR RATE MAXIMIZATION WITH QOS/QOE PROVISIONING IN WIRELESS NETWORKS

Thesis presented to the Graduate Program in Teleinformatics Engineering of the Federal University of Ceará as a partial requisite to obtain the Ph.D. degree in Teleinformatics Engineering. Concentration Area: Signals and Systems.

Aproved on: 09/07/2018.

EXAMINING COMMITTEE

Prof. Dr. Tarcisio Ferreira Maciel (Advisor)
Universidade Federal do Ceará

Prof. Dr. Francisco Rafael Marques Lima (Co-Advisor)
Universidade Federal do Ceará - Campus Sobral

Prof. Dr. Francisco Rodrigo Porto Cavalcanti
Universidade Federal do Ceará

Prof. Dr. Renato da Rocha Lopes
Universidade de Campinas

Prof. Dr. Danilo Silva
Universidade Federal de Santa Catarina

To my beloved wife, Nélia, and my little daughter, Clarice.

ACKNOWLEDGEMENTS

First of all, I would like to thank God for all the blessings I have been granted in my life, and for providing me with the necessary wisdom and strength to reach the end of the doctorate course whilst reconciling work and studies.

I would also like express my gratitude to my parents, José Maria and Aparecida, for their love and care. Even though they did not understand well my field of study or the content of this doctoral thesis, they have always encouraged and supported me, celebrating each one of my achievements. They were always committed to providing me with the best education possible; thus, I would not have gotten where I am now without them.

I am very thankful to my beloved wife, Nélia Braga, who has tolerated me since the beginning of my undergraduate course (more than ten years ago). She has always stood by my side with her love, understanding and patience, especially because I had to sacrifice a lot of our time together due to my research. Besides all that, at the very end of the doctorate course, she also gifted me with my little daughter, Clarice, who came to fill my life with joy. Therefore, she was undoubtedly extremely important to the accomplishment of this doctoral thesis.

I am also grateful to my supervisor, Professor Tarcisio F. Maciel, for his great supervision, his patience, and for believing in me even when I myself did not. Thanks also to my co-supervisor, Professor Rafael Lima, for the insightful discussions, comments and contributions to this work.

I would like to thank Professor Rodrigo Cavalcanti who gave me the opportunity to join the GTEL projects, which in turn provided me with the necessary resources to conduct my research. In this context, I am thankful to the IT personnel of the GTEL laboratory, especially to João Pinheiro, who has always been of great help.

At last but not least, I want to express my sincere gratitude to my friends from GTEL, who have in some way contributed to the accomplishment of this work. I am thankful to Victor Farias for our fruitful research partnership and important discussions. A special thanks goes to Roberto Antonioli, who, in spite of being extremely busy with his internship in Sweden, has always found some time to help me when I needed, in addition to performing numerous text reviews in my doctoral thesis. Another special thanks goes to Igor Guerreiro, who has dedicated a lot of his precious time to help me accomplish this work, providing me with some of his academic experience as well as with several insightful discussions, lessons and text reviews.

"I have fought a good fight, I have finished my course, I have kept the faith."

(Holy Bible, 2 Tim. 4:7, King James Version)

RESUMO

Atualmente, a quinta geração (5G) de comunicações móveis está sendo intensivamente discutida para finalizar a nova padronização. Sabe-se que os sistemas 5G devem suportar uma ampla variedade de aplicações, além de prover conexão a um grande número de dispositivos. Um dos casos de uso do 5G é a banda larga móvel aprimorada (do inglês, *Enhanced Mobile Broadband* (eMBB)), que consiste em uma melhoria nos atuais serviços de banda larga de quarta geração (4G). O eMBB está focado em prover, entre outras coisas, altas capacidades, altos picos de taxa, e uma boa vazão de dados por usuário. Uma estratégia para atingir os objetivos do eMBB é fazer uso de métodos apropriados de alocação de recursos de rádio (do inglês, *Radio Resource Allocation* (RRA)). Deste modo, o trabalho desenvolvido nesta tese estudou métodos de RRA com o intuito de maximizar a taxa total do sistema, garantindo uma certa taxa de satisfação por serviço em cenários com um ou mais serviços. Tais problemas de RRA consideram requerimentos diferentes por serviço e um compromisso entre alta eficiência espectral e satisfação dos usuários. Tal compromisso pode ser gerenciado pelo operador do sistema, o que torna este estudo bastante relevante, principalmente para os operadores de redes móveis. Os métodos de RRA foram estudados em três contextos distintos. O primeiro problema considera que a estação rádio-base divide a potência disponível igualmente entre todos os blocos de recurso (do inglês, *Resource Blocks* (RBs)) e somente estes são alocados pelo RRA. Além disso, os usuários devem ter seus requisitos atendidos em um único *slot* de tempo. Este problema é inicialmente descrito como um problema de otimização e, a partir de sua análise, foi proposta uma nova heurística subótima e de baixa complexidade. O algoritmo proposto atinge resultados melhores que a heurística do estado da arte. Ademais, diferentemente do estado da arte, quando não há solução viável para o problema, o algoritmo proposto é capaz de prover resultados próximos ao desejado. Logo depois, o problema de RRA é estendido para atender os requisitos dos usuários dentro de um intervalo de tempo. Neste contexto, o algoritmo proposto anteriormente foi estendido para escalonar os usuários ao longo do tempo. Tal escalonador é comparado com algoritmos de RRA baseados em utilidade com objetivos similares. Os resultados de simulação mostraram que o escalonador proposto apresenta valores de satisfação e taxa total consideravelmente melhores que as soluções de referência. Por fim, o problema de RRA é novamente estudado considerando um único *slot* de tempo, contudo desta vez alocando ambos potência e RBs. De modo similar ao primeiro problema estuado nesta tese, o RRA é primeiramente escrito como um problema de otimização. Usando a mesma estrutura de solução adotada na análise do primeiro problema tratado nesta tese, uma nova heurística subótima é proposta. Simulações computacionais mostraram que a solução proposta supera o algoritmo do estado da arte. Além disso, a heurística proposta provê soluções próximas ao desejado, quando o problema não possui solução viável. Enquanto isso, nestes casos, a algoritmo do estado da arte não é capaz de prover uma solução realizável.

Palavras-chave: alocação de recursos de rádio, maximização de taxa, multi-serviço, satisfação de usuários.

ABSTRACT

Currently, the Fifth Generation (5G) of mobile communications is under intensive discussions in order to setup the new standardization. It is already known that 5G systems must provide support to a large variety of applications besides handling a higher number of devices connected to the network. One of the use cases of the 5G is the Enhanced Mobile Broadband (eMBB), which consists in an improvement of the existing Fourth Generation (4G) broadband service. The eMBB focuses on providing, among other characteristics, high system capacity, high peak data rate and user experienced data rate. One possible strategy to achieve the eMBB goals is to properly use Radio Resource Allocation (RRA) methods to increase the efficiency of the spectrum usage and the Quality of Service (QoS) perceived by the users. Therefore, the work developed in this thesis studies methods of RRA aiming at maximizing the overall system throughput, constrained by guaranteeing a certain satisfaction rate per service in single and multi-service scenarios. The RRA problems addressed in this thesis deal with different service requirements and a trade-off between high spectral efficiency and users satisfaction. This trade-off can be managed by the system, which makes the study performed in this thesis very relevant, mainly to the mobile network operators. The RRA methods are studied in three different contexts. The first problem considers that the Base Station (BS) employs an Equal Power Allocation (EPA) among Resource Blocks (RBs) and only the RB assignment is addressed by the RRA. Besides, the users shall meet their requirements in a single time slot, i.e., on a Transmission Time Interval (TTI) basis. This problem is initially described mathematically and, from the analysis of the optimization problem formulation, a new suboptimal low complexity heuristic is proposed. By means of computational simulations, it is shown that the proposed algorithm outperforms the state-of-the-art heuristic, achieving near optimal results. Moreover, in contrast to the state-of-the-art literature, the proposed algorithm is capable of providing near feasible solutions in infeasible instances of the problem. Thereafter, this RRA problem is extended to address the users' requirements over a given timespan. In this context, the heuristic earlier proposed is extended to schedule the users over time. The proposed heuristic is compared with utility-based benchmark algorithms with similar objectives. Simulation results show that proposed scheduler considerably outperforms the benchmark solutions in terms of both satisfaction and overall system throughput. Lastly, the RRA problem is once again studied on a TTI basis, however allocating both power and RBs. Like the first problem studied in this thesis, the RRA is firstly stated as an optimization problem. Using the same solution framework adopted with the first problem, a new suboptimal heuristic is proposed. Computational simulations show that the proposed heuristic outperforms the state-of-the-art algorithm. Additionally, the proposed heuristic is capable of providing near feasible solutions in infeasible instances of the RRA problem, while the state-of-art literature does not provide a practical solution.

Keywords: radio resource allocation, rate maximization, multi-service, user satisfaction.

LIST OF FIGURES

Figure 2.1 – Scenario overview. 29
Figure 2.2 – Frequency-time grid of RBs. 30
Figure 2.3 – 3D view. 34
Figure 2.4 – Geometrical model for 3D one-ring with scatterers around the UE. 34
Figure 2.5 – SINR to BLER mapping for the MCSs in the LTE standard. 37
Figure 2.6 – Simulation flowchart. 39
Figure 3.1 – Flowchart of the RMEC Algorithm. 46
Figure 3.2 – Bipartite Graph with solid edges denoting the maximum matching. 52
Figure 3.3 – System performance for a single service scenario with $\xi_1 = 100\%$ of U. . . . 56
Figure 3.4 – System performance varying the percentage of satisfied UEs in a single service scenario with $U = 30$ UEs. 58
Figure 3.5 – Outage probability for $\xi_1 = 100\%$ and $U = 10, 20$ and 30 UEs. 59
Figure 3.6 – Outage probability considering $\xi_1 = 80\%, 90\%$ and 100% of $U = 30$ UEs. . . 60
Figure 3.7 – CDF of the satisfaction and throughput considering $U = 20$ UEs, a minimum MOS equal to 4.4 and $\xi = 100\%$ of the UEs. 61
Figure 3.8 – Satisfaction and throughput considering $U = 30$ UEs, minimum satisfaction requirement equal to $\xi = 100\%$ of the UEs. 62
Figure 3.9 – CDF of the satisfaction and throughput considering $U = 20$ UEs, a minimum MOS equal to 4.4 and $\xi = 100\%$ of the UEs. 64
Figure 3.10–System performance considering $U = 30$ UEs and $S = 2$ service plans, where $U_1 = 20$ and $U_2 = 10$ UEs, $\Omega_1^{\text{target}} = 500$ kbps and $\Omega_2^{\text{target}} = 1.5$ Mbps. 66
Figure 3.11–CDF of the satisfaction of each service plan and the overall system throughput in a scenario considering $U = 40$ UEs and $S = 3$ service plans, where $U_1 = 20$, $U_2 = 15$ and $U_2 = 5$ UEs, and $\Omega_1^{\text{target}} = 500$ kbps, $\Omega_2^{\text{target}} = 300$ kbps and $\Omega_3^{\text{target}} = 1.5$ Mbps. 68
Figure 4.1 – Flowchart of the TRMEC Algorithm. 73
Figure 4.2 – System performance for a single service scenario with $\xi_1 = 100\%$ of U. . . . 80
Figure 4.3 – System performance for a single service scenario with a minimum MOS $\Omega_1^{\text{target}} = 4$. 83
Figure 4.4 – System performance for a single service scenario with a minimum MOS $\Omega_1^{\text{target}} = 4$ and $\xi_1 = 100\%$ of U. 85
Figure 4.5 – System performance considering $U = 60$ UEs and $S = 2$ service plans, where $U_1 = 40$ and $U_2 = 20$ UEs, $\Omega_1^{\text{target}} = 500$ kbps and $\Omega_2^{\text{target}} = 1.5$ Mbps. 88
Figure 5.1 – System performance for a single service scenario with $\xi_1 = 100\%$ of U. . . . 106
Figure 5.2 – System performance varying the percentage of satisfied UEs in a single service scenario with $U = 30$ UEs. 110

Figure 5.3 – CDF of the satisfaction and throughput considering $U = 30$ UEs and $\xi = 100\%$ of the UEs. 113
Figure 5.4 – System performance considering $U = 30$ UEs and $S = 2$ service plans, where $U_1 = 20$ and $U_2 = 10$ UEs, $\Omega_1^{\text{target}} = 500$ kbps and $\Omega_2^{\text{target}} = 1.5$ Mbps. 116

LIST OF TABLES

Table 2.1 – Mapping between CQI and MCS in the LTE standard. 35
Table 3.1 – Simulation parameters. 55
Table 4.1 – Parameters of ATES and ASC algorithms. 77
Table 4.2 – Simulation parameters. 79

LIST OF ABBREVIATIONS AND ACRONYMS

3GPP	3rd Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
ASC	Adaptive Satisfaction Control
ATES	Adaptive Throughput-based Efficiency-Satisfaction Trade-Off
AWGN	Additive White Gaussian Noise
BB	Branch and Bound
BC	Branch and Cut
BER	Bit Error Rate
BLER	BLock Error Rate
BS	Base Station
CBR	Constant Bit Rate
CDF	Cumulative Distribution Function
CNR	Channel-to-Noise Ratio
CQI	Channel Quality Indicator
CSI	Channel State Information
eMBB	Enhanced Mobile Broadband
EPA	Equal Power Allocation
FTP	File Transfer Protocol
GAP	Generalized Assignment Problem
GAP-MQ	Generalized Assignment Problem with Minimum Quantities
HH	Hughes-Hartogs
IID	Independent and Identically Distributed
IJRAPA	Improved JRAPA
ILP	Integer Linear Problem
IP	Internet Protocol
ITU	International Telecommunication Union
JRAPA	Joint RB Assignment and Power Allocation
KPI	Key Performance Indicator
LP	Linear Programming
LTE	Long Term Evolution
MCS	Modulation and Coding Scheme
MILP	Mixed Integer Linear Programming
MIMO	Multiple Input Multiple Output
mMTC	Massive Machine-Type Communications
MOS	Mean Opinion Score

MU-MIMO	Multi-User Multiple Input Multiple Output
OFDMA	Orthogonal Frequency Division Multiple Access
PRARMEC	Power and Resource Allocation for RMEC
QAM	Quadrature Amplitude Modulation
QoE	Quality of Experience
QoS	Quality of Service
QPSK	Quadri-Phase Shift Keying
RAISES	Reallocation-based Assignment for Improved Spectral Efficiency and Satisfaction
RAT	Radio Access Technology
RB	Resource Block
RMEC	Rate Maximization under Experience Constraints
RRA	Radio Resource Allocation
RRM	Radio Resource Management
SINR	Signal to Interference-plus-Noise Ratio
SNR	Signal to Noise Ratio
TDMA	Time Division Multiple Access
TRMEC	Temporal RMEC
TTI	Transmission Time Interval
UE	User Equipment
ULA	Uniform Linear Array
URA	Uniform Rectangular Array
URLLC	Ultra-Reliable and Low-Latency Communications
VoIP	Voice over IP
ZMCSCG	Zero Mean Circularly Symmetric Complex Gaussian

LIST OF SYMBOLS

$(\cdot)^{\mathrm{T}}$	Vector/Matrix transposition operator
$(\cdot)^{(\mathrm{dB})}$	Convert the argument to dB scale
$\odot$	Hadamard product operator
$*$	Khatri-Rao product operator
$\otimes$	Kronecker product operator
$\lvert\mathcal{A}\rvert$	Cardinality of the set $\mathcal{A}$
$\lceil\cdot\rceil$	Nearest integer greater than or equal to the argument
$\mathbf{0}_{a\times b}$	Matrix with dimension $a\times b$ composed by 0's
$\mathbf{0}_a$	Column vector with length a with all elements equal to 0
$\mathbf{1}_a$	Column vector with length a with all elements equal to 1
$A_b\left(\theta,\phi\right)$	Radiation power pattern of each individual antenna element of the BS b for a given pair of zenith and azimuth angles of departure (θ,ϕ)
$A_u\left(\theta,\phi\right)$	Radiation power pattern of the antenna of the UE u for a given pair of zenith and azimuth angles of arrival (θ,ϕ)
$\mathcal{A}$	Auxiliary set of UEs
$\mathcal{B}$	Set of BSs
B	Number of BSs
$\mathcal{E}$	Set of edges of G
$\mathbf{f}_{b,k}[t]$	Transmission filter of BS b in the RB k at TTI t
$G(\mathcal{V},\mathcal{K},\mathcal{E})$	Bipartite graph with partitions $\mathcal{V}$ and $\mathcal{K}$ connected by the edges in the set $\mathcal{E}$
$H\left(a,b\right)$	Heaviside step function, which return 1 if $a\geq b$ and 0 otherwise
$I[t]$	ZMCSCG variable which models the perceived interference
$\mathbf{I}_a$	Identity matrix with order a
$\mathcal{K}$	Set of RBs
K	Number of RBs
$L_{b,n,u,k,z}$	Large-scale fading of the z-th ray the n-th antenna element of the BS b and the antenna of the UE u in the RB k
$\mathcal{L}$	Set containing the UEs u that must be satisfied
$\mathcal{M}$	Set containing the edges from $\mathcal{E}$ which represents the minimum weighted matching of the bipartite graph $G(\mathcal{V},\mathcal{K},\mathcal{E})$
M	Number of MCSs
N	Number of antennas of the BS
$\mathcal{O}\left(\cdot\right)$	Complexity order
$PL_{b,n,u,k,z}$	Average path loss suffered by the z-th ray between the n-th antenna element of the BS b and the antenna of the UE u in the RB k

P_{total}	Total available power in the BS
$\overline{\mathbf{P}}$	Power tensor with dimensions $U \times K \times M$ composed by the elements $p_{u,k,m}$
Q_{sub}	Number of subcarriers of one RB
Q_{sym}	Number of symbols of one RB
$\mathbf{Q}$	Binary association matrix with each element (s,u) equal to $q_{s,u}$
$R_u[t]$	Total achieved rate of a UE u at a TTI t
R_{sys}	Overall system rate
$R^{\text{avg}}[t]$	Average rate of a UE u at TTI t
$\mathbf{R}$	Rate matrix composed by the elements $r_{u,k}$
$\mathcal{S}$	Set of services
S	Number of services
T_{tti}	Duration of one TTI
$\mathcal{T}$	Sequence of consecutive TTIs of a timespan
T	Number of TTIs
$\mathcal{U}_s^{\text{sat}}$	Set of UEs subscribing the service s with their QoS/QoE requirements fulfilled
$\mathcal{U}_s$	Set of UEs subscribing the service s
$\mathcal{U}$	Set of UEs
U	Number of UEs
$\mathcal{V}$	Set of virtual UE nodes
$\mathcal{X}_u$	Set of the RBs allocated to each UE u
$\tilde{\mathbf{X}}$	Fractional assignment matrix composed by the elements $\tilde{x}_{u,k}$
$\mathbf{X}_{\text{sat}}$	Binary assignment matrix composed by the rows $u \in \mathcal{L}$ of $\mathbf{X}$
$\mathbf{X}$	Binary assignment matrix with each element (u,k) equal to $x_{u,k}$
$\tilde{\overline{\mathbf{X}}}$	Fractional assignment tensor composed by the elements $\tilde{x}_{u,k,m}$
$\overline{\mathbf{X}}_{\text{sat}}$	Binary assignment tensor composed by the rows $u \in \mathcal{L}$ of $\overline{\mathbf{X}}$
$\overline{\mathbf{X}}$	Binary assignment tensor composed by the elements $x_{u,k,m}$
Z	Number of scatterers in the 3D one-ring channel model
Δf	Space between two adjacent subcarriers in a RB
Ω_s^{target}	QoS/QoE target value for a service s
$\Omega_u^{\dagger}(\omega)$	Function that maps a given QoS/QoE value ω into a rate value for a UE u
$\Omega_u(r)$	Function that maps a given rate value r into a QoS/QoE value for a UE u
$\Phi_{b,n,u,z}$	Initial random phase uniformly distributed between $-\pi$ and π of the z-th component of the channel between the n-th antenna element of the BS b and the antenna of the UE u
Υ_s	Satisfaction rate of a service s
$\text{vec}\{\cdot\}$	Vectorization operator
α_{PL}	Linear coefficient of $PL_{b,n,u,k,z}^{(\text{dB})}$, which describes among other effects, the path loss in a reference distance and other system losses

β_{PL}	Angular coefficient of $PL^{(\mathrm{dB})}_{b,n,u,k,z}$, which models among other effects, the path loss exponent of the environment
$\chi_{b,u}$	Shadowing coefficient of the link between the BS b and the UE u
η_k	MCS assigned to be used in the RB k
$\boldsymbol{\eta}$	Vector MCS composed by the elements η_k
$\gamma_{b,u,k}[t]$	SINR between the BS b and the UE u in the RB k at TTI t
$\tilde{\gamma}_{u,k}[t]$	Estimated SINR of the UE u in the RB k at TTI t
$\inf\{\cdot\}$	Infimum operator
λ_k	Wave length of the central subcarrier of the RB k
$\mathrm{unit}(\theta,\phi)$	Function that converts a pair of zenith and azimuth (θ,ϕ) in spherical coordinates into a unit column vector with the corresponding Cartesian coordinates
ν_u	Estimated number of RBs needed for the UE u to get satisfied
$\boldsymbol{\nu}_{b,n}$	Coordinate of the n-th antenna element of the BS b
$\check{\phi}_{u,z}$	Azimuth of arrival of the signal in the antenna of the UE u from the z-th scatterer
$\hat{\phi}_{b,n,z}$	Azimuth of departure of the signal from the n-th antenna element from the BS b to the z-th scatterer
ϕ_{sp}	Azimuth spread which characterizes the ellipsoid of scatterers in the 3D one-ring channel model
ψ_u	Minimum rate requirement of the UE u
ρ_u	Slack binary variable which indicates if the UE u must be satisfied
$\boldsymbol{\rho}$	Vector of slack variables composed by the elements ρ_u, for $u \in \mathcal{U}$
σ_I^2	Variance of the perceived interference
σ_n^2	Variance of the AWGN
σ_χ	Shadowing standard deviation
$\tau_{b,u,n,z}$	Delay associated to the signal traveling from the n-th antenna element of the BS b to the antenna of the UE u through the z-th scatterer
θ_{sp}	Zenith spread which characterizes the ellipsoid of scatterers in the 3D one-ring channel model
$\check{\theta}_{u,z}$	Zenith of arrival of the signal in the antenna of the UE u from the z-th scatterer
$\hat{\theta}_{b,n,z}$	Zenith of departure of the signal from the n-th antenna element from the BS b to the z-th scatterer
$\vartheta_{u,k}$	CNR of the UE u in the RB k
$\boldsymbol{\psi}$	Vector of ψ_u
ξ_s	Minimum number of UEs that should be satisfied in the service s
$\boldsymbol{\xi}$	Vector of ξ_s
$\arg\max\limits_{x\in\mathcal{A}}\{f(x)\}$	Maximum argument $x \in \mathcal{A}$ of a function $f(x)$

$\arg\min_{x \in \mathcal{A}} \{f(x)\}$	Minimum argument $x \in \mathcal{A}$ of a function $f(x)$
b	BS index
$c^{(rx)}_{b,u,k}[t]$	Signal received by the UE u by the BS b in the RB k at TTI t
$c^{(tx)}_{b,u,k}[t]$	Signal transmitted from the BS b to the UE u in the RB k at TTI t
$\hat{c}^{(rx)}_{b,u,k}[t]$	Estimated signal that should be received by the UE u by the BS b in the RB k at TTI t
$d_{\phi_{sp}}$	Azimuth scattering radius of the ellipsoid of scatterers in the 3D one-ring channel model
$d_{b,u}$	Distance between the BS b and the UE u
$d_{\theta_{sp}}$	Zenith scattering radius of the ellipsoid of scatterers in the 3D one-ring channel model
f_k	Frequency the central subcarrier of the RB k
$f^{\text{BLER}}_{adapt}(\gamma, m)$	Function that maps a pair of SINR and MCS, γ and m, into a BLER
$f^{\text{CQI}}_{adapt}(\gamma)$	Function that maps a given SINR value γ into a CQI/MCS index
$f^{\text{MCS}}_{adapt}(m)$	Function that maps a given MCS value m into a rate value
$f^{\text{SINR}}_{adapt}(\gamma)$	Function that maps a given SINR value γ into a CQI/MCS index
$g_{u,k}[t]$	Reception filter of UE u in the RB k at TTI t
$h_{b,n,u,k}[t]$	Channel coefficient between the n-th antenna element of the BS b and the UE u in the RB k at TTI t
$\tilde{\mathbf{h}}_{b,u,k}[t]$	CSI vector between the BS b and the UE u in the RB k at TTI t
$\mathbf{h}_{b,u,k}[t]$	Channel vector between the BS b and the UE u in the RB k at TTI t
k	RB index
$m_{u,k}[t]$	MCS of a UE u in a RB k at a TTI t
$n_{b,u,k}[t]$	AWGN coefficient in the link between the BS b and the UE u in the RB k at TTI t
$p_{b,k}[t]$	Power allocated by the BS b to the RB k at TTI t
$q_{s,u}$	Binary variable which indicates if the UE u subscribes the service s
$r_{u,k}[t]$	Instantaneous rate of a UE u in a RB k at TTI t
$\mathbf{r}$	Rate vector of elements r_m
s	Service index
t	TTI index
u	UE index
$v_{u,n}$	Virtual UE node from set $\mathcal{V}$, which is uniquely associated to an RB
$\mathbf{v}_u$	Speed vector of UE u
w	Priority value associated to a UE or RB in the RRA algorithm
$\mathbf{w}$	Priority vector composed by the elements w
$\check{\mathbf{w}}_{b,n,z}$	Outgoing wave plane from the n-th antenna element of the BS b to the z scatterer
$\hat{\mathbf{w}}_{u,z}$	Incoming wave plane from the z scatterer to the antenna of the UE u

$x_{u,k}$	Binary variable which indicates if the RB k is assigned to the UE u
$\tilde{x}_{u,k}$	Variable which indicates the fraction of the RB k must be scheduled to the UE u
z	Scatterer index

SUMMARY

1	**INTRODUCTION**	**21**
1.1	**State-of-the-art**	**22**
1.2	**Objectives and Thesis Structure**	**24**
1.3	**Scientific Contributions**	**26**
2	**GENERALIZED SYSTEM MODELING**	**29**
2.1	**Scenario Overview**	**29**
2.2	**Signal Modeling**	**30**
2.3	**Channel Modeling**	**31**
2.4	**System Level Simulation Modeling**	**34**
2.5	**Performance Metrics**	**38**
2.6	**Simulation Flowchart**	**39**
3	**RESOURCE MANAGEMENT FOR RATE MAXIMIZATION WITH QOE/QOS PROVISIONING IN WIRELESS NETWORKS**	**41**
3.1	**Problem Formulation**	**41**
3.2	**Optimal Solution**	**42**
3.3	**State-of-the-art algorithm**	**45**
3.4	**Proposed suboptimal solution**	**46**
3.4.1	***Step 1: User Selection***	***47***
3.4.2	***Step 2: Initial User Assignment***	***48***
3.4.3	***Step 3: Reallocation***	***52***
3.5	**Performance Analysis**	**54**
3.6	**Chapter Summary**	**69**
4	**QOS/QOE-AWARE SCHEDULING ALGORITHM FOR RATE MAXIMIZATION IN WIRELESS NETWORKS**	**71**
4.1	**Problem Formulation**	**71**
4.2	**Optimal Solution**	**72**
4.3	**Suboptimal Solution**	**73**
4.4	**Benchmark Algorithms**	**76**
4.5	**Performance Analysis**	**77**
4.6	**Chapter Summary**	**90**
5	**POWER AND RESOURCE MANAGEMENT FOR RATE MAXIMIZATION WITH QOS/QOE PROVISIONING IN WIRELESS NETWORKS**	**92**
5.1	**Problem Formulation**	**92**
5.2	**Optimal Solution**	**93**
5.3	**Proposed suboptimal solution**	**97**
5.3.1	***Step 1: User Selection***	***97***

5.3.2 Step 2: Initial RB and Power Allocation 98
5.3.3 Step 3: RB and Power Reallocation 101
5.4 State-of-the-art algorithm 102
5.5 Improvement on the JRAPA algorithm 103
5.6 Performance Analysis 104
5.7 Chapter Summary 118
6 CONCLUSIONS AND FUTURE PERSPECTIVES 120
REFERENCES 124

1 INTRODUCTION

A wide variety of new services and applications is expected to emerge and be present in the upcoming era of wireless communications. Predicted to be commercially deployed beyond 2020, the Fifth Generation (5G) of wireless communications has been the main area of research for the telecommunication industry and academia over the last years. A new body of technologies and techniques (e.g., use of higher frequencies of the spectrum, massive Multiple Input Multiple Output (MIMO) antennas, multi-connectivity, cell densification by the deployment of small cells and different numerologies for frequency and time domains [1]) have been proposed in the literature towards serving the increasing amount of devices in the wireless networks and meeting the diverse requirements of the 5G era applications.

Regarding the wide range of new applications expected for the 5G era, the International Telecommunication Union (ITU) has categorized them into three broad use cases, namely Enhanced Mobile Broadband (eMBB), Ultra-Reliable and Low-Latency Communications (URLLC) and Massive Machine-Type Communications (mMTC) [2]. In terms of requirements, the services and applications in the URLLC use case demand very high reliability and availability as well as very low latency. Examples of URLLC applications are autonomous cars, remote robotics and medical surgery. Meanwhile, eMBB applications require very high throughputs and large bandwidths, and can be exemplified by 4K video and augmented reality applications. Furthermore, services related to smart cities and smart homes, which are included in the mMTC use cases, are predicted to demand low bandwidth, high connection density, enhanced coverage and low energy consumption. Therefore, one can easily see the diversified set of requirements that will be present in the 5G networks.

A widely used approach in the literature to quantify the service experience of mobile users is by measuring Quality of Service (QoS) metrics such as throughput, delay, jitter, battery life packet loss rate, among others [3, 4]. Then, after the QoS measurement, the QoS metrics are compared to some minimum QoS requirement to determine how satisfied the user is, where each user might experience a different level of satisfaction even when considering the same QoS parameter. Consequently, due to the wide and diversified range of requirements envisioned for the 5G era, it might be difficult for network operators to define optimum values for some QoS metrics since each specific service has its particular QoS demands. Thus, a common way of measuring the user experience regardless of the technical requirements of the application being used might be necessary to ease the comparison between the level of satisfaction of different services.

One way of achieving such a goal of shifting from the conventional and numerous network centric metrics to a more unified approach is by using Quality of Experience (QoE) concept. According to what has been discussed in [5], QoE models aim to provide a more subjective measurement of the service quality (i.e., the user experience) by abstracting network centric QoS metrics. In [6], the ITU defined QoE as the perception of the acceptability of a service by the user. The QoE experienced by users is often evaluated in the literature in terms of Mean Opinion Score (MOS), which varies from 1 to 5 and consists of measurements of the QoE subjectively perceived by the users [7].

In [8, 9, 10], generic models presenting a mathematical relationship between QoS and QoE are proposed. More specifically, in [8], the IQX hypothesis (exponential interdependency of QoE and QoS) is presented, which consists of an exponential relationship between QoS metrics and QoE. On the other hand, based on the Weber-Fechner law, which states that the human perception of a certain phenomena diminishes with the increasing magnitude of the stimuli, a logarithmic relationship between QoS and QoE is presented in [9, 10]. Other works also studied and proposed service-specific models to relate QoS and QoE, such as [11, 12, 13, 14], which proposed utility functions that map QoS metrics into QoE for web browsing, Voice over IP (VoIP), video streaming and 3D video traffic, respectively. In [15], the author presented QoS-QoE mapping functions for VoIP, File Transfer Protocol (FTP), video streaming and web-browsing services.

In terms of improving the end-user experience by delivering high-quality content, the QoE concept allows us to incorporate the user demands into optimization problems in a more unified and holistic manner beyond the traditional QoS concept. A direct application of such an approach is on formulating optimization problems for modeling Radio Resource Allocation (RRA) strategies where the user requirements are represented by means of QoE demands. In this context, the main objective of this thesis is to propose RRA strategies, including frequency resource assignment and power allocation, targeting the maximization of the total system throughput while guaranteeing that a minimum number of users have their QoS/QoE demands met.

1.1 State-of-the-art

Several types of RRA algorithms have been proposed in the literature. According to the surveys presented in [3, 16], RRA algorithms can be classified into several different categories, such as opportunistic algorithms, spectral efficient algorithms, fair algorithms, channel-unaware or -aware algorithms, delay-based, throughput-based, queue-based, QoS-unaware or -aware algorithms, multi-class or multi-service algorithms, among others. In [17], a more recent study presented a survey on another category of RRA algorithms, QoE-aware algorithms for wireless networks. Finally, [4] presents an extensive survey on RRA algorithms that take into account the MIMO technology. The purpose of this thesis is not to perform a complete or exhaustive survey about RRA algorithms, for which the readers are directed to the works in [3, 4, 16, 17] for more detailed surveys. In the following, some works from the literature on RRA techniques that have some similarity to the work on this thesis are highlighted.

Since the wireless resources are becoming scarcer and more expensive, operators are concerned about using them efficiently to achieve high transmission rates while providing high satisfaction to the users; this task is becoming more challenging with the increasing number of system users. In this context, the maximization of the total system rate is a topic well discussed in the literature. In fact, it is well-known that the rate of systems with orthogonal resources is maximized when the resources are assigned to users that have better channel quality on each resource. However, this solution usually favors users close to the base station and leads cell-edge users to starvation [3].

In order to guarantee that all users in the system receive resources, some works study optimum resource allocation problems aiming at maximizing the total data rate maximization subject to QoS constraints [18, 19, 20]. In [18], the optimal solution is provided as an Integer Linear Problem (ILP) and a low complexity heuristic is proposed. In [19], a scenario with non-real-time and real-time services is considered. Therein, only users subscribed in the real-time service have minimum QoS constraints given in terms of maximum packet delay and loss probability. In [20], the optimal solution of the maximization of the weighted sum of the users' rates subject to minimum individual QoS requirements is provided in a scenario where the system operator offers a video service to the users. Moreover, the users' QoS is given in terms of the minimum Bit Error Rate (BER) therein. In [21], the authors proposed a suboptimal approach of firstly assigning sub-bands giving priority to the users that need more power to achieve the minimum QoS requirement; after assigning all sub-bands, the remaining power is allocated in order to maximize the system capacity.

Notice that in [18, 19, 20], the network operator intends to satisfy all users. However, the network operator, in general, requires that at least a certain fraction of the users be satisfied per service due to resources scarcity and/or economic reasons [22, 23]. In this context, in [24] the problem of maximizing the total system rate considering a multi-service scenario has been analyzed, where each service must have at least a certain number of users with their QoS requirement satisfied. The users' QoS is measured in terms of their individual throughput and the requirement is defined by the service the user subscribed to. Therein, the optimal solution for this problem is modeled and a low complexity heuristic, called Reallocation-based Assignment for Improved Spectral Efficiency and Satisfaction (RAISES) is proposed. Later, the authors extended their work to a Multi-User Multiple Input Multiple Output (MU-MIMO) case in [25]. Another extension of [24] is presented in [26], where the authors evaluated the achievable performance gains using the joint optimization of adaptive power allocation and frequency resource assignment aiming at maximizing the spectral efficiency.

The works discussed so far take into account QoS measurements as the main criteria for performing the RRA. However, as mentioned before, the concept of QoE allows us to shift from the conventional and numerous QoS metrics to a more unified approach more focused on the user experience. Considering this shift from QoS to QoE, the authors in [27] compared the performance of three QoS-based RRA algorithms (namely max rate, max-min rate and proportional fair) in terms of QoE metrics (geometric mean QoE and average QoE). The conclusion drawn from [27] was that the performance in terms of QoE still needed to be enhanced.

Some works are found in the literature proposing opportunistic algorithms with QoE considerations. For example, in [28], the RRA problem is modeled considering a bounded optimization problem with the objective of achieving the maximum overall QoE taking into account a constraint for the total transmit power at the transmitter. Also, the authors proposed in [29] a power allocation technique for video services in MIMO wireless systems. The main objective in [29] was to maximize the QoE and the problem was decomposed into sub-problems, while a bisection search algorithm is used to provide the upper bound solution by computing the optimum values. Finally, in [30], a multi-cell wireless system is consider during the formulation of an RRA algorithm for interference mitigation and overall QoE maximization, where the multiple transmitters play a cooperative game with peer transmitters when scheduling users and allocating power. However, as expected from opportunistic algorithms, the works in [28, 29, 30] sometimes degrade the user experience when the users undergo poor link conditions.

A common way to avoid the starvation of users experiencing poor link conditions is by considering fairness during the RRA process. In [31], the authors proposed proportional fair RRA algorithms considering not only the users' QoE maximization but also the fairness among users. The authors in [32] considered the problem of maximizing the minimum MOS (which objectively quantifies the users' QoE) among the users subject to a minimum number of satisfied users. More specifically, the work in [32] proposed a snapshot-based RRA scheme, considering resource assignment and power allocation, relying on a heuristic approach and showed using system level simulations that their proposal performed close to the optimum solution. Later, in [33], an extension of [32] was proposed considering that the users should be satisfied on average over a certain timespan instead of being satisfied every snapshot as in [32]. Another work focusing on maximizing the minimum MOS is presented in [34], where the authors converted a intractable Mixed Integer Linear Programming (MILP) into an equivalent convex form and then developed a fast algorithm to efficiently solve the problem. Other works also aiming at maximizing the minimum MOS of the system are also presented in [35, 36]

Considering QoE measurements, besides maximizing the maximum overall QoE and maximizing the minimum MOS, some other works from the literature intend to maximize the number of users having their minimum QoE requirements satisfied [37] or maximize the energy efficiency of the system [38]. However, as far as the author's knowledge goes, no previous work discusses the rate maximization problem subject to satisfaction constraints based on QoS/QoE measurements, which is the main focus of this thesis.

1.2 Objectives and Thesis Structure

The main objective of this thesis is to propose new methods of RRA for multi-service scenarios. The proposed methods aim at maximizing the overall system throughput, while ensuring the fulfillment of QoS/QoE requirements for a fraction of the users of each individual service. The outline of this thesis is structured in the sequel.

In Chapter 2, the system model adopted in the following chapters is described. In this context, the general characteristics of the wireless communication system are explained, along with the channel models and the performance metrics. Besides, the simulation chain used to benchmark the algorithms proposed in the next chapters is also explained.

Chapter 3 addresses the resource assignment problem of maximizing the overall system rate, guaranteeing that a minimum number of User Equipments (UEs) of each service plan meet their QoS/QoE requirements. It is considered in this study that the power is equally divided among all the frequency resources, i.e., it is not part of the studied problem. The resource assignment addressed in this chapter is performed and evaluated instantaneously for each individual time slot. The studied problem is initially formulated as an optimization problem, which is further reformulated in a more tractable form. Nevertheless, solving the resource assignment problem optimally is impracticable due to its excessive computational burden. Therefore, a low complexity heuristic is proposed and its performance is evaluated by means of computational simulations and compared against the state-of-the-art algorithm as well as the optimal solution.

Chapter 4 extends the study performed in Chapter 3 by considering that the QoS/QoE requirements must be achieved in a given timespan, instead of a single snapshot. This consideration implies the need of considering the temporal evolution of the system performance metrics. Like in Chapter 3, here, this problem is mathematically formulated as an optimization problem. However, besides the high computational complexity, in order to assess the optimal solution, the system must have the knowledge of the users' channel conditions during the entire timespan, which is often an unrealistic assumption. Therefore, from the study of the optimization problem formulation, a low complexity suboptimal algorithm is proposed to tackle the problem, extending the heuristic proposed in Chapter 3. The suboptimal algorithm proposed in this chapter is evaluated by means of computational simulations and compared against two benchmark algorithms that have similar goals as the studied problem.

In Chapter 5, the problem of allocating the radio resources aiming at maximizing the overall system rate, while ensuring that a minimum number of users of each service meet their QoS/QoE requirements, originally studied in Chapter 3, is revisited. Furthermore, the allocation problem addressed in this chapter is also analyzed instantaneously, i.e., the users' QoS/QoE requirements must be met in a single time slot, similarly as done in Chapter 3. However, differently from Chapter 3, here, the available power and the frequency resources are jointly allocated, hardening the allocation problem. In a similar manner as done in Chapters 3 and 4, the radio resource allocation problem is mathematically formulated as an optimization problem and it is further rewritten in an equivalent form that can be solved by standard methods present in the literature. Due to the high computational complexity, solving this problem optimally is prohibitively time-consuming for a real time system. Thus, in this chapter two algorithms are proposed. The first one employs the solution framework also adopted in Chapters 3 and 4, and the second one stands for an improvement over the state-of-the-art heuristic which does not increase its worst case computational complexity. Both heuristics proposed in this chapter are compared against the state-of-the-art algorithm and the optimal solution by means of computational simulations.

Finally, Chapter 6 presents the main conclusions and some future perspectives of this thesis.

1.3 Scientific Contributions

Currently, the content of this thesis has been partially published with the following bibliographic information:

- SOUSA, D. A.; MONTEIRO, V. F.; MACIEL, T. F.; LIMA, F. R. M.; CAVALCANTI, F. R. P. Resource Management for Rate Maximization with QoE Provisioning in Wireless Networks. **Journal of Communication and Information Systems (JCIS)**, v. 31, n. 1, p. 290–303, 2016. ISSN 1980-6604. DOI: `10.14209/jcis.2016.25`
- SOUSA, D. A.; MAURÍCIO, W. V. F.; ANTONIOLI, R. P.; MACIEL, T. F.; LIMA, F. R. M. Improved Joint Resource and Power Allocation Algorithm with QoS Provisioning. In: PROCEEDINGS OF THE BRAZILIAN TELECOMMUNICATIONS SYMPOSIUM (SBrT), Sept. 2018, Campina Grande, Brazil. **Proceedings [...]** Campina Grande, Brazil: [s.n.], Sept. 2018. P. 1–5

Alongside the work in the Ph.D. program, the author has been working on research projects. From the work developed in these projects, the author collaborated in the following scientific publications:

Journal Papers

- ANTONIOLI, R. P.; PARENTE, G. C.; SILVA, C. F. M. e.; SOUSA, D. A.; RODRIGUES, E. B.; MACIEL, T. F.; CAVALCANTI, F. R. P. Dual Connectivity for LTE-NR Callular Networks: Challenges and Open Issues. **Journal of Communication and Information Systems (JCIS)**, 2018
- MONTEIRO, V. F.; SOUSA, D. A.; MACIEL, T. F.; CAVALCANTI, F. R. P.; SILVA, C. F. M. e.; RODRIGUES, E. B. Distributed RRM for 5G Multi-RAT Multi-Connectivity Networks. **IEEE Systems**, p. 1–13, 2018. ISSN 1932-8184. DOI: `10.1109/JSYST.2018.2838335`
- ANTONIOLI, R. P.; RODRIGUES, E. B.; MACIEL, T. F.; SOUSA, D. A.; CAVALCANTI, F. R. P. Adaptive resource allocation framework for user satisfaction maximization in multi-service wireless networks. **Telecommunication Systems**, v. 68, n. 2, p. 259–275, June 2018. ISSN 1572-9451. DOI: `10.1007/s11235-017-0391-3`

- COSTA NETO, F. H.; RODRIGUES, E. B.; SOUSA, D. A.; MACIEL, T. F.; CAVALCANTI, F. R. P. QoS-aware scheduling algorithms to enhance user satisfaction in OFDMA systems. **Transactions on Emerging Telecommunications Technologies**, v. 28, n. 10, p. 1–15, 2017. ISSN 1541-8251. DOI: `10.1002/ett.3165`

- MAURÍCIO, W. V. F.; LIMA, F. R. M.; SOUSA, D. A.; MACIEL, T. F.; CAVALCANTI, F. R. P. Joint Resource Block Assignment and Power Allocation Problem for Rate Maximization With QoS Guarantees in Multiservice Wireless Systems. **Journal of Communication and Information Systems (JCIS)**, v. 31, n. 1, p. 211–223, 2016. ISSN 1980-6604. DOI: `10.14209/jcis.2016.19`

- MONTEIRO, V. F.; SOUSA, D. A.; MACIEL, T. F.; LIMA, F. R. M.; CAVALCANTI, F. R. P. A QoE-Aware Scheduler for OFDMA Networks. **Journal of Communication and Information Systems (JCIS)**, v. 31, n. 1, p. 41–48, Nov. 2016. ISSN 1980-6604. DOI: `10.14209/jcis.2016.3`

- MONTEIRO, V. F.; SOUSA, D. A.; MACIEL, T. F.; LIMA, F. R. M.; RODRIGUES, E. B.; CAVALCANTI, F. R. P. Radio resource allocation framework for quality of experience optimization in wireless networks. **IEEE Network**, v. 29, n. 6, p. 33–39, Nov. 2015. ISSN 0890-8044. DOI: `10.1109/MNET.2015.7340422`

Conference Papers

- ANTONIOLI, R. P.; RODRIGUES, E. B.; MACIEL, T. F.; SOUSA, D. A.; CAVALCANTI, F. R. P. Alocação de Recursos Adaptativa para Maximização da Satisfação dos Usuários em Redes Celulares. In: PROCEEDINGS OF THE BRAZILIAN TELECOMMUNICATIONS SYMPOSIUM (SBrT), Sept. 2017, São Pedro, Brazil. **Proceedings [...]** São Pedro, Brazil: [s.n.], Sept. 2017. P. 1–5

- DE PAULA, T. D.; SOUSA, D. A.; MACIEL, T. F. Estudo do Compromisso Entre Eficiência Espectral e Justiça na Alocação de Recursos Rádio. In: PROCEEDINGS OF THE BRAZILIAN TELECOMMUNICATIONS SYMPOSIUM (SBrT), Sept. 2017, São Pedro, Brazil. **Proceedings [...]** São Pedro, Brazil: [s.n.], Sept. 2017. P. 1–5

- MONTEIRO, V. F.; SOUSA, D. A.; MACIEL, T. F.; LIMA, F. R. M.; CAVALCANTI, F. R. P. Alocação de Recursos em Redes Sem Fio Baseada na Qualidade de Experiência do Usuário. In: PROCEEDINGS OF THE BRAZILIAN TELECOMMUNICATIONS SYMPOSIUM (SBrT), Sept. 2015, Juiz de Fora, Brazil. **Proceedings [...]** Juiz de Fora, Brazil: [s.n.], Sept. 2015. P. 1–5

- MONTEIRO, V. F.; SOUSA, D. A.; NETO, F. H. C.; RODRIGUES, E. B.; MACIEL, T. F.; CAVALCANTI, F. R. P. Throughput-Based Satisfaction Maximization for a

Multi-Cell Downlink OFDMA System Considering Imperfect CSI. In: PROCEEDINGS OF THE BRAZILIAN TELECOMMUNICATIONS SYMPOSIUM (SBrT), Sept. 2015, Juiz de Fora, Brazil. **Proceedings [...]** Juiz de Fora, Brazil: [s.n.], Sept. 2015. P. 1–5

- OSTERNO, I. S.; SOUSA, D. A.; FERNANDES, C. E. R. On supervised channel estimation techniques for very large MIMO communication systems. In: INTERNATIONAL TELECOMMUNICATIONS SYMPOSIUM, Sept. 2014, São Paulo, Brazil. **Proceedings [...]** São Paulo, Brazil: [s.n.], Sept. 2014. P. 1–5. DOI: `10.1109/ITS.2014.6948023`

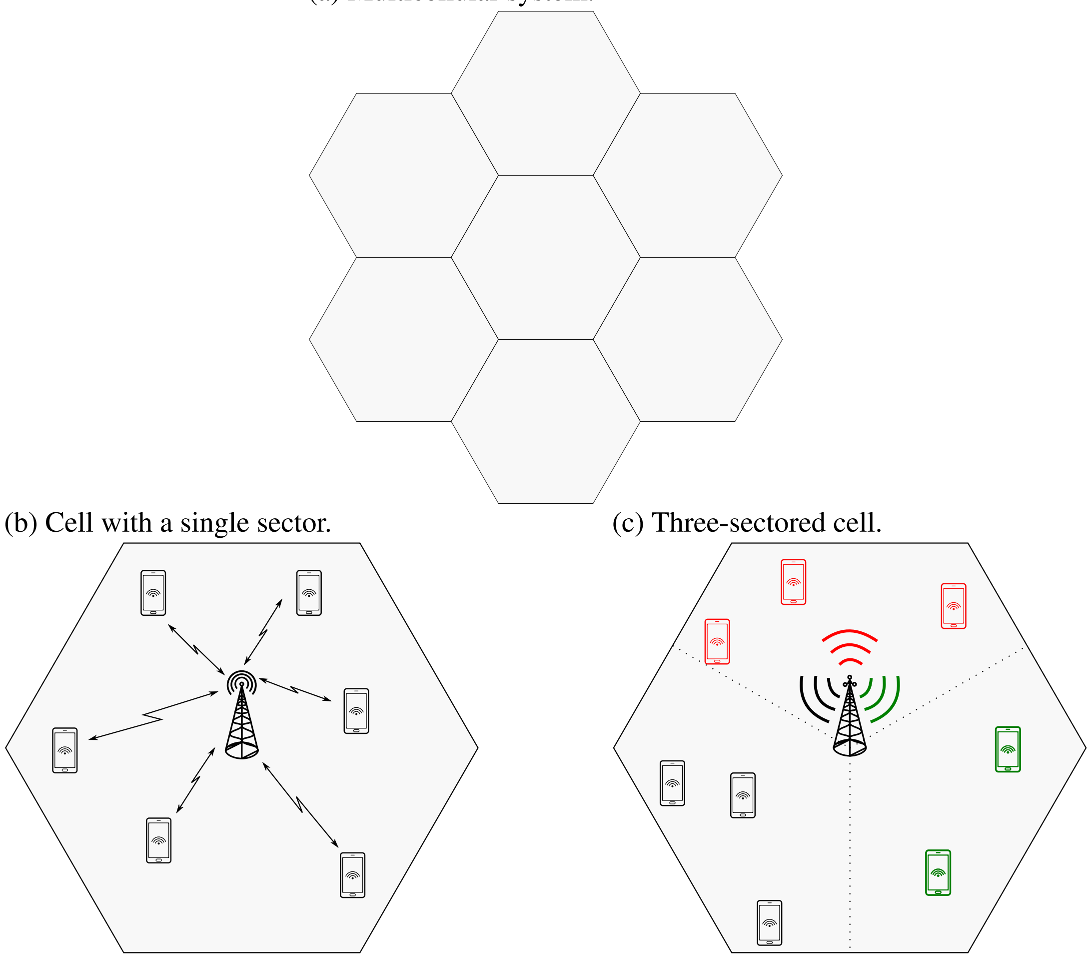
2 GENERALIZED SYSTEM MODELING

The system model used in the simulations of this thesis is detailed in this chapter. In Section 2.1, the general characteristics of the system are presented. Sections 2.2 and 2.3 specify the signal and channel modeling, respectively. In Section 2.4, the system level simulation framework is detailed. Finally, Section 2.5 describes the performance evaluation metrics.

2.1 Scenario Overview

In this thesis, a multicellular system is considered, where each cell may have one or multiple sectors. Each sector has a Base Station (BS) $b \in \mathcal{B} = \{1, 2, \ldots, B\}$ positioned on the center of the cell where the sector belongs to. Moreover, each BS b serves a set $\mathcal{U}_b = \{1, 2, \ldots, U_b\}$ of UEs distributed on its coverage area. The multicellular scenario is depicted in Fig. 2.1a. Fig. 2.1b illustrates the case where the cell has a single sector, while Fig. 2.1c shows an example of a cell composed of three sectors.

Figure 2.1 – Scenario overview.



(a) Multicellular system.

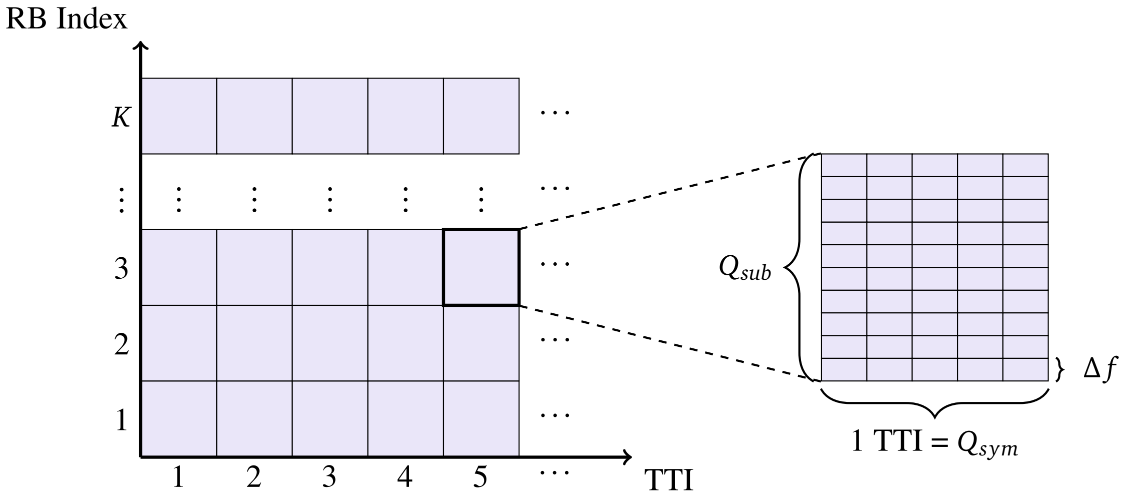
(b) Cell with a single sector.

(c) Three-sectored cell.

Source: Created by the author.

The BSs employ as multiple access scheme a combination of Orthogonal Frequency Division Multiple Access (OFDMA) and Time Division Multiple Access (TDMA), allocating a set $\mathcal{K} = \{1, 2, \ldots, K\}$ of time-frequency Resource Blocks (RBs) to the UEs. Scheduling subcarriers individually would require a signaling overhead to the system. Consequently, an RB is considered the minimum allocable resource [39]. Each RB consists of a set of Q_{sub} adjacent subcarriers spaced of Δf Hz in frequency-domain and a set of Q_{sym} consecutive symbols in time-domain, whose total duration corresponds to one Transmission Time Interval (TTI), as presented in Fig. 2.2. The RBs are designed to have a bandwidth equal to $Q_{sub} \cdot \Delta f$ that is smaller than the system's coherence bandwidth. Moreover, it is assumed that the coherence time of the system is larger than a TTI. It means that the channel response of an RB can be considered flat during its Q_{sym} symbols on its Q_{sub} subcarriers.

Figure 2.2 – Frequency-time grid of RBs.



Source: Created by the author.

Due to the diversity of applications with distinct requirements, the UEs may be separated into different mobile service subscription plans. In the considered system, each UE can subscribe only to a single service plan from a set $\mathcal{S} = \{1, 2, \ldots, S_b\}$, with $\mathcal{U}_{b,s}$ corresponding to the set of subscribers of the service plan $s \in \mathcal{S}$, where $\bigcup_{s\in\mathcal{S}} \mathcal{U}_{b,s} = \mathcal{U}_b$ and $\bigcap_{s\in\mathcal{S}} \mathcal{U}_{b,s} = \emptyset$.

2.2 Signal Modeling

The BSs are equipped with an array of N antennas elements vertically polarized and the UEs have only a single antenna with the same polarization. Furthermore, each RB can be assigned to only one UE per TTI. Note that as the RBs are orthogonal to each other in time and frequency and there is no resource reuse inside a sector, there is no intra-cell interference among the UEs. However, the UEs might experience inter-cell interference from other sectors that share the same resources.

Considering the downlink, in order to transmit data to a UE $u \in \mathcal{U}_b$ in the TTI t, the BS b must assign an RB k to the UE and allocate a transmit power $p_{b,k}[t]$ to the RB k, where the powers allocated to the RBs are constrained by the total power P_{total} available at BS b, i.e., $\sum_{k\in\mathcal{K}} p_{b,k}[t] \leq P_{\text{total}}, \forall b \in \mathcal{B}$. Therefore, the signal received by UE u from BS b in RB k at TTI t, $c^{(rx)}_{b,u,k}[t]$, can be modeled as

$$c^{(rx)}_{b,u,k}[t] = \underbrace{g_{u,k}[t]\mathbf{h}^{\mathrm{T}}_{b,u,k}[t]\mathbf{f}_{b,k}[t]\sqrt{p_{b,k}[t]}c^{(tx)}_{b,u,k}[t]}_{\text{Desired signal}} + \underbrace{g_{u,k}[t]\sum_{b'\in\mathcal{B}\setminus\{b\}}\mathbf{h}^{\mathrm{T}}_{b',u,k}[t]\mathbf{f}_{b',k}[t]\sqrt{p_{b',k}[t]}c^{(tx)}_{b',u',k}[t]}_{\text{Inter-cell Interference}} + \underbrace{n_{b,u,k}[t]}_{\text{Noise}}, \quad (2.1)$$

where $c^{(tx)}_{b,u,k}[t] \in \mathbb{C}$ is the transmitted signal with unit average power, i.e., $\mathbb{E}\left\{\left|c^{(tx)}_{b,u,k}[t]\right|^2\right\} = 1$, in which $\mathbb{E}\{\cdot\}$ denotes the expectation operator and $|\cdot|$ returns the absolute value. The vector $\mathbf{h}_{b,u,k}[t] = \begin{bmatrix} h_{b,1,u,k}[t] & h_{b,2,u,k}[t] & \cdots & h_{b,N,u,k}[t] \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$ with each element $h_{b,n,u,k}[t]$ representing the channel response coefficient of the link between the n-th antenna element of the BS b and the antenna of the UE u in RB k. The vector $\mathbf{f}_{b,k}[t] \in \mathbb{C}^{N\times 1}$ and the scalar $g_{u,k}[t] \in \mathbb{C}$ are the unit transmission and reception filters, respectively, i.e., $\left\|\mathbf{f}_{b,k}[t]\right\| = \left\|g_{u,k}[t]\right\| = 1$, where $\|\cdot\|$ represents the ℓ^2-norm operator (Euclidean norm). The term $n_{b,u,k}[t]$ denotes the Additive White Gaussian Noise (AWGN) that is modeled as a Zero Mean Circularly Symmetric Complex Gaussian (ZMCSCG) with variance σ_n^2.

2.3 Channel Modeling

Several channel models exist in the literature [40, 41, 42, 43], where each one tries to capture more precisely some aspects of the channel, such as spatial consistency, cross polarization, frequency-time correlation, among others.

The simplest model present in the literature consists in modeling the channel coefficients $h_{b,n,u,k}[t]$ as Independent and Identically Distributed (IID) ZMCSCG variables with variance equal to the large-scale fading, which is detailed later in this section. This model does not take into account either the spatial correlation between the antennas, or the temporal correlation of the channel samples. Nevertheless, even with its simplicity, it is widely used in the literature [40, 44, 45]. In this thesis, this channel is considered for snapshot simulations (a single TTI is simulated) and when the BSs employ a single antenna.

Another class of channel model is the stochastic-geometric. In this model, the antenna geometry of both BS and UE are considered during the channel generation. Moreover, the channel for each link between one antenna element of the BS and one of the UE is geometrically generated by summing the contributions of Z individual scatterers with specific propagation parameters, such as delay, path loss, angle of arrival and angle of departure. Therefore, the channel coefficient between an antenna element n of the BS b and the antenna of the UE u in an RB k at TTI t, considering a stochastic-geometric model, can be modeled as

$$h_{b,n,u,k}[t] = \frac{1}{\sqrt{Z}} \sum_{z=1}^{Z} \underbrace{\sqrt{L_{b,n,u,k,z}}}_{\text{Large-scale fading}} \cdot \underbrace{\exp\left(j\Phi_{b,n,u,z}\right)}_{\text{Initial random phase shift}} \cdot \underbrace{\exp\left(j2\pi \frac{\hat{\mathbf{w}}_{u,z}^{\mathrm{T}} \boldsymbol{\nu}_u}{\lambda_k}\right)}_{\text{Receiver steering direction}} \cdot \underbrace{\exp\left(j2\pi \frac{\check{\mathbf{w}}_{b,n,z}^{\mathrm{T}} \boldsymbol{\nu}_{b,n}}{\lambda_k}\right)}_{\text{Transmitter steering direction}} \cdot \underbrace{\exp\left(j2\pi \frac{\hat{\mathbf{w}}_{u,z}^{\mathrm{T}} \mathbf{v}_u}{\lambda_k} T_{\mathrm{tti}} t\right)}_{\text{Doppler effect}} \cdot \underbrace{\exp\left(-j2\pi f_k \tau_{b,u,n,z}\right)}_{\text{Fourier transform}}. \tag{2.2}$$

Note that in (2.2), $h_{b,n,u,k}[t]$ is modeled as the contribution of Z scatterers, which characterizes a multipath fading channel. Each multipath has an initial random phase $\Phi_{b,n,u,z}$, which is modeled as a random variable uniformly distributed between $-\pi$ and π. The receiver steering direction describes the phase experienced by the incident incoming wave from the direction $\hat{\mathbf{w}}_{u,z}$ in the antenna of the UE u. The direction $\hat{\mathbf{w}}_{u,z}$ denotes a unit vector $\hat{\mathbf{w}}_{u,z} = \mathrm{unit}\left(\check{\theta}_{u,z}, \check{\phi}_{u,z}\right)$, where $\check{\theta}_{u,z}$ and $\check{\phi}_{u,z}$ correspond to the zenith and azimuth angles of arrival, respectively, and the $\mathrm{unit}(\theta, \phi)$ function is defined as

$$\mathrm{unit}(\theta, \phi) = \begin{bmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix}. \tag{2.3}$$

The vector $\boldsymbol{\nu}_u$ represents the coordinate of the antenna of the UE u and λ_k is the wavelength of the central subcarrier, with frequency f_k, of the RB k. In a similar way, the transmitter steering direction models the phase experienced by an outgoing wave from the z-th antenna element of the BS b with direction $\check{\mathbf{w}}_{b,n,z}$. The direction $\check{\mathbf{w}}_{b,n,z}$ is also a unit vector defined as $\check{\mathbf{w}}_{b,n,z} = \mathrm{unit}\left(\hat{\theta}_{b,n,z}, \hat{\phi}_{b,n,z}\right)$, where $\hat{\theta}_{b,n,z}$ and $\hat{\phi}_{b,n,z}$ correspond to the zenith and azimuth angles of departure, respectively. The vector $\boldsymbol{\nu}_{b,n}$ represents the coordinate of the n-th antenna element of BS b. The Doppler effect depends on the speed vector of the UE $\mathbf{v}_u$ and is responsible for the time varying channel component.

Since OFDMA is assumed as the multiple access scheme, the channel coefficients used herein are represented in the frequency domain. Due to this fact, there is a Fourier transform component in the channel response, in which $\tau_{b,u,n,z}$ denotes the associated delay of the signal traveling from the n-th antenna of the BS b to the antenna of UE u by the z-th scatterer.

The term $L_{b,n,u,k,z}$ corresponds to the large-scale fading and can be expanded as

$$L_{b,n,u,k,z} = PL_{b,n,u,k,z}^{-1} \cdot \chi_{b,u} \cdot A_u\left(\check{\theta}_{u,z}, \check{\phi}_{u,z}\right) \cdot A_b\left(\hat{\theta}_{b,n,z}, \hat{\phi}_{b,n,z}\right), \tag{2.4}$$

where $PL_{b,n,u,k,z}$ is the average path loss experienced by each ray z from the n-th antenna of the BS b to the UE u. $\chi_{b,u}$ corresponds to the shadowing coefficient in the link between b and u.

$A_b(\cdot)$ and $A_u(\cdot)$ are functions modeling the radiation power pattern of the antennas of BS b and the UE u, respectively.

There are many formulations in the literature modeling the path loss $PL_{b,n,u,k,z}$ [40, 42, 43]. These formulations may depend on the system frequency carrier, the heights of the BS and UE, the environment, the distance between BS and UE, among other parameters. However, once the system is deployed, the only relevant parameter that varies with the time in most of these models is the distance $d_{b,u}$ between the BS b and UE u, given by

$$d_{b,u} = \|\boldsymbol{\nu}_u - \boldsymbol{\nu}_b\|, \tag{2.5}$$

where $\boldsymbol{\nu}_b$ is the coordinate of the center of the antenna array of BS b. Therefore, for a given scenario, the path loss can be modeled in dB scale as

$$PL^{(\text{dB})}_{b,n,u,k,z} = \alpha_{PL} + \beta_{PL}\log_{10}(d_{b,u}), \tag{2.6}$$

where $PL^{(\text{dB})}_{b,n,u,k,z} = 10\log_{10}\left(PL_{b,n,u,k,z}\right)$, and the coefficients α_{PL} and β_{PL} characterize the environment. α_{PL} describes among other effects, the path loss in a reference distance and other system losses. Analogously, the coefficient β_{PL} models among other effects, the path loss exponent, which is dependent of the distance $d_{b,u}$ [40].

The shadowing coefficient $\chi_{b,u}$ models the impact of blockages in the environment and in this thesis it is modeled as a log-normal random variable with zero mean and standard deviation σ_χ [41].

In the literature, many channel models propose how to define coefficients of azimuth and zenith of arrival and departure, as well as the delays. In this thesis, when multiple antennas are employed in the BS, i.e., $N > 1$, the channel model that will be considered is a generalization of the one ring channel model presented in [41] to a 3D scenario. The model presented in [41] is characterized only for a 2D coordinate system. In a 3D version of the one-ring model, the scatterers are uniformly positioned in the surface of a 3D ellipsoid centered in the UE position, as depicted in Fig. 2.3, where the points s_1, s_2, s_3 and s_4 are examples of scatterers.

The ellipsoid in Fig. 2.3 is characterized by a pair of zenith and azimuth spreading angles (θ_{sp}, ϕ_{sp}), which describe how the scatterers are positioned around the UE. In other words, they define the aperture of the ellipsoid where the scatterers will be placed on.

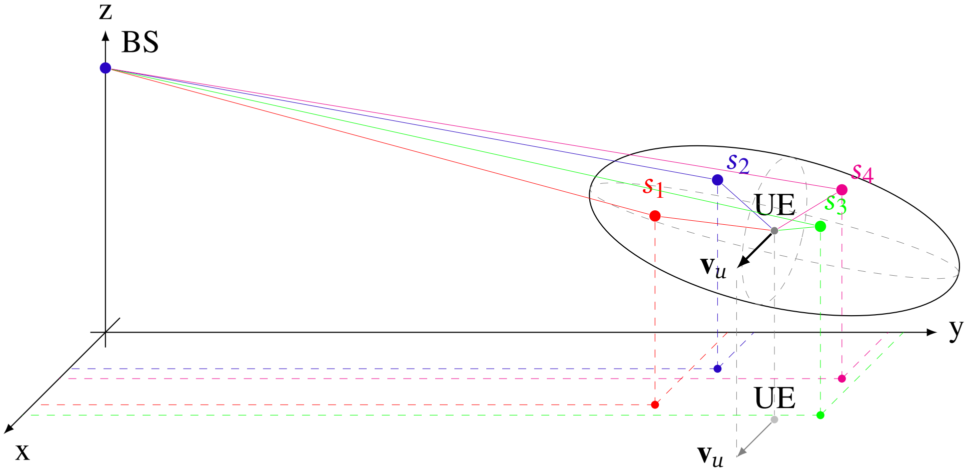
Fig. 2.4 presents the XY and YZ cuts of the model presented in Fig. 2.3. In Fig. 2.4a, the XY-plane view of the 3D model is presented. The azimuth scattering radius $d_{\phi_{sp}}$ of the ellipsoid is related to the distance $d_{b,u}$ and a given azimuth spreading angle, ϕ_{sp}, thus

$$d_{\phi_{sp}} = d_{b,u}\tan\left(\frac{\phi_{sp}}{2}\right). \tag{2.7}$$

In a similar manner, in Fig. 2.4b, the YZ-plane view of the 3D one-ring channel model is depicted. The zenith scattering radius $d_{\theta_{sp}}$ of the ellipsoid can be calculated based on the distance $d_{b,u}$ and a given zenith scattering angle, θ_{sp}, i.e.,

$$d_{\theta_{sp}} = d_{b,u}\tan\left(\frac{\theta_{sp}}{2}\right). \tag{2.8}$$

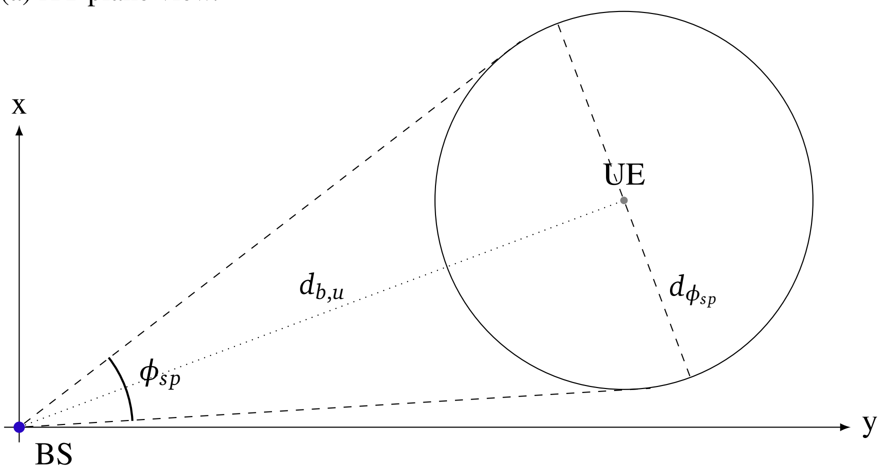
Figure 2.3 – 3D view.



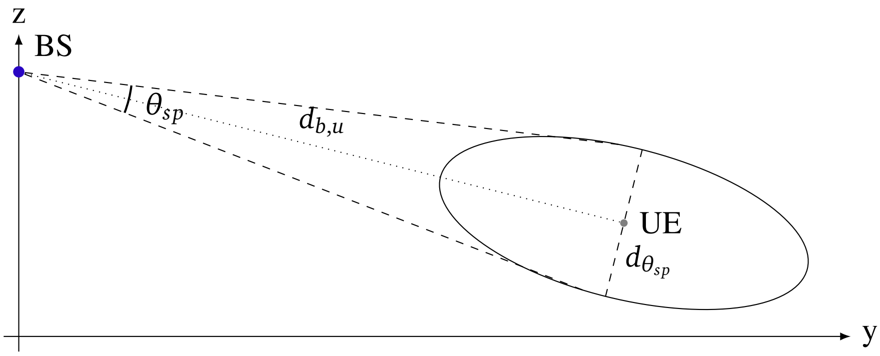
Source: Created by the author.

Figure 2.4 – Geometrical model for 3D one-ring with scatterers around the UE.

(a) XY-plane view.



(b) YZ-plane view.



Source: Created by the author.

2.4 System Level Simulation Modeling

The simulation results presented in this thesis are obtained using a system level simulator. In this kind of simulations, the transmission of the information itself (bit-by-bit or symbol-by-symbol) is not implemented. Instead, the success or not of a transmission is abstracted by the value of the link's Signal to Interference-plus-Noise Ratio (SINR) and the Channel Quality

Indicator (CQI) considered for the transmission. The SINR is obtained from (2.1) and is given by

$$\gamma_{b,u,k}[t] = \frac{\left|g_{u,k}[t]\mathbf{h}^{\mathrm{T}}_{b,u,k}[t]\mathbf{f}_{b,k}[t]\right|^2 p_{b,k}[t]}{\sum\limits_{b'\in\mathcal{B}\setminus\{b'\neq b\}} \left|g_{u,k}[t]\mathbf{h}^{\mathrm{T}}_{b',u,k}[t]\mathbf{f}_{b',k}[t]\right|^2 p_{b',k}[t] + \sigma_n^2}, \tag{2.9}$$

where it is assumed that the interfering signals are summed coherently, which implies a worst case SINR value.

The CQI value maps the channel quality into a scalar value, which is reported by the UEs to the BS to which they are connected to. Each CQI value is associated to a Modulation and Coding Scheme (MCS), which indicates to the system which modulation and coding schemes should be used during the transmission. Each MCS provides a trade-off between susceptibility to transmission errors and data rate. In this thesis, the CQI table of the Long Term Evolution (LTE) standard, presented in Table 2.1, is considered.

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

CQI	**Modulation**	**Code Rate** [÷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: [46].

Since the CQIs are obtained from the SINR values, in order to execute the Radio Resource Management (RRM) algorithms, prior information about the SINR must be considered. By looking at eq. (2.9), in order to perfectly estimate the SINR, all the BSs in the system would need to share their scheduling decisions taken by their RRM algorithms. Besides that, the BSs would need to share for each RB, the Channel State Information (CSI) of the links between them and all UEs scheduled in it. In summary, obtaining prior inter-cell interference values is prohibitive due to the enormous amount of information that should be exchanged among all the BSs in the system.

The CQI is a metric useful for the decisions taken by the RRM algorithms which are responsible for important tasks, such as:

- RRA: Schedule which UE will use a given RB during the current TTI;
- Power allocation: Allocate the power $p_{b,k}[t]$ that will be used in the RBs, constrained by a maximum value P_{total};
- Precoding: Define the precoder and decoder that will be used during the transmission.

Due to the stochastic behavior of wireless data traffic, the interference modeling and its mitigation in packet-switched systems are challenging issues. In the literature, some works propose methods to mitigate the interference, such as coordination among BSs and interference alignment [47, 48, 49, 50, 51]. On the other hand, for the sake of simplicity, some papers treat the inter-cell interference as part of the AWGN added to the received signal in the UE, which is valid when the number of UEs and BSs increases [52]. Since a more detailed/sophisticated interference modeling and mitigation is out of the scope of this thesis, it is considered that the inter-cell interference can be modeled as a ZMCSCG variable $I[t]$ with variance σ_I^2, which can be incorporated to the noise. Moreover, instead of the actual channel coefficients $\mathbf{h}_{b,u,k}[t]$, the BS b makes use of the available CSI, $\tilde{\mathbf{h}}_{b,u,k}[t]$. Therefore, the estimated received signal, $\hat{c}^{(rx)}_{b,u,k}[t]$, can be obtained from (2.1) as

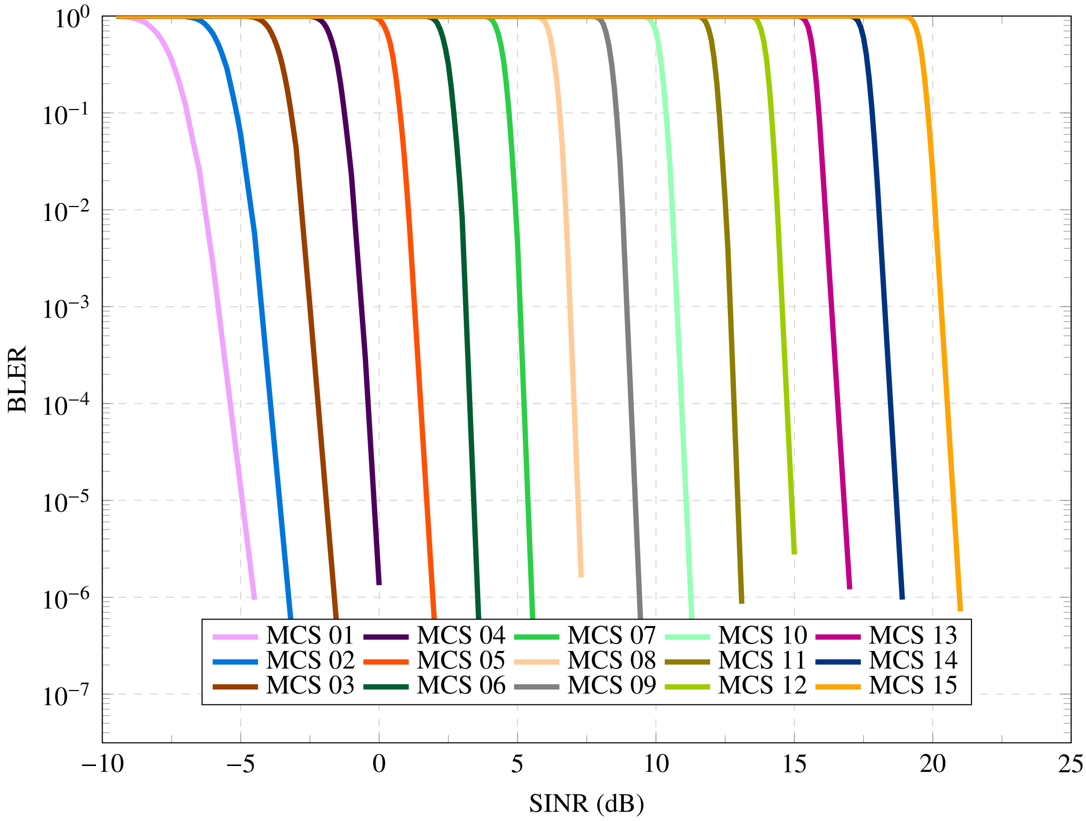
$$\hat{c}^{(rx)}_{b,u,k}[t] = g_{u,k}[t]\tilde{\mathbf{h}}^{\mathrm{T}}_{b,u,k}[t]\mathbf{f}_{b,k}[t]\sqrt{p_{b,k}[t]}c^{(tx)}_{b,u,k}[t] + I[t] + n_{b,u,k}[t]. \tag{2.10}$$

Since the RRM algorithms in this thesis are performed in each BS independently, in order to simplify the notation, from this point on, the index of the BS will be omitted without loss of generality. Moreover, since the impact of imperfections at the CSI estimation are out of the scope of this thesis, a perfect CSI is considered, i.e., $\tilde{\mathbf{h}}_{u,k}[t] = \mathbf{h}_{u,k}[t]$. Therefore, the estimated SINR $\tilde{\gamma}_{u,k}[t]$ that is used by the RRM algorithms can be obtained from (2.10) as

$$\tilde{\gamma}_{u,k}[t] = \frac{p_k[t]\left|g_{u,k}[t]\tilde{\mathbf{h}}^{\mathrm{T}}_{u,k}[t]\mathbf{f}_k[t]\right|^2}{\sigma_I^2 + \sigma_n^2}. \tag{2.11}$$

The casting of the estimated SINR to CQI measurement is performed by the link adaptation procedure $f^{\text{CQI}}_{adapt}(\cdot)$. In this thesis, it is considered that the CQIs are chosen considering a fixed target BLock Error Rate (BLER), equal to 10^{-4}. It means that the chosen CQI is the highest one with estimated BLER smaller than the target BLER, for a given estimated SINR

Figure 2.5 – SINR to BLER mapping for the MCSs in the LTE standard.



Source: Adapted from [53].

$\tilde{\gamma}_{u,k}[t]$. Since each CQI corresponds solely to a MCS, the considered MCS for a UE u in RB k at TTI t can be estimated as

$$m_{u,k}[t] = f_{adapt}^{\mathrm{CQI}}\left(\tilde{\gamma}_{u,k}[t]\right). \tag{2.12}$$

The BLER estimation for a given SINR measurement is obtained from link level curves for all available MCSs. All simulations performed in this thesis consider the link level curves presented in Fig. 2.5, from [53].

After the execution of the RRM algorithms, the signal is transmitted. As mentioned before, the signal transmission is not modeled. Instead, a random variable, ς, uniformly distributed between 0 and 1 is taken and if its value is greater than the BLER value, it implies that the transmission succeed. In case of success, it is assumed that all the bits carried by an RB using the given MCS are successfully received by the UE. Otherwise, all the information contained in this RB is lost. Therefore, the instantaneous data rate received by an UE u in RB k at TTI t is given by

$$r_{u,k}[t] = \begin{cases} f_{adapt}^{\mathrm{MCS}}\left(m_{u,k}[t]\right) & ; \text{if } \varsigma > f_{adapt}^{\mathrm{BLER}}\left(\gamma_{u,k}[t], m_{u,k}[t]\right) \\ 0 & ; \text{otherwise} \end{cases}, \tag{2.13}$$

where $f_{adapt}^{\mathrm{BLER}}(\gamma, m)$ is a function that returns the BLER value from the link level curve of the MCS m for an SINR γ, and $f_{adapt}^{\mathrm{MCS}}(m)$ returns the achieved rate in a single RB using the MCS m.

The total instantaneous data rate allocated to a UE u in a TTI t is given by

$$R_u[t] = \sum_{k \in \mathcal{K}_{u,t}} r_{u,k}[t], \tag{2.14}$$

where $\mathcal{K}_{u,t} \subset \mathcal{K}$ is the subset of RBs allocated to the UE u in the TTI t.

2.5 Performance Metrics

This section explains the performance metrics used in this thesis analyses, which are basically three: overall system throughput, satisfaction rate per service and outage probability per service.

The overall system throughput, R_{sys}, is the summation of the achieved rate of all UEs divided by the number of TTIs considered in the analysis, T, i.e.,

$$R_{sys} = \frac{1}{T} \sum_{t=1}^{T} \sum_{u \in \mathcal{U}} R_u[t]. \tag{2.15}$$

The UEs' satisfaction may be expressed in terms of their QoS or QoE measurements. In this thesis, the QoS metric considered is the UE overall rate, $R_u^{\text{avg}}[T]$, where

$$R_u^{\text{avg}}[t] = \frac{1}{t} \sum_{t'=1}^{t} R_u[t'] \tag{2.16}$$

denotes the average rate of a UE u at a TTI t.

The UEs' QoE measurements are given in terms of their MOSs, which depend on the multimedia applications used by the UEs. In this thesis analyses, it is considered that the MOS of a UE u can be obtained from the UE's data rate, i.e., $\Omega_u\left(R_u^{\text{avg}}[t]\right)$, where $\Omega_u(\cdot)$ is an increasing function that maps the achieved rate of a UE u into a MOS value.

Note that, the QoE's framework is more general than the QoS's. When $\Omega_u\left(R_u^{\text{avg}}[t]\right) = R_u^{\text{avg}}[t]$, the QoE metric reduces to a QoS one. Therefore, a UE u is considered satisfied when

$$\Omega_u\left(R_u^{\text{avg}}[T]\right) \geq \Omega_u^{\text{target}}, \tag{2.17}$$

where Ω_u^{target} represents the minimum QoS/QoE requirement of the UE u.

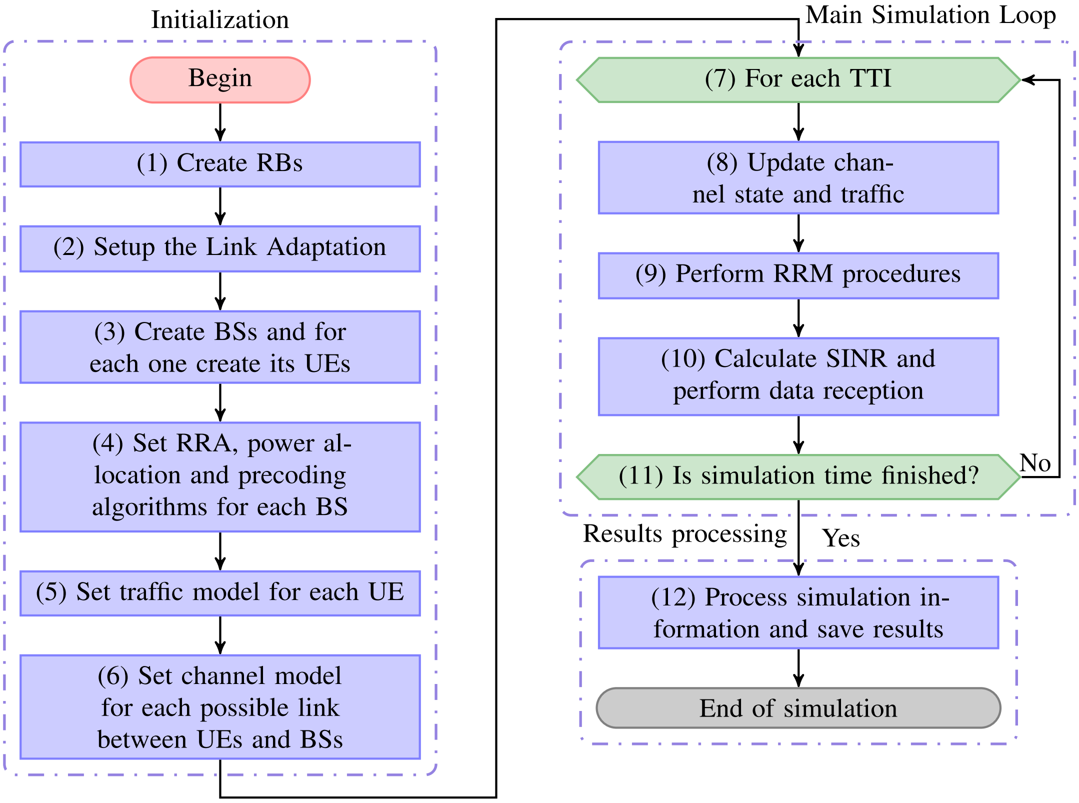
The satisfaction rate for a service s is given by

$$\Upsilon_s = \frac{\left|\mathcal{U}_s^{\text{sat}}\right|}{\left|\mathcal{U}_s\right|}, \tag{2.18}$$

where $\mathcal{U}_s^{\text{sat}} = \left\{u \in \mathcal{U}_s \mid \Omega_u\left(R_u^{\text{avg}}[T]\right) \geq \Omega_u^{\text{target}}\right\}$.

The outage probability of the service s is the chance that an outage event occurs. In its turn, an outage event occurs when the number of satisfied UEs $\left|\mathcal{U}_s^{\text{sat}}\right|$ of the service s is less than the minimum target number of UEs that are required to be satisfied ξ_s of the service s, i.e., $\left|\mathcal{U}_s^{\text{sat}}\right| < \xi_s$.

Figure 2.6 – Simulation flowchart.



Source: Created by the author.

2.6 Simulation Flowchart

In this section, the summary of how the simulations in this thesis are performed is presented. In Fig. 2.6, a flowchart with the sequence of functions executed during the simulation is depicted.

The simulator can be split into three major steps: the initialization, the main loop and the results processing. In the initialization part, all the structures needed by the simulation is created and the simulation is setup. The initialization blocks are:

- Create RBs objects and setup the link adaptation (blocks 1 and 2): The first step is to create a list of RBs and setup the link adaptation, configuring the link level curves and the SINR thresholds for each CQI;

- Create BSs and for each one create its UEs (block 3): The BSs and the UEs are created and their main characteristics are defined, such as: position, height, transmission power, antenna model. In this block, each UE is associated to one BS;

- Set RRA, power allocation and precoders for each BS (block 4): The RRM algorithms adopted by the BSs are configured;

- Set traffic model for each UE (block 5): The characteristics of the data that will be transmitted/received by each UE is defined;

- Set channel model for each possible link between UEs and BSs (block 6): The channel generator is configured for each link BS-UE.

The core of the simulator is within the main loop. In this part the system evolves according to a predefined time step, which is equal to the duration of a TTI. In each iteration, some tasks are performed:

- Update channel state and traffic (block 8): A new sample of the channel is generated for each link BS-UE and new packets are generated.

- Perform RRM procedures (block 9): The RRM algorithms set for each BS are executed in this block, scheduling the RBs, allocating power and setting the precoders.

- Calculation of SINR and data reception (block 10): 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. 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 by the user during the reception.

Finally, in the result processing, some measurements are extracted from the simulation and processed in order to obtain statistics that will be useful to analyze the system performance.

3 RESOURCE MANAGEMENT FOR RATE MAXIMIZATION WITH QOE/QOS PROVISIONING IN WIRELESS NETWORKS

In this chapter, the problem of maximizing the overall system rate, while ensuring that a minimum number of UEs of each service plan meet their QoS/QoE requirements is addressed. This problem extends the one studied in [24], since here the UE's requirements may be written in a more holistic manner beyond the traditional QoS.

The main contributions of this chapter are:

- Study of the problem of maximizing the overall system rate in a multi-service scenario, considering that a fraction of the users of each service must have their QoE requirements met;
- The reformulation of this problem as an ILP and its solution using standard algorithms;
- The proposal of a low-complexity suboptimal solution that has near optimal performance and presents high scalability in terms of the size of the problem input. Moreover, differently of [24], the algorithm proposed in this chapter also treats infeasible instances of the problem producing near feasible solutions by relaxing the problem constraints.

The rest of this chapter is divided as follows. In Section 3.1, the problem addressed in this chapter is mathematically formulated as an optimization problem. In Sections 3.2 and 3.3, the mathematical formulation developed in Section 3.1 is rewritten as an ILP, which can be solved using standard numerical algorithms from the literature, such as Branch and Bound (BB), and the state-of-the-art suboptimal algorithm that solves the problem is described. In Sections 3.4 and 3.5, a new low-complexity suboptimal algorithm is proposed to solve the problem stated in Section 3.1 and a performance analysis of this proposal against the optimal solution and the existing state-of-the-art heuristic is performed, respectively. Finally, the main conclusions of this chapter are presented in Section 3.6.

3.1 Problem Formulation

The problem of allocating the available RBs in order to maximize the overall system rate while ensuring that a minimum number ξ_s of UEs in service plan s meet their QoS/QoE requirements was initially proposed by [24]. It is assumed in this problem, that the BS distribute the total power available, P_{total}, over the RBs using an Equal Power Allocation (EPA).

Let $\mathbf{X} \in \{0,1\}^{U\times K}$ be the assignment matrix, where each element $x_{u,k}$ is equal to 1 if the RB k is allocated to the UE u and equal to 0 otherwise. The problem addressed in this chapter can be written as an optimization problem as follows:

$$\max_{\mathbf{X}} \sum_{u\in\mathcal{U}} \sum_{k\in\mathcal{K}} r_{u,k} x_{u,k}, \tag{3.1a}$$

$$\text{s.t.} \sum_{u\in\mathcal{U}} x_{u,k} = 1, \forall k \in \mathcal{K}, \tag{3.1b}$$

$$\sum_{u\in\mathcal{U}_s} H\left(\Omega_u\left(\sum_{k\in\mathcal{K}} r_{u,k} x_{u,k}\right), \Omega_s^{\text{target}}\right) \geq \xi_s, \forall s \in \mathcal{S}, \tag{3.1c}$$

$$x_{u,k} \in \{0,1\}, \forall u \in \mathcal{U} \text{ and } \forall k \in \mathcal{K}, \tag{3.1d}$$

where $H(a,b)$ denotes the Heaviside step function, which assumes the value 1 when $a \geq b$ and 0 otherwise, and Ω_s^{target} is the minimum MOS value required by a UE u of service s to be satisfied.

The problem stated in (3.1) aims at finding the optimal resource assignment that maximizes the achievable total system rate in the objective function (3.1a). Constraints (3.1b) and (3.1d) guarantee that each RB is assigned to a single UE. Furthermore, (3.1c) requires that a minimum number ξ_s of UEs should be satisfied for each service plan s.

Since this problem is solved in a single snapshot, i.e., $T = 1$ TTI, the TTI index will be omitted in this chapter in order to ease the notation.

3.2 Optimal Solution

It is worth noting that (3.1) is a combinatorial optimization problem with a nonconvex constraint (3.1c), which has a prohibitive computational complexity [54]. In this section, the problem stated in (3.1) is hence reformulated into a more tractable form. In other words, the problem (3.1) is rewritten as an ILP optimization problem, which can be solved by standard methods presented in the literature [55].

The function $\Omega_u(\cdot)$ which maps rate into MOS values defined in Section 2.5 is an increasing function. However, it does not ensure its linearity. In fact, in most of the cases, functions which map rate into MOS values are nonlinear [8, 10]. To address this issue, the constraint (3.1c) must be rewritten in such a way that $\Omega_u(\cdot)$ is not applied over the variables of the optimization problem. Following the definition of the Heaviside step function, for any invertible function $f(\cdot)$, it follows that

$$H(f(a), b) = H\left(a, f^{-1}(b)\right).$$

In the context of the constraint (3.1c), this is equivalent to convert the minimum MOS constraint into a minimum UE's rate requirement. However, $\Omega_u(\cdot)$ is an increasing function, so its invertibility cannot be guaranteed, unless it is strictly increasing. To cope with that, the concept of generalized inverse function for increasing functions stated in [56] can be adopted. Let $\Omega^{\dagger}(\cdot)$ be a function which maps MOS into rate defined as

$$\Omega^{\dagger}\left(\Omega^{\text{target}}\right) = \inf\left\{R \in \mathbb{R} : \Omega(R) \geq \Omega^{\text{target}}\right\},\ \Omega^{\text{target}} \in \mathbb{R}, \tag{3.2}$$

where $\inf\{\cdot\}$ is the infimum operator, which denotes the greatest lower bound of a set [54].

Therefore, the data rate, ψ_u, that UE u requires to achieve its MOS requirement Ω_s^{target} of the service plan s is given by

$$\psi_u = \Omega_u^{\dagger}\left(\Omega_s^{\text{target}}\right), \quad u \in \mathcal{U}_s, \quad \forall s \in \mathcal{S}. \tag{3.3}$$

This property ensures that the QoE constraint (3.1c) can be simplified to a QoS constraint as

$$\sum_{u \in \mathcal{U}_s} H\left(\sum_{k \in \mathcal{K}} r_{u,k} x_{u,k}, \psi_u\right) \geq \xi_s. \tag{3.4}$$

Consider $\boldsymbol{\rho} \in \{0,1\}^{U \times 1}$ as a vector, where each element ρ_u is a binary variable that assumes the value 1 if the UE u is selected to get satisfied and 0 otherwise. Using $\boldsymbol{\rho}$, (3.4) can be rewritten as two new constraints, (3.5c) and (3.5d), and the problem (3.1) can be restated as follows

$$\max_{\mathbf{X},\boldsymbol{\rho}} \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} r_{u,k} x_{u,k}, \tag{3.5a}$$

$$\text{s.t.} \sum_{u \in \mathcal{U}} x_{u,k} = 1, \ \forall k \in \mathcal{K}, \tag{3.5b}$$

$$\sum_{k \in \mathcal{K}} r_{u,k} x_{u,k} \geq \psi_u \rho_u, \ \forall u \in \mathcal{U}, \tag{3.5c}$$

$$\sum_{u \in \mathcal{U}} q_{s,u} \rho_u \geq \xi_s, \forall s \in \mathcal{S}, \tag{3.5d}$$

$$x_{u,k} \in \{0,1\}, \ \forall u \in \mathcal{U} \text{ and } \forall k \in \mathcal{K}, \tag{3.5e}$$

$$\rho_u \in \{0,1\}, \ \forall u \in \mathcal{U}, \tag{3.5f}$$

where $q_{s,u}$ assumes value 1 if the UE u subscribes the service plan s and 0 otherwise.

It is noteworthy that in (3.1c), the functions $H(\cdot)$ and $\Omega_u(\cdot)$ are applied over the optimization variables $x_{u,k}$. Since $H(\cdot)$ is neither convex nor concave, and $\Omega_u(\cdot)$ is usually nonlinear, the optimal solution of (3.1) becomes harder to find. Differently, the constraints (3.5c) and (3.5d) are linear, thus presenting the desired effect of simplifying the problem structure.

Besides that, the problem stated in (3.5) can be rewritten in a compact form, where the variables are organized in matrices and vectors, which often simplify solving the problem with commercial tools. Consider that the terms $r_{u,k}$ are organized into a matrix $\mathbf{R}$ with dimensions $U \times K$. Also consider that the variable ξ_s, for $s \in \mathcal{S}$, is arranged into a column vector $\boldsymbol{\xi}$ with length S and let $\boldsymbol{\psi} = \begin{bmatrix} \psi_1 & \psi_2 & \dots & \psi_U \end{bmatrix}^{\text{T}}$ be a column vector containing the rate requirements of all U UEs in the system. Moreover, the terms $q_{s,u}$ are also grouped into a matrix $\mathbf{Q}$ with dimension

$S \times U$. Using these definitions, the problem (3.5) can be rewritten as

$$\max_{\mathbf{R},\boldsymbol{\rho}} \ \text{vec}^{\text{T}}\{\mathbf{R}\}\,\text{vec}\{\mathbf{X}\}, \tag{3.6a}$$

$$\text{s.t.} \ \left(\mathbf{I}_K \otimes \mathbf{1}_U^{\text{T}}\right)\text{vec}\{\mathbf{X}\} = \mathbf{1}_K, \tag{3.6b}$$

$$\left(\mathbf{R}^{\text{T}} * \mathbf{I}_U\right)^{\text{T}}\text{vec}\{\mathbf{X}\} \geq \left(\left(\boldsymbol{\psi} \otimes \mathbf{1}_U^{\text{T}}\right) \odot \mathbf{I}_U\right)\boldsymbol{\rho}, \tag{3.6c}$$

$$\mathbf{Q}\boldsymbol{\rho} \geq \boldsymbol{\xi}, \tag{3.6d}$$

$$\mathbf{X} \in \{0,1\}^{U\times K} \tag{3.6e}$$

$$\boldsymbol{\rho} \in \{0,1\}^{U\times 1}, \tag{3.6f}$$

where $\mathbf{1}_a$ is a column vector with length a composed by 1's and $\mathbf{I}_a$ denotes the identity matrix with order a. The operator $\text{vec}\{\cdot\}$ is defined as $\text{vec}\{\mathbf{X}\} = \begin{bmatrix}\mathbf{x}_1^{\text{T}} & \mathbf{x}_2^{\text{T}} & \dots & \mathbf{x}_K^{\text{T}}\end{bmatrix}^{\text{T}}$, where $\mathbf{x}_k$ denotes the k^{th} column of the matrix $\mathbf{X}$. The operator $\odot$ is the Hadamard product, which consists in a element-wise matrix multiplication. The operators $\otimes$ and $*$ are the Kronecker and the Khatri-Rao products, respectively, where the Khatri-Rao product consists in a column-wise Kronecker product.

At this point, the optimization variables can be rearranged into a single vector $\mathbf{y} = \left[\begin{array}{c|c}\text{vec}^{\text{T}}\{\mathbf{X}\} & \boldsymbol{\rho}^{\text{T}}\end{array}\right]^{\text{T}}$. By defining $\mathbf{A} = \left[\begin{array}{c|c}\mathbf{I}_{UK} & \mathbf{0}_{UK\times U}\end{array}\right]$ and $\mathbf{B} = \left[\begin{array}{c|c}\mathbf{0}_{U\times UK} & \mathbf{I}_U\end{array}\right]$, where $\mathbf{0}_{a\times b}$ is a matrix with dimensions $a \times b$ composed by zeros, the variables $\mathbf{X}$ and $\boldsymbol{\rho}$ can be obtained from $\mathbf{y}$ by making $\text{vec}\{\mathbf{X}\} = \mathbf{A}\mathbf{y}$ and $\boldsymbol{\rho} = \mathbf{B}\mathbf{y}$. Thus, (3.6) can be rewritten as

$$\max_{\mathbf{y}} \ \text{vec}^{\text{T}}\{\mathbf{R}\}\,\mathbf{A}\mathbf{y}, \tag{3.7a}$$

$$\text{s.t.} \ \left(\mathbf{I}_K \otimes \mathbf{1}_U^{\text{T}}\right)\mathbf{A}\mathbf{y} = \mathbf{1}_K, \tag{3.7b}$$

$$\left(\mathbf{R}^{\text{T}} * \mathbf{I}_U\right)^{\text{T}}\mathbf{A}\mathbf{y} \geq \left(\left(\boldsymbol{\psi} \otimes \mathbf{1}_U^{\text{T}}\right) \odot \mathbf{I}_U\right)\mathbf{B}\mathbf{y}, \tag{3.7c}$$

$$\mathbf{Q}\mathbf{B}\mathbf{y} \geq \boldsymbol{\xi}, \tag{3.7d}$$

$$\mathbf{y} \text{ is a binary vector.} \tag{3.7e}$$

Notice that in (3.7), the optimization variables were reduced to a single vector $\mathbf{y}$, in opposite to (3.6), where the optimization variables are $\mathbf{X}$ and $\boldsymbol{\rho}$. In order to further simplify the problem (3.7), it can be expressed in a more compact form as

$$\max_{\mathbf{y}} \ \mathbf{c}^{\text{T}}\mathbf{y}, \tag{3.8a}$$

$$\text{s.t.} \ \mathbf{D}\mathbf{y} \leq \mathbf{w}, \tag{3.8b}$$

$$\mathbf{F}\mathbf{y} = \mathbf{1}_K, \tag{3.8c}$$

$$\mathbf{y} \text{ is a binary vector,} \tag{3.8d}$$

where

$$\mathbf{c} = \mathbf{A}^{\mathrm{T}} \operatorname{vec}\{\mathbf{R}\}, \tag{3.9}$$

$$\mathbf{D} = \begin{bmatrix} \left((\boldsymbol{\psi} \otimes \mathbf{1}_U^{\mathrm{T}}) \odot \mathbf{I}_U\right) \mathbf{B} - \left(\mathbf{R}^{\mathrm{T}} * \mathbf{I}_U\right)^{\mathrm{T}} \mathbf{A} \\ -\mathbf{Q}\mathbf{B} \end{bmatrix}, \tag{3.10}$$

$$\mathbf{w} = \left[\begin{array}{c|c} \mathbf{0}_{UK}^{\mathrm{T}} & -\boldsymbol{\xi}^{\mathrm{T}} \end{array}\right]^{\mathrm{T}}, \tag{3.11}$$

and

$$\mathbf{F} = \left(\mathbf{I}_K \otimes \mathbf{1}_U^{\mathrm{T}}\right) \mathbf{A}. \tag{3.12}$$

Finally, the initial optimization problem presented in (3.1) is reformulated as the standard ILP in (3.8) to which standard methods to solve ILPs, such as BB and Branch and Cut (BC), can be directly applied. These methods have much lower average complexity than the brute force solution, i.e., the complete enumeration of all possible assignments [55]. However, this class of ILPs is known to be NP-Hard (unless NP = P), i.e., these problems cannot be solved in polynomial time since their complexity increases exponentially with the problem dimensions. Thus, an approach relying on optimally solving an ILP might not be adequate for solving problems in real-time systems, such as RRA in cellular communication systems. Therefore, low-complexity and efficient suboptimal methods to solve (3.8) are highly desired.

3.3 State-of-the-art algorithm

The state-of-the-art heuristic was proposed by [24] and it is called RAISES. However, it is important to highlight that the RAISES does not deal with QoE requirements.

The RAISES algorithm is divided into three stages:

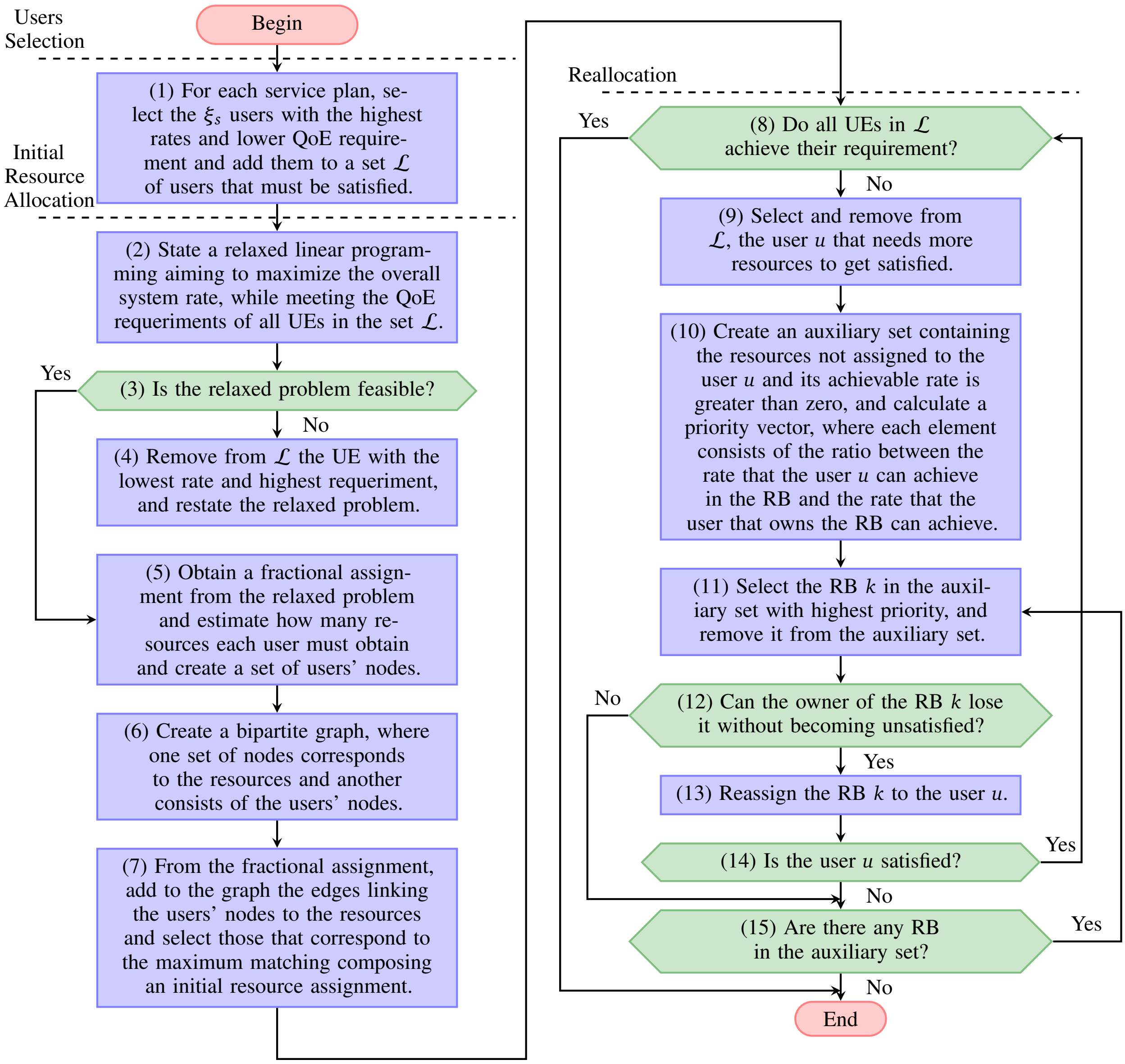
i. Disregard from the users' set $\mathcal{U}$ a certain number of UEs which will not be satisfied;

ii. Calculate an initial assignment;

iii. Reallocate the RBs between the users in order to ensure that the problem constraints are met.

At step i, the RAISES algorithm determines the $U - \sum_{s \in \mathcal{S}} \xi_s$ UEs with worse channel conditions and requiring more resources to get satisfied to be disregarded. At step ii, the initial solution is obtained by allocating the RBs to the UEs that were not disregarded based on the max-rate criteria, i.e., the resource will be allocated to the UE with best channel conditions. Finally, step iii consists in reallocating the resources in order to address the UEs' QoS requirement. Here, the UEs are divided into two sets: donors and receivers. The donors are UEs that got satisfied in the previous step and the receivers are the UEs that still need resources to get satisfied. The RBs are reallocated from the donors to the receivers until the receiver set gets empty or there are no more resources available to donate. Readers are directed to [24] for more details about the RAISES algorithm.

Figure 3.1 – Flowchart of the RMEC Algorithm.



Source: Created by the author.

3.4 Proposed suboptimal solution

In this section, a low complexity suboptimal solution to the problem described in Section 3.1 is proposed, which is called Rate Maximization under Experience Constraints (RMEC).

The proposed heuristic was inspired on the state-of-the-art algorithm and it is divided into the same three stages of the RAISES algorithm. A general overview of the suboptimal algorithm is shown as a flowchart in Fig. 3.1. The three steps of the RMEC algorithm are detailed in the rest of this section, with reference to each block of the flowchart in Fig. 3.1.

3.4.1 Step 1: User Selection

This first step of the algorithm is represented by (3.1c) in the optimization problem (3.1). The criteria used for selecting the UEs that will compete for resources in the next step is the same adopted by RAISES on its first step. For the sake of completeness of the proposed algorithm, the selection criteria is described in the sequel.

As presented in Section 2.4, the rate of a user u in RB k depends on its Signal to Noise Ratio (SNR) $\gamma_{u,k}$. Due to the large-scale fading, in most cases, the users often present similar values of SNR in all RBs, and consequently similar rates. It is also noteworthy that the higher the QoS/QoE requirement of a UE is, the harder it is to satisfy it since the UE needs more RBs to get satisfied.

The main goal is to maximize the transmit rate with the constraint of satisfying at least ξ_s users of each set $\mathcal{U}_s$, for all $s \in \mathcal{S}$. The idea of this phase of the proposed suboptimal solution is to select exactly ξ_s users from each set $\mathcal{U}_s$. Consider a set $\mathcal{L}$ of users that must be satisfied, which is initially empty. For each service $s \in \mathcal{S}$, an auxiliary set $\mathcal{A}$ initially equal to $\mathcal{U}_s$ is created. Then, the UEs with lowest transmit rate (considering the summation of all achievable rates in all RBs) and higher rate requirement are iteratively removed from this set, until $|\mathcal{A}| = \xi_s$. The criterion for disregarding a UE can be written as

$$u' = \arg\min_{u \in \mathcal{A}} \left\{ \frac{\sum_{k \in \mathcal{K}} r_{u,k}}{\psi_u} \right\}, \tag{3.13}$$

where u' denotes the index of the user to be disregarded. The choice of this criterion is reasonable since users with higher rates and lower rate requirements are easier to satisfy. Moreover, by selecting more than ξ_s users on each service, more users would be satisfied than the necessary, however the total transmit rate would be lower since the same number of RBs would be distributed to more users with worse channel conditions. After disregarding $|\mathcal{U}_s| - \xi_s$ users, the remaining ξ_s UEs are moved from $\mathcal{A}$ to the set $\mathcal{L}$, as illustrated in block (1) of Fig. 3.1. At the end of this step of the proposed suboptimal solution, $|\mathcal{L}| = \sum_{s \in \mathcal{S}} \xi_s$. In Algorithm 3.1, the procedure of building the user set $\mathcal{L}$ is presented.

Algorithm 3.1 User Selection

```
1: L ← ∅                                        ▷ Initialize the set L as an empty set
2: for all s ∈ S do
3:     A ← U_s                                  ▷ Create an auxiliary set equal to U_s
4:     while |A| > ξ_s do
5:         u' ← argmin_{u∈A} { Σ_{k∈K} r_{u,k} / ψ_u }   ▷ Select the user that will be disregarded
6:         A ← A \ {u'}                         ▷ Remove the selected user from the auxiliary set
7:     end while
8:     L ← L ∪ A                                ▷ Add the ξ_s selected UEs to L
9: end for
```

Observe that this phase of RMEC heuristic consists in determining a value for $\boldsymbol{\rho}$, where $\rho_u = 1$ if $u \in \mathcal{L}$ and $\rho_u = 0$ otherwise. Therefore, the optimization problem stated in (3.5) reduces to

$$\max_{\mathbf{X}_{\text{sat}}} \sum_{u\in\mathcal{L}} \sum_{k\in\mathcal{K}} r_{u,k} x_{u,k}, \tag{3.14a}$$

$$\text{s.t.} \sum_{u\in\mathcal{L}} x_{u,k} = 1, \quad \forall k \in \mathcal{K}, \tag{3.14b}$$

$$\sum_{k\in\mathcal{K}} r_{u,k} x_{u,k} \geq \psi_u, \quad \forall u \in \mathcal{L}, \tag{3.14c}$$

$$x_{u,k} \in \{0, 1\}, \quad \forall u \in \mathcal{L} \text{ and } \forall k \in \mathcal{K}, \tag{3.14d}$$

where $\mathbf{X}_{\text{sat}}$ is a matrix composed by the rows $u \in \mathcal{L}$ of the assignment matrix $\mathbf{X}$.

3.4.2 Step 2: Initial User Assignment

The problem stated in (3.14) has a similar structure to the one studied in [57], called Generalized Assignment Problem with Minimum Quantities (GAP-MQ), which is a variant of the classical Generalized Assignment Problem (GAP) [58]. In order to improve the understanding of this step of the algorithm, a brief description of both GAP and GAP-MQ is presented in the following.

The GAP consists in a problem of packing n_{items} items into n_{bins} bins, where each item $i \in [1, n_{\text{items}}]$ has a profit $g_{i,j}$ and a size $s_{i,j}$ when packed into the bin $j \in [1, n_{\text{bins}}]$. Its goal is to maximize the total profit of packing the n_{items} items into the n_{bins} bins, constrained by the capacity B_j of each bin j [58, 59]. The GAP framework has been applied in a considerable number of applications, such as routing, scheduling and task assignment problems [60]. It is important to mention that the GAP is NP-Hard [60] and has a 2-approximation algorithm [58]. An α-approximation is defined as an algorithm that provides in a polynomial-time a solution at least $1/\alpha$ of the optimal solution of any instance of a maximization problem. Furthermore, the GAP is also APX-Hard [57], which means that finding a polynomial-time α-approximation algorithm with $\alpha < 2$ is NP-Hard [61].

Considering the same framework of GAP, the GAP-MQ can be defined as a problem of maximizing the total profit of packing a subset of items into bins such that the total space used in each bin j is either zero, if the bin is not opened, or at least q_j and at most $B_j \geq q_j$. That is, differently from the GAP, it imposes a minimum capacity for each nonempty bin. For further details of GAP-MQ, see [57].

In this section, a particular case of the GAP-MQ is considered, where the number of bins is fixed, i.e., all the n_{bins} bins are opened. This problem can be related with (3.14) by considering that bins are users, items are RBs, and the minimum capacity of a bin u is related to the minimum UE u rate requirement. Consequently, $n_{\text{bins}} = |\mathcal{L}|$, $n_{\text{items}} = |\mathcal{K}|$, and $q_u = \psi_u$. Moreover, there is no capacity limits, i.e, $B_u \to \infty$, for $u \in \mathcal{L}$. The profit and the size in the underlying problem are equal to the user's achievable rate in a RB, i.e, $s_{u,k} = g_{u,k} = r_{u,k}$, where $k \in \mathcal{K}$. In [57], the authors present a polynomial time dual approximation algorithm to solve this particular case of the GAP-MQ, where the number of bins is fixed.

Here, the algorithm presented in [57] is adapted in order to find an initial resource allocation. Initially, in block (2) of Fig. 3.1, the optimization problem stated in (3.14) is relaxed, by modifying the binary constraint (3.14d) into a relaxed one, i.e, $0 \leq \tilde{x}_{u,k} \leq 1$, and replacing all $x_{u,k}$ by $\tilde{x}_{u,k}$, for all $u \in \mathcal{L}$ and $k \in \mathcal{K}$. Therefore, the relaxed version of problem (3.14) consists in a Linear Programming (LP) problem, whose solution provides a fractional assignment matrix $\tilde{\mathbf{X}}$ composed of the terms $\tilde{x}_{u,k}$. This LP can be written as

$$\max_{\tilde{\mathbf{X}}} \sum_{u \in \mathcal{L}} \sum_{k \in \mathcal{K}} r_{u,k} \tilde{x}_{u,k}, \tag{3.15a}$$

$$\text{s.t.} \sum_{u \in \mathcal{L}} \tilde{x}_{u,k} = 1, \quad \forall k \in \mathcal{K}, \tag{3.15b}$$

$$\sum_{k \in \mathcal{K}} r_{u,k} \tilde{x}_{u,k} \geq \psi_u, \quad \forall u \in \mathcal{L}, \tag{3.15c}$$

$$0 \leq \tilde{x}_{u,k} \leq 1, \quad \forall u \in \mathcal{L} \text{ and } \forall k \in \mathcal{K}. \tag{3.15d}$$

This problem can be efficiently solved by many algorithms proposed in the literature, such as the simplex and interior point methods [54].

Notice that if the problem (3.15) is infeasible, so is the problem (3.14), which suggests that the original problem (3.1) has a high probability of being infeasible since (3.14) is an approximation of (3.1). In order to deal with the infeasibility, in blocks (3) and (4) of Fig. 3.1, if (3.15) presents no feasible solution, then we proposed to remove one UE of the set $\mathcal{L}$ using the same criterion (3.13) of the previous step of our suboptimal solution and we proposed to repeat this process until (3.15) presents a feasible solution. Notice that if we need to withdraw any user of $\mathcal{L}$ in this step, the solution that will be obtained for our algorithm violates the constraint of the minimum number of satisfied users, however we still provide a near feasible solution to the problem. In summary, differently from previous works, our proposed solution takes into account a strategy to deal with potential no feasibility of the considered optimization problem.

The following steps of this section intend to round the fractional solution of (3.15) yielding an initial user assignment. One way of doing this is to simply get, for each $k \in \mathcal{K}$, the user u with the greatest value $\tilde{x}_{u,k}$. However, this method is well-known to yield a solution often far from the original ILP's optimal solution [62], which in this case is presented in (3.14). The rounding adopted here is the bipartite graph-based technique presented in [58] and adapted by [57] to the GAP-MQ. This rounding method is explained in the following.

In order to create the bipartite graph used in the rounding, in block (5) of Fig. 3.1, the amount ν_u of resources that each user $u \in \mathcal{L}$ demands can be estimated from $\tilde{\mathbf{X}}$ as

$$\nu_u = \left\lceil \sum_{k \in \mathcal{K}} \tilde{x}_{u,k} \right\rceil, \tag{3.16}$$

where $\lceil a \rceil$ returns the smallest integer greater than or equal to a.

In block (6) of Fig. 3.1, from $\tilde{\mathbf{X}}$ a bipartite graph $G(\mathcal{V}, \mathcal{K}, \mathcal{E})$ is created. In one side of G, each node represents an RB from the set $\mathcal{K}$. On the other side of G, for each user u, there are ν_u nodes, denoted by $\upsilon_{u,n} \in \mathcal{V}$, where $n \in [1, \nu_u]$. The edges linking both sides $(\upsilon_{u,n}, k) \in \mathcal{E}$ are added according to the following explanation.

In block (7) of Fig. 3.1, the edges linking both sides of the bipartite graph $G(\mathcal{V}, \mathcal{K}, \mathcal{E})$ are added. For each user $u \in \mathcal{L}$, a counter $c = 0$ and the index of the user's node $n = 1$ are initialized. Moreover, in order to ensure that RBs with higher rates are prioritized in this process, the indices of RBs are sorted in non-increasing order of $r_{u,k}$, i.e., $r_{u,1} \geq r_{u,2} \geq \cdots \geq r_{u,K}$. For each k and $\tilde{x}_{u,k} > 0$, the values of $\tilde{x}_{u,k}$ are accumulated into the counter c and a new edge $(\upsilon_{u,n}, k)$ is created with weight equal to $r_{u,k}$. When the counter c becomes greater or equal to one, it means that the user's node $\upsilon_{u,n}$ reaches its maximum capacity. Therefore, the value of the counter c is decremented by one and pass the residual capacity to the next user's node, i.e., $n = n + 1$. If c is still greater than zero, another edge $(\upsilon_{u,n}, k)$ is added with weight equal to $r_{u,k}$.

In the problem addressed in this chapter, obeying the QoS/QoE constraints is more important than achieving the optimal system rate since the cell operator aims to satisfy a certain quantity of users. Therefore, after the construction of the bipartite graph, instead of obtaining the maximum matching, as in [57], the minimum weighted matching is computed. This modification in the algorithm proposed in [57] ensures that users with worse channel conditions are selected to receive resources.

The minimum weighted matching of a bipartite graph consists in finding a subset of edges which yields a minimum cost attending the following constraints: all nodes of the graph are connected by at most one edge of this subset and the number of selected edges for this subset must be maximum (maximum cardinality). In the context of this chapter, the edges that produce the minimum sum rate are selected in a way that each node of $\mathcal{V}$ is connected by at most one edge and each node from the set of resources $\mathcal{K}$ in the graph G is connected by one edge.

A classic method used to find the minimum weighted matching in a bipartite graph is the Hungarian algorithm [63]. This method basically creates a cost matrix representing the bipartite graph and selects the maximum number of elements, where at most one element per row and column must be selected. The number of rows and columns of the matrix is equal to the number of nodes of each partition of the graph. If there is an edge between two nodes on the graph, the corresponding element of the matrix is equal to the edge weight, otherwise it is equal to infinity.

Consider the subset $\mathcal{M} \subset \mathcal{E}$ of edges that compose the minimum weighed matching. In order to find $\mathcal{M}$, the Hungarian algorithm [63] is applied over G. Finally, all elements of $\mathbf{X}_{\text{sat}}$ are set to zero, then for each edge $(\upsilon_{u,n}, k) \in \mathcal{M}$ set $x_{u,k} = 1$, for $u \in \mathcal{L}$, $k \in \mathcal{K}$ and $n \in (1, \nu_u)$. The entire description of this phase of the proposed suboptimal solution is presented in Algorithm 3.2.

The process of obtaining the initial user assignment is extremely important in the algorithm proposed in this chapter. Therefore, in order to ease the understanding of this step of the proposed algorithm and illustrate how it works, a numerical example of how to acquire the initial solution is presented in the sequel.

Consider that 3 users that must be satisfied are competing for 5 RBs and all of them

Algorithm 3.2 Initial Resource Assignment

1: **while** (3.15) has no feasible solution **do**
2: $\quad u' \leftarrow \arg\min_{u \in \mathcal{L}} \left\{ \frac{\sum_{k \in \mathcal{K}} r_{u,k}}{\psi_u} \right\}$
3: $\quad \mathcal{L} \leftarrow \mathcal{L} \setminus \{u'\}$
4: **end while**
5: Solve (3.15) and find a fractional assignment $\tilde{\mathbf{X}}$
6: Create a bipartite graph $G(\mathcal{V}, \mathcal{K}, \mathcal{E})$
7: **for all** $u \in \mathcal{L}$ **do**
8: $\quad c \leftarrow 0$ ▹ Initialize a counter
9: $\quad n \leftarrow 1$
10: $\quad$ **for all** $k \in \mathcal{K}$ sorted in non-increasing order of $r_{u,k}$ **do**
11: $\quad\quad$ **if** $\tilde{x}_{u,k} > 0$ **then**
12: $\quad\quad\quad c \leftarrow c + \tilde{x}_{u,k}$
13: $\quad\quad\quad$ Add edge $(v_{u,n}, k) = r_{u,k}$ to G
14: $\quad\quad\quad$ **if** $c \geq 1$ **then**
15: $\quad\quad\quad\quad c \leftarrow c - 1$
16: $\quad\quad\quad\quad n \leftarrow n + 1$
17: $\quad\quad\quad\quad$ **if** $c > 0$ **then**
18: $\quad\quad\quad\quad\quad$ Add edge $(v_{u,n}, k) = r_{u,k}$ to G
19: $\quad\quad\quad\quad$ **end if**
20: $\quad\quad\quad$ **end if**
21: $\quad\quad$ **end if**
22: $\quad$ **end for**
23: **end for**
24: Find the minimum weighed matching $\mathcal{M}$ of G using the Hungarian algorithm
25: Construct $\mathbf{X}_{\text{sat}}$ from $\mathcal{M}$

have the same MOS requirement that implies in a rate requirement of 512 kbps. Consider also that the users achievable rates in the RBs are given by the 5×3 rate matrix $\mathbf{R}$ as follows

$$\mathbf{R} = \begin{bmatrix} 655 & 248 & 248 & 39 & 147 \\ 655 & 321 & 25 & 25 & 558 \\ 63 & 458 & 197 & 759 & 933 \end{bmatrix} \text{ kbps.}$$

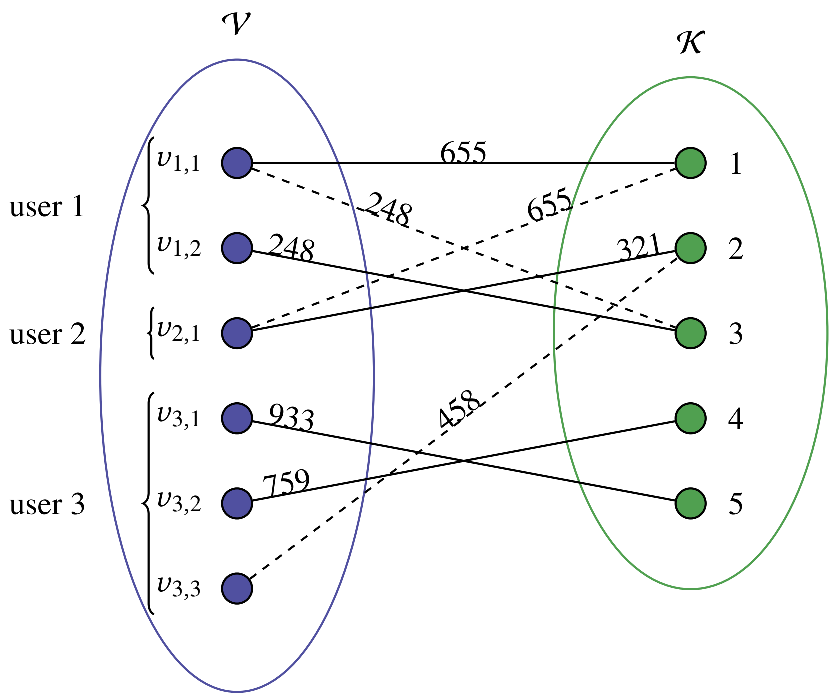
Solving (3.15), the fractional assignment obtained is given by

$$\tilde{\mathbf{X}} = \begin{bmatrix} 0.4029 & 0 & 1 & 0 & 0 \\ 0.5971 & 0.3762 & 0 & 0 & 0 \\ 0 & 0.6238 & 0 & 1 & 1 \end{bmatrix}.$$

From the fractional assignment the bipartite graph G is constructed. Notice that applying (3.16) on $\tilde{\mathbf{X}}$, the number of nodes that each user will hold in the bipartite graph is calculated, namely $\nu_1 = 2$, $\nu_2 = 1$ and $\nu_3 = 3$. Therefore, the bipartite graph is created as depicted in Fig. 3.2 with 6 user nodes and 5 resource nodes. The edges are added to G accordingly to lines 7-23 of the Algorithm 3.2. For user 1, c is initialized with value 0 and the indices of the RBs are sorted in non-increasing order of achievable rate, yielding the list $\{1\ 2\ 3\ 5\ 4\}$. Taking the first RB in the sorted list, $\tilde{x}_{1,1} = 0.4029 > 0$, so an edge $(v_{1,1}, 1) = r_{1,1} = 655$ is added and $\tilde{x}_{1,1}$ is added to c, resulting in $c = 0.4029$. The next resource on the list has $\tilde{x}_{1,2} = 0$, thus the algorithm goes further to the next RB, 3, which has $\tilde{x}_{1,3} = 1$. A new edge $(v_{1,1}, 3) = r_{1,2} = 248$ is added to

$\mathcal{E}$ and $\tilde{x}_{1,3}$ is added to c, resulting in $c = 1.4029$. As $c > 1$, it means that the node $v_{1,1}$ reaches its maximum capacity, so the algorithm go to the next node, $v_{1,2}$, and decrement the value of c by one, i.e., $c = c - 1 = 0.4029$. As $c > 0$, a new edge $(v_{1,2}, 3) = r_{1,3} = 248$ is added to $\mathcal{E}$. The next RBs in the sorted list has no assigned portion, therefore the algorithm pass to the next user. The edge addition to the users 2 and 3 follows the same idea as for user 1. After the creation of the bipartite graph, the edges that compose the minimum weighted matching are computed using the Hungarian algorithm, denoted by the solid edges in Fig. 3.2.

Figure 3.2 – Bipartite Graph with solid edges denoting the maximum matching.



Source: Created by the author.

Neglecting the edges $(v_{u,n}, k) \notin \mathcal{M}$, notice that all RB nodes are linked to only one user node, meeting the constraint (3.5b). From Fig. 3.2, the initial allocation is given by

$$\mathbf{X}_{\text{sat}} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

3.4.3 Step 3: Reallocation

In this step of RMEC, a resource reallocation is proposed in order to respect the constraint (3.14c) rather than achieve higher overall system rates.

As depicted in block (8) of Fig. 3.1, if all UEs in $\mathcal{L}$ are already satisfied, then the initial solution provided in the previous step is used. Otherwise, the algorithm starts by removing from $\mathcal{L}$ all users that have already achieved their QoS/QoE. After that, the UE $u \in \mathcal{L}$ that needs more resources to achieve its requirement is chosen. Considering the user u, $\mathcal{J}$ is defined as an auxiliary set containing all resources' indices that may aggregate rate to user u ($r_{u,k} > 0$), but are not allocated to it ($x_{u,k} = 0$), i.e., all $k \in \mathcal{K} \setminus \{\{x_{u,k} = 1\} \cup \{r_{u,k} = 0\}\}$, as shown in block (9) of Fig. 3.1. After that, in block (10) of Fig. 3.1, a priority vector $\mathbf{w} = [w_1\ w_2\ \dots\ w_K]^{\mathrm{T}}$ is created, where each element w_k, for $k \in \mathcal{J}$, is equal to the ratio between the rate of user u in RB k and the achievable rate of the user that got the referred resource in the initial allocation. Then, in block (11) of Fig. 3.1, the index of the RB $k \in \mathcal{J}$ with the highest value w_k is removed from the auxiliary set $\mathcal{J}$. Consider $i \in \mathcal{L}$ the user that owns resource k. In blocks (12) and (13) of Fig. 3.1, if the user i achieves the requirement ψ_i even without the RB k, then we transfer the resource k from the user i to user u. If the user u gets satisfied with the received resource, then the algorithm go further to the next unsatisfied user, if any. Otherwise, if the user u do not get satisfied, then the process continues until removing all resources from $\mathcal{J}$. These steps are depicted in blocks (14) and (15) of Fig. 3.1. This step of the proposed algorithm finishes when all unsatisfied users are parsed. The algorithm description is presented in Algorithm 3.3.

Algorithm 3.3 Reallocation

while $|\mathcal{L}| > 0$ **do**
  $u \leftarrow \underset{u \in \mathcal{L}}{\arg\min} \left\{\psi_u - \sum_{k \in \mathcal{K}} r_{u,k} x_{u,k}\right\}$
  $\mathcal{L} \leftarrow \mathcal{L} \setminus \{u\}$
  if $\sum_{k \in \mathcal{K}} r_{u,k} x_{u,k} < \psi_u$ **then** ▹ If user u is not satisfied
    $\mathcal{J} \leftarrow \mathcal{K} \setminus \{\{x_{u,k} = 1\} \cup \{r_{u,k} = 0\}\}$
    $w_k \leftarrow \dfrac{r_{u,k}}{\sum_{i \in \mathcal{U}} r_{i,k} x_{i,k}}\ \forall k \in \mathcal{K}$
    while $|\mathcal{J}| > 0$ **do**
      $k \leftarrow$ the resource index $k \in \mathcal{J}$ with highest w_k
      $\mathcal{J} \leftarrow \mathcal{J} \setminus \{k\}$
      $i \leftarrow$ the user that owns the resource j
      if $\sum_{j \in \mathcal{K}} r_{i,j} x_{i,j} - r_{i,k} > \psi_i$ **then**
        $x_{i,k} \leftarrow 0$
        $x_{u,k} \leftarrow 1$
        if $\sum_{k \in \mathcal{K}} r_{u,j} x_{u,j} > \psi_u$ **then**
          break
        end if
      end if
    end while
  end if
end while

This step of the proposed algorithm shares many similarities to the reallocation process present in the RAISES algorithm, however there are two differences between them. In the reallocation process in the RAISES algorithm, the receiver that is chosen firstly to receive resources is the one with worse channel conditions and still unsatisfied. In the proposed heuristic, the first UE that will be candidate to receive resources is the one that needs more resources to get satisfied after the initial solution. Another difference between the algorithms is that when an unsatisfied UE that gets satisfied in the reallocation process in the RAISES algorithm, it is removed from the receiver set and its resources will not be available for reallocation. In the proposed algorithm, if one unsatisfied UE gets satisfied, it will still be able to donate resources to another UE, if this action does not make it unsatisfied again.

The complexity of the proposed algorithm is bounded by the solution of the LP stated in (3.15). This LP can be solved efficiently using the well known simplex algorithm, which has a polynomial-time average-case complexity [54]. Nevertheless, the LP can also be solved using the Karmarkar's algorithm, which solves LP in the problem in polynomial time with a complexity of $O\left(U^{3.5}K^{3.5}\right)$ [64].

3.5 Performance Analysis

In this section, the performance of the algorithm proposed in Section 3.4 is evaluated by comparing it to the optimal solution (3.8) provided in Section 3.2 and to the algorithm presented in [24].

In the following performance evaluation, the simulations considered a BS located on the center of an hexagonal cell with a 800 m radius. The system operates at a frequency of 3.5 GHz with a downlink bandwidth of 20 MHz, which consists of $K = 100$ RBs in the LTE standard. Each RB is composed by $Q_{sub} = 12$ adjacent subcarriers spaced of $\Delta f = 15$ kHz and by $Q_{sym} = 14$ consecutive symbols. The channel modeling considers path loss, shadowing and IID small-scale fading, as described in Section 2.3 when considering a single cell scenario. A summary of the system parameters is presented in Table 3.1.

All QoE measurements considered in the following simulations are given in terms of MOS and all services are assumed to be web browsing based [11] and have their MOS given in terms of rate by the following relationship

$$\Omega\left(R_u\right) = 5 - \frac{578}{1 + \left(\dfrac{R_u + 541.1}{45.98}\right)^2}, \tag{3.17}$$

where R_u denotes the UE rate given in kbps. It is also assumed that the UEs are always demanding traffic, therefore, the traffic to all UEs are modeled as full buffer.

In the following analyses, the algorithm proposed in this chapter is compared with the optimal solution, obtained by solving (3.8), and with the RAISES heuristic provided by [24]. The comparisons are performed in terms of average satisfaction, total system rate and outage probability.

Notice that RAISES intends to solve a problem similar to the one treated in this chapter, however considering QoS instead of QoE constraints. In order to compare the results of the RMEC algorithm with those provided by the algorithm proposed in [24], the QoS metrics are displayed in terms of QoE metrics using the mapping function $\Omega^{\dagger}(\cdot)$.

In Fig. 3.3, the satisfaction and the total system rate are presented as function of the required MOS considering three different number of served UEs, namely $U = 10$, 20 and 30. In this analysis, a single service is considered and all users subscribing to it need to be satisfied, i.e., $\xi_1 = 100\% \cdot U$. In order to perform a fair comparison, the analyses are conducted considering only the cases where the optimization problem stated in (3.1) has a feasible solution.

Table 3.1 – Simulation parameters.

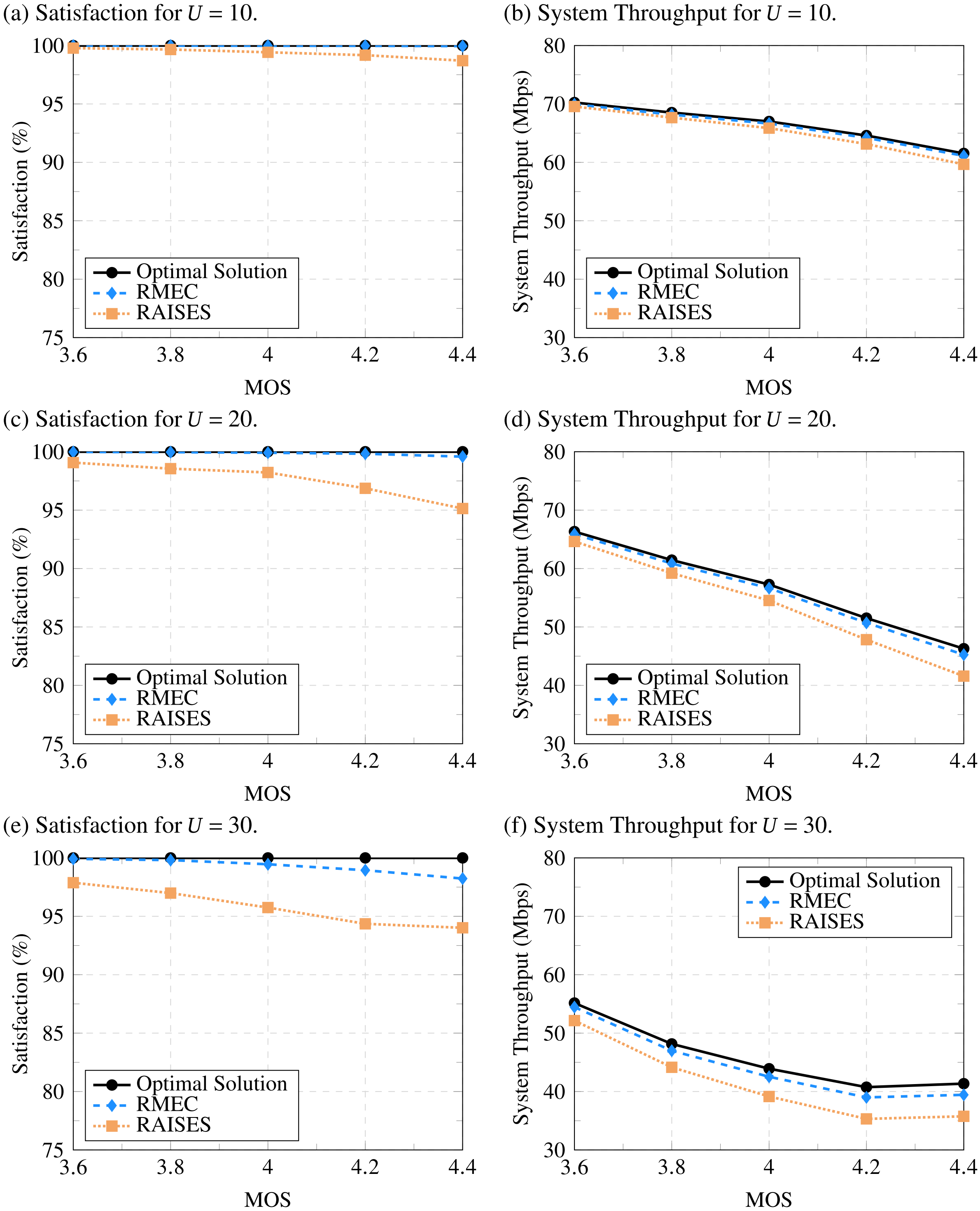
Parameter	Value
Maximum BS transmit power (P_{total})	49 dBm [65]
BS antenna radiation pattern	Omnidirectional
Cell radius	800 m
UE speed	3 km/h [39]
Carrier frequency	3.5 GHz [39]
System bandwidth	20 MHz [65]
Subcarrier bandwidth (Δf)	15 kHz
Number of RBs (K)	100
Number of subcarriers per RB (Q_{sub})	12
Number of symbols per RB (Q_{sym})	14
Path loss	$34.5 + 35\log_{10}(d_{b,u})$ [66]
Log-normal shadowing standard deviation	8 dB [39]
Small-scale fading	IID
AWGN power per sub-carrier	-123.24 dBm
Noise figure	9 dB
Link adaptation	Link level curves from [53]
Traffic model	Full buffer
Transmission Time Interval	1 ms
Number of snapshots	10000
Confidence interval	95%

Source: Created by the author.

In Fig. 3.3a, for $U = 10$ UEs, the satisfaction slightly varies with the required MOS for both RMEC and RAISES algorithms. In this case, both heuristics achieve an average satisfaction rate very close to the optimal solution, with a difference of less than 0.05% and 1.3% for RMEC and RAISES, respectively. The good performance of both heuristics stands also when the overall system throughput is analyzed in Fig. 3.3b. Here, RMEC achieves an average system throughput less than 1% below the optimal solution, while RAISES also yields a good result achieving a efficiency loss of around 3%.

In Figs. 3.3c and 3.3d, the same analysis is done for $U = 20$ UEs. In this case, it is possible to observe that RMEC remains near optimal in both considered Key Performance Indicators (KPIs), namely average satisfaction and the overall throughput, while the satisfaction rate achieved by RAISES distantiates from the target with the increasing MOS. The satisfaction rate of RMEC is far from the target by less than 0.5%, while RAISES presents a gap of 4.9% to the optimal solution. Regarding the overall system rate, the RMEC algorithm provides near

Figure 3.3 – System performance for a single service scenario with $\xi_1 = 100\%$ of U.



Source: Created by the author.

optimal results, with an efficiency loss of at most 2.3% when compared to the optimal solution. On the other hand, the RAISES yields an average throughput 10.2% lower than the optimal result, for the highest analyzed MOS requirement.

Observe that the gap between the solution provided by RAISES and RMEC increases with the increasing number of UEs from $U = 10$ to $U = 20$. While RMEC presents near optimal satisfaction results, RAISES presents a lack of scalability with the increasing number of UEs in the system which demands to be satisfied. Moreover, notice that the average RMEC satisfaction slightly varies with the increasing minimum MOS requirement, differently of RAISES that presents a considerable increase in the gap between its average satisfaction rate and the required target.

The results presented in Figs. 3.3e and 3.3f endorse these conclusions. Note that in Fig. 3.3e, due to the increasing number of UEs, the gap between the optimal solution and RMEC increases reaching around 1.8% for a MOS of 4.4 while RAISES presents a gap of 6%. In terms of system throughput, the proposed algorithm still leads to a maximum optimality gap of about 4.6% for a MOS equal to 4.4. Nevertheless, the throughput gap of the RAISES algorithm with respect to the optimal solution increases, reaching 5.5% when the MOS is equal to 3.6 and going up to 13.5% for a MOS equal to 4.4. Comparing the results regarding the overall system rate from Figs. 3.3b, 3.3d and 3.3f, notice that the throughput gap to the optimal solution presented by RAISES increases with the number of UEs. Meanwhile, the RMEC algorithm presented results closer to the optimal one. These observations reinforce that the proposed algorithm scales better with the increasing number of UEs.

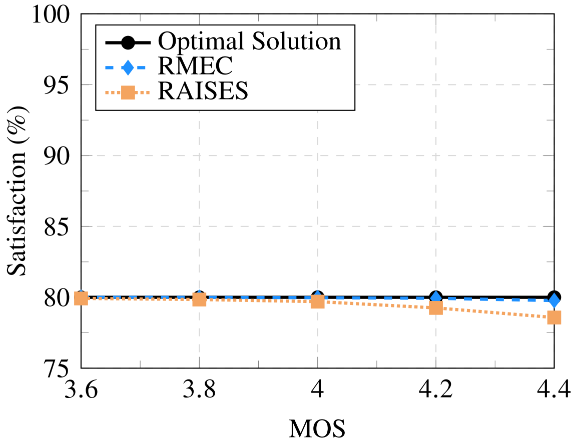
In Fig. 3.4 the impact of requiring different values of minimum number of satisfied UEs is analyzed considering $U = 30$. Again in this analysis, in order to perform a fair comparison of the algorithms, only the feasible instances of the problem were considered.

For a minimum requirement of satisfying 80% of the UEs, both algorithms present good results, as shown in Figs. 3.4a and 3.4b. In this case, for the highest value of MOS considered, the gap of satisfaction with respect to the optimal solution was around 0.23% for RMEC and 1.4% for RAISES. Regarding the system throughput, the proposed heuristic has a loss of at most 2.2% compared to the optimal solution when a MOS of 4.4 is considered, while RAISES reaches a throughput 6.4% lower than the optimal one. Observe that both RAISES and RMEC achieve results close to the optimal solution regarding the satisfaction. This happens due to the increase of the UE diversity, i.e., in this case the algorithms are free to neglect at most 6 UEs (80% of 30 UEs) with poor channel conditions, increasing the chances of achieving a feasible result.

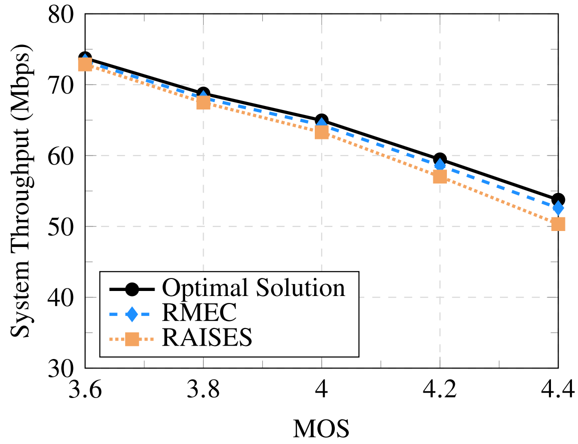
Observe that in Figs. 3.4c and 3.4d, the difference of performance between RMEC and RAISES becomes more evident. Since in this case the algorithms can neglect at most 3 UEs, the UE diversity diminishes. Observe that, in terms of average satisfaction rate, RMEC achieves a result at most 0.84% below the target for a MOS of 4.4, while RAISES presents an average satisfaction rate 3.6% below the target for the same value of MOS. In terms of rate, the distinction between RMEC and RAISES is clearer. For a MOS target of 4.4, while RMEC achieves a system throughput 3.6% below the optimal solution, RAISES has a loss of 11%, i.e., a difference of 7.4% between the algorithms, relative to the optimal result. The results presented in Fig. 3.3e and 3.3f, for $\xi_1 = 100\%$ previously discussed confirm the better robustness of RMEC against the state-of-the-art solution with respect to the minimum number of UEs that should be satisfied.

Figure 3.4 – System performance varying the percentage of satisfied UEs in a single service scenario with $U = 30$ UEs.

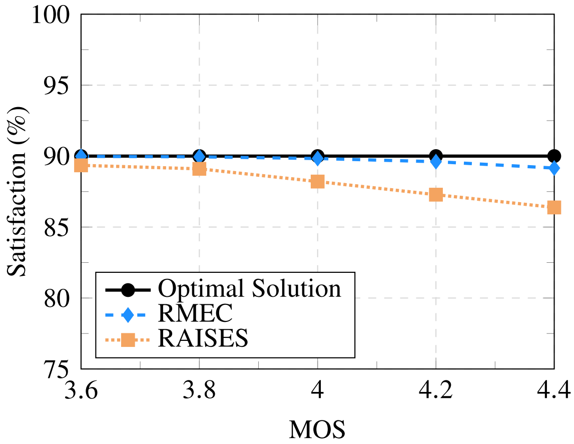
(a) Satisfaction for $\xi_1 = 80\%$.



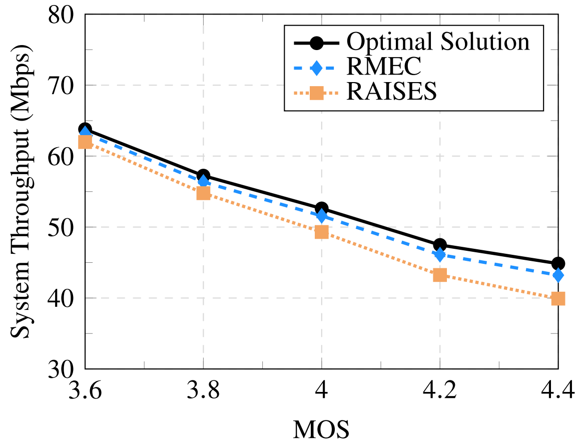
(b) System Throughput for $\xi_1 = 80\%$.



(c) Satisfaction for $\xi_1 = 90\%$.



(d) System Throughput for $\xi_1 = 90\%$.



Source: Created by the author.

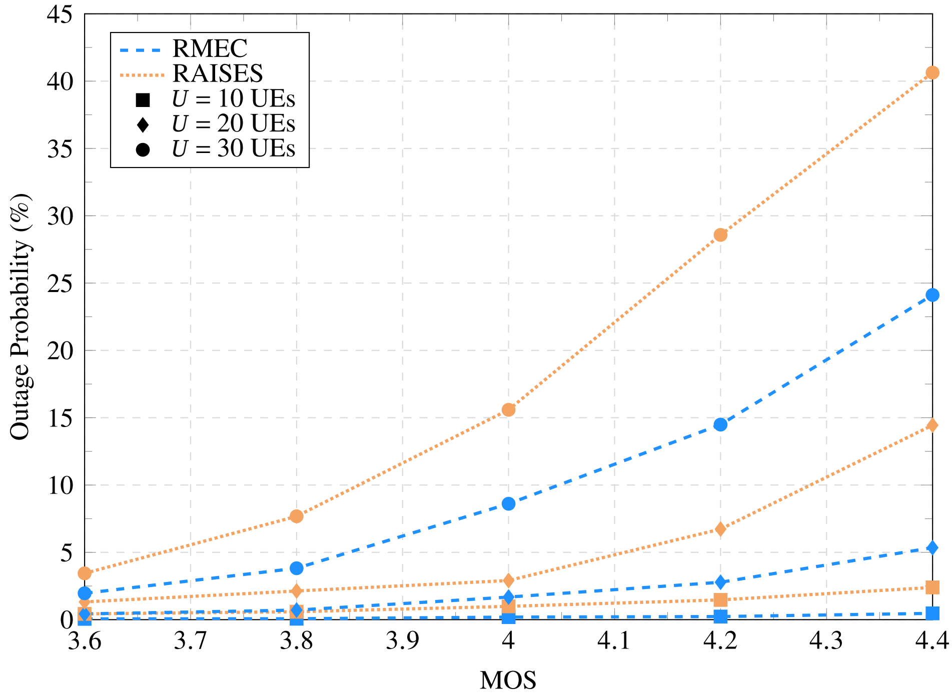
Besides achieving a higher average satisfaction rate, the RMEC heuristic also achieves a feasible solution more often than the state-of-the-art algorithm. This fact can be observed in Figs. 3.5 and 3.6, where the outage probability is analyzed. An outage event occurs when the satisfaction rate Υ_s of the service s is lesser than the satisfaction target ξ_s, i.e., $\Upsilon_s < \xi_s$.

Fig. 3.5 presents the outage probability for both analyzed algorithms, namely RMEC and RAISES, varying the number of UEs in the system and considering a satisfaction target of 100%, similarly to the setup considered in Fig. 3.3.

Observe that for $U = 10$, RMEC presents an outage probability of at most 0.47% for the highest MOS value considered in this analysis. This result implies that the proposed algorithm achieves the satisfaction target in almost all feasible instances of this scenario. The state-of-the-art algorithm presents an outage probability of 0.43% for a required MOS of 3.6, achieving 2.4% for a MOS of 4.4. Although the outage probability of both algorithms are close, notice that the proposed algorithm provides feasible solutions more often than the RAISES.

When the number of UEs increases to 20, the difference between RMEC and RAISES becomes more evident. In this scenario, the proposed algorithm is not able to find a feasible

Figure 3.5 – Outage probability for $\xi_1 = 100\%$ and $U = 10, 20$ and 30 UEs.

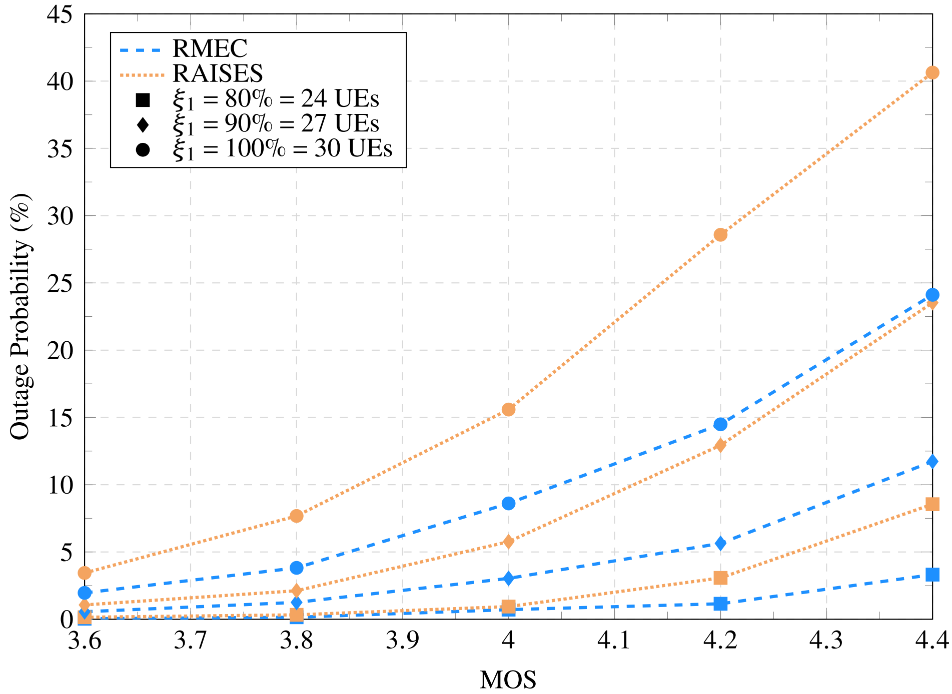


Source: Created by the author.

solution in at most 5.3% of the cases for a MOS of 4.4. Meanwhile, for the same MOS requirement RAISES algorithm fails in finding a feasible solution in 14.4% of the cases, i.e., a 9.1% difference compared to the proposed heuristic. These results reinforce the robustness of the RMEC, which in addition to achieving better results than the state-of-the-art algorithm regarding both analyzed KPIs, also finds feasible solutions more often. The case when $U = 30$ UEs ratifies these conclusions. In this case, for a MOS of 3.6, the RMEC has an outage probability of 2%, against 3.4% of RAISES. However, the gap between the algorithms increases with the increasing MOS, and for a required MOS of 4.4, the gap between the outage probability of RMEC and RAISES is 16.5%. Comparing the outage results for $U = 10$, 20 and 30, it can be observed that the slope of the outage probability curves of RMEC is smaller than the RAISES, which means that the gap between the algorithms grows with the increasing MOS. A similar analysis is done considering $U = 30$ UEs and varying the percentage of UEs that should be satisfied, which is presented in Fig. 3.6, enriching the previous analyses depicted in Fig. 3.4.

For $\xi_1 = 80\% = 24$ out of 30 UEs, both algorithms present a very low outage probability of 0.15% for a MOS equal to 3.6, however, the robustness of the proposed heuristic becomes evident when the MOS requirement is greater than 4. Observe that for a MOS equal to 4.2, the gap between RAISES and RMEC is 1.9%, and increases to 5.2% for a MOS equal to 4.4. When the algorithms have the requirement of satisfying at least 90% of the UEs, i.e., 27 out of 30 UEs, the difference between RAISES and RMEC is notorious. Notice that for a MOS equal to 4.4, the proposed heuristic fails in finding a feasible solution in 11.7% of the cases. Nevertheless, an outage event occurs in 23.5% of the cases for the state-of-the-art algorithm, i.e., a gap of 11.8%. Finally, for $\xi_1 = 100\% = 30$ UEs, as already mentioned in the analyzes

Figure 3.6 – Outage probability considering $\xi_1 = 80\%, 90\%$ and 100% of $U = 30$ UEs.



Source: Created by the author.

of Fig. 3.5, the gap between RAISES and RMEC reaches 16.5% for a MOS equal to 4.4. It is possible to observe in these analyzes that, similarly to Fig. 3.4, the outage probability of RAISES grows more rapidly with the increasing MOS value, compared to RMEC, for all the minimum requirement values considered, increasing the gap between the outage probability of both algorithms. Indeed, it is important to highlight that for MOS values equal to 4.2 and 4.4, the outage probability registered by RAISES for $\xi_1 = 90\%$ is less than 2% below the one achieved by RMEC considering $\xi_1 = 100\%$.

The better results of RMEC in comparison to RAISES relies mainly on the process of obtaining an initial solution. Both RMEC and RAISES share the same structure:

i. the selection of the users that will be satisfied,

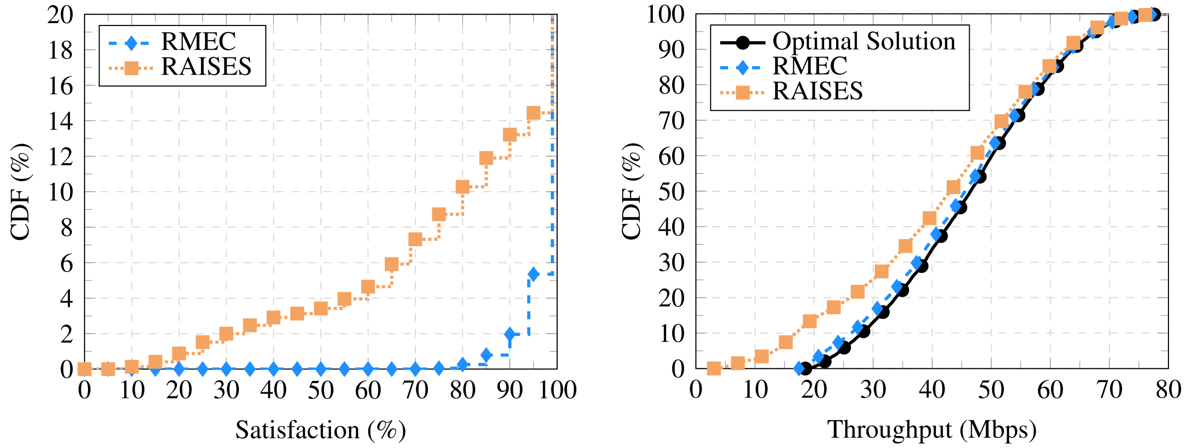
ii. an initial user assignment and,

iii. a reallocation process.

The choice of an initial allocation is very important for both algorithms since the reallocation process works as a "fine-tuning" for the final solution. The RAISES algorithm starts from the solution of the max rate scheduler, i.e., first it allocates the resources to the users aiming to maximize the overall system rate, without considering the users' requirements. This initial allocation is often far from a feasible solution. Therefore, when the number of resources and users increase, the capability of finding a feasible solution in the reallocation process decreases. On the other hand, the RMEC heuristic considers as initial allocation a graph-based rounding of

Figure 3.7 – CDF of the satisfaction and throughput considering $U = 20$ UEs, a minimum MOS equal to 4.4 and $\xi = 100\%$ of the UEs.

(a) CDF of the satisfaction. (b) CDF of the throughput.



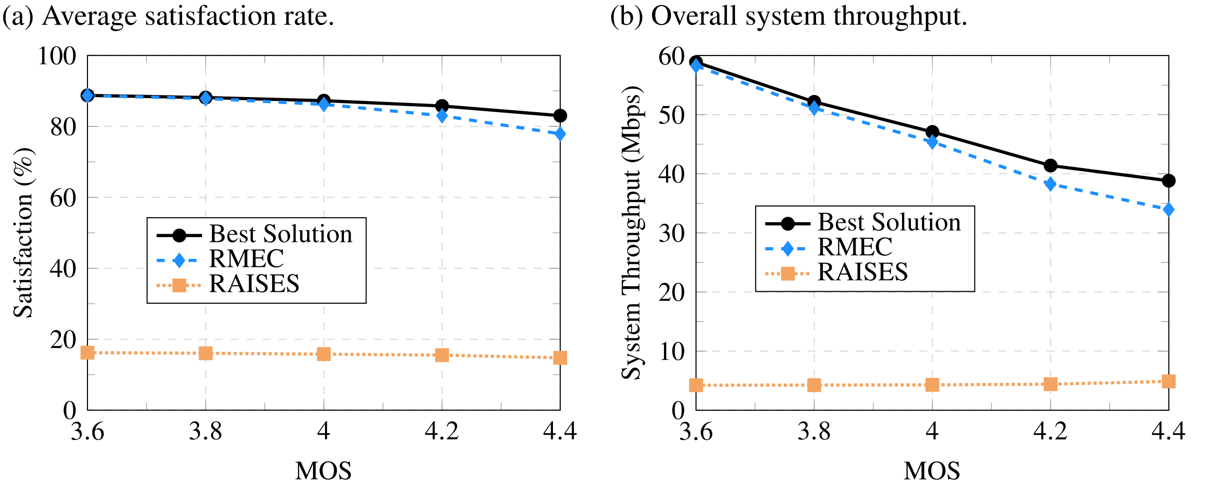
Source: Created by the author.

the upper-bound solution (relaxed solution of the problem), which yields an initial allocation that is in fact much closer to a feasible solution than the one employed by RAISES. Therefore, in the reallocation process, the proposed algorithm finds a near-optimal feasible solution in most of the cases where it exists. This explains why RMEC presents itself as a more scalable and robust heuristic.

Until now the proposed heuristic was compared to the state-of-the-art algorithm regarding the average satisfaction rate, the overall system rate and the outage probability, and in all these analyses, the proposed algorithm outperforms the existing state-of-the-art. However, since RMEC is a polynomial-time suboptimal heuristic, it does not ensure to find a feasible solution. Nevertheless, it is highly desirable that when the algorithm does not achieve the minimum satisfaction requirement, it provides a near feasible solution, i.e., it is important to provide a solution as close as possible to the desired target requirements. After all, in a communication system, achieving the minimum QoS/QoE requirement is usually more important than reaching a higher system rate. In order to analyze the quality of the non-feasible solutions of both algorithms, the Cumulative Distribution Function (CDF) of the satisfaction and the overall system throughput are depicted in Fig. 3.7. This analysis considers a scenario with $U = 20$ UEs, with a minimum MOS target equal to 4.4 and in which all UEs should be satisfied, i.e., $\xi_1 = 20$ UEs. Similarly to the previous results, it was considered only instances of the problem where a feasible solution exists.

Observe in Fig. 3.7a that the proposed algorithm presents much more satisfactory results than the RAISES. As already mentioned, in this case the RMEC algorithm finds a feasible solution in 94.7% (i.e., 5,3 % of fail) of the cases in opposite to RAISES, which failed in finding a feasible solution in 14.4% of the simulations. Moreover, while RMEC reaches the minimum satisfaction target at the 5.3%-ile, the RAISES heuristic ensures a satisfaction rate only of 65% at the same percentile. It is also important to highlight that the satisfaction rate achieved by the

Figure 3.8 – Satisfaction and throughput considering $U = 30$ UEs, minimum satisfaction requirement equal to $\xi = 100\%$ of the UEs.



Source: Created by the author.

proposed algorithm was at least 75%, while RAISES reaches in one simulation a satisfaction rate of 10%. Concerning the throughput, observe that RMEC achieves a CDF closer to the optimal solution, while RAISES presents a larger gap, mainly for smaller percentiles. Note that for the 10%-ile the proposed heuristic has a gap of 8% to the optimal solution. Meanwhile the RAISES algorithm has a gap of 39.9%. Observe that the lower the percentiles, the harder it is to find a feasible solution in the given scenario. Therefore, the results presented in Fig. 3.7 show that RMEC provides better results when the feasible solution is harder to find. It means that the proposed heuristic is able to provide a better satisfaction rate to UEs closer to the cell edge, in addition to achieving a higher overall system rate than the state-of-the-art algorithm.

This characteristic of providing a better QoS/QoE in harder scenarios is highly desired in a RRA algorithm with satisfaction constraints. Moreover, when the scenario does not have a feasible solution, an important feature that a QoS/QoE constrained RRA algorithm should seek is to provide a good result within the presented circumstances. In order to further evaluate the performance of the proposed algorithm against the state-of-the-art, the next analyses consider only results where there is no feasible solution available. Therefore, the “best solution” is obtained as following:

i. Try to solve the optimization problem stated in (3.1);

ii. If a feasible solution is found, then the “best solution” is found, otherwise relax the optimization problem by reducing the number of UEs that should be satisfied by one, i.e., $\xi_1 = \xi_1 - 1$, and go back to step i.

The results presented in Fig. 3.8 depict the average satisfaction and the overall system throughput of the proposed algorithm compared to the “best solution” and the RAISES heuristic considering only cases that yield infeasible instances of the problem (3.1). This analysis considers a scenario with $U = 30$ UEs and a minimum satisfaction requirement equal to $\xi = 100\%$ of the UEs.

Notice that in Fig. 3.8a, the average satisfaction rate of the proposed algorithm is very close to the "best solution". Indeed, for the highest MOS value analyzed, the proposed algorithm presented an average satisfaction rate 5.1% below the best result. On the other hand, the state-of-the-art algorithm does not deal properly with the infeasibility, yielding results very far from the expected. The same conclusions apply to the overall system throughput presented in Fig. 3.8b. Observe that in this case, the gap between the proposed heuristic and the "best result" is at most 12.9%. Meanwhile, the RAISES heuristic achieves very low values of system throughputs, reaching up at most 12.6% of the best expected throughput.

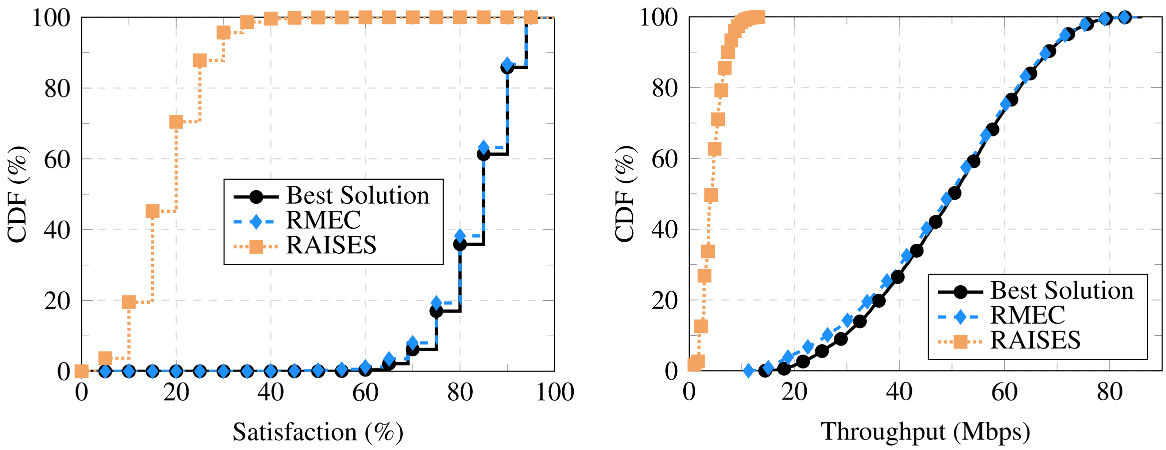
The main reason behind this huge difference in the results presented by RMEC and RAISES algorithms is also caused by the initial solution. As already discussed, the starting point considered by the state-of-the-art algorithm is the output of the max-rate scheduler, i.e., all the RBs are allocated to the best possible user. However, if no UE is satisfied after this initial allocation, it means that even the UE with the best channel conditions were not able to become satisfied, probably receiving all the available RBs. In these cases, the state-of-the-art returns this initial allocation as its best achievable solution. Nevertheless, in order for such event to happen, the scenario must be extremely hard such that no UE can be satisfied. On the contrary, if some UE is satisfied then the RAISES algorithm tries to reallocate the RBs in order to meet the UEs' requirements. In a hard scenario, the usual initial solution is that all RBs are allocated to only one or at most a very few UEs. Therefore, in the reallocation process, the resources will be redistributed to other UEs even if the other UEs do not get satisfied after the reallocation process. In the end, the resources are redistributed from the UEs with better channel conditions to those that may not perform a good use of these resources, which causes a drastic reduction in the overall system throughput. On the other hand, the starting point of the initial solution of RMEC algorithm is the result of the LP, presented in (3.15). In hard scenarios, it is probable that no feasible solution exists for (3.15). In these cases, this LP is relaxed by disabling the QoS/QoE constraint of the UE considered harder to satisfy, adopting the same criterion used in the UE selection step, explained in Section 3.4.1. This relaxation is performed until a feasible solution to (3.15) is reached. Therefore, the UEs that are harder to satisfy are disregarded from the initial solution and they will not compete for resources in the reallocation process. Moreover, as already explained, compared to the RAISES, the initial solution of the RMEC algorithm is closer to the desirable, since it is obtained from a graph-based rounding of the LP solution.

In order to evaluate the quality of the solutions reached by the proposed heuristic in the infeasible scenarios, the CDF of the satisfaction and the overall system throughput are presented in Fig. 3.9 considering 20 UEs constrained by satisfying all UEs with a minimum MOS requirement of 4.4.

Observe that in Fig. 3.9a, the proposed algorithm reaches satisfaction results very close to the "best solution", which corroborates the high robustness of the RMEC algorithm. Indeed, the highest satisfaction value achieved by RAISES is 65%, against 95% achieved by RMEC. Moreover, regarding the 10%-ile, i.e., the 10% of the harder scenarios, the proposed

Figure 3.9 – CDF of the satisfaction and throughput considering $U = 20$ UEs, a minimum MOS equal to 4.4 and $\xi = 100\%$ of the UEs.

(a) CDF of the satisfaction. (b) CDF of the throughput.



Source: Created by the author.

algorithm ensures a satisfaction rate of at least 70% of the UEs, while RAISES only guarantees to satisfy 10% of the UEs. When comparing the results presented in Figs. 3.7a and 3.9a, the problem of the initial solution of RAISES becomes more evident. Observe that even when a feasible solution is guaranteed to exist, the RAISES algorithm presents very low satisfaction results when the feasible solution is hard to find. When a feasible solution does not exist, the satisfaction results provided by RAISES are far from the target.

Regarding the overall system throughput in Fig. 3.9b, the proposed algorithm also achieves results very close to the "best solution". In the 10%-ile, the "best solution" achieves a throughput of 29.95 Mbps. Meanwhile the proposed algorithm achieves a throughput of 26.3 Mbps. On the other hand, the RAISES heuristic only reaches a throughput of 1.96 Mbps at the 10%-ile. Moreover, the state-of-the-art algorithm reaches, in this analysis, at most 13.72 Mbps, which is only slightly higher than the minimum throughput achieved by the proposed algorithm, corresponding to 11.28 Mbps.

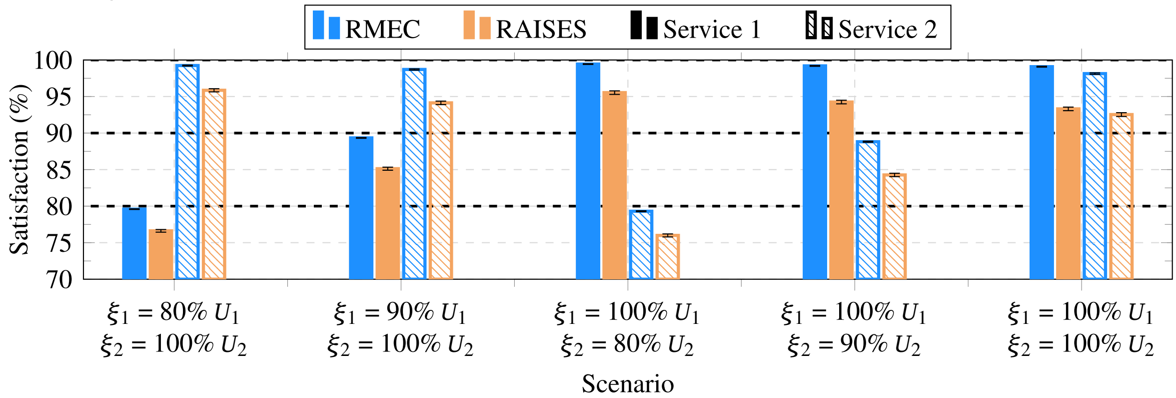
Until now, all the performed analyses showed that the proposed algorithm outperforms the state-of-the-art heuristic, namely RAISES, besides reaching near optimal results. Moreover, even in scenarios where there is no feasible solution, the RMEC algorithm provides good results, achieving near feasible solutions. However, these analyses considered that all UEs subscribed the same service plan, varying the number of UEs in the system, U, the minimum MOS requirement, Ω_1^{target}, and the minimum number of UEs that should be satisfied by the allocation algorithm, ξ_1. The next analyses consider multi-service scenarios, i.e., the UEs are divided into distinct groups depending on the service plan they subscribe to, which may have distinct requirements. Moreover, in order to evaluate the proposed algorithm using QoS constraints, the minimum requirement of the services are given in terms of minimum throughput instead of minimum MOS. In Fig. 3.10, the proposed algorithm is compared against the optimal solution and the state-of-the-art heuristic, namely RAISES, in five different scenarios, considering two service plans and $U = 30$ UEs. The first service plan consists in a high-quality skype video call, which has a recommended minimum throughput of 500 kbps [67], i.e., $\Omega_1^{\text{target}} = 500$ kbps. The second service plan models a high definition skype video call, which recommends a minimum throughput of 1.5 Mbps [67], i.e, $\Omega_2^{\text{target}} = 1.5$ Mbps. It is also considered that 20 UEs subscribe the first service plan and 10 UEs subscribe the second one, i.e., $U_1 = 20$ UEs and $U_2 = 10$ UEs. During the following analyses, these service plans will be referred as service 1 and 2, respectively. The difference between the analyzed scenarios relies on the minimum number of UEs that should be satisfied in each service plan, namely, ξ_1 and ξ_2. The results presented in Fig. 3.10 considered only feasible instances of the problem (3.1).

In Fig. 3.10a, the average satisfaction is depicted. In the first scenario, when $\xi_1 = 80\% \cdot U_1$ and $\xi_2 = 100\% \cdot U_2$, the average satisfaction rate provided by RMEC algorithm is 0.41% and 0.76% below the target for services 1 and 2, respectively. Meanwhile the state-of-the-art solution presents a gap from the target of services 1 and 2 of 3.38% and 4.14%, respectively. In this scenario, the joint throughput required by both services is $\xi_1\Omega_1^{\text{target}} + \xi_2\Omega_2^{\text{target}} = 80\% \cdot 20 \cdot 500\text{ kbps} + 100\% \cdot 10 \cdot 1.5\text{ Mbps} = 23\text{ Mbps}$. Observe that when the minimum number of UEs that should be satisfied by the service 1 increases to $90\% \cdot U_1$ and the target for service 2 is kept in $100\% \cdot U_2$, the required joint throughput is equal to 24 Mbps. In this case, which is a more demanding scenario compared to the previous case, the impact on the gap between the average satisfaction rate and the target presented by the proposed algorithm is negligible. On the other hand, the RAISES algorithm presented a significant increase on the gap between the average satisfaction rate and the minimum target required by both services, namely a gap of 4.87% for service 1 and 5.87% for service 2. This difference between the average satisfaction rate provided by RMEC and RAISES grows when $\xi_1 = U_1$ and $\xi_2 = U_2$, where the overall required throughput is 25 Mbps. Herein, the RMEC algorithm achieves satisfaction rates of 99.08% and 98.98% for in service 1 and 2, respectively. However, the RAISES algorithm presents a considerable gap to the satisfaction target for both service plans, which are equal 6.70% and 7.49% for service 1 and 2, respectively. With the increasing minimum number of UEs that should be satisfied in service 1, ξ_1, the UE diversity diminishes, i.e., the algorithms are free to neglect a smaller number of UEs to be satisfied. Nevertheless, observe that the proposed algorithm achieves a near optimal result with an average satisfaction 1% below the target for service 1 and 2% below for service 2 in the worst case scenario, where all UEs should be satisfied, i.e., $\xi_1 = U_1$ and $\xi_2 = U_2$. On the other hand, as already discussed in the single service scenario, RAISES scales worse than RMEC, degrading its average satisfaction rate more rapidly with the decreasing UE diversity. A similar analysis can be performed considering $\xi_1 = U_1$ and varying the minimum number of UEs that should be satisfied in service 2, ξ_2. Observe that the proposed algorithm outperforms the state-of-art solution providing a near optimal average satisfaction results in all evaluated scenarios.

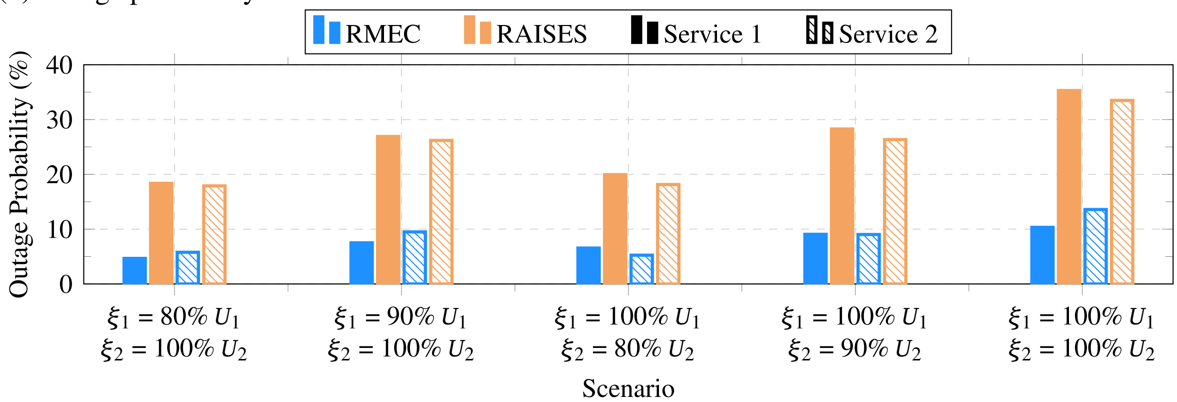
Regarding the outage probability presented in Fig. 3.10b, it is possible to observe in all the scenarios that the proposed algorithm finds a feasible solution more often than the

Figure 3.10 – System performance considering $U = 30$ UEs and $S = 2$ service plans, where $U_1 = 20$ and $U_2 = 10$ UEs, $\Omega_1^{\text{target}} = 500$ kbps and $\Omega_2^{\text{target}} = 1.5$ Mbps.

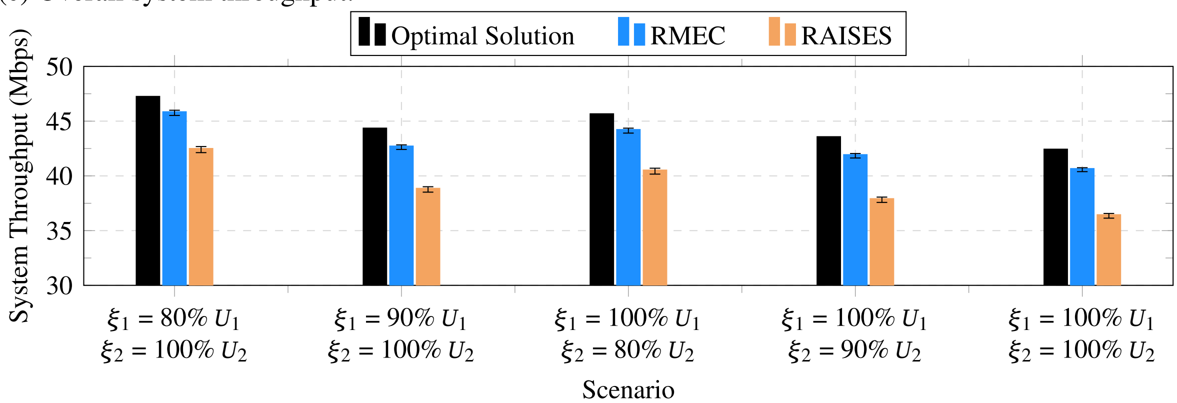
(a) Average satisfaction rate.



(b) Outage probability.



(c) Overall system throughput.



Source: Created by the author.

RAISES heuristic. Moreover, in almost all cases, the outage probability achieved by RMEC is below 10%. Indeed, in the worst evaluated scenario, where $\xi_1 = U_1$ and $\xi_2 = U_2$, the outage probability of services 1 and 2 achieved by RMEC is 10.39% and 13.70%, respectively. On the other hand, the state-of-the-art algorithm fails in finding a feasible solution much more often, with a probability of 35.04% for service 1 and 33.17% for service 2. Comparing these results with those obtained in the single service scenario in Fig. 3.6, it is possible to observe that the gap between the outage probability per service of the proposed and the state-of-the-art algorithms is higher in the multi-service scenario, even when the same number of UEs is considered, i.e., when there are 30 UEs in the single service case. In the scenario where $\xi_1 = U_1$ and $\xi_2 = U_2$, the required total throughput is equal to 25 Mbps. In its turn, the required total throughput in the single service scenario with 30 UEs and a required MOS of 4.4 (equivalent to a throughput requirement of 885.27 kbps) is equal to 26.59 Mbps and the gap between the outage probability of the RMEC and RAISES is equal to 16.5%. Observe that although in the single service case the throughput required by the UEs in the system is greater than in the multi-service scenario, the gap between the outage probability per service plan of RMEC and RAISES is higher in the latter case. This happens because the algorithm must divide the same amount of resources between different services, which implies that the UEs from each service plan will dispute less RBs. This result shows that the proposed algorithm has a better scalability with the increasing number of service plans offered by the system operator, which is also an effect of the initial solution provided by the proposed algorithm and was already discussed in the previous analyses.

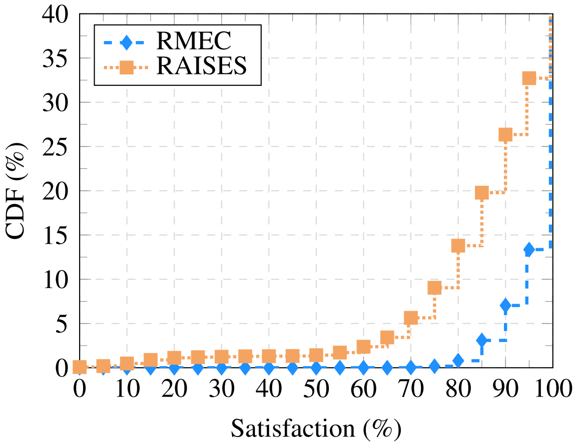
Comparing the results of outage probability in all scenarios, it is possible to observe that the proposed solution deals better with the UE diversity than the RAISES algorithm, as already discussed in the results presented in Fig. 3.6. Indeed, in the scenario where $\xi_1 = 80\% \cdot U_1$ and $\xi_2 = 100\% \cdot U_2$, the outage probability of the proposed algorithm for services 1 and 2 are equal to 4.66% and 5.78%, respectively. On the other hand, when the RAISES algorithm is employed, an outage event occurs in 18.36% of the times for the service 1 and 17.89% for the service 2. When the scenario where $\xi_1 = 100\% \cdot U_1$ and $\xi_2 = 80\% \cdot U_2$ is considered, the results of outage probability achieved by the proposed algorithm for services 1 and 2 are 6.55% and 5.25%, respectively, against 19.95% and 18.14% reached by RAISES. This result reinforces the better scalability of the RMEC algorithm, showing that the proposed solution takes better advantage of the UE diversity.

In the scenario where $\xi_1 = 80\% \cdot U_1$ and $\xi_2 = 100\% \cdot U_2$, the required throughput is 23 Mbps. On the other hand, the required throughput in the scenario where $\xi_1 = 100\% \cdot U_1$ and $\xi_2 = 80\% \cdot U_2$ is equal to 22 Mbps. However, although in the latter scenario the required throughput is lower than the first one, the average service outage probability is slightly higher than in the first case, specially for RAISES algorithm. This effect appears more relevantly in Fig. 3.10c, where the overall system throughput is analyzed. Observe that, when the $\xi_1 = 80\% \cdot U_1$ and $\xi_2 = 100\% \cdot U_2$, the throughput of the optimal solution is equal to 47.16 Mbps and for $\xi_1 = 100\% U_1$ and $\xi_2 = 80\% \cdot U_2$, the optimal solution achieves a throughput 3.37% smaller. This results can be explained by the UE diversity, which is slightly different in both cases. In the scenario where $\xi_1 = 80\% \cdot U_1$ and $\xi_2 = 100\% \cdot U_2$, the system can neglect at most 4 UEs of the service 1. On the other hand, in the scenario where $\xi_1 = 100\% \cdot U_1$ and $\xi_2 = 80\% \cdot U_2$, the system must satisfy all UEs subscribing service 1 and 8 out of 10 UEs from service 2. Therefore, even thought in the first case the system requires more throughput, it can disregard more UEs with bad channel conditions, easing the resource allocation.

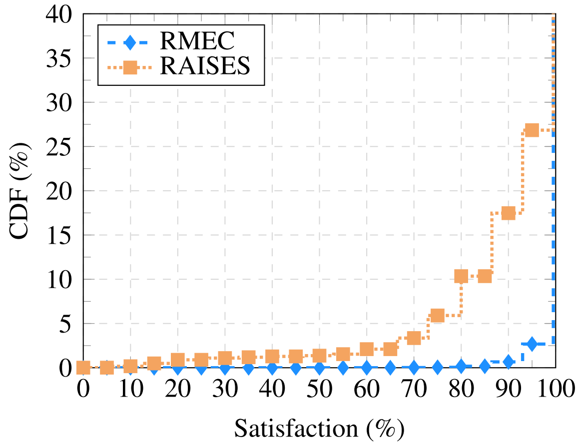
Moreover, regarding the system throughput in Fig. 3.10c, the proposed algorithm presented results close to the optimal solution. In fact, the highest performance loss is registered

Figure 3.11 – CDF of the satisfaction of each service plan and the overall system throughput in a scenario considering $U = 40$ UEs and $S = 3$ service plans, where $U_1 = 20$, $U_2 = 15$ and $U_2 = 5$ UEs, and $\Omega_1^{\text{target}} = 500$ kbps, $\Omega_2^{\text{target}} = 300$ kbps and $\Omega_3^{\text{target}} = 1.5$ Mbps.

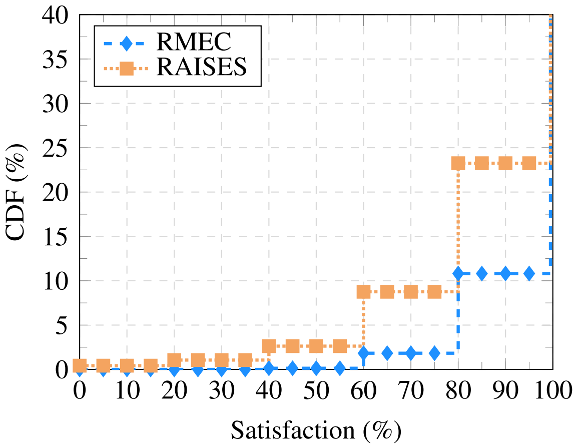
(a) CDF of the satisfaction of service 1.



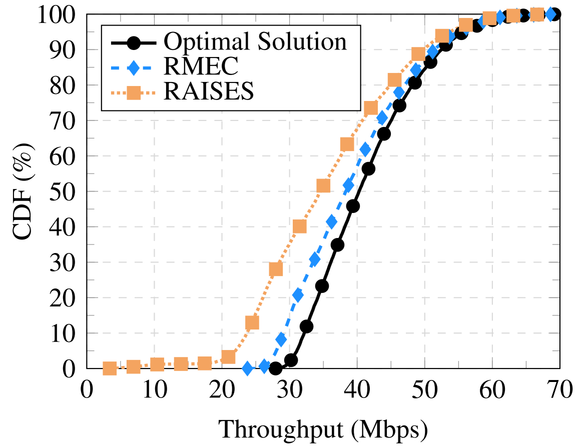
(b) CDF of the satisfaction of service 2.



(c) CDF of the satisfaction of service 3.



(d) CDF of the throughput.



Source: Created by the author.

in the hardest scenario with $\xi_1 = U_1$ and $\xi_2 = U_2$, where the proposed algorithm achieves a throughput 4.22% below the optimal solution. Nevertheless, in this scenario the state-of-the-art algorithm reaches a throughput 14.19% below the optimal solution.

In order to evaluate the quality of the non-feasible solutions of both algorithms in a multi-service scenario, as done for the single service case in Fig. 3.7, the CDF of the satisfaction rate of each service plan and the overall system throughput are presented in Fig. 3.11. Here, a scenario with $U = 40$ UEs divided into subscribing to 3 different service plans is considered. The service 1 consists in a high-quality skype video call, i.e., $\Omega_1^{\text{target}} = 500$ kbps, and has 20 subscribers. The service 2 has 15 subscribers and it is equivalent to a standard video call with a minimum recommended throughput of $\Omega_2^{\text{target}} = 300$ kbps [67]. Finally, the service 3 consists in a high definition skype video call, which recommends a minimum throughput of $\Omega_3^{\text{target}} = 1.5$ Mbps.

Observe in Figs. 3.11a, 3.11b and 3.11c that the quality of the results regarding the satisfaction rate achieved by the proposed heuristic are better than the ones provided by the state-of-the-art algorithm, as already observed in the single service scenario. Indeed, for the service 1, the proposed algorithm ensures in 99% of the times a satisfaction rate of a least 80%, meanwhile the RAISES algorithm only ensures 80% of satisfaction in 86.20% of the cases. Moreover, observe that the RMEC algorithm satisfies all UEs subscribing this service at the 13.35%-ile with a gap of 19.37% from the RAISES algorithm. Regarding service 2, observe in Fig. 3.11b that the difference between the CDF of the satisfaction rate of RMEC and RAISES algorithms is more significant. In fact, the proposed solution manages to satisfy all UEs of service 2 at the 2.67%-ile. On the other hand, RAISES only satisfy all UEs of service 2 in 73.15% of the times. Comparing this result with the CDF of the satisfaction of service 1, it is possible to observe the robustness of the proposed algorithm to the increasing number of service plans managed by the system. Although the service 2 serves a smaller number of UEs and each UE subscribing it requires a lower throughput, the state-of-the-art algorithm does not allocate the resources in a more proper manner, achieving a high outage rate even for a less demanding service. In Fig. 3.11c, the satisfaction rate of service 3 is depicted. Here, the proposed solution is able to satisfy all UEs subscribed to service 3 in 89.19%, 12.44% more than the state-of-the-art algorithm. Besides that, the RMEC algorithm ensures satisfying at least 3 out of 5 UEs of service 3, in the 2%-ile, and RAISES at the 8.76%-ile. Completing the analysis, the CDF of the overall system throughput is depicted in Fig. 3.11d. Notice that the spectral efficiency of the proposed algorithm is close to the optimal solution. Moreover, as already discussed in the previous analyses, the RAISES algorithm presented a poor performance mainly in the harder instances. Indeed, at the 10%-ile, the RMEC and RAISES algorithms achieve a throughput 9.46% and 26.33% below the optimal solution, respectively. Besides the larger gap at the lower percentiles of the CDF, this result also reinforces the fact that the RAISES algorithm does not deal properly with infeasible solutions. Note that the RAISES algorithm presents very low system throughput results in some instances of this scenario. In fact, the minimum overall system throughput achieved by RAISES in the presented CDF is 3.22 Mbps, much less than the minimum value achieved by RMEC, which is 23.64 Mbps.

3.6 Chapter Summary

In this chapter, the problem of maximizing the overall system rate, subject to meeting the QoS/QoE requirements of at least a minimum number of UEs per service has been studied, considering both single and multi-service cellular scenarios. It is worth to note that this QoS/QoE constraint is very important to the mobile network operators, in order to ensure the customers' minimum requirements, and consequently their satisfaction.

The problem was reformulated as an ILP, which can be solved by standard methods, like BB or BC. However, the computational complexity to obtain the optimal solution is prohibitive in real-time systems. Therefore, a low-complexity suboptimal algorithm, called RMEC, was proposed.

In the analysis performed in Section 3.5, the RMEC algorithm presented a near optimal behavior in terms of achieved system throughput and average satisfaction. Furthermore, the proposed algorithm was also analyzed in scenarios were the constraints of the problem are impossible to be met. These analyses showed that besides of the near optimal results, the RMEC algorithm also provides near feasible solutions, i.e, it reaches a solution that is close to the best one available.

During all the analyses, the RMEC algorithm outperforms the state-of-art heuristic that intends to solve the same problem, but considering only QoS constraints. However, it is important to emphasize that the better performance of the proposed algorithm comes with the cost of a higher computational complexity. Therefore, the RAISES algorithm remains as a good choice as a RRA at non challenging scenarios, i.e., systems with UEs with good channel conditions and low rate requirements. On the other hand, the RMEC algorithm is more suitable in challenging scenarios, mainly when a feasible solution is hard to find.

4 QOS/QOE-AWARE SCHEDULING ALGORITHM FOR RATE MAXIMIZATION IN WIRELESS NETWORKS

In the previous chapter, the problem of maximizing the overall system rate while ensuring that a minimum number of UEs of each service plan gets their QoS/QoE requirements met was discussed. However, the problem treated in Chapter 3 consists of a snapshot optimization problem solved on a TTI basis, i.e., the resources are scheduled at each TTI considering no information from the previous allocations.

This approach has two inherent limitations:

1. The maximum number of UEs that can be satisfied by the scheduler is limited by the number of available RBs;
2. It does not consider the memory of the previous resource allocations in the system, which may impact significantly in the satisfaction and throughput results.

Aware of these problems, this chapter addresses the scheduling problem discussed in previous chapter, but considering here that the UE satisfaction is measured during a timespan, instead of a single TTI. The main contributions of this chapter are:

- Study of the problem of maximizing the overall system rate in a multi-service scenario, considering that a fraction of the users of each service must have their QoE requirements met during a timespan;
- Reformulation of this problem as an ILP and solving it using standard algorithms;
- Proposal of a low-complexity suboptimal solution that has near optimal performance and presents high scalability in terms of the problem inputs.

The remainder of this chapter is organized as follows. In Sections 4.1 and 4.2, the problem addressed in this chapter is formulated as an optimization problem and then rewritten as an ILP, which has a more tractable form. In Section 4.3 a low complexity scheduler is proposed to solve the problem stated in Section 4.1. In Section 4.4, the benchmark algorithms to be compared to the low complexity heuristic proposed in Section 4.3 are briefly described. In Section 4.5, the performance of the suboptimal algorithm is evaluated, by comparing it against benchmarking algorithms from the literature. Finally, the chapter remarks are presented in Section 4.6.

4.1 Problem Formulation

This section addresses the problem of maximizing the overall system throughput constrained by ensuring that a minimum number of UEs ξ_s per service plan s meet their QoS/QoE requirements during a given timespan, which is formulated as an optimization problem in what follows below.

Specifically, let $\mathcal{T}$ be a sequence of T consecutive TTIs of a timespan and $\mathbf{X}^{(T)}$ be a $U \times K \times T$ assignment tensor, where each element $x_{u,k}[t]$ is equal to 1 if the RB k is allocated to the UE u in the t^{th} TTI, or equal to 0 otherwise. Therefore, the studied optimization problem can be written as follows

$$\max_{\mathbf{X}^{(T)}} \quad \frac{1}{T}\sum_{u\in\mathcal{U}}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}} r_{u,k}[t]x_{u,k}[t], \tag{4.1a}$$

$$\text{s.t.} \quad \sum_{u\in\mathcal{U}} x_{u,k}[t] = 1, \forall k \in \mathcal{K} \text{ and } t \in \mathcal{T}, \tag{4.1b}$$

$$\sum_{u\in\mathcal{U}_s} H\left(\Omega_u\left(\frac{1}{T}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}} r_{u,k}[t]x_{u,k}[t]\right), \Omega_s^{\text{target}}\right) \geq \xi_s, \forall s \in \mathcal{S}, \tag{4.1c}$$

$$x_{u,k}[t] \in \{0,1\}, \forall u \in \mathcal{U}, \forall k \in \mathcal{K} \text{ and } t \in \mathcal{T}, \tag{4.1d}$$

where $r_{u,k}[t]$, $H(\cdot)$ and $\Omega_u(\cdot)$ were defined in Section 3.1.

The optimization problem stated in (4.1) aims at finding the optimal resource assignment that maximizes the total system throughput in the objective function (4.1a). Constraints (4.1b) and (4.1d) guarantee that each RB is assigned to a single UE per TTI. Furthermore, (4.1c) requires that a minimum number ξ_s of UEs should be satisfied for each service plan s.

4.2 Optimal Solution

In a similar manner to the problem stated in Section 3.1 of the previous chapter, (4.1) denotes a combinatorial optimization problem with a nonconvex constraint (4.1c), hence it has a prohibitive computational complexity [54].

Observe that the constraint (4.1c) is very similar to (3.1c) in Section 3.2. Therefore, in an analogous way as it was done in Section 3.2, the minimum MOS requirement can be converted into a rate requirement, using the function $\Omega^{\dagger}(\cdot)$, defined in (3.2). Moreover, the entire constraint (4.1c) can be rewritten into two new linear constraints, yielding

$$\max_{\mathbf{X}^{(T)},\boldsymbol{\rho}} \quad \frac{1}{T}\sum_{u\in\mathcal{U}}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}} r_{u,k}[t]x_{u,k}[t], \tag{4.2a}$$

$$\text{s.t.} \quad \sum_{u\in\mathcal{U}} x_{u,k}[t] = 1, \forall k \in \mathcal{K} \text{ and } t \in \mathcal{T}, \tag{4.2b}$$

$$\frac{1}{T}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}} r_{u,k}[t]x_{u,k}[t] \geq \psi_u\rho_u, \ \forall u \in \mathcal{U}, \tag{4.2c}$$

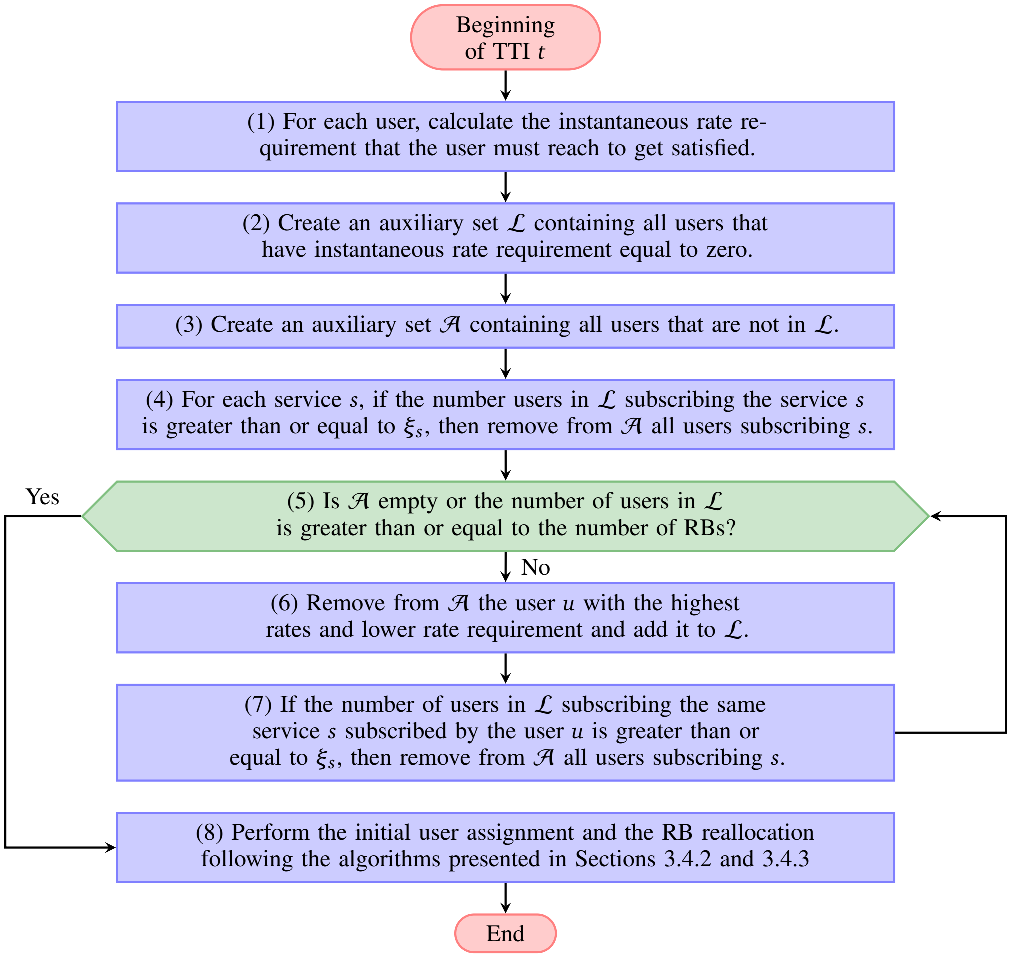
$$\sum_{u\in\mathcal{U}} q_{s,u}\rho_u \geq \xi_s, \forall s \in \mathcal{S}, \tag{4.2d}$$

$$x_{u,k}[t] \in \{0,1\}, \forall u \in \mathcal{U}, \forall k \in \mathcal{K} \text{ and } t \in \mathcal{T}, \tag{4.2e}$$

$$\rho_u \in \{0,1\}, \forall u \in \mathcal{U}. \tag{4.2f}$$

Notice that (4.2) is an ILP, which has a more friendly structure and can be solved using standard algorithms, such as BB. However, even for small instances of the problem, the

Figure 4.1 – Flowchart of the TRMEC Algorithm.



Source: Created by the author.

optimal resource scheduling requires the knowledge of all $r_{u,k}[t]$, during the entire timespan, which is not a realistic assumption for a real system, depending on the channel coherence time.

4.3 Suboptimal Solution

In this section a low complexity suboptimal solution to the problem described in Section 4.1 is proposed. It is a temporal extension, called Temporal RMEC (TRMEC), of the RMEC algorithm presented in the previous chapter.

The procedures followed by TRMEC are very similar to the ones presented by RMEC. Indeed, the TRMEC differs from RMEC only in the first step of the algorithm, i.e., the user selection. The last two steps, namely, the initial user assignment and the reallocation procedures are the same in both algorithms. As will be presented at the performance analyses of this chapter, there are significant improvements at the first step of the proposed algorithm when compared to the first step of RMEC, presented in Section 3.4.1. A general overview of the TRMEC algorithm is depicted in the flowchart of Fig. 4.1 and its detailed description is presented in the remainder of this section.

The proposed algorithm is meant to be executed at each TTI, therefore, it must deal with the time dimension properly. Consider $R_u^{\text{avg}}[t]$ as the average data rate of user u until the t^{th} TTI, as defined in Section 2.5. In order to solve the scheduling over a given timespan, as depicted in block (1) of Fig. 4.1, an instantaneous user requirement $\psi'_u[t]$ is defined as

$$\psi'_u[t] = \begin{cases} \max\left(\dfrac{t\psi_u - (t-1)R_u^{\text{avg}}[t-1]}{t}, 0\right), & \text{for } t \text{ greater than the first TTI of } \mathcal{T} \\ \psi_u, & \text{otherwise,} \end{cases} \tag{4.3}$$

which corresponds to the total rate that the UE u must achieve in order to fulfill its rate requirement in the current TTI t. It is important to observe that in the formulation presented in (4.3), the value of $R_u^{\text{avg}}[t]$ comes from a iterative mean formula, as stated in (2.16). However, the value of $R_u^{\text{avg}}[t]$ adopted in this formulation could be obtained by other methods, such as the exponential moving average [68].

Besides the instantaneous requirement $\psi'_u[t]$, the selection variable ρ_u must also be restated in order to be defined on each TTI. Consider a new instantaneous selection variable $\boldsymbol{\rho}'[t] \in \{0,1\}^{U\times 1}$, where each element $\rho'_u[t]$ defines whether the UE u is selected to get satisfied in the current TTI t. Therefore, a new optimization problem is stated from (4.2) as:

$$\max_{\mathbf{X}^{(T)}, \boldsymbol{\rho}'[t]} \quad \sum_{u\in\mathcal{U}} \sum_{k\in\mathcal{K}} r_{u,k}[t] x_{u,k}[t], \tag{4.4a}$$

$$\text{s.t.} \quad \sum_{u\in\mathcal{U}} x_{u,k}[t] = 1, \forall k \in \mathcal{K}, \tag{4.4b}$$

$$\sum_{k\in\mathcal{K}} r_{u,k}[t] x_{u,k}[t] \geq \psi'_u[t] \rho'_u[t],\ \forall u \in \mathcal{U}, \tag{4.4c}$$

$$\frac{1}{T} \sum_{u\in\mathcal{U}} \sum_{t'\in\mathcal{T}} q_{s,u} \rho'_u[t'] \geq \xi_s, \forall s \in \mathcal{S}, \tag{4.4d}$$

$$x_{u,k}[t] \in \{0,1\}, \forall u \in \mathcal{U} \text{ and } \forall k \in \mathcal{K}, \tag{4.4e}$$

$$\rho'_u[t] \in \{0,1\}, \forall u \in \mathcal{U} \text{ and } t \in \mathcal{T}. \tag{4.4f}$$

Notice that (4.4) is now an optimization problem similar to (3.5) treated in Section 3.4, except for constraint (4.4d), which is the only equation depending on the entire timespan $\mathcal{T}$. In its turn, the remaining equations of the optimization problem (4.4) rely only on information of current TTI t. Therefore, in order to select the fraction of users that will get satisfied on each service, the $\rho'_u[t]$ variables must be estimated at each TTI t.

In most cases, due to the large-scale fading, the SNR in all RBs of a specific user u present similar values, which leads to similar rates. Furthermore, the higher the QoE/QoS requirement of a UE is, the harder it is to satisfy it, since the UE requires a higher rate and, consequently, more RBs to get satisfied. Moreover, since the objective is to maximize the total system rate, it is plausible to satisfy the easiest UEs first. In order to do it, an auxiliary set $\mathcal{L}$ is created, initially containing all users for which $\psi'_u[t] = 0$, i.e., the users that do not require resources at the current TTI, as illustrated in block (2) of Fig. 4.1. After that, in block (3) of Fig. 4.1, one defines set $\mathcal{A}$ that initially contains the users who were not yet selected to get satisfied, i.e., $\mathcal{A} = \mathcal{U} \setminus \mathcal{L}$.

After this initial selection, some services may have their minimum number of satisfied UEs already fulfilled. For these services, there is no need to satisfy the requirements of any more users. Therefore, as depicted in block (4) of Fig. 4.1, all the users subscribing services with its minimum number of satisfied UEs fulfilled are removed from $\mathcal{A}$, i.e., all UEs $u \in \{\mathcal{U}_s \cap \mathcal{A} \mid |\mathcal{U}_s \cap \mathcal{L}| \geq \xi_s, \forall s \in \mathcal{S}\}$. Then, in blocks (4)-(7) of Fig. 4.1, the users are iteratively moved from $\mathcal{A}$ to $\mathcal{L}$ based on the following criterion

$$u' = \arg\max_{u \in \mathcal{A}} \left\{ \frac{\sum_{k \in \mathcal{K}} r_{u,k}[t]}{\psi'_u[t]} \right\}. \tag{4.5}$$

Considering $s \in \mathcal{S}$ the service where $q_{s,u'} = 1$, if $|\mathcal{U}_s \cap \mathcal{L}| \geq \xi_s$, all UEs $u \in \mathcal{U}_s \bigcap \mathcal{A}$ are taken out from $\mathcal{A}$. This process is repeated until $\mathcal{A}$ becomes empty or $|\mathcal{L}| \geq K$. Finally, the values of $\rho'_u[t]$ are defined as $\rho'_u[t] = 1$ for all UEs $u \in \mathcal{L}$ and $\rho'_u[t] = 0$, otherwise.

At the end of the estimation of $\rho'_u[t]$ for all users $u \in \mathcal{U}$, the constraint (4.4d) can be removed from (4.4), since $\boldsymbol{\rho}'[t]$ is no longer a variable of the optimization problem. Thus, (4.4) can be restated as

$$\max_{\mathbf{X}^{(T)}} \quad \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} r_{u,k}[t] x_{u,k}[t], \tag{4.6a}$$

$$\text{s.t.} \quad \sum_{u \in \mathcal{U}} x_{u,k}[t] = 1, \quad \forall k \in \mathcal{K}, \tag{4.6b}$$

$$\sum_{k \in \mathcal{K}} r_{u,k}[t] x_{u,k}[t] \geq \psi'_u[t] \rho'_u[t], \quad \forall u \in \mathcal{U}, \tag{4.6c}$$

$$x_{u,k}[t] \in \{0, 1\}, \quad \forall u \in \mathcal{U} \text{ and } \forall k \in \mathcal{K}. \tag{4.6d}$$

Notice that (4.6) is a time independent optimization problem which can be solved on each TTI. Furthermore, at this point, the users that should be satisfied at the TTI t are already selected, i.e., $\rho'_u[t] = 1$.

Observe that the proposed algorithm reaches a point similar to the end of the first step of RMEC. In fact, problem (4.6) is similar to the one stated (3.14), which corresponds to the resulting relaxed optimization problem after the user selection step of RMEC algorithm, presented in Section 3.4.1. Therefore, the rest of the suboptimal heuristic proposed in this chapter follows the steps 2 and 3 of the RMEC, detailed in Sections 3.4.2 and 3.4.3, as illustrated in block (8) of Fig. 4.1.

The main improvement of TRMEC with respect to RMEC is the indirect exploitation of the previous allocations. Differently of RMEC, the TRMEC adjusts the UEs rate requirement on each TTI, which impacts on the their priority calculation at the user selection. Therefore, although in Section 3.4.1 the UEs are selected following the same criteria (4.5), the proposed solution takes advantage of the users' current throughput. It is important to highlight that the worst-case computational complexity of the TRMEC on each TTI is equal to the RMEC's, which is $\mathcal{O}\left(U^{3.5} K^{3.5}\right)$.

4.4 Benchmark Algorithms

In order to evaluate the performance of the proposed algorithm, it will be compared against the RMEC, presented in the previous chapter, and two other benchmark algorithms, namely Adaptive Throughput-based Efficiency-Satisfaction Trade-Off (ATES) [69, 70] and Adaptive Satisfaction Control (ASC) [70]. As far as the author's knowledge goes, these algorithms are the ones with the objective and constraints more closely related to the herein proposed heuristic. Compared to TRMEC, although the benchmark algorithms are capable of addressing different QoS requirements to each UE, they were designed to address the constraint of a minimum number of required UEs to be satisfied in a single service scenario. ATES and ASC are briefly described in the sequel.

Both benchmark algorithms work based on the utility theory, which is well explained in [70]. In summary, utility-based algorithms follow the same general 4 steps, which are executed for each RB $k \in \mathcal{K}$:

1. For each UE $u \in \mathcal{U}$, calculate the priority w_u based on a marginal utility function taking the UE's KPIs as inputs;
2. Schedule the RB k to the UE $u^* = \underset{u \in \mathcal{U}}{\arg\max} \left\{ w_u \cdot r_{u,k} \right\}$;
3. Update the KPIs of the UE u^* considering that the transmission over the RB k will succeed, and be error-free;
4. Update the utility function parameters.

The marginal utility function used to calculate the priority of both ATES and ASC is the shifted log-logistic marginal utility [70], which is given by

$$w_u = \frac{\dfrac{1}{\lambda_{\text{scale}}} \left(1 + \dfrac{\lambda_{\text{shape}} \left(R_u^{\text{avg}}[t] - \psi_u \right)}{\lambda_{\text{scale}}} \right)^{-1 - \frac{1}{\lambda_{\text{shape}}}}}{\left(1 + \left(1 + \dfrac{\lambda_{\text{shape}} \left(R_u^{\text{avg}}[t] - \psi_u \right)}{\lambda_{\text{scale}}} \right)^{-\frac{1}{\lambda_{\text{shape}}}} \right)^2}, \tag{4.7}$$

where the parameters λ_{scale} and λ_{shape} are used to adapt the scale and the shape of the marginal utility function.

The difference between ATES and ASC are the way in which they update the utility function parameters aiming to achieve a certain satisfaction rate, i.e., ξ_1/U. The ATES algorithm adapts the scale parameter, λ_{scale}, keeping the shape parameter fixed. The adaptation of the scale parameter based on the current satisfaction rate, Υ_1, is given by

$$\lambda_{\text{scale}} = \lambda_{\text{scale}} - \eta \left(\Upsilon_1 - \frac{\xi_1}{U} \right), \tag{4.8}$$

where η denotes the step size that determines the speed of the parameter adaptation. This adaptation provides to ATES a trade-off between throughput and satisfaction rate. Additionally, the scale parameter $\lambda_{\text{scale}} \in \left[\lambda_{\text{scale}}^{\min}, \lambda_{\text{scale}}^{\max}\right]$, where $\lambda_{\text{scale}}^{\min}$ and $\lambda_{\text{scale}}^{\max}$ represents the minimum and the maximum values of scale, respectively.

On the other hand, the ASC algorithm controls only the shape parameter of the utility function. Moreover, the adaptation of the shape parameter is similar to that of the scale parameter in the ATES algorithm, and is given by

$$\lambda_{\text{shape}} = \lambda_{\text{shape}} - \eta \left(\Upsilon_1 - \frac{\xi_1}{U} \right). \tag{4.9}$$

Similarly to the scale parameter, the shape is also bounded between a minimum and a maximum values, $\lambda_{\text{shape}}^{\min}$ and $\lambda_{\text{shape}}^{\max}$, respectively, i.e., $\lambda_{\text{shape}} \in \left[\lambda_{\text{shape}}^{\min}, \lambda_{\text{shape}}^{\max}\right]$. Differently of ATES, the ASC aims to fit the satisfaction rate at the desired target, hence, it deals with the UEs' satisfaction more efficiently, however it provides a throughput usually lower than ATES. For further details about ATES and ASC, see [70].

The parametrization of the ATES and the ASC algorithms adopted in this thesis are present in Table 4.1.

Table 4.1 – Parameters of ATES and ASC algorithms.

Parameter	**Value**
Scale (λ_{scale})	0.1088
Minimum scale parameter $\left(\lambda_{\text{scale}}^{\min}\right)$	-50 dB
Maximum scale parameter $\left(\lambda_{\text{scale}}^{\max}\right)$	-10 dB
ASC Shape (λ_{shape})	-0.5
ATES Shape (λ_{shape})	10^{-6}
Minimum shape parameter $\left(\lambda_{\text{shape}}^{\min}\right)$	-5
Maximum shape parameter $\left(\lambda_{\text{shape}}^{\max}\right)$	-0.5
ATES step size (η)	0.10
ASC step size (η)	0.01

Source: Created by the author.

4.5 Performance Analysis

In this section, the performance of the TRMEC algorithm, proposed in Section 4.3, is evaluated by comparing it to the RMEC, described in Section 3.4, and two other benchmark algorithms, namely, ATES and ASC. All the comparisons performed in this section are presented in terms of satisfaction rate and system throughput. Moreover, all the QoE measurements adopted in this section are given in terms of MOS, as in previous chapter. The relationship between MOS and rate is given by (3.3) and (3.17).

The simulations performed in this section considered a BS located on the center of a tri-sectored hexagonal cell, as described in Section 2.1, with a 500 m radius. The BS transmits at a central frequency of 3.5 GHz with a bandwidth of 20 MHz, which is equivalent to $K = 100$ RBs in the LTE standard, characterized in the same way as in Section 3.5. The UEs are deployed uniformly over the sector of the hexagonal cell and the channel between the BS and the UEs are modeled using the one ring channel model for 3D scenarios, described in Section 2.3, with azimuth and zenith angular spreads equal to $\phi_{sp} = 65°$ and $\theta_{sp} = 9°$, respectively. When not specified, it is considered that the BS is equipped with 4 antennas disposed as an Uniform Linear Array (ULA), parallel to the ground. A summary of the system parameters is presented in Table 4.2.

In Fig. 4.2, the satisfaction and the system throughput are presented in terms of the number of UEs demanding resources of the BS. In this analysis, it is considered that all UEs subscribe the same service, which requires that all UEs should be satisfied, i.e., $\xi_1 = 100\% \cdot U$. The RRA algorithms are compared considering three different minimum MOS targets, namely, $\Omega_1^{\text{target}} = 3.6$, 4 and 4.4.

In Fig. 4.2a, for a minimum required MOS equal to 3.6, the proposed heuristic achieved a satisfaction rate equal to $100\% \cdot U$ for all simulated loads, outperforming the benchmark algorithms. On the other hand, the RMEC algorithm attains a satisfaction rate of $100\% \cdot U$ for $U \leq 100$, which is the number of available RBs. However, for higher loads, its satisfaction rate decreases. Since RMEC works as a snapshot-based algorithm, it is expected that the number of UEs that RMEC is capable of satisfying is at most equal to the number of available RBs, which implies in one RB assigned to each UE. Regarding the benchmark algorithms, the ASC reaches a satisfaction rate considerably higher than ATES, satisfying all UEs while $U \leq 120$ UEs. Meanwhile, ATES achieves a satisfaction rate of $100\% \cdot U$ for at most a load of $U = 70$ UEs. For the highest simulated load, i.e., $U = 150$ UEs, the difference between ASC and the proposed heuristic is of 10.41%. Regarding the overall system throughput, presented in Fig. 4.2b, observe that besides satisfying all UEs in all simulated loads, TRMEC also presents the highest throughput results, regardless of the number of UEs served by the BS. In its turn, the RMEC algorithm yields an overall system throughput close to the one achieved by TRMEC. Nevertheless, the gap between the throughput achieved by TRMEC and RMEC increases when the number of UEs becomes larger. Indeed, for $U = 30$ UEs, the overall system throughput achieved by RMEC is 1.68% lower than the one achieved by the proposed algorithm. This gap increases up to 7.35%, for $U = 150$ UEs. Notice that, although small, the TRMEC presents a throughput gain over RMEC even for a low number of UEs served by the BS. This can be explained by the fact that the heuristic proposed in this chapter, differently of RMEC, considers the UEs' KPIs to convert the minimum rate requirement that must be achieved by the UEs during the entire

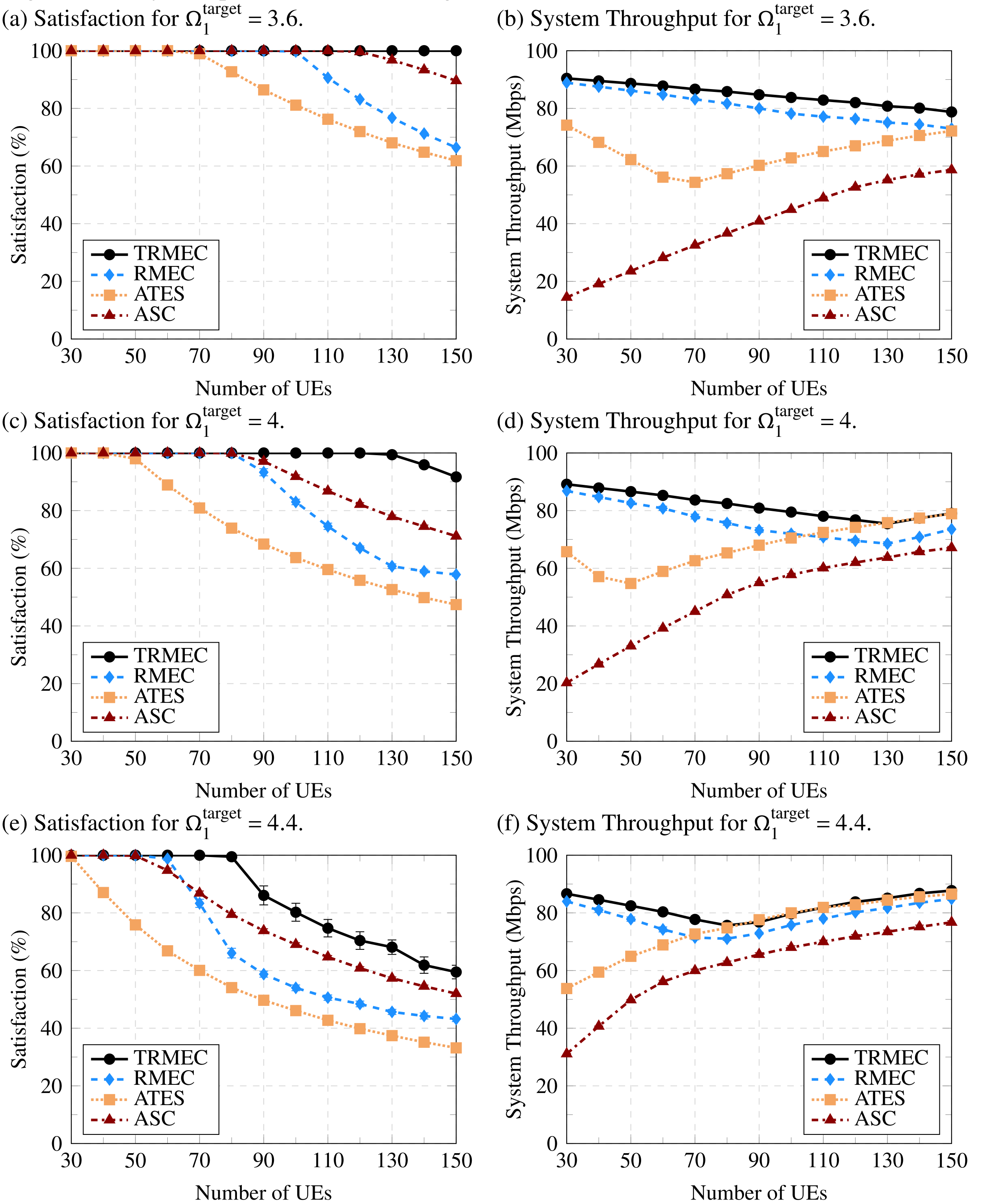
Table 4.2 – Simulation parameters.

Parameter	Value
Maximum BS transmit power (P_{total})	49 dBm [65]
BS antenna type	ULA with 4 elements (unless specified otherwise)
BS antenna element radiation pattern	3GPP 3D [43]
Cell radius	500 m
UE speed	3 km/h [39]
Carrier frequency	3.5 GHz [39]
System bandwidth	20 MHz [65]
Subcarrier bandwidth (Δf)	15 kHz
Number of RBs (K)	100
Number of subcarriers per RB (Q_{sub})	12
Number of symbols per RB (Q_{sym})	14
Path loss	$34.5+35\log_{10}(d_{b,u})$ [66]
Log-normal shadowing standard deviation	8 dB [39]
Small-scale fading	One ring model for 3D scenarios
Azimuth spread (ϕ_{sp})	65°
Zenith spread (θ_{sp})	9°
AWGN power per sub-carrier	-123.24 dBm
Noise figure	9 dB
Link adaptation	Link level curves from [53]
Traffic model	Full buffer
Transmission Time Interval	1 ms
Simulation duration	1 s
Number of simulations	50
Confidence interval	95%

Source: Created by the author.

session into instantaneous requirements. Therefore, TRMEC makes better use of the RBs over time, achieving a higher system throughput. Regarding the ASC algorithm, it is noteworthy that its overall system throughput increases along with the number of UEs. In fact, the overall system throughput of the proposed heuristic is 6.26 times greater than the one achieved by ASC, for $U = 30$ UEs. On the other hand, for $U = 150$ UEs, the throughput achieved by the TRMEC algorithm is 34% higher than ASC's. As mentioned in Section 4.4 and detailed in [70], the ASC algorithm aims at reaching the satisfaction target, ξ_1, in this case $\xi_1 = 100\% \cdot U$, regardless of the system throughput. Therefore, when the amount of available resources is much greater than the

Figure 4.2 – System performance for a single service scenario with $\xi_1 = 100\%$ of U.



Source: Created by the author.

necessary to meet requirements the UEs, the ASC algorithm does not avail to improve the system rate, in opposite to TRMEC. Instead, after meeting the UEs' requirements, the ASC algorithm distributes the spare RBs to the UEs without aiming at the throughput maximization. The ATES algorithm in turn aims at a trade-off between overall system throughput and satisfaction rate, instead of prioritizing the satisfaction rate over the throughput, as the other RRA algorithms. Indeed, notice that the throughput of ATES algorithm decreases until $U = 70$ UEs, which is the maximum number of UEs where the algorithm was capable of satisfying all the UEs served by the BS. For $U > 70$, the overall system throughput increases, highlighting the trade-off between satisfaction and overall system throughput. Besides that, it is important to emphasize that the TRMEC yields a better trade-off than the ATES algorithm, since it achieves higher satisfaction rates as well as higher throughput. In fact, besides satisfying all UEs in all analyzed cases, the proposed heuristic presents a gain of 38.14% in terms of satisfaction rate, when compared to ATES algorithm, for $U = 150$ UEs. Moreover, it also yields an overall system throughput at least 10% higher than the ATES's.

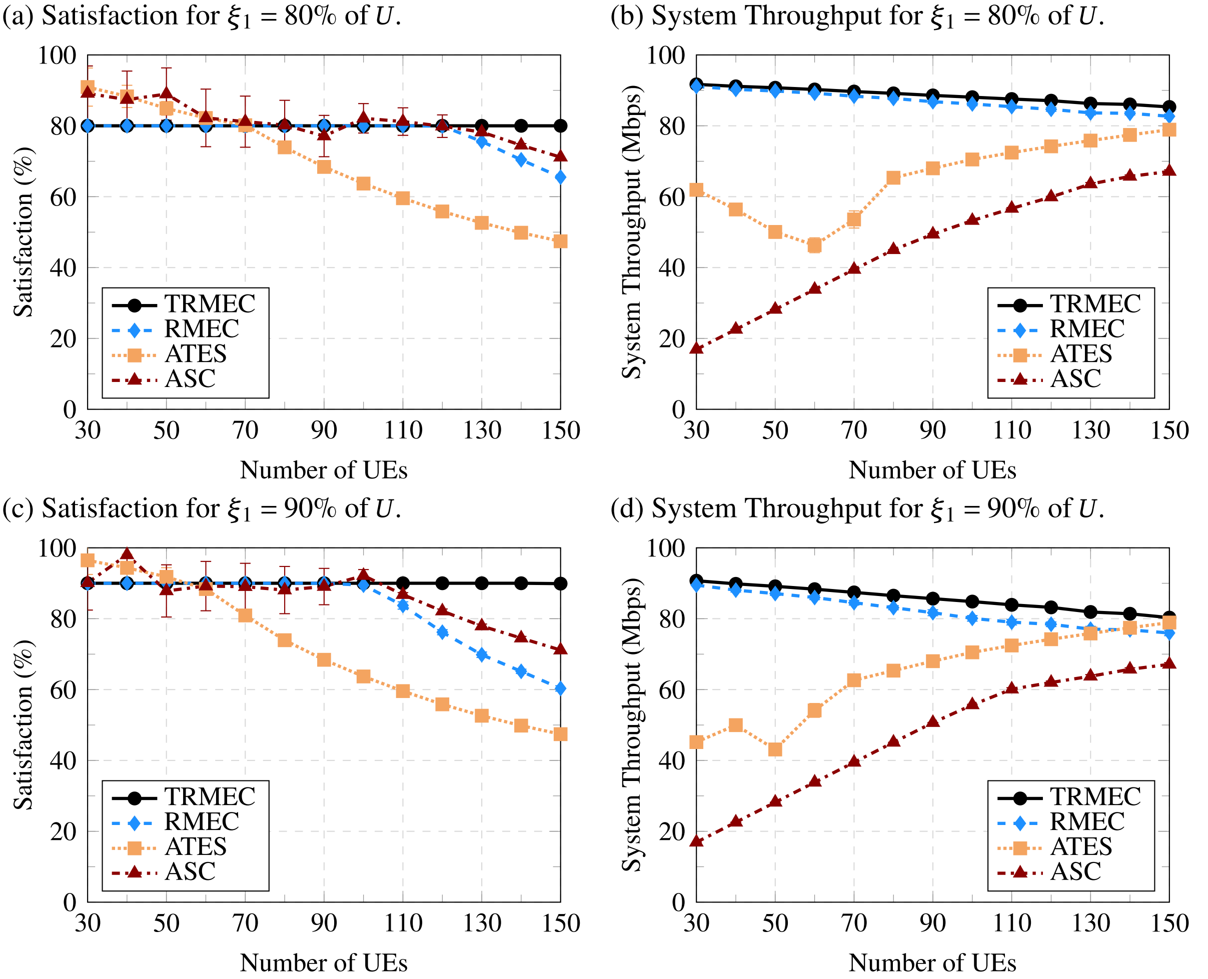
In Figs. 4.2c and 4.2d, the satisfaction rate and the overall system throughput considering a minimum MOS value equal to 4 are depicted. In this analysis, the proposed heuristic was capable of satisfying all UEs until $U = 130$. Additionally, due to the near feasibility characteristic inherited from RMEC, the proposed heuristic tries to ensure a high satisfaction rate. In fact, for $U = 150$, the proposed heuristic is capable of satisfying more than 90% of the UEs. Moreover, observe that when the TRMEC algorithm is not capable of satisfying all the UEs, for $U > 130$, it compensates by increasing the system throughput, providing a good trade-off between satisfaction and system rate. The RMEC algorithm, in its turn, is capable of satisfying all UEs, while $U \leq 80$. Comparing this result with the case where $\Omega_1^{\text{target}} = 3.6$, here, the maximum load for which the RMEC algorithm ensures the satisfaction target $\xi_1 = 100\%$ of the UEs is not limited by the number of available RBs. In this case, the better performance of the proposed algorithm, compared to RMEC, is mainly due to the rate requirement adaptation over time, presented in (4.3). It enables the proposed algorithm to schedule the RBs more properly, achieving higher satisfaction rates, besides a higher throughput. Furthermore, observe that for $U \geq 130$ UEs, the satisfaction rate of RMEC decreases more slowly, furthermore, its overall system rate increases. This happens due to the high UE diversity, i.e., the RMEC heuristic is satisfying the UEs with better channel conditions, which requires fewer RBs. Although RMEC satisfies all UEs at a load 38.46% lower than the proposed heuristic, its achieved throughput is at most 9.43% lower than TRMEC's. As for the ASC algorithm, in this analysis, as RMEC, it also ensures a satisfaction rate of $100\% \cdot U$ until $U = 80$ UEs. However, for $U > 80$, the ASC algorithm is capable of satisfying more UEs than the RMEC heuristic. When compared with the algorithm proposed in this chapter, the ASC is considerably outperformed. In fact, for $U = 80$, both algorithms satisfy all UEs, however, the proposed heuristic reaches a throughput 62.37% higher than ASC's. Moreover, for $U = 130$, the ASC algorithm satisfies 21.48% less UEs than the TRMEC heuristic, in addition of reaching a throughput 15.51% lower. The worst results presented in this analyses are obtained by the ATES algorithm, which is capable of satisfying all UEs for at most $U = 50$ UEs. Additionally, for $U = 50$ UEs, the proposed algorithm reaches a throughput 58.21% higher than the one achieved by the ATES heuristic. Comparing the performance of the ATES algorithm against the proposed heuristic for $U = 130$ UEs, the satisfaction gap is equal to 46.78%. However, due to the trade-off between system rate and satisfaction, ATES achieves an overall throughput similar to TRMEC. It is important to highlight that the ATES is able of satisfying less UEs than RMEC, even for a large number of UEs. As a matter of fact, for $U = 150$, the RMEC algorithm is capable of satisfying 10.41% more UEs than the ATES heuristic, even though it is a snapshot-based algorithm.

Finally, in Figs. 4.2e and 4.2f, the satisfaction rate and the overall system throughput considering a minimum MOS requirement equal to 4.4 are depicted, respectively. Due to the high MOS requirement demanded by the UEs, the maximum number of UEs that each algorithm is capable of satisfying decreases. However, observe that the heuristic proposed in this chapter outperforms all the benchmark algorithms in terms of satisfaction rate. Moreover, it achieves the highest throughput values in most of the simulated loads. Here, the proposed algorithm is capable of satisfying all UEs while $U \leq 80$ UEs, which is 33.33% more UEs than RMEC. As in the previous analyses, the satisfaction rate of RMEC is not limited by the number of RBs, but because it does not consider the results of previous allocations on each TTI. On the other hand, notice that the RMEC heuristic is capable of meeting the satisfaction requirement of $\xi_1 = 100\% \cdot U$ for a larger number of UEs than the ASC algorithm. However, the satisfaction rate provided by the ASC algorithm decreases more slowly than the RMEC's. It happens because the ASC algorithm, on contrary of RMEC, considers the current KPIs of the UEs on each allocation. Regarding the ATES algorithm, once more it presents the worst result in terms of satisfaction rate. In fact, the proposed algorithm can satisfy all UEs for a load 2.66 times higher than ATES. On the other hand, notice that the overall system throughput provided by the ATES algorithm is higher than the one achieved by ASC. Additionally, for $U \geq 80$, the throughput reached by ATES is almost the same as the one achieved by the proposed heuristic. However, its satisfaction rate is considerably lower. Indeed, for $U = 80$ UEs, the satisfaction rate of the ATES algorithm is around 54.6%, while the proposed heuristic satisfies all UEs.

In the next analysis, depicted in Fig. 4.3, the algorithms are evaluated varying the satisfaction target ξ_1. Here, it is considered a scenario where all UEs subscribe the same service, requiring a minimum MOS $\Omega_1^{\text{target}} = 4$.

In Fig. 4.3a, the satisfaction rate of the algorithms are presented considering a target $\xi_1 = 80\% \cdot U$. Notice that both TRMEC and RMEC are capable of locking into the satisfaction target of $\xi_1 = 80\% \cdot U$. However, the heuristic proposed in this chapter is capable of meeting this constraint for all simulated loads, in opposite to RMEC, which ensures a satisfying $80\% \cdot U$ for at most $U = 120$ UEs. In fact, as already explained in the previous analyses, due to the limitation of the number of RBs, RMEC is capable of satisfying at most 100 UEs. Therefore, the RMEC algorithm is able to satisfy $\xi_1 = 80\% \cdot U$ when the number of UEs served by the BS is at most $U = 125$ UEs. Regarding the utility-based algorithms, the ASC algorithm meets the satisfaction target ξ_1 roughly up to 130 UEs. Meanwhile, the ATES algorithm ensures satisfying $\xi_1 = 80\% \cdot U$ for $U = 70$ UEs. Differently from the heuristics proposed in this thesis until here, namely TRMEC and RMEC, the ASC and ATES algorithms do not lock into the satisfaction target. Indeed, the satisfaction rate of ATES is a strictly descending curve, which crosses the target ξ_1 when $U = 70$ UEs. Furthermore, ASC satisfaction curve considerably varies around the satisfaction target $\xi_1 = 80\% \cdot U$, until it starts to strictly descend at $U = 130$ UEs. This fluctuation

Figure 4.3 – System performance for a single service scenario with a minimum MOS $\Omega_1^{\text{target}} = 4$.



Source: Created by the author.

can be observed in Fig. 4.3a, by looking at the confidence interval range. Comparing these satisfaction results to the overall system throughput depicted in Fig. 4.3b, as already observed in the previous analyses, besides of the proposed heuristic meeting the satisfaction target ξ_1 more often, it also yields the highest values of throughput among the analyzed algorithms. The RMEC algorithm also presents high values of system rate, which are at most 3.08% below than the throughput achieved by TRMEC. Comparing these results with those presented in Fig. 4.2d, for $\xi_1 = 100\% \cdot U$, observe that the overall system throughputs achieved by TRMEC and RMEC are higher when $\xi_1 = 80\% \cdot U$. Since the algorithms are required to satisfy less UEs than the total served by the BS, the TRMEC and RMEC algorithms take advantage of the UE diversity satisfying those UEs with the best channel conditions, hence achieving higher throughputs. In fact, for $\xi_1 = 80\% \cdot U$, the TRMEC reaches throughput values between 2.86% and 13.42% higher than the case where all UEs must be satisfied. For the RMEC heuristic, this gap is even higher, achieving a gain of at most 22.05% for $\xi_1 = 80\% \cdot U$ over the scenario where the satisfaction target is $\xi_1 = 100\% \cdot U$. On the other hand, ASC and ATES do not provide the same trade-off between satisfaction and throughput. Indeed, both utility-based benchmark algorithms present a loss in terms of system throughput when compared to the case where $\xi_1 = 100\% \cdot U$. As already mentioned, the ASC algorithm does not aim at maximizing the system rate, but it prioritizes to reach a satisfaction rate as close as it seems possible to the desired target ξ_1. The loss of system rate for the case where $\xi_1 = 80\% \cdot U$ when compared to the throughput registered in Fig. 4.2d, for $\xi_1 = 100\% \cdot U$, is around 16.6% for $U = 30$ UEs. This loss diminishes when the number of UEs served by the BS becomes larger. Besides, for $U \geq 130$, the throughput achieved by the ASC algorithm is the same for $\xi_1 = 80\%$ and 100% of the UEs. Regarding the ATES algorithm, although its goal is to provide a trade-off between satisfaction and system throughput, when the desired satisfaction target $\xi_1 = 80\% \cdot U$, this algorithm presents a loss in terms of throughput which reaches up to 21.41% when compared to the case where the system intends to satisfy all UEs. As already explained in previous analyses, the ATES algorithm presents a trade-off between satisfaction and throughput that can be perceived with the increasing number of UEs. However, when the number of UEs required to get satisfied, ξ_1, decreases, the ATES algorithm does not take advantage of this fact to increase the overall system throughput.

The results presented in Figs. 4.3c and 4.3d depict the satisfaction rate and the overall system throughput for a satisfaction target $\xi_1 = 90\% \cdot U$. These results corroborate the conclusions already obtained, regarding Figs. 4.3a and 4.3b. Comparing the results varying the satisfaction target to $\xi_1 = 80\%$, 90% and 100% of the UEs, it is possible to infer that the algorithms proposed in this thesis, namely TRMEC and RMEC, are considerably more robust than the utility-based benchmark algorithms. Moreover, in all analyses performed so far, TRMEC provided the best results in terms of satisfaction and throughput.

Until now, the analyses have shown the better robustness of the proposed heuristic in terms of minimum MOS requirement, Ω_1^{target}, minimum number of UEs required to be satisfied, ξ_1, and the total number of UEs served by the BS. However, all these analyses have considered that the UEs are deployed in a scenario under the same channel conditions, i.e., the statistical distribution of the UEs' SNR is the same for all analyses. Recall that for all the analyses of this chapter, it is considered that the antenna array is configured to provide a spatial diversity gain. It means that only one data stream is transmitted per RB, implying an SNR increase. Therefore, one way to improve the link quality is to increase the number of antenna elements.

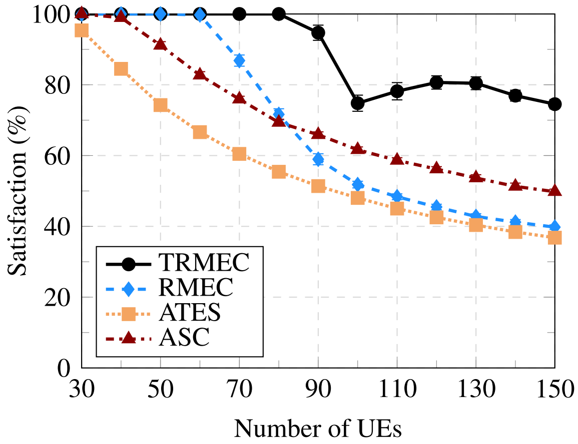
In order to evaluate the robustness of the proposed algorithm with respect to the link quality between the BS and the UEs, in the next analyses, depicted in Fig. 4.4, it is considered that the BS is equipped with two different number of antennas elements, namely 1 and 16. It is also considered that all UEs subscribe the same service which requires that all UEs be satisfied with a minimum MOS requirement equal to 4, i.e., $\Omega_1^{\text{target}} = 4$ and $\xi_1 = 100\% \cdot U$.

In Figs. 4.4a and 4.4b, the BS is equipped with a single antenna element, while the previous analyses considered a BS with an ULA with 4 antenna elements. Hence, since the number of antennas decreases, the SNR of the link between the BS and the UEs becomes worse. In spite of this, observe that the RRA algorithm proposed in this chapter considerably outperforms the other algorithms.

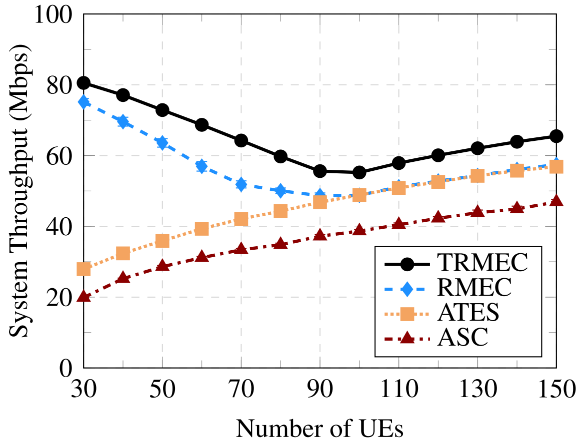
Notice that the satisfaction curve of the proposed algorithm drastically decreases

Figure 4.4 – System performance for a single service scenario with a minimum MOS $\Omega_1^{\text{target}} = 4$ and $\xi_1 = 100\%$ of U.

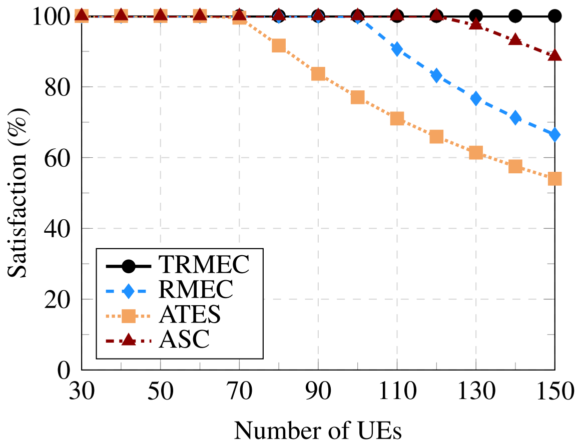
(a) Satisfaction considering a BS equipped with a single antenna.



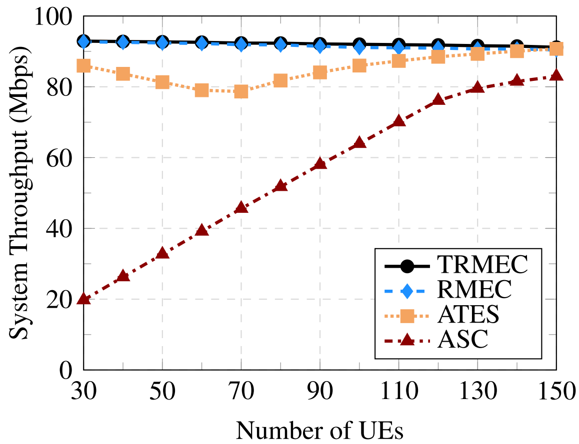
(b) System Throughput considering a BS equipped with a single antenna.



(c) Satisfaction considering a BS equipped with a 4×4 URA.



(d) System Throughput considering a BS equipped with a 4×4 URA.



Source: Created by the author.

from 94.7%, for $U = 90$, to 74.79%, for $U = 100$ UEs. Furthermore, the satisfaction rate increases again up to 80.65%, for $U = 120$ UEs, and only then it decreases normally as expected. The sudden decrease of the satisfaction rate of the proposed algorithm happens due to cases where the LP generated by the relaxation of the binary constraint (4.6d) has a feasible solution, but problem (4.6) does not. In these situations, the algorithm seeks for a feasible solution that does not exist and ends up satisfying fewer UEs than it could satisfy if any more UEs were disregarded before the LP relaxation. On the other hand, since it is not possible to satisfy more UEs than the number of available RBs in each TTI, the proposed algorithm limits the number of UEs that can compete for resources by at most the number of available RBs, as explained in Section 4.3. Therefore, for $U \geq K$, the proposed algorithm takes advantage of the UE diversity, mitigating the probability of the LP returning a "false feasible" fractional assignment. Indeed, for $100 \leq U \leq 120$ UEs, the satisfaction rate increases, yielding an inflection point at $U = 100$ UEs. This satisfaction increase shows that the proposed algorithm would be able to present a higher satisfaction rate between 90 and 120 UEs if the "false feasible" fractional assignment yielded by the LP in the initial assignment step did not happen. The effects of the "false feasible" fractional assignment are also observed in the overall system throughput. Notice that the satisfaction rate starts decreasing for $U > 80$, however the throughput yielded by the proposed algorithm increases only for $U > 100$. It means that for $80 < U \leq 100$ UEs the TRMEC does not present the usual trade-off between satisfaction and throughput observed in the previous analyses. Nevertheless, one way to mitigate the effects of the "false feasible" fractional assignment is to verify at the end of the algorithm if the solution met the constraint of satisfying all UEs selected in the initial steps of the algorithm, i.e., all UEs where $\rho'_u[t] = 1$ at the TTI t. If not, the UE u with the lowest rates and highest instantaneous rate requirement would be disregarded, i.e., $\rho'_u[t] = 0$, and then return to the initial assignment step of the algorithm. This workaround would solve the effect of the "false feasible" fractional assignment, but it would increase the complexity of the proposed algorithm. Since this effect appears only in some corner situations, as the one presented in the analysis of Figs. 4.4a and 4.4b, it may not justify the additional complexity. Moreover, even with this drawback, the proposed algorithm still achieves results regarding the satisfaction rate and overall system throughput considerably better than the benchmark algorithms. Indeed, in this scenario, the proposed heuristic is capable of satisfying all UEs for at most $U = 80$ UEs, which corresponds to 33.33% and 100% more UEs than RMEC and ASC, respectively. The ATES algorithm in turn was not capable of satisfying all UEs for any of the number of UEs considered in the analyses. Comparing these analyses with those presented in Figs. 4.2c and 4.2d, it is possible to observe that when a single antenna is considered in the BS, the throughput gap between RMEC and TRMEC increases. In fact, for a single antenna BS, the system throughput presented by the RMEC algorithm is up to 19.3% lower than that of TRMEC. On the other hand, for the case where the BS is equipped with 4 antennas, the throughput loss of the RMEC algorithm with respect to TRMEC is at most 9.47%. It means that the algorithm proposed in this chapter is capable of providing better results even in scenarios with low SNR values. This fact ratifies the importance of the initial steps of the TRMEC where the minimum rate requirements are dynamically adapted on each TTI.

The results presented in Figs. 4.4a and 4.4b consider that the BS is equipped with 16 antenna elements disposed as a 4×4 Uniform Rectangular Array (URA). In this analysis, the performance of the RRA algorithms is evaluated considering links with higher quality between the BS and UEs. The proposed algorithm was capable of satisfying all UEs for all simulated loads. Moreover, due to the high quality links, the overall system throughput achieved by TRMEC is at most 2.29% smaller than the system maximum capacity, which is 93.3 Mbps. Although the throughput results reached by the RMEC algorithm are similar to the ones achieved by TRMEC, the RMEC heuristic is not capable of satisfying all the UEs for $U > 100$, i.e., satisfy more UEs than the number of available RBs, regardless of the link quality. Regarding the utility-based algorithms, the ASC and ATES methods achieved a satisfaction rate of $100\% \cdot U$ for $U = 120$ and $U = 70$ UEs, respectively. When compared to the cases where the BS is equipped with 1 and 4 antennas, both utility-based algorithms presented better results with the improvement of the link quality. However, only ATES was able to reach a higher throughput, close to the one achieved by the proposed algorithm. This behavior stresses the trade-off aimed by the ATES algorithm, between satisfaction and throughput.

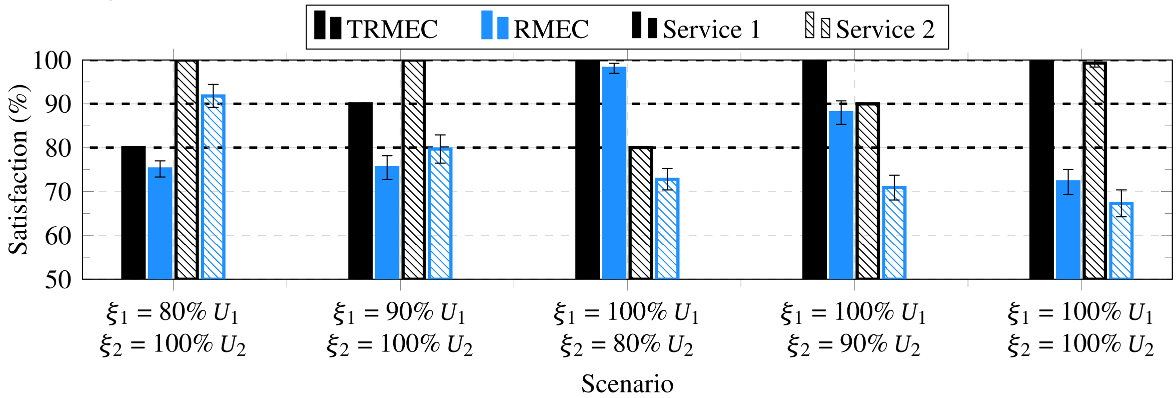
The ASC heuristic, in its turn, did not take advantage of the better link quality to improve system throughput. In fact, comparing the results presented in Figs. 4.2c and 4.4b, observe that for $U \leq 80$ UEs, the overall system throughput achieved by the ASC algorithm is roughly the same, regardless of the number of antennas considered. As already explained, the ASC algorithm aims at reaching the desired satisfaction target, in this analysis $\xi_1 = 100\% \cdot U$. When this goal is accomplished, the ASC algorithm distributes the spare RBs to the UEs without aiming at the throughput maximization.

All the results presented until this point show that the algorithm proposed in this chapter substantially outperforms all the benchmark algorithms. However, the analyses heretofore consider that all UEs served by the BS subscribe to a single service. Therefore, in order to complete the benchmarking of the algorithm proposed in this chapter, exploiting all parameters of the heuristic proposal, in the following analysis, presented in Fig. 4.5, the UEs are divided into two different services. Similarly to the multi-service analysis performed in Section 3.5, the two service plans consist of a high-quality and a high definition skype video calls, which recommend a minimum throughput of 500 kbps and 1.5 Mbps [67], respectively. It was considered that the BS equipped with a ULA with 4 antenna elements serves $U = 60$ UEs. From these UEs, it is considered that the 40 UEs are demanding a high-quality skype video call and the rest of the 20 UEs are using a high definition skype video call. In other words, $U_1 = 40$, $U_2 = 20$, $\Omega_1^{\text{target}} = 500$ kbps and $\Omega_2^{\text{target}} = 1.5$ Mbps. The analyses presented in Fig. 4.5 consider 5 different pairs of minimum number of UEs that should be satisfied per service, i.e., ξ_1 and ξ_2. Moreover, the algorithm proposed in this chapter is compared solely against the RMEC heuristic. The results considering the utility-based algorithms were not simulated because, as already mentioned in Section 4.4, these algorithms were not designed to support different satisfaction targets per service.

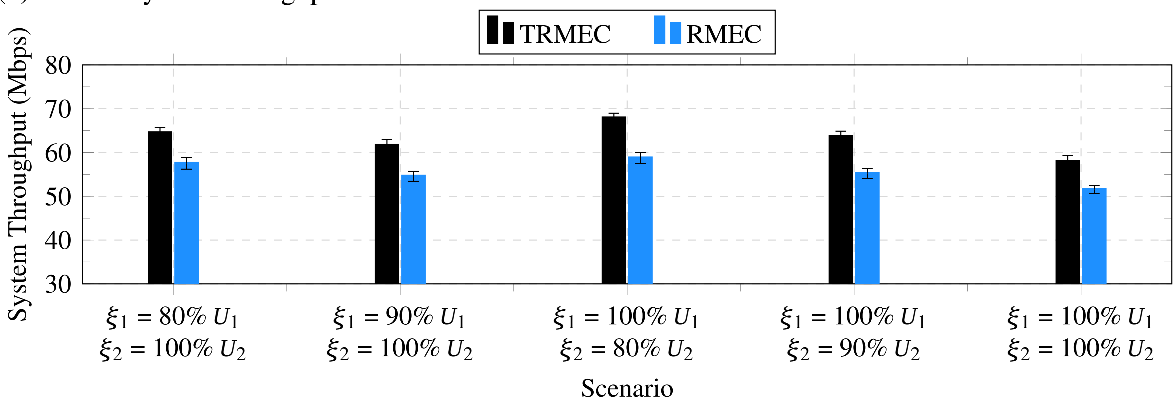
The first analyzed scenario considers that the minimum number of UEs that should be satisfied for each service are $\xi_1 = 80\% \cdot U_1 = 32$ and $\xi_2 = 100\% \cdot U_2 = 20$ UEs. Here, the algorithm proposed in this chapter was capable of satisfying the minimum number of UEs required by each service, as depicted in Fig. 4.5a. On the other hand, the RMEC algorithm presents a satisfaction rate of 75.15% and 91.8% of the UEs subscribing the services 1 and 2, respectively. In this scenario, the minimum necessary throughput required to meet both services constraints is $\xi_1\Omega_1^{\text{target}} + \xi_2\Omega_2^{\text{target}} = 80\% \cdot 40 \cdot 500 \text{ kbps} + 100\% \cdot 20 \cdot 1.5 \text{ Mbps} = 46$ Mbps. From Fig. 4.5b, observe that the throughput achieved by RMEC in this scenario is equal to 57.52 Mbps, which is a rate 25.05% higher than the required one. Therefore, comparing the results of satisfaction and throughput, it is possible to infer that the RMEC was not capable of properly distributing the RBs to the UEs. As already explained in previous analyses, since the RMEC algorithm was designed as an snapshot-based heuristic, it does not consider the current KPIs of

Figure 4.5 – System performance considering $U = 60$ UEs and $S = 2$ service plans, where $U_1 = 40$ and $U_2 = 20$ UEs, $\Omega_1^{\text{target}} = 500$ kbps and $\Omega_2^{\text{target}} = 1.5$ Mbps.

(a) Average satisfaction rate.



(b) Overall system throughput.



Source: Created by the author.

the UEs in the scheduling process. Therefore, for RMEC, all UEs require to transmit the same amount of data in each TTI. In its turn, the TRMEC algorithm meets the minimum required satisfaction target of both services, in addition of being capable of reaching a throughput 12.08% higher than RMEC. This result reinforces the relevance of the selection of the UEs that will compete for resources, as well as the rate requirement adaptation of the algorithm proposed in Section 4.3.

In the second scenario, it is considered that the at least $\xi_1 = 90\% \cdot U_1 = 36$ UEs should be satisfied in service 1 and $\xi_2 = 100\% \cdot U_2 = 20$ UEs in service 2. Once more, the TRMEC algorithm met the requirements ξ_1 and ξ_2 imposed by each service. Regarding the overall system throughput, the proposed algorithm achieved a system rate of 61.63 Mbps, which is 4.4% smaller compared to the first scenario. Although the number of UEs is the same in both scenarios, the minimum number of UEs that should be satisfied is greater in the second scenario. Therefore, in order to satisfy more UEs, the TRMEC algorithm has given up a higher throughput. On the other hand, the RMEC algorithm achieves a satisfaction rate equal to 75.45% in service 1 and 79.7% in service 2. Comparing these results with the ones achieved in the first scenario, observe that the satisfaction rate of the service 1 barely changed, while for service 2, the percentage of satisfied UEs is considerably lower. Moreover, the achieved throughput is 5.14% lower than the one reached in the first scenario. In other words, the RMEC algorithm did not present the usual trade-off between satisfaction rate and throughput. This fact can be explained by the problem of the "false feasible" fractional assignment, already explained in the previous analyses in the context of TRMEC. This drawback of the proposed algorithm was in fact inherited from the RMEC heuristic. In the case of RMEC, the initial fractional assignment comes from the LP presented in (3.15). In the context of the RMEC algorithm, the "false feasible" fractional assignment problem appears when the LP (3.15) has a feasible solution, but the original optimization problem (3.1) does not. Therefore, the "false feasible" fractional assignment may lead the algorithm to satisfy less UEs than it could usually satisfy if the "false feasible" fractional assignment yielded by the LP in the initial assignment step did not happen.

The third scenario considers that the minimum number of UEs that should be satisfied for each service is $\xi_1 = 100\% \cdot U_1 = 40$ and $\xi_2 = 80\% \cdot U_2 = 16$ UEs. Likewise in the two previous scenarios, the TRMEC algorithm reaches the minimum satisfaction target ξ_1 and ξ_2 of both service plans. In this scenario, the number of UEs that shall meet their requirements is greater than in the first scenario. In fact, here, at least 56 out of 60 UEs are required to be satisfied, against the first scenario, which requires to satisfy at least 52 out of 60 UEs. Nevertheless, in this scenario, both algorithms present better results. The throughput achieved by the TRMEC algorithm is 67.87 Mbps, which is 5.27% higher than the system rate achieved in the first scenario. Meanwhile, the RMEC algorithm reaches a throughput 2.1% higher than the first scenario, namely 58.73 Mbps. This better performance of the algorithms can be explained by the minimum necessary throughput required jointly by the services, which in this scenario is $\xi_1\Omega_1^{\text{target}} + \xi_2\Omega_2^{\text{target}} = 100\% \cdot 40 \cdot 500\text{ kbps} + 80\% \cdot 20 \cdot 1.5\text{ Mbps} = 44\text{ Mbps}$, while in the first scenario it is 46 Mbps. Regarding the RMEC algorithm, it achieves a satisfaction rate equal to 98.1% and 72.8% in services 1 and 2, respectively.

Differently from the previous analyzed scenarios, here, the RMEC heuristic almost met the minimum number of satisfied UEs in service 1. In the first scenario, all UEs subscribing the service 2 are required to get satisfied, meanwhile only 80% of the UEs subscribing service 1 shall meet their requirements. Here, the percentage of UEs that should be satisfied by each service plan is inverted, i.e., all UEs from service 1 and only 80% of the UEs shall be satisfied. Although the number of UEs that can be disregarded in the first scenario is higher than in this scenario, the UEs subscribing service 2 are harder to satisfy than those subscribing service 1. In fact the UEs that subscribe the service 1 require a minimum rate of 500 kbps, while those subscribing the service 2 demand a minimum throughput of 1.5 Mbps. Therefore, the UE diversity is better exploited in this scenario, since UEs with worst channel conditions subscribing to the more resource demanding service can be disregarded. This fact also explains the better performance of both algorithms in this scenario when compared with the first one.

In the fourth scenario, at least $\xi_1 = 100\% \cdot U_1 = 40$ UEs should be satisfied in service 1 and $\xi_2 = 90\% \cdot U_2 = 18$ UEs in service 2. Notice that the proposed algorithm achieves the minimum satisfaction target ξ_1 and ξ_2 of both services. Moreover, it also reaches an overall system throughput equal to 63.60 Mbps, which is 6.29% less than the throughput achieved in the third scenario. This loss in terms of system throughput is explained by the trade-off between satisfaction and overall system rate pursued by TRMEC. Thus, since in this scenario there are two additional UEs required to meet their requirements, the UE diversity decreases, making the throughput loss perfectly understandable. On the other hand, observe that with the increasing number of UEs that should be satisfied subscribing the service plan 1, RMEC algorithm yields a satisfaction rate lower than the one presented in the third scenario for both services. Moreover, the throughput achieved by the RMEC algorithm in this scenario is 6.05% lower than its system rate in the third scenario. This means that, unlike the TRMEC algorithm, in this scenario, the RMEC heuristic does not present a trade-off between satisfaction and throughput. This fact can be justified by the “false feasible” fractional assignment problem, already explained in previous analyses.

Finally, in the last scenario, all the 60 UEs served by the BS should be satisfied, which means that there is no UE diversity, i.e., no UE can be disregarded by the RRA scheduler. Differently from the previous scenarios, here, the proposed algorithm was not capable of satisfying all UEs. However, it still provides a high satisfaction rate equal to 99.55% and 99.3% for services 1 and 2, respectively. Besides, the TRMEC algorithm presents a high system throughput equal to 57.92 Mbps, which is 12.34% higher than RMEC’s. On the other hand, the RMEC algorithm achieves a satisfaction rate of 72.2% and 67.3% for services 1 and 2, respectively. The poor performance of RMEC is mainly due to the fact that it is a snapshot-based algorithm. Additionally, the “false feasible” fractional assignment problem explained in previous analyses.

4.6 Chapter Summary

In this chapter, the problem of scheduling users aiming at maximizing the overall system rate is studied. It is considered that at least a certain fraction of the users should be satisfied in a multi-service scenario.

Since the optimal solution requires a high computational effort, a suboptimal algorithm with low complexity is provided here, namely TRMEC. The proposed heuristic is an extension of the low complexity algorithm proposed in Chapter 3. However, it considers previous UEs’ information to improve the scheduling on each TTI.

The computational simulations presented in Section 4.5 show that the proposed heuristic outperforms the benchmark algorithms, namely ASC and ATES, meeting the satisfaction rate constraint for higher loads and harder QoS/QoE requirements. Additionally, the proposed solution overcomes the RMEC limitation of only satisfying a number of users equal to or lower than the number of RBs. Besides that, the TRMEC algorithm presented a high scalability regarding all studied parameters, which are: number of UEs in the system, UEs minimum QoS/QoE requirement, link quality and multi-service requirements. Furthermore, besides the higher robustness and scalability, when the algorithm proposed in this chapter can not satisfy the minimum number of UEs required by the service plans served by the BS, it aims at satisfying as many UEs as possible, providing a near-feasible solution. This characteristic was inherited from its predecessor, the RMEC algorithm.

It was also observed that the proposed algorithm inherits a drawback from RMEC, which was referred in Section 4.5 as the "false feasible" fractional assignment problem. However, this issue only happens in some corner situations, which discourage a workaround to this problem that will significantly increases the heuristic complexity. Moreover, even in results where this problem was evident, the proposed algorithm still achieved a better performance than the benchmark algorithms, including the RMEC heuristic.

Although the algorithm proposed in this chapter is significantly more complex than the benchmark algorithms, the performance gain presented over the benchmark algorithms is quite expressive for all analyzed scenarios.

5 POWER AND RESOURCE MANAGEMENT FOR RATE MAXIMIZATION WITH QOS/QOE PROVISIONING IN WIRELESS NETWORKS

In this chapter, the same problem addressed in Chapter 3 is revisited, which is to maximize the overall system rate, while ensuring that a minimum number of UEs of each service plan meet their QoS/QoE requirements. However, in Chapter 3, the RBs are allocated to the UEs considering that the power is divided equally among all RBs. Here, the power is also considered in the allocation process, i.e., RBs and power are jointly allocated.

The main contributions of this chapter are:

- Study of the problem of allocating RBs and power to the UEs aiming at maximizing the overall system rate in a multi-service scenario, considering that a fraction of the users of each service must have their QoE requirements met;
- Reformulation of this problem as an ILP and its solution using standard algorithms;
- Proposal of a low-complexity suboptimal solution that has near optimal performance and presents high scalability in terms of the size of the problem input.
- Proposal of an improvement over the state-of-the-art to deal with infeasible instances of the RRA and provide better results when the system is required to satisfy a number of UEs smaller than the total.

The rest of this chapter is divided as follows. In Sections 5.1 and 5.2, the problem addressed in this chapter is mathematically formulated as an optimization problem and it is rewritten as an ILP, which can be solved using standard numerical algorithms from the literature. In Section 5.3, a new low-complexity suboptimal algorithm is proposed to solve the problem stated in Section 5.1. In Section 5.4, the state-of-the-art suboptimal algorithm that solves the problem stated in 5.1 is described. In Section 5.5, a improved version of the state-of-the-art is proposed to overcome its inherent limitations. In Section 5.6, a performance analysis of the algorithm proposed in Section 5.3 against the optimal solution, the existing state-of-the-art heuristic and its improvement is performed. Finally, the main conclusions of this chapter are presented in Section 5.7.

5.1 Problem Formulation

In this section, the problem of jointly allocating the available RBs and power in order to maximize the overall system rate while ensuring that a minimum number ξ_s of UEs in service plan s meet their QoS/QoE requirements is described as an optimization problem. Another constraint of the problem is that the summation of the power allocated to all RBs can not exceed the maximum power available at the BS. Analogously to Chapter 3, this problem is solved in a single snapshot, i.e., $T = 1$ TTI. Therefore, in order to ease the notation, the TTI index will be omitted in the rest this chapter. It is important to highlight that this problem was already addressed in the literature by [26]. However, in [26], the authors considered only QoS requirements.

Similarly to Section 4.1, consider an assignment matrix $\mathbf{X} \in \{0, 1\}^{U \times K}$, where each element $x_{u,k}$ is equal to 1 if the RB k is allocated to the UE u and equal to 0 otherwise. In addition, consider a vector $\mathbf{p} \in \mathbb{R}_{+}^{K \times 1}$, where each element p_k is the power allocated to the RB k. The problem addressed in this chapter can be written as an optimization problem as follows

$$\max_{\mathbf{X},\mathbf{p}} \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} \mathfrak{R}_{u,k}(p_k)\, x_{u,k}, \tag{5.1a}$$

$$\text{s.t.} \sum_{k \in \mathcal{K}} p_k \leq P_{\text{total}}, \tag{5.1b}$$

$$\sum_{u \in \mathcal{U}} x_{u,k} \leq 1, \forall k \in \mathcal{K}, \tag{5.1c}$$

$$\sum_{u \in \mathcal{U}_s} H\left(\Omega_u\left(\sum_{k \in \mathcal{K}} \mathfrak{R}_{u,k}(p_k)\, x_{u,k}\right), \Omega_s^{\text{target}}\right) \geq \xi_s, \forall s \in \mathcal{S}, \tag{5.1d}$$

$$x_{u,k} \in \{0, 1\}, \forall u \in \mathcal{U} \text{ and } \forall k \in \mathcal{K}, \tag{5.1e}$$

$$p_k \geq 0, \forall k \in \mathcal{K}, \tag{5.1f}$$

where $\mathfrak{R}_{u,k}(p)$ denotes the rate achieved by the UE u in the RB k transmitting with a power p.

The problem stated in (5.1) aims at finding the optimal power and resource assignment that maximizes the achievable total system rate in the objective function (5.1a). Constraints (5.1b) and (5.1f) guarantee that the total power allocated does not exceed the total available power, P_{total}, at the BS and the power allocated to an RB is non negative. The constraint (5.1d) states that a minimum number ξ_s of UEs should be satisfied for each service plan s, i.e., at least ξ_s UEs per service s must meet their QoS/QoE requirements. Finally, constraints (5.1c) and (5.1e) ensure that each RB is assigned to at most one UE.

5.2 Optimal Solution

Notice that the problem stated in (5.1) is a mixed binary optimization problem with a nonconvex constraint (5.1d). Moreover, (5.1) is more complex than the problem stated in (3.1). Therefore, as (3.1), (5.1) also has a prohibitive computational complexity. In this section, the problem (5.1) will be rewritten as an ILP, which can be solved by standard methods presented in the literature [55].

One important consideration in this thesis is that the SINR values are mapped into rate values using a link adaptation scheme, as described in Section 2.4. In other words, there is a finite number of possible rate values, which is equal to the number of available MCSs. Therefore, considering that the system has M possible MCSs, the set of achievable rates can be represented by a vector $\mathbf{r} \in \mathbb{R}^{M \times 1}$ with elements $r_m = f_{adapt}^{\text{MCS}}(m)$ denoting the achievable rate at MCS m.

Moreover, since the number of possible rate values is finite, it makes sense that the values of power allocated to each RB are calculated *a priori* as the minimum power necessary to achieve the desired MCS. From (2.13), observe that the rate is given by the MCS chosen at the transmission, which in turn is calculated using the estimated SINR. Thus, from (2.11), it is possible to rewrite the expression of the estimated SINR of an UE u in a RB k as

$$\tilde{\gamma}_{u,k} = p_k \vartheta_{u,k}, \tag{5.2}$$

where $\vartheta_{u,k}$ denotes the estimated Channel-to-Noise Ratio (CNR) of the UE u on the RB k, given by

$$\vartheta_{u,k} = \frac{\left| g_{u,k} \tilde{\mathbf{h}}_{u,k}^{\mathrm{T}} \mathbf{f}_k \right|^2}{\sigma_I^2 + \sigma_n^2}. \tag{5.3}$$

Now, from (2.12), using the same concept of generalized inverse functions from [56] adopted in Chapter 3, it is possible to define γ_m^* as the minimum SINR value necessary to achieve the MCS m, i.e.,

$$\gamma_m^* = f_{adapt}^{\mathrm{SINR}}(m), \tag{5.4}$$

where $f_{adapt}^{\mathrm{SINR}}(\cdot)$ denotes the generalized inverse of $f_{adapt}^{\mathrm{CQI}}(\cdot)$, defined in Section 2.4.

In order to rewrite problem (5.1) considering the pre-calculated power values, the variables of the problem must be restated. Consider $\overline{\mathbf{X}} \in \{0,1\}^{U \times K \times M}$ as a binary tensor, where each element $x_{u,k,m}$ is equal to 1 if the UE u is select to transmit in the RB k using the MCS m and 0 otherwise. Moreover, let $\overline{\mathbf{P}} \in \mathbb{R}_+^{U \times K \times M}$ be a tensor, where each element $p_{u,k,m}$ denotes the power needed by the UE u on the RB k to transmit using the MCS m, i.e,

$$p_{u,k,m} = \frac{\gamma_m^*}{\vartheta_{u,k}}. \tag{5.5}$$

Finally, the optimization problem stated in (5.1) can be rewritten as

$$\max_{\overline{\mathbf{X}}} \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} \sum_{m=1}^{M} r_m x_{u,k,m}, \tag{5.6a}$$

$$\text{s.t.} \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} \sum_{m=1}^{M} p_{u,k,m} x_{u,k,m} \leq P_{\text{total}}, \tag{5.6b}$$

$$\sum_{u \in \mathcal{U}} \sum_{m=1}^{M} x_{u,k,m} \leq 1, \forall k \in \mathcal{K}, \tag{5.6c}$$

$$\sum_{u \in \mathcal{U}_s} H\left(\Omega_u \left(\sum_{k \in \mathcal{K}} \sum_{m=1}^{M} r_m x_{u,k,m} \right), \Omega_s^{\text{target}} \right) \geq \xi_s, \forall s \in \mathcal{S}, \tag{5.6d}$$

$$x_{u,k,m} \in \{0,1\}, \forall u \in \mathcal{U}, \forall k \in \mathcal{K} \text{ and } m \in \{0,1,\ldots,M\}. \tag{5.6e}$$

The objective function (5.6a) aims at maximizing the overall system rate. The constraint (5.6b) indicates that the summation of the allocated power can not exceed the limit P_{total}. Constraints (5.6c) and (5.6e) require that the number of scheduled UEs is at most one per RB. Furthermore, only one MCS can be selected. Finally, the (5.6d) indicates that at least ξ_s UEs per service s must have their requirements met.

The difference between (5.6) and (5.1) relies on the fact that now the possible power values are known and that the assignment variable indicates which MCS is allocated, besides the UE-RB association. However, they are essentially the same optimization problem. Observe that, in (5.6) the power is not a variable of the optimization, in opposite to (5.1). It means that (5.6) is a purely binary optimization problem, but still with a noncovex constraint (5.6d).

The satisfaction constraint (5.6d) is analogous to (3.1c) in Section 3.2 and also to (4.1c) in Section 4.2. Both constraints (3.1c) and (4.1c) relate to the satisfaction constraint in the previous chapters where joint RB and power allocation is out of the scope. However, these constraints share the same structure of (5.6d), ergo the same linearization technique adopted in Sections 3.2 and 4.2 can be employed here. Moreover, the minimum MOS requirement Ω_s^{target} per service s can be converted into a rate requirement ψ_u per UE u, using the function $\Omega^{\dagger}(\cdot)$, defined in (3.2). Therefore, the problem (5.6) can be rewritten as

$$\max_{\overline{\mathbf{X}},\boldsymbol{\rho}} \sum_{u\in\mathcal{U}}\sum_{k\in\mathcal{K}}\sum_{m=1}^{M} r_m x_{u,k,m}, \tag{5.7a}$$

$$\text{s.t.} \sum_{u\in\mathcal{U}}\sum_{k\in\mathcal{K}}\sum_{m=1}^{M} p_{u,k,m} x_{u,k,m} \leq P_{\text{total}}, \tag{5.7b}$$

$$\sum_{u\in\mathcal{U}}\sum_{m=1}^{M} x_{u,k,m} \leq 1, \forall k \in \mathcal{K}, \tag{5.7c}$$

$$\sum_{k\in\mathcal{K}}\sum_{m=1}^{M} r_m x_{u,k,m} \geq \psi_u \rho_u,\ \forall u \in \mathcal{U}, \tag{5.7d}$$

$$\sum_{u\in\mathcal{U}} q_{s,u}\rho_u \geq \xi_s, \forall s \in \mathcal{S}, \tag{5.7e}$$

$$x_{u,k,m} \in \{0,1\}, \forall u \in \mathcal{U}, \forall k \in \mathcal{K} \text{ and } m \in \{0,1,\ldots,M\}, \tag{5.7f}$$

$$\rho_u \in \{0,1\},\ \forall u \in \mathcal{U}. \tag{5.7g}$$

As already explained in Section 3.2, the linearization of the constraint (5.6d) yields two new equivalent constraints, namely, (5.7d) and (5.7e). In addition to these new constraints, a new variable $\boldsymbol{\rho} \in \{0,1\}^{U\times 1}$ was added in the optimization problem. The vector of slack variables $\boldsymbol{\rho} \in \{0,1\}^{U\times 1}$ consists of elements ρ_u which indicates whether the UE u will meet its requirements. Moreover, $q_{s,u}$ is equal to 1 if the UE u subscribes the service s and 0 otherwise.

The optimization problem (5.7) can now be defined as an ILP with only binary variables. Nevertheless, analogously as done in Section 3.2, problem (5.7) can be rewritten in a compact matrix form, which is more suitable to be implemented into many computational solvers. The objective and each constraint of (5.7) can be rewritten in a matrix form as

$$\max_{\overline{\mathbf{X}},\boldsymbol{\rho}} \ (\mathbf{r}\otimes\mathbf{1}_{UK})^{\mathrm{T}}\,\mathrm{vec}^{\mathrm{T}}\left\{\overline{\mathbf{X}}\right\} \tag{5.8a}$$

$$\text{s.t. } \mathrm{vec}^{\mathrm{T}}\left\{\overline{\mathbf{P}}\right\}\mathrm{vec}^{\mathrm{T}}\left\{\overline{\mathbf{X}}\right\} \le P_{\text{total}} \tag{5.8b}$$

$$\left(\mathbf{1}_M^{\mathrm{T}}\otimes\mathbf{I}_K\right)\left(\mathbf{I}_{KM}\otimes\mathbf{1}_U^{\mathrm{T}}\right)\mathrm{vec}^{\mathrm{T}}\left\{\overline{\mathbf{X}}\right\} \le \mathbf{1}_K \tag{5.8c}$$

$$\left(\left(\mathbf{1}_{KM}^{\mathrm{T}}\otimes\mathbf{I}_U\right)\odot\left(\left(\mathbf{r}\cdot\mathbf{1}_U^{\mathrm{T}}\right)\otimes\mathbf{1}_{UK}\right)^{\mathrm{T}}\right)\mathrm{vec}^{\mathrm{T}}\left\{\overline{\mathbf{X}}\right\} \ge \left(\left(\boldsymbol{\psi}\otimes\mathbf{1}_U^{\mathrm{T}}\right)\odot\mathbf{I}_U\right)\boldsymbol{\rho}, \tag{5.8d}$$

$$\mathbf{Q}\boldsymbol{\rho} \ge \boldsymbol{\xi}, \tag{5.8e}$$

$$\overline{\mathbf{X}} \in \{0,1\}^{U\times K\times M}, \tag{5.8f}$$

$$\boldsymbol{\rho} \in \{0,1\}^{U\times 1}, \tag{5.8g}$$

In order to simplify the problem (5.8), the optimization variables are rearranged into a single vector $\mathbf{y} = \left[\ \mathrm{vec}^{\mathrm{T}}\left\{\overline{\mathbf{P}}\right\}\ \middle|\ \boldsymbol{\rho}^{\mathrm{T}}\ \right]^{\mathrm{T}}$ and two separation matrices, $\mathbf{A}$ and $\mathbf{B}$, are defined in such a way that $\mathrm{vec}^{\mathrm{T}}\left\{\overline{\mathbf{X}}\right\} = \mathbf{A}\mathbf{y}$ and $\boldsymbol{\rho} = \mathbf{B}\mathbf{y}$. The two matrices that satisfy these conditions are $\mathbf{A} = \left[\ \mathbf{I}_{UKM}\ \middle|\ \mathbf{0}_{UKM\times U}\ \right]$ and $\mathbf{B} = \left[\ \mathbf{0}_{U\times UMK}\ \middle|\ \mathbf{I}_U\ \right]$. Thus, the optimization problem in matrix form (5.8) can be rewritten as

$$\max_{\mathbf{y}} \ (\mathbf{r}\otimes\mathbf{1}_{UK})^{\mathrm{T}}\mathbf{A}\mathbf{y} \tag{5.9a}$$

$$\text{s.t. } \mathrm{vec}^{\mathrm{T}}\left\{\overline{\mathbf{P}}\right\}\mathbf{A}\mathbf{y} \le P_{\text{total}} \tag{5.9b}$$

$$\left(\mathbf{1}_M^{\mathrm{T}}\otimes\mathbf{I}_K\right)\left(\mathbf{I}_{KM}\otimes\mathbf{1}_U^{\mathrm{T}}\right)\mathbf{A}\mathbf{y} \le \mathbf{1}_K \tag{5.9c}$$

$$\left(\left(\mathbf{1}_{KM}^{\mathrm{T}}\otimes\mathbf{I}_U\right)\odot\left(\left(\mathbf{r}\cdot\mathbf{1}_U^{\mathrm{T}}\right)\otimes\mathbf{1}_{UK}\right)^{\mathrm{T}}\right)\mathbf{A}\mathbf{y} \ge \left(\left(\boldsymbol{\psi}\otimes\mathbf{1}_U^{\mathrm{T}}\right)\odot\mathbf{I}_U\right)\mathbf{B}\mathbf{y}, \tag{5.9d}$$

$$\mathbf{Q}\mathbf{B}\mathbf{y} \ge \boldsymbol{\xi}, \tag{5.9e}$$

$$\mathbf{y} \in \{0,1\}^{(UKM+U)\times 1}, \tag{5.9f}$$

which can be also expressed in a standard ILP form as

$$\max_{\mathbf{y}} \mathbf{c}^{\mathrm{T}}\mathbf{y}, \tag{5.10a}$$

$$\text{s.t. } \mathbf{D}\mathbf{y} \le \mathbf{w}, \tag{5.10b}$$

$$\mathbf{y} \in \{0,1\}^{(UKM+U)\times 1}, \tag{5.10c}$$

where

$$\mathbf{c} = (\mathbf{r}\otimes\mathbf{1}_{UK})\,\mathbf{A}, \tag{5.11}$$

$$\mathbf{D} = \begin{bmatrix} \mathrm{vec}^{\mathrm{T}}\left\{\overline{\mathbf{P}}\right\}\mathbf{A} \\ \left(\mathbf{1}_M^{\mathrm{T}}\otimes\mathbf{I}_K\right)\left(\mathbf{I}_{KM}\otimes\mathbf{1}_U^{\mathrm{T}}\right)\mathbf{A} \\ \left(\left(\boldsymbol{\psi}\otimes\mathbf{1}_U^{\mathrm{T}}\right)\odot\mathbf{I}_U\right)\mathbf{B} - \left(\left(\mathbf{1}_{KM}^{\mathrm{T}}\otimes\mathbf{I}_U\right)\odot\left(\left(\mathbf{r}\cdot\mathbf{1}_U^{\mathrm{T}}\right)\otimes\mathbf{1}_{UK}\right)^{\mathrm{T}}\right)\mathbf{A}, \\ -\mathbf{Q}\mathbf{B} \end{bmatrix}, \tag{5.12}$$

and

$$\mathbf{w} = \left[\begin{array}{c|c|c|c} P_{\text{total}} & \mathbf{1}_K^{\mathrm{T}} & \mathbf{0}_U^{\mathrm{T}} & -\boldsymbol{\xi}^{\mathrm{T}} \end{array} \right]^{\mathrm{T}}. \tag{5.13}$$

5.3 Proposed suboptimal solution

In this section, a polynomial time heuristic, called Power and Resource Allocation for RMEC (PRARMEC), is proposed to solve the joint power and resource allocation problem described in Section 5.1. The algorithm proposed in this section utilizes the same framework as RMEC, detailed in Chapter 3, following the same three general steps, which are:

i. Selection of the $\sum_{s\in\mathcal{S}} \xi_s$ UEs from $\mathcal{U}$ that should meet their requirements;

ii. Calculation of an initial assignment;

iii. Reallocation of the RBs between the users in order to ensure that the problem constraints are met.

These steps are detailed in the remainder of this section, presenting the similarities and the differences between the proposed solution and the original RMEC from Chapter 3.

5.3.1 Step 1: User Selection

This first step of PRARMEC is very similar to the UEs selection of RMEC, described in Section 3.4.1. Since the objective of the proposed heuristic is to maximize the overall system throughput, like Section 3.4.1, it seems plausible to select the UEs that are more probable to achieve higher rates, meeting their requirements more easily.

Therefore, consider a set $\mathcal{L}$, initially empty, which will contain the UEs selected to compete for resources. After that, for each service $s \in \mathcal{S}$, an auxiliary set $\mathcal{A}$, initially equal to $\mathcal{U}_s$, is created. In Section 3.4.1, the UEs are removed from $\mathcal{A}$ following the criterion presented in Section 3.13, which corresponds to removing the UE with lowest transmit rate and highest QoS/QoE requirement. However, in the problem addressed by this chapter, the rate achieved by the UE depends on the power that will be allocated to it. Thus, it turns out that the transmit rate itself is not a suitable criterion. To overcome this issue, the removal criterion is slightly changed to consider a measurement that characterizes the UE's channel condition. Thus, the UEs with worst channel conditions and highest rate requirements are iteratively removed from this set, until $|\mathcal{A}| = \xi_s$. This criterion can be mathematically written as

$$u' = \arg\min_{u\in\mathcal{A}} \left\{ \frac{\sum_{k\in\mathcal{K}} \vartheta_{u,k}^{(\text{dB})}}{\psi_u} \right\}, \tag{5.14}$$

where $\vartheta^{(\text{dB})} = 10\log_{10}(\vartheta)$ denotes the estimated CNR in dB scale. The choice of the estimated CNR in dB scale is due to the logarithmic relationship between SINR and rate, which can be verified by the well-known Shannon's capacity formula [40]. Likewise, the relationship between the SINR and the rate for the MCS-based link adaptation adopted in this thesis can be logarithmically approximately, as shown in [71].

After the removal of the $|\mathcal{U}_s| - \xi_s$ UEs from $\mathcal{A}$, the remaining ξ_s UEs are moved from $\mathcal{A}$ to the set $\mathcal{L}$. This process is repeated for each service $s \in \mathcal{S}$. In the same way as for RMEC, the selection of the UEs corresponds to set the values of the association variables $\boldsymbol{\rho}$ of the optimization problem (5.7). Therefore, the resulting optimization problem after this first step can be written as

$$\max_{\overline{\mathbf{X}}_{\text{sat}}} \sum_{u \in \mathcal{L}} \sum_{k \in \mathcal{K}} \sum_{m=1}^{M} r_m x_{u,k,m}, \tag{5.15a}$$

$$\text{s.t.} \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} \sum_{m=1}^{M} p_{u,k,m} x_{u,k,m} \leq P_{\text{total}}, \tag{5.15b}$$

$$\sum_{u \in \mathcal{L}} \sum_{m=1}^{M} x_{u,k,m} \leq 1, \forall k \in \mathcal{K}, \tag{5.15c}$$

$$\sum_{k \in \mathcal{K}} \sum_{m=1}^{M} r_m x_{u,k,m} \geq \psi_u, \ \forall u \in \mathcal{L}, \tag{5.15d}$$

$$x_{u,k,m} \in \{0,1\}, \forall u \in \mathcal{U}, \forall k \in \mathcal{K} \text{ and } m \in \{0,1,\ldots,M\}, \tag{5.15e}$$

where $\overline{\mathbf{X}}_{\text{sat}}$ denotes a tensor containing only the rows $u \in \mathcal{L}$ of the original assignment tensor $\overline{\mathbf{X}}$.

5.3.2 Step 2: Initial RB and Power Allocation

Once the UEs that should be satisfied are chosen, the next step of the proposed heuristic is to provide an initial assignment as close as possible to the feasible set and the optimal solution. In order to provide an initial solution to the joint power and RB assignment problem studied in this chapter, the RMEC's solution framework, explained in Section 3.4.2, is adopted with minor modifications. Initially, the binary variables $x_{u,k,m}$ of the optimization problem (5.15) are relaxed into fractional ones, i.e., $0 \leq \tilde{x}_{u,k,m} \leq 1$, turning the ILP (5.15) into a LP, given by

$$\max_{\tilde{\mathbf{X}}} \sum_{u \in \mathcal{L}} \sum_{k \in \mathcal{K}} \sum_{m=1}^{M} r_m \tilde{x}_{u,k,m}, \tag{5.16a}$$

$$\text{s.t.} \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} \sum_{m=1}^{M} p_{u,k,m} \tilde{x}_{u,k,m} \leq P_{\text{total}}, \tag{5.16b}$$

$$\sum_{u \in \mathcal{L}} \sum_{m=1}^{M} \tilde{x}_{u,k,m} \leq 1, \forall k \in \mathcal{K}, \tag{5.16c}$$

$$\sum_{k \in \mathcal{K}} \sum_{m=1}^{M} r_m \tilde{x}_{u,k,m} \geq \psi_u, \ \forall u \in \mathcal{L}, \tag{5.16d}$$

$$0 \leq \tilde{x}_{u,k,m} \leq 1, \forall u \in \mathcal{U}, \forall k \in \mathcal{K} \text{ and } m \in \{0,1,\ldots,M\}. \tag{5.16e}$$

If the problem (5.16) is infeasible, so is (5.15). Furthermore, it is highly probable that (5.1) is also infeasible. On the other hand, if (5.16) has no feasible solution, then the UEs are iteratively removed from $\mathcal{L}$ following the criterion (5.14), until (5.16) becomes feasible. Observe that if the LP (5.16) is initially infeasible, so that it becomes necessary to remove any UE from $\mathcal{L}$, it is already known that the solution produced by PRARMEC will violate the minimum number of UEs that should be satisfied per service. However, like RMEC, the proposed algorithm in this chapter seeks at finding a solution as close as it seems possible, i.e., the PRARMEC tries to satisfy as many UEs as possible.

Observe that, until now, this step of the proposed algorithm is similar to RMEC. However, it is important to highlight that the rounding technique employed by the RMEC algorithm in Section 3.4.2 was designed for a 2-dimensional assignment, i.e., only UE-RB. On the other hand, the problem addressed in (5.16) consists of a 3-dimensional assignment, i.e., UE-RB-MCS. Therefore, in order to use the same solution framework of RMEC, the rounding technique must be adapted to provide a reliable initial solution. Hereafter, the rounding of the fractional solution of (5.16) is presented.

First, the fractional assignment tensor $\tilde{\tilde{\mathbf{X}}}$ is compressed into a $U \times K$ matrix $\tilde{\mathbf{X}}$, where each element $\tilde{x}_{u,k}$ of the matrix is equal to the summation of the elements of the tensor in the MCS dimension of $\tilde{\tilde{\mathbf{X}}}$, i.e.,

$$\tilde{x}_{u,k} = \sum_{m=1}^{M} \tilde{x}_{u,k,m} \tag{5.17}$$

With the fractional compressed matrix $\tilde{\mathbf{X}}$, the minimum number of RBs, ν_u, required by each UE u can be estimated by using (3.16). The next step is to create a bipartite graph $G(\mathcal{V},\mathcal{K},\mathcal{E})$ in the same way as done in Section 3.4.2. The difference between this step of the algorithm proposed in this chapter and the equivalent one of RMEC relies on the weights of the edges $(v_{u,n}, k) \in \mathcal{E}$ of the bipartite graph. While in RMEC, the edge $(v_{u,n}, k)$ is weighed by the achievable rate of the UE u in the RB k, here, the edge $(v_{u,n}, k)$ is weighed by the CNR in dB scale, $\vartheta_{u,k}^{(\mathrm{dB})}$ of the link between the BS and the UE u in the RB k. As in the previous step, since the rate depends on the power allocated in each RB and the logarithmic relationship between the SINR and the rate, the choice of the $\vartheta_{u,k}^{(\mathrm{dB})}$ as weight to bipartite graph seems reasonable.

Once the bipartite graph is built, the subset $\mathcal{M} \subset \mathcal{E}$ of edges that composes the minimum weighed matching of the bipartite graph $G(\mathcal{V},\mathcal{K},\mathcal{E})$ is selected, the same way as done in Section 3.4.2. The minimum weighed matching implies that the links with worse channel quality will be prioritized in this initial solution. Although it seems inconsistent with the objective of problem (5.1), likewise in RMEC, here obeying the QoS/QoE constraints is more important than achieving the optimal system rate. Therefore, selecting UEs with worse channel increases the chances of these UEs becoming satisfied at the end of the algorithm. On the other hand, UEs with better channel conditions usually need less resources as well as lower transmit power. Thus, even if UEs with better channel conditions do not get satisfied in this initial assignment, they meet their requirements more easily during the reallocation process than UEs with worse channel conditions.

After calculating the minimum matching $\mathcal{M}$, a set $\mathcal{X}_u$ containing the RBs allocated to each UE u is created, i.e., $\mathcal{X}_u = \{k \in \mathcal{K} \mid (v_{u,n}, k) \in \mathcal{M}\}$. Notice that the constraint (5.16c) of the LP does not require the allocation of all RBs to some UE. In fact, some RB may not be allocated to any UE when solving the LP (5.16), depending on the channel conditions. Therefore, if after the rounding process, some RB k^* is not allocated to any UE, then it will be assigned to the UE u^* with the best channel condition on it, i.e., $\mathcal{X}_{u^*} = \mathcal{X}_{u^*} \cup k^*$, where $u^* = \arg\max_{u \in \mathcal{L}} \left\{\vartheta_{u,k^*}^{(\mathrm{dB})}\right\}$.

At this moment, the proposed algorithm provides an initial allocation of the RBs to the UEs, however, the power was not yet allocated between the RBs. The final step of the initial UE assignment consists of allocating the total available power P_{total} among the RBs.

In the context of discrete power allocation, the Hughes-Hartogs (HH) bit-loading iterative algorithm, proposed in [72], was designed to optimally allocate the available power over the allocated RBs considering discrete powers levels with a low computational complexity in a point-to-point communication [73]. However, in general, the HH algorithm takes several iterations, due to the bit-by-bit loading. Therefore, in this thesis, a HH-based algorithm is employed, where instead of increasing the rate bit-by-bit, it will be increased MCS-by-MCS. In other words, the power and rate steps of the algorithm are determined by the MCSs employed by the system. This adaptation of the HH algorithm does not ensure optimality in the power allocation, but yields satisfactory results with fewer iterations. This kind of adaptation was also employed by [26]. Consider a vector $\boldsymbol{\eta} \in \{0, M\}^{K\times 1}$, where each element η_k corresponds to the current MCS set in the RB k. The implementation of the HH-based algorithm is quite simple. Given a certain available power, P_{avail}, and a target rate ψ, the algorithm will iteratively select the RB that needs the least power to step up by one MCS, and allocate the required power. The algorithm stops when one of three conditions is met: there is no more power, P_{avail}, to increase the MCS of an RB; the total rate achieved by the RBs is greater than or equal to the target ψ; or all the RBs reached the maximum MCS. Notice that neither the original HH algorithm nor the adaptation described above deals with multi user requirements. In other words, these power allocation algorithms do not distinguish groups of RBs of different UEs. Nevertheless, this HH-based algorithm will be adopted in the next steps of the proposed algorithm, as explained in the following. Hereafter, the HH-based algorithm described earlier will be referred, without loss of generality, simply as HH algorithm.

Consider an auxiliary set $\mathcal{A}$ containing all UEs that should be satisfied, i.e., all $u \in \mathcal{L}$, and consider that $\eta_k = 0$, for all $k \in \mathcal{K}$. In addition, consider an auxiliary variable, P_{avail}, that stores the remaining power that was not allocated initially equal to the total power, i.e., $P_{\mathrm{avail}} = P_{\mathrm{total}}$. After that, select the UE $u^* \in \mathcal{A}$ which needs the least amount of power to meet its requirements, which is usually the one with best channel conditions. Hence, it makes sense to select the UEs using the CNR criterion, i.e.,

$$u^* = \arg\max_{u \in \mathcal{A}} \left\{ \sum_{k \in \mathcal{K}} \vartheta_{u,k}^{(\mathrm{dB})} \right\}. \tag{5.18}$$

Once the UE u^* is selected, the HH algorithm [72] is employed to allocate the remaining power, P_{avail}, among the RBs $\mathcal{X}_{u^*}$ assigned to the UE u^* until its minimum rate requirement, ψ_{u^*}, is achieved or power P_{avail} is exhausted. At the end of the HH algorithm, the current MCS values of the RBs assigned to the UE u^* are updated and the UE u^* is removed from $\mathcal{A}$. Moreover, the remaining power is updated $P_{\text{avail}} = P_{\text{avail}} - \sum_{k \in \mathcal{X}_{u^*}} p_{u^*,k,\eta_k}$. This process is repeated until $\mathcal{A}$ becomes empty.

If at the end of the power allocation all UEs that should be satisfied meet their requirements, the remaining power P_{avail} is allocated among all RBs aiming at maximizing the overall system rate, i.e., until there is no more power or all RBs reached the maximum MCS. In this case, the solution of the proposed algorithm is already found. Thus, each element of the assignment tensor $\overline{\mathbf{X}}$ is given by

$$x_{u,k,m} = \begin{cases} 1, & \text{if } k \in \mathcal{X}_u \text{ and } m = \eta_k; \\ 0, & \text{otherwise.} \end{cases} \tag{5.19}$$

On the other hand, if not all UEs $u \in \mathcal{L}$ meet their requirements, then the proposed algorithm starts the reallocation process, explained in the next subsection.

5.3.3 Step 3: RB and Power Reallocation

In this step of the proposed algorithm, the radio resources, power and RBs, are reallocated in order to meet the UEs' constraints. As already mentioned, this step is executed only if the initial solution is not feasible. Although the idea of this step is similar to the one described in Section 3.4.3, the reallocation process here is very different, since here, the power must be reallocated as well.

Consider a priority vector $\mathbf{w} \in \mathbb{R}^{|\mathcal{L}| \times 1}$, where each element w_u denotes de priority of the UE calculated as

$$w_u = \sum_{k \in \mathcal{K}} \vartheta_{u,k}^{(\text{dB})}. \tag{5.20}$$

Consider also an auxiliary set $\mathcal{A}^{(\text{R})}$ containing UEs that are the potential resource receivers. The set $\mathcal{A}^{(\text{R})}$ initially contains all UEs that should be satisfied, i.e., $\mathcal{A}^{(\text{R})} = \mathcal{L}$. After that, the receiver UE $u^* \in \mathcal{A}^{(\text{R})}$ with highest priority w_{u^*} is chosen. Once the receiver is chosen, an auxiliary donor set, $\mathcal{A}^{(\text{D})}$, is created, containing all UEs that should be satisfied, except the receiver UE u^*, i.e., $\mathcal{A}^{(\text{D})} = \mathcal{L} \setminus \{u^*\}$. In addition, the donor UE u is selected as the one with highest priority w_u. Now that both donor and receiver are selected, calculate the current power consumed by the donor and the receiver UEs together, $p^{(\text{current})} = \sum_{k \in \mathcal{X}_u} p_{u,k,\eta_k} + \sum_{k \in \mathcal{X}_{u^*}} p_{u^*,k,\eta_k}$. Select the RB k assigned to the donor UE u, i.e., $k \in \mathcal{X}_u$, where the receiver UE u^* presents the highest CNR. Then, the RB k is taken from the set of RBs assigned to the donor UE, and it is assigned to the receiver UE u^*, i.e., $\mathcal{X}_u = \mathcal{X}_u \setminus \{k\}$ and $\mathcal{X}_{u^*} = \mathcal{X}_{u^*} \cup \{k\}$. Once the RB reallocation is done, the HH algorithm is applied over the donor's and the receiver's RBs, $\mathcal{X}_u$ and $\mathcal{X}_{u^*}$. In both cases, the HH algorithm will allocate power until both donor and receiver meet their requirements. After that, the new power consumed by both UEs together, $p^{(\text{new})}$, is recalculated. Now, in order to set the power and the RB reallocation, one of two conditions must be met:

1. If before the reallocation process the UE u^* was already satisfied, then the new joint consumed power of donor and receiver must be less than the current one, i.e., $p^{(\text{new})} < p^{(\text{current})}$;

2. If before the reallocation process the UE u^* was not satisfied, then the new joint consumed power of donor and receiver must be less than the current one plus the remaining available power, i.e., $p^{(\text{new})} < p^{(\text{current})} + P_{\text{avail}}$.

If one of the two conditions above is satisfied, then the RB and the power reallocation are confirmed. Therefore, the MCSs, η_k, of the donor's and receiver's RBs, i.e., $k \in \mathcal{X}_u \cup \mathcal{X}_{u^*}$, are updated. Furthermore, the current available power is also updated, by making $P_{\text{avail}} = P_{\text{avail}} + p^{(\text{new})} - p^{(\text{current})}$. If after the reallocation, the receiver UE u^* has met its requirements, then u^* is removed from the receiver set $\mathcal{A}^{(\text{R})}$ and the reallocation process is restarted. If all UEs from the set $\mathcal{L}$ are already satisfied, then all UEs from $\mathcal{A}^{(\text{R})}$ are removed. On the other hand, if neither of the conditions mentioned above is met, all the changes are reversed and the donor UE u is removed from the donor set $\mathcal{A}^{(\text{D})}$ and a new donor UE is selected, restarting the process. If there is no more donors in the set $\mathcal{A}^{(\text{D})}$, the receiver u^* is removed from $\mathcal{A}^{(\text{R})}$. The reallocation process finishes when $\mathcal{A}^{(\text{R})}$ becomes empty or all UEs $u \in \mathcal{L}$ are already satisfied.

At the end of the reallocation process, if there is still remaining power to allocate, then it is distributed among all RBs using the HH algorithm. Like the previous step, this power allocation aims at maximizing the overall system throughput. After this, the assignment tensor $\overline{\mathbf{X}}$ is filled according to (5.19), finishing the PRARMEC algorithm.

The overall computational complexity of the suboptimal algorithm proposed in this chapter follows the same trend of the solution framework adopted also in the previous chapters, being bounded by the complexity of solving the LP (5.16). The LP in turn can be solved using the Karmarkar's algorithm, which solves LP in polynomial time with a complexity of $O\left(U^{3.5}K^{3.5}M^{3.5}\right)$ [64].

5.4 State-of-the-art algorithm

The state-of-the-art low-complexity algorithm that addresses the problem described in Section 5.1 was proposed in [26], and is further referred in this thesis as Joint RB Assignment and Power Allocation (JRAPA) algorithm. In its turn, the JRAPA algorithm was inspired by the RAISES heuristic [24], briefly described in this thesis in Section 3.3. The first step of the JRAPA algorithm is to select which UEs will compete for resources in the next steps of the heuristic. Similarly to RAISES, the criterion adopted is to select, for each service $s \in \mathcal{S}$, a number ξ_s of UEs with the highest ratio between average throughput over the RBs and rate requirement. The JRAPA heuristic estimates the average throughput by considering mean SNR over all the RBs of each UE. Moreover, since the power is also a resource that should be allocated by the RRA algorithm, it is also considered in [26] that the power is equally divided among all the RBs. In other words, the algorithm proposed in [26] selects the ξ_s UEs with highest priority w_u, given by

$$w_u = \frac{f^{\text{MCS}}_{adapt}\left(f^{\text{CQI}}_{adapt}\left(\frac{P_{\text{total}}}{K}\sum_{k\in\mathcal{K}}\vartheta_{u,k}\right)\right)}{\psi_u}, \tag{5.21}$$

for $u \in \mathcal{U}_s$, for all services $s \in \mathcal{S}$.

Once the UEs that should be satisfied by the JRAPA algorithm are selected, an initial RB assignment is performed. Firstly, the JRAPA algorithm estimates the minimum amount of RBs that each UE should receive, considering that the RBs will be able to transmit at the maximum MCS. In other words, the minimum number of RBs that each UE should receive is given by $\lceil \psi_u / r_M \rceil$. Iteratively, the algorithm selects the UE u with the lowest priority w_u that should be satisfied and has not received the minimum number of RBs. Then, the algorithm allocates to that UE u the best RB that has not been assigned yet. The initial assignment stops when all UEs that should be satisfied received the estimated minimum number of RBs or when all RBs were assigned. If the later condition is firstly achieved, then it means that there is no feasible solution and the algorithm stops.

Considering that all UEs have already received the minimum amount of RBs, the JRAPA heuristic performs the HH-based algorithm (explained in Section 5.3.2) over each UE u considering a minimum target rate equal to ψ_u and that there is no power limit. If the total amount of power necessary to satisfy all the UEs previously selected is greater than the total power, P_{total}, available at the BS, the JRAPA algorithm iteratively selects the UE capable of receiving resources which has the lowest priority w_u and allocate to it the best RB k that was not assigned yet. After that, the HH-based algorithm is reapplied over the RBs of the UE u aiming the minimum rate ψ_u. If the total amount of power necessary to satisfy the selected UEs did not decrease, the RB k is deallocated from the UE u and this UE is forbidden to receive any more RBs. If all RBs are assigned and the necessary power is still higher than P_{total}, it means that no feasible solution is found and the JRAPA algorithm stops. If a feasible solution is found, then the JRAPA algorithm assigns the remaining RBs to the UEs with the best channel quality. Lastly, the remaining power that was not allocated is distributed over all RBs using the HH algorithm, in order to maximize the overall system rate.

5.5 Improvement on the JRAPA algorithm

In this section, two improvements are proposed on the JRAPA algorithm in order to improve the result achieved by [26] without increasing its computational complexity. This new version of the JRAPA heuristic is called Improved JRAPA (IJRAPA) algorithm. The first improvement is in the calculation of the priority w_u of each UE u, presented in (5.21). There, the priority calculation considers an estimation of the throughput of each UE by mapping its mean SNR over all RBs into rate. However, in usual communication systems, there are a finite number of possible rate values that can be employed by the system given by the number of existing MCSs, as adopted in this thesis. In other words, here, the mapping between SNR into rate is done by a surjective increasing function, i.e., there is a continuous range of SNR values that lead to each MCS value. It means that a UE u_1 with a better channel quality than another UE u_2 and demanding the same minimum rate requirement may have the same priority. Due to this fact, the selection of the UEs that will compete for radio resources following the priority stated in (5.21) may result into a high outage as it will be shown in Section 5.6. Therefore, in order to prioritize UEs with better channel conditions, instead of using the estimated average throughput, in the improved version of the algorithm proposed by [26], the priority w_u is given by the ratio between the average CNR over all RBs and the minimum rate requirement, i.e.,

$$w_u = \frac{\frac{1}{K}\sum_{k\in\mathcal{K}} \vartheta_{u,k}}{\psi_u}, \tag{5.22}$$

for $u \in \mathcal{U}_s$, for all services $s \in \mathcal{S}$.

Another issue of the JRAPA heuristic proposed by [26] is that it does not deal with infeasibility, like the RAISES algorithm. Notwithstanding, differently of the RAISES algorithm, it may lead to non practical solutions, i.e., the algorithm may allocate more power than the total available. In practice, this is a major drawback of the JRAPA algorithm. The second improvement proposed over the state-of-art algorithm is to add the capability of dealing with infeasible instances of the problem. If the algorithm detects that there is no feasible solution, the remaining RBs that were not allocated are assigned to the UEs with better channel conditions. After that, the total available power, P_{total}, is iteratively distributed to the RBs of each UE u in descending order of priority w_u using the HH algorithm, until it meets its minimum rate requirement ψ_u or there is no more available power to increase the rate of the UE u. If at the end of the power distribution, there is some remaining power, it will be distributed over all UEs using the HH algorithm. Besides always providing an RB and power allocation that can be employed by the BS, this second improvement also ensures that the maximum number of UEs will be satisfied given the algorithm possibilities.

5.6 Performance Analysis

In this section, the performance of the RRA algorithm proposed in Section 5.3, namely PRARMEC, is evaluated by comparing it to the optimal solution (5.10) provided in Section 5.2. Additionally, the algorithm proposed in this chapter is compared against the JRAPA heuristic [26], briefly described in Section 5.4, and with an improved version of JRAPA, called IJRAPA, proposed in Section 5.5. The following analyses are done considering the same scenario where the RMEC algorithm was evaluated, which was described in Section 3.5.

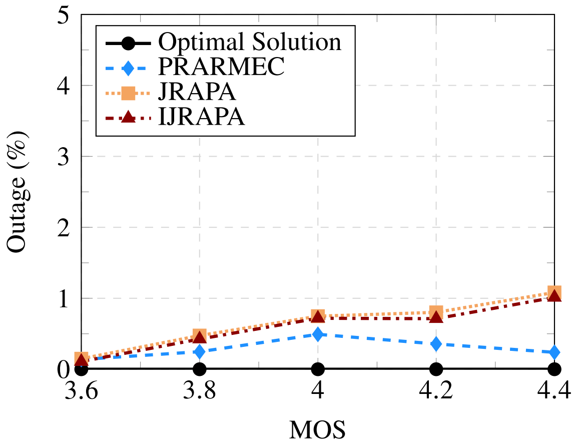
In Fig. 5.1, the outage probability and the overall system throughput are depicted varying the minimum MOS required by all UEs subscribing to the same service. In this analysis, three different number of UEs are considered, namely, $U = 10, 20$ and 30 UEs, and the system is required to meet the QoS/QoE requirements of all UEs, i.e., $\xi_1 = 100\%$ of U. In order to perform a fair comparison between the algorithms, only feasible instances of the problem (5.1) are considered. As explained in Section 5.4, the JRAPA algorithm does not provide a useful solution when it is not capable of meeting the requirement ξ_1 of minimum number of satisfied UEs, i.e., when an outage event happens. Moreover, the throughput results presented in Figs. 5.1b, 5.1d and 5.1f consider only instances of the problem where all algorithms yield a feasible solution.

In Figs. 5.1a, 5.1c and 5.1e, observe that the highest outage probability yielded by the proposed algorithm is equal to 4.36% for $U = 30$ UEs and a minimum MOS required equal to $\Omega_1^{\text{target}} = 4.4$. It is worth to mention that this is the worst case scenario considered during the simulation analyses, where the system was overloaded with high demanding UEs and the radio resource distribution needed to be wisely conducted. In the rest of the analyzed cases, the outage probability achieved by PRARMEC algorithm is less than 1%. It is important to highlight that the results presented by the proposed heuristic are very similar to the ones achieved by the comparison algorithms. In fact, comparing the results achieved in all analyzed scenarios, the gaps between the outage probability of PRARMEC and the benchmark algorithms, JRAPA and IJRAPA, are at most 1.92% and 1.71%, respectively. When comparing these results of outage probability with those presented in Fig. 3.5 of Section 3.5, it is possible to observe that the outage probability achieved by the algorithms analyzed in this chapter is considerably lower. This can be explained by the fact that the algorithms analyzed in this chapter have an additional resource to allocate: the available power, P_{total}, which in Chapter 3 is equally divided among the RBs. When the RRA algorithm relies only on the RB allocation, if the heuristic takes a bad decision in the assignment of one RB, an outage event may happen because this is the only type of resource into play for the allocation. On the other hand, when the power is also a resource that can be explored by the RRA algorithm, the damage caused by a bad RB assignment can be minimized during the power allocation process so that less outage events happen. In other words, when the RRA algorithm deals with the RB assignment and the power allocation jointly, it has an additional degree of freedom to explore and can correct some small problems in the RB assignment using that new degree of freedom. In terms of the optimization problem (5.1), it has a greater feasible set than the problem addressed in (3.1) since now the power allocated to each RB can assume values other than P_{total}/K. However, despite the fact that the JRAPA and IJRAPA algorithms reach a low outage probability, they yield an overall system throughput far from the optimal solution, as shown in the next analyses.

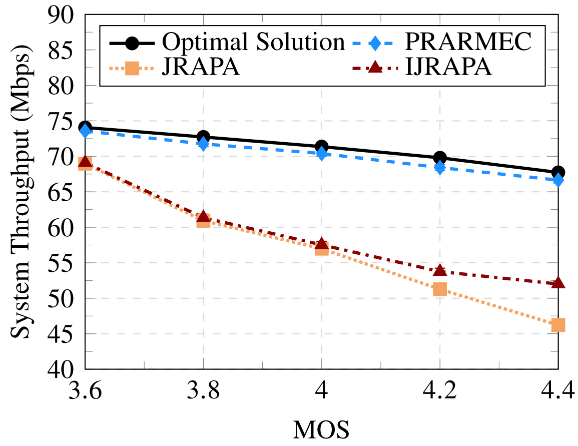
In Fig. 5.1b, for $U = 10$ UEs, the system throughput achieved by the algorithm proposed in this chapter is near optimal, for all considered minimum MOS requirement, Ω_1^{target}. Indeed, the PRARMEC algorithm achieves a throughput at most 1.98% below the optimal solution. On the other hand, the state-of-the-art algorithm, namely JRAPA, reaches a throughput 6.89% lower than the optimal solution when the UEs require a minimum MOS equal to 3.6. Additionally, when the minimum MOS required by the UEs is equal to 4.4, the efficiency

Figure 5.1 – System performance for a single service scenario with ξ_1 = 100% of U.

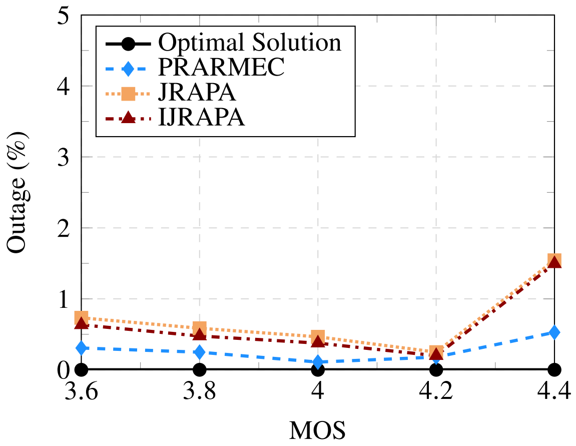
(a) Outage Probability for $U = 10$.



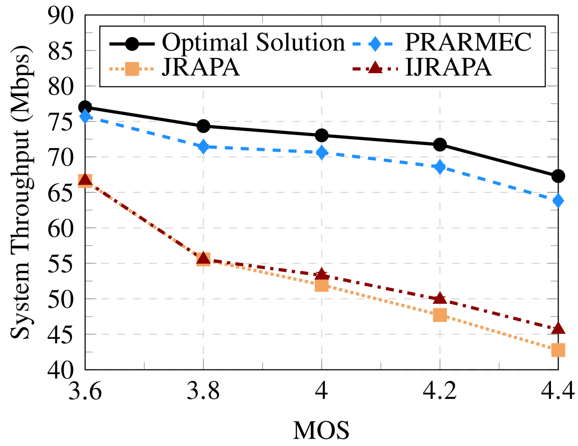
(b) System Throughput for $U = 10$.



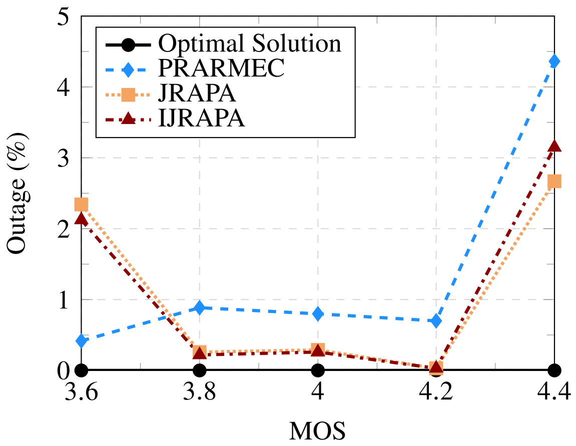
(c) Outage Probability for $U = 20$.



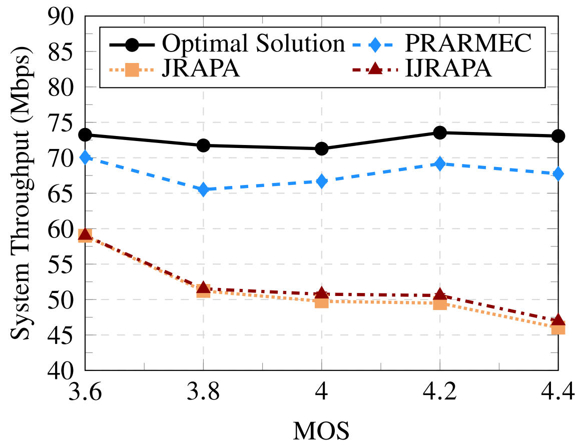
(d) System Throughput for $U = 20$.



(e) Outage Probability for $U = 30$.



(f) System Throughput for $U = 30$.



Source: Created by the author.

loss increases up to 31.76%. Regarding the improvement over the state-of-the-art algorithm, proposed in Section 5.5, it achieves an overall system throughput slightly better than the JRAPA. Moreover, the gain becomes more evident with the increasing minimum MOS requirement. In fact, for $\Omega^{\text{target}} = 4.4$, the IJRAPA algorithm presents a system throughput 12.52% higher than the JRAPA's.

As already mentioned, the number of possible solutions that meet the constraints of the joint power and RB allocation problem presented in (5.1) is larger compared to the case when an EPA was considered. Due to this fact, the proposed heuristic, as well as the benchmark algorithms, often provide feasible solutions. However, these solutions, although in the feasible set, may be far from the optimal solution. In other words, even though a solution meets the UEs' satisfaction constraints, the overall system throughput may be far from the optimal one. This fact can be observed in the results provided by the JRAPA and IJRAPA algorithms, which despite of being feasible, yield a system throughput considerably lower than the optimal, as seen in Figs. 5.1e and 5.1f. In order to achieve a system rate close to the optimal solution, the RRA algorithm must wisely allocate both RBs and power. As it can be observed in the previous analysis, the PRARMEC algorithm is capable of achieving throughput results closer to the optimal solution with a low outage probability. The main reason for the good performance of the PRARMEC algorithm relies on the initial RB assignment. Both JRAPA and IJRAPA algorithms initially estimate the amount of RBs necessary by each UE to meet its QoS/QoE requirements. During this estimation, these algorithms consider that the BS has enough available power to ensure that the UE is capable of transmitting data over the RBs using the highest MCS. However, depending on the channel conditions, this assumption may lead to a bad RB assignment, which may result in a solution far from the optimal one, regardless of the power allocation. On the other hand, similarly to the RMEC algorithm, explained in Chapter 3, the initial RB assignment of the PRARMEC algorithm is the result of a graph-based rounding of the fractional solution of the relaxed problem (5.16), which corresponds to the upper bound solution of the joint power and RB allocation problem, presented in (5.15). This method of acquiring the initial assignment leads to a RB allocation closer to the optimal solution, since it takes into consideration the available power and the UEs' channel quality.

In order to explain the better performance of the IJRAPA over the state-of-the-art algorithm, recall that one of the differences between the JRAPA and the IJRAPA is in the calculation of the UEs' priority. The JRAPA algorithm prioritizes the UEs according to (5.21), i.e., the UE's estimated rate of the average SNR of the entire bandwidth considering that the power is equally divided among all RBs. On its turn, the IJRAPA algorithm prioritizes the UEs by their average CNR, as presented in (5.22). In these analyses, all UEs are required to meet their requirements, therefore, both algorithms initially assign the minimum number of RBs needed by each UE considering that the highest MCS is employed. The RBs are iteratively assigned to the UEs, in ascending order of priority, as explained in Section 5.4. The difference in the UEs' prioritization has no considerable effect in the performance of this step of the algorithms, since the channel quality of the RBs is in general mostly dependent of the UE path loss. After this initial assignment, the power is allocated to RBs assigned to each UE ensuring that its rate requirement is met. In the considered scenario, the power needed to meet all the UEs requirements considering only the minimum number of necessary RBs is usually higher than the total available, especially when the required rate is high. Therefore, in the second step of the JRAPA and IJRAPA algorithms, the remaining RBs are iteratively distributed to the UEs until the power necessary to meet the UEs requirements is less than or equal to the available power in the BS, as explained in Section 5.4. The first UE to receive an additional RB is the one with the lowest priority and whose RB assignment causes a decrease in the total required power. Here, the difference between the algorithms prioritization is more relevant to the results. Due to the logarithmic relationship between SINR and rate, the amount of power needed to increase by one the MCS value used by an RB increases exponentially the higher the MCS is. Due to this fact, the UEs with poor channel conditions usually save more power when they receive an additional RB, since rather than using a few RBs employing high MCS values to transmit data, it is preferable to transmit over a larger number of RBs using lower MCS values. Therefore, in order to save more power, it is preferable that the UEs with poorest channel conditions receive additional RBs first. As already explained in Section 5.5, due to the priority adopted by the JRAPA algorithm, a UE with better channel condition may be selected to get additional RBs before a UE with worse channel conditions. Due to this, the JRAPA may spend more RBs to achieve a transmission power that meets the BS constraint. When the algorithms find an RB and power allocation that meets the power constraint and the UEs' requirements, they assign the rest of the RBs and allocate the remaining power aiming exclusively at maximizing the system throughput. Since the JRAPA algorithm usually needs more RBs to find a feasible solution to the RRA problem than the IJRAPA heuristic, it is natural that the overall system throughput achieved by the IJRAPA is higher than that of JRAPA.

In Fig. 5.1d, the overall system throughput is analyzed, considering a BS serving $U = 20$ UEs. In this scenario, the throughput achieved by the proposed algorithm remains close to the optimal solution. In fact, for a minimum required MOS equal to 3.6, the throughput achieved by the PRARMEC algorithm is 1.66% below the optimal solution. This gap increases up to 5.15% when the highest analyzed value of minimum MOS is considered. On the other hand, the JRAPA algorithm reaches a throughput 13.49% lower than the optimal solution when the lowest minimum MOS requirement is considered. On its turn, for a target MOS value equal to 4.4, the JRAPA algorithm achieves a system throughput 36.44% below the one reached by the optimal solution. Like in the previous analysis, the IJRAPA algorithm achieves throughput values similar to the ones reached by the JRAPA heuristic. Furthermore, the performance gap between the algorithms becomes more evident with the increasing minimum MOS requirement. For $\Omega^{\text{target}} = 4.4$, IJRAPA yields a throughput 6.77% higher than the JRAPA algorithm, described in [26].

Comparing the results presented in Figs. 5.1b and 5.1d, observe that the gap between the system throughput achieved by the PRARMEC algorithm and the optimal solution slightly increases when the number of UEs served by the BS becomes larger. Indeed, with the increasing number of UEs, the same amount of RBs must be properly divided among more UEs, which makes the RRA problem more challenging. In these cases, the initial RB and power allocation, explained in Section 5.3.2, is often infeasible, which may require more RB reallocations in the further step of the proposed algorithm. Recall that the initial RB and power allocation seeks an initial solution close to the optimal, however it may not be feasible, i.e., it may not meet all the UEs constraints. In its turn, in the reallocation process, explained in Section 5.3.3, the proposed algorithm performs a “fine-tunning” of the initial solution that is not yet feasible. However, since the reallocation process does not seek optimality, the final solution provided by the PRARMEC may be slightly far from the optimal. Besides, the odds of an infeasible initial solution increase when the RRA problem becomes harder, i.e., when there are more demanding UEs with non favorable channel conditions. Nevertheless, the gap between the optimal solution and the proposed algorithm does not significantly increase and remain considerable smaller than the optimality gap yielded by the benchmark algorithms.

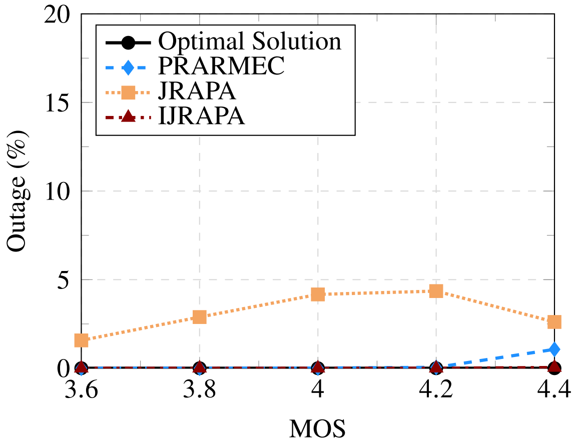
On the other hand, notice that the gap between the overall system throughput achieved by the IJRAPA when compared the one reached by the JRAPA heuristic diminishes when the number of UEs becomes larger. In fact, when the number of UEs increases, the number of RBs that both RRA algorithms need to meet the UEs requirements constrained by the available power increases. It means that, at the end of both algorithms, the available resources to perform the throughput maximization will be scarcer.

The results regarding overall system throughput considering a BS serving $U = 30$ UEs are depicted in Fig. 5.1f. These results corroborate with the previous analyses. Notice that the gap between the throughput achieved by the proposed algorithm and the optimal solution increases. Indeed, for a minimum MOS requirement equal to 3.6, the PRARMEC algorithm reaches a throughput 4.36% below the optimal. This gap increases when the UEs demand more rate. For a minimum MOS equal to 4.4, the PRARMEC algorithm reaches a throughput gap of 7.29% compared to the optimal solution. Nevertheless, besides the throughput achieved by the proposed algorithm distantiates from the optimal solution, its results are considerably better than the benchmark algorithms. In fact, the throughput achieved by the JRAPA and the IJRAPA algorithms are respectively 19.46% and 19.41% below the optimal solution when the minimum MOS required by the UEs is equal to 3.6. Moreover, this optimality gap increases up to 36.99% and 35.71% for the JRAPA and IJRAPA, respectively. As already explained, with the increasing number of UEs, the gap between the overall system throughput achieved by the JRAPA and the IJRAPA heuristics decreases.

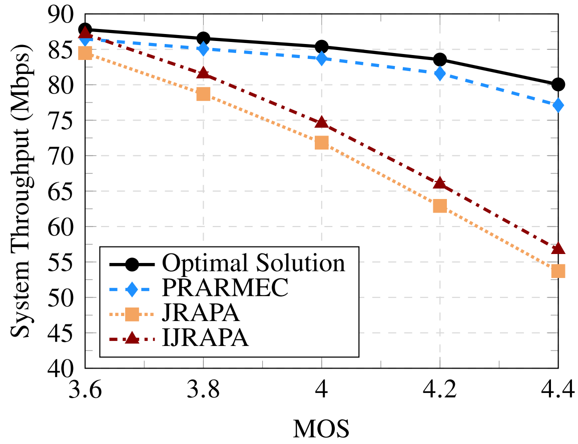
Until now, the performed analyses showed that compared to the benchmark algorithms, the joint RB and power RRA heuristic proposed in this chapter is more robust and scales better with the number of UEs served by the system and with the minimum MOS required by the UEs. Notwithstanding, in order to further evaluate the performance of the PRARMEC algorithm, in Fig. 5.2, the impact of varying the minimum number of UEs that should be satisfied is analyzed in terms of outage probability and overall system throughput. In this analysis, it is considered that the BS serves $U = 30$ UEs subscribing the same service. Again, in order to perform a fair comparison between the algorithms, in the outage analyses, only feasible instances of the problem (5.1) are considered. Additionally, the results of throughput consider only instances of the problem where all algorithms yield a feasible solution.

Figure 5.2 – System performance varying the percentage of satisfied UEs in a single service scenario with $U = 30$ UEs.

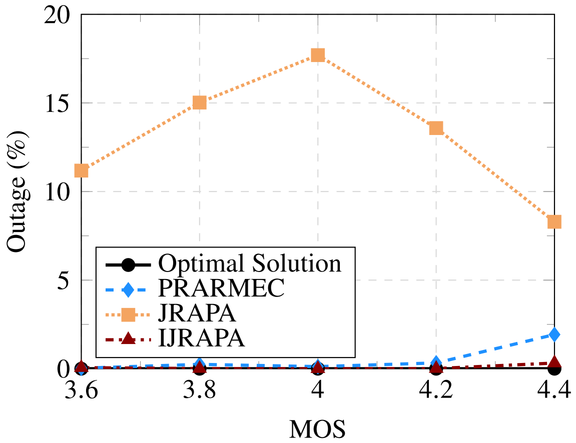
(a) Outage Probability for $\xi_1 = 80\%$.



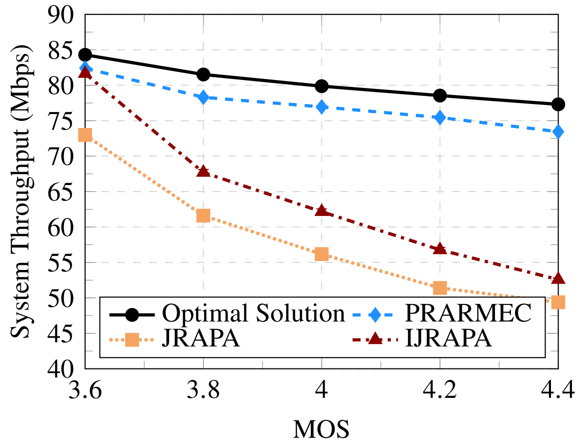
(b) System Throughput for $\xi_1 = 80\%$.



(c) Outage Probability for $\xi_1 = 90\%$.



(d) System Throughput for $\xi_1 = 90\%$.



Source: Created by the author.

In Figs. 5.2a and 5.2b, the results of outage probability and the overall system throughput are depicted when the BS is required to satisfy at least 80% of the UEs, i.e., $\xi_1 = 80\% \cdot U = 24$ UEs. In this scenario, the proposed heuristic presents a near optimal performance. Indeed, the PRARMEC algorithm reaches an outage probability of at most 1.07%, which happens when the highest minimum MOS requirement is considered, i.e., $\Omega_1^{\text{target}} = 4.4$. Moreover, in this case, the throughput achieved by the proposed algorithm is 3.67% below the optimal solution. Regarding the benchmark algorithms, notice that for a minimum MOS equal to 3.6, they are capable of achieving high throughput values. In fact, the system rate achieved by the IJRAPA algorithm is 0.82% higher than the one reached by the proposed algorithm. However, when the UEs become more demanding, the throughput achieved by the benchmark algorithms drastically decreases. For a minimum MOS equal to 4.4, the throughput values achieved by the JRAPA and IJRAPA algorithms are 32.90% and 29.14% below the optimal solution, respectively. These results show that the proposed algorithm is more capable of taking advantage of the UE diversity, since in this case the algorithms are free to neglect at most 6 UEs with poor channel conditions, increasing the chances of achieving a feasible result and higher throughput values.

Concerning the benchmark algorithms, observe that, when compared to the results presented in Figs. 5.1e and 5.1f, for $\xi_1 = 100\% \cdot U = 30$ UEs, the gap between the performance of the JRAPA and the IJRAPA algorithms considerably increase. Regarding the outage probability, notice that the JRAPA fails at finding a feasible solution in 4.35% of the times when the UEs require a MOS equal to 4.2. Meanwhile, in this scenario, the IJRAPA algorithm was capable of finding the feasible solution in all the simulated instances. Besides the gap in the outage probability, the IJRAPA algorithm also achieves higher throughput values than its predecessor, whereas in the previous analyses, for $\xi_1 = 100\% \cdot U$, the throughput gap between the JRAPA and IJRAPA algorithms was negligible in scenarios with low demanding UEs. Here, the system throughput achieved by the IJRAPA algorithm is at least 3.19% higher than the one reached by the JRAPA heuristic. The reason behind the low performance of the JRAPA algorithm, when compared to the IJRAPA, once again, relies on how the heuristic calculates the UEs' priority. In this scenario, the first step of both algorithms, JRAPA and IJRAPA, consists in neglecting the $U - \xi_1 = 6$ UEs with lowest priority. As already explained, several UEs with distinct channel conditions may have the same priority in the JRAPA. Therefore, an UE that could have its requirements met consuming less power may be neglected by the JRAPA algorithm to the detriment of another UE with the same priority, but harder to satisfy, i.e., with poorer channel conditions. On the other hand, the priority employed by the IJRAPA does not present this problem with UEs with equal priority, since it relies directly on channel quality.

The results presented in Figs. 5.2c and 5.2d depict the outage probability and the overall system throughput for a satisfaction target $\xi_1 = 90\% \cdot U = 27$ UEs. Compared to the previous analysis, for a minimum satisfaction target equal to 80% of the UEs, the outage probability of the proposed algorithm slightly increases, achieving 1.93% for a minimum MOS equal to 4.4. Moreover, in this scenario, the optimality gap of the PRARMEC algorithm also increases up to 4.99%. These results reinforce the conclusions made in the previous analysis, where the BS is required to satisfy 24 out of 30 UEs. In fact, instead of 6 UEs, here, the RRA algorithms are free to neglect at most 3 UEs, i.e., the UE diversity is smaller. The same behavior can also be verified in the results yielded by the benchmark algorithms.

Comparing the results regarding the outage probability considering a satisfaction target equal to $\xi_1 = 80\%$, 90% and 100% of the UEs presented in Figs. 5.2a, 5.2c and 5.1e, respectively, notice that for $\xi_1 = 90\%$, an outage event happens considerably more often than for the other analyzed cases when the JRAPA algorithm is employed. Indeed, in this case, the outage probability achieved by the JRAPA algorithm goes up to 17.70% of the instances when the minimum MOS required by the UEs is equal to 4. Moreover, the throughput gap between the JRAPA and the IJRAPA algorithms increases, with the later reaching throughput values up to 10.65% higher than the state-of-art algorithm. As already mentioned in the previous analysis, for $\xi_1 = 80\% \cdot U = 24$ UEs, this lack of scalability presented by the JRAPA algorithm can also be explained by its metric of UEs' prioritization. In addition to, when the minimum number of UEs that must meet their requirements increases to $\xi_1 = 90\% \cdot U = 27$ UEs, the UE diversity is smaller. This small UE diversity implies at a higher chance of the JRAPA disregarding an UE that could have its requirements met, to the detriment of another UE with the same priority, but with a channel quality that prevent it of being satisfied.

All the analyses until now have considered only feasible instances of the problem. However, as already explained in Chapter 3, when the scenario is challenging, either because the system is overloaded or UEs have poor channel quality, or even when they are requiring higher throughput values, the optimization problem (5.1) may not have a feasible solution. In these scenarios, an important feature that a QoS/QoE constrained RRA algorithm should seek is to provide a good result within the presented circumstances. In the next analyses, depicted in Fig. 5.3, the performance of the proposed algorithm is evaluated in terms of satisfaction rate and overall system throughput considering only infeasible instances of the problem (5.1). The PRARMEC heuristic is compared against the IJRAPA algorithm, proposed in Section 5.5, and the best possible achievable solution. The "best solution" is obtained similarly to the one described in the analysis of infeasibility performed in Section 3.5:

i. Try to solve the optimization problem stated in (5.1);

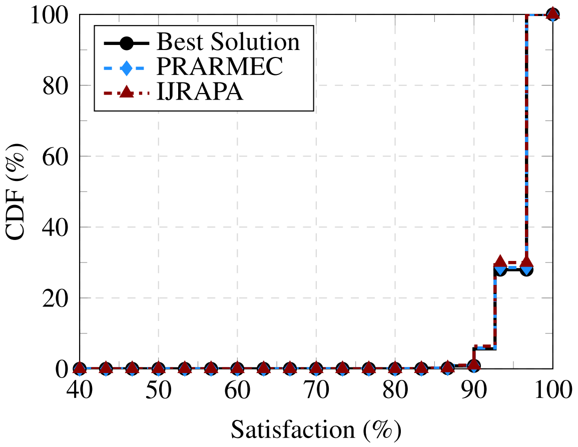
ii. If a feasible solution is found, then the "best solution" is found, otherwise relax the optimization problem by reducing the number of UEs that should be satisfied by one, i.e., $\xi_1 = \xi_1 - 1$, and go back to step i.

Recall that the JRAPA algorithm was not designed to deal with infeasible instances of the RRA problem. That is why it is not used as a benchmark algorithm in the results depicted in Fig. 5.3. These analyses consider a scenario with $U = 30$ UEs and a minimum satisfaction requirement equal to $\xi = 100\%$ of the UEs. Moreover, the algorithms are evaluated considering two different minimum MOS requirement, namely, $\Omega^{\text{target}} = 3.6$ and 4.4.

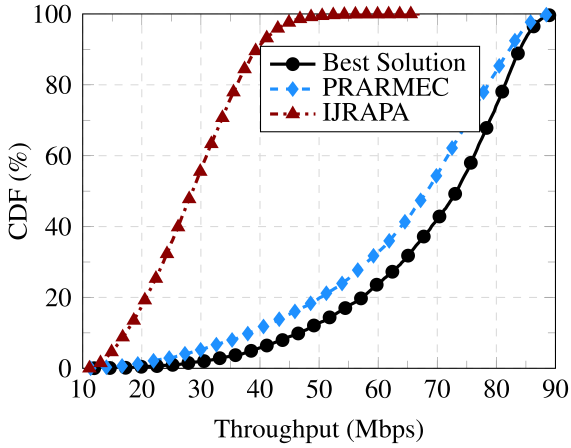
When the UEs require a minimum MOS equal to 3.6, both PRARMEC and IJRAPA algorithms reach almost the same satisfaction levels as the "best solution", as depicted in Fig. 5.3a. Moreover, the algorithms reach a satisfaction rate greater than or equal to 90% in at least 93.6% of the instances. In its turn, the overall system throughput achieved by the PRARMEC algorithm is close to the best solution, besides being considerably higher than the one reached by the IJRAPA heuristic, as shown in Fig. 5.3b. Indeed, the algorithm proposed in this chapter ensures the maximum throughput achieved by the benchmark heuristic in 51.89% of the cases. Additionally, at the 50%-ile, the throughput reached by the PRARMEC and the IJRAPA algorithms are 6.52% and 60.75% lower than the best solution, respectively. The reason of the better performance of the proposed algorithm compared to the IJRAPA heuristic relies on the initial RB and power allocation. When the IJRAPA is not capable of finding a feasible solution, the available power is distributed among the UEs in descending order of priority considering the already obtained RB assignment. This method of dealing with the infeasibility is rather inefficient when the main goal of the RRA algorithm is to maximize the overall system throughput, due to the fact that at end of the algorithm, many RBs may be assigned to UEs with extremely poor channel conditions

Figure 5.3 – CDF of the satisfaction and throughput considering $U = 30$ UEs and $\xi = 100\%$ of the UEs.

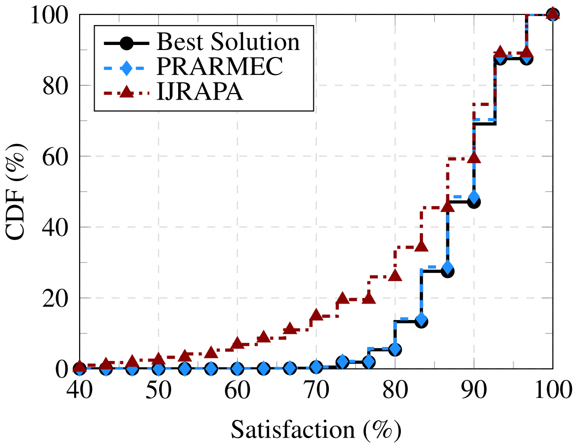
(a) CDF of the satisfaction for a minimum MOS equal to 3.6.



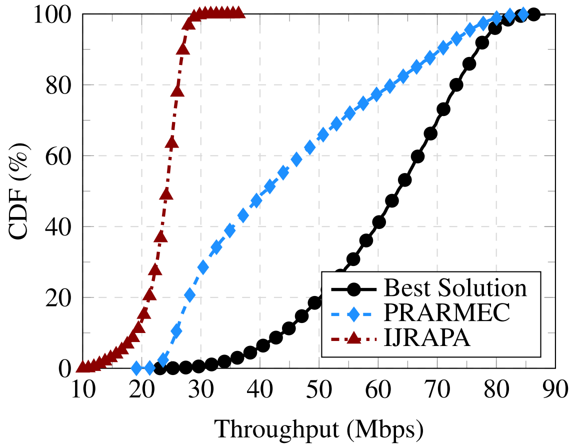
(b) CDF of the throughput for a minimum MOS equal to 3.6.



(c) CDF of the satisfaction for a minimum MOS equal to 4.4.



(d) CDF of the throughput for a minimum MOS equal to 4.4.



Source: Created by the author.

that will not meet their requirements. On the other hand, recall that the first step of obtaining the initial solution for the algorithm proposed in this chapter is to solve the relaxed LP (5.16), which denotes the upper bound solution of the joint power and RB allocation problem, presented in (5.15). If the LP is not feasible, the proposed algorithm iteratively disregards from the RRA problem the UE with worst channel condition and higher rate requirement until the relaxed LP becomes feasible. Therefore, besides the good quality of the initial RB and power allocation, previously discussed, the proposed heuristic is also capable of estimating if the RRA problem yields a feasible solution or not by assessing the feasibility of the relaxed LP.

Nevertheless, it is important to highlight that the feasibility of the relaxed LP does not ensure that the problem (5.15) is feasible. In fact, in some cases, the relaxed LP has a feasible solution capable of meeting the QoS/QoE requirements of a larger number of UEs than the “best solution”. This problem is referred in this thesis as the “false feasible” fractional assignment and was already discussed in Section 4.5. In summary, this is the main drawback of the framework adopted by all the heuristics proposed along this thesis. In the context of the

PRARMEC algorithm, when the LP yields a "false feasible" fractional assignment, the proposed algorithm tries to satisfy all the UEs that were satisfied by the fractional assignment provided by the LP, however, it is not possible. In these cases, UEs that do not meet their requirements at the end of the algorithm, which are usually UEs with poor channel conditions, may still receive some RBs and power. Hence, the overall system throughput reached by the proposed algorithm may present a significant loss compared to the "best solution".

The effects of the "false feasible" fractional assignment can still be observed in Fig. 5.3b, by analyzing the throughput results at the 10%-ile, i.e., the 10% of the harder scenarios. Here, the overall system throughput reached by the PRARMEC and the IJRAPA algorithms are 17.69% and 62.93% lower than the best solution, respectively. Compared to the results at the 50%-ile, notice that, while the throughput loss of the benchmark heuristic with respect to the "best solution" slightly changes, the PRARMEC algorithm throughput loss increases significantly. This performance loss is correlated with the satisfaction rate in these cases. At the 50%-ile, the proposed algorithm is capable of satisfying 29 out of 30 UEs. On the other hand, at the 10%-ile, the proposed algorithm is capable of satisfying two less UEs than what was required. Therefore, as previously explained, due to the "false feasible" fractional assignment, part of the radio resources are inefficiently allocated, decreasing the system throughput.

In order to ratify the conclusions made in the previous analysis, in Figs. 5.3c and 5.3d, the CDF of the satisfaction rate and the overall system throughput are respectively depicted considering that all UEs require a minimum MOS equal to 4.4. In this scenario, the CDF of satisfaction rate achieved by the proposed algorithm is almost equal to the one achieved by the "best solution". On the other hand, observe that the satisfaction rate achieved by the IJRAPA algorithm distantiates from the "best solution" the lower is the percentile. It means that, for challenging scenarios, the PRARMEC algorithm scales better than the benchmark heuristic, satisfying a number of UEs considerably higher. In fact, at the 10%-ile, the PRARMEC and IJRAPAs algorithms are capable of satisfying 24 and 20 out of 30 UEs, respectively.

Regarding the overall system throughput, the proposed algorithm outperforms the benchmark heuristic. Indeed, at the 50%-ile, the PRARMEC and the IJRAPA algorithms present a loss of throughput equal to 35.37% and 61.80% compared to the "best solution", respectively. Comparing this result with those presented in Fig. 5.3b, where the UEs require a minimum MOS equal to 3.6, notice that the gap between the proposed algorithm and the "best solution" rather increases. The reason behind this performance loss is due to "false feasible" fractional assignment. Considering the results at the 10%-ile, the throughput achieved by the PRARMEC algorithm is 41.24% lower than the "best solution", while the loss of throughput of the benchmark algorithm is 56.73%. As already observed in the previous analysis, the throughput loss of the proposed algorithm with respect to the "best solution" increases when the scenario becomes more challenging. On the other hand, the gap between the throughput achieved by the benchmark heuristic and the "best solution" decreases. This fact can be explained by the lower satisfaction rate achieved by IJRAPA compared to the "best solution". Nevertheless, the proposed algorithm is capable of ensuring the maximum throughput achieved by the benchmark heuristic in 57.66% of the cases.

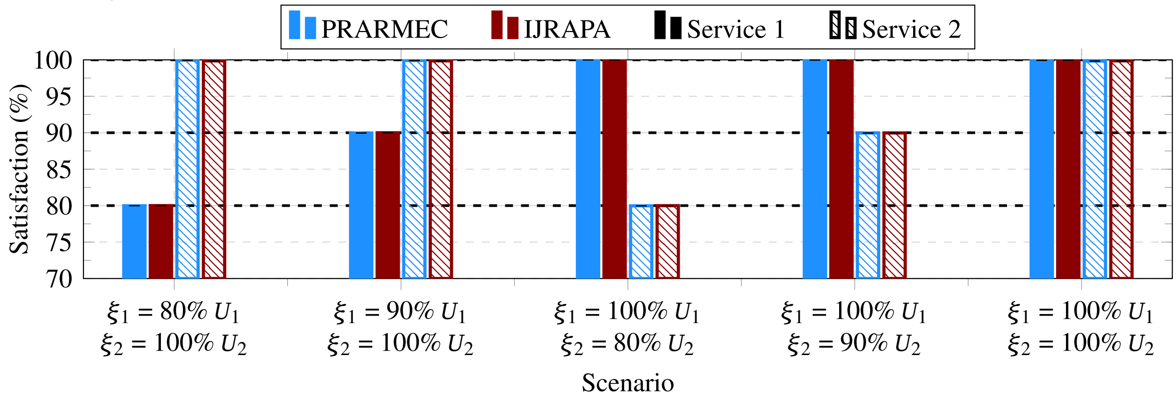
As previously mentioned, the "false feasible" fractional assignment is a drawback of the solution framework adopted by both RMEC and PRARMEC algorithms. From an infeasibility analysis perspective, the proposed heuristic for joint power and RB allocation clearly presents a higher throughput loss compared to the "best solution", as depicted in Fig. 5.3, than in the case where the power is equally divided among the RBs, as can been in Figs. 3.8 and 3.9. Thus, it turns out that the PRARMEC heuristic is more susceptible to the "false feasible" issue than RMEC. The reason for this comes from the additional degree of freedom of the RRA algorithm, i.e., the power allocation. As already explained, due to this additional degree of freedom, the RRA problem solved by the PRARMEC algorithm has a larger feasible set than the problem studied in Chapter 3. Furthermore, recall that both PRARMEC and RMEC algorithms solve an LP that derives from the relaxation of their original RRA problems. Therefore, it is expected that the LP solved by the PRARMEC has a larger convex hull than the one solved by the RMEC, which consequently ends up providing a solution that may be outside the feasible set of the original problem. In such cases, it is likely that some resources are allocated to UEs with unfavorable channel conditions, then decreasing the throughput.

In order to complete the benchmarking of the proposed algorithm, the next analysis, depicted in Fig. 5.4, evaluate the performance of the PRARMEC algorithm in multi-service scenarios comparing it against the optimal solution and the IJRAPA heuristic, in terms of average satisfaction rate per service and overall system throughput. Due to incapability of the JRAPA algorithm of providing a useful solution when an outage event happens, it was left out of this analysis. Here, it is considered the same setup as in the multi-service analyses in Section 3.5, where the BS serves the 30 UEs, divided into two different service plans. The first service plan has 20 subscribes, i.e., $U_1 = 20$ UEs, and consists of a high-quality skype video call, which has a recommended minimum throughput of $\Omega_1^{\text{target}} = 500$ kbps [67]. The remaining 10 UEs subscribe the second service plan, i.e., $U_2 = 10$ UEs, which models a high definition skype video call, recommending a minimum throughput of $\Omega_2^{\text{target}} = 1.5$ Mbps [67]. The results presented in Fig. 5.4 consider only feasible instances of the problem (5.1) in five different scenarios, varying the minimum number of UEs that should be satisfied in each service plan, namely, ξ_1 and ξ_2.

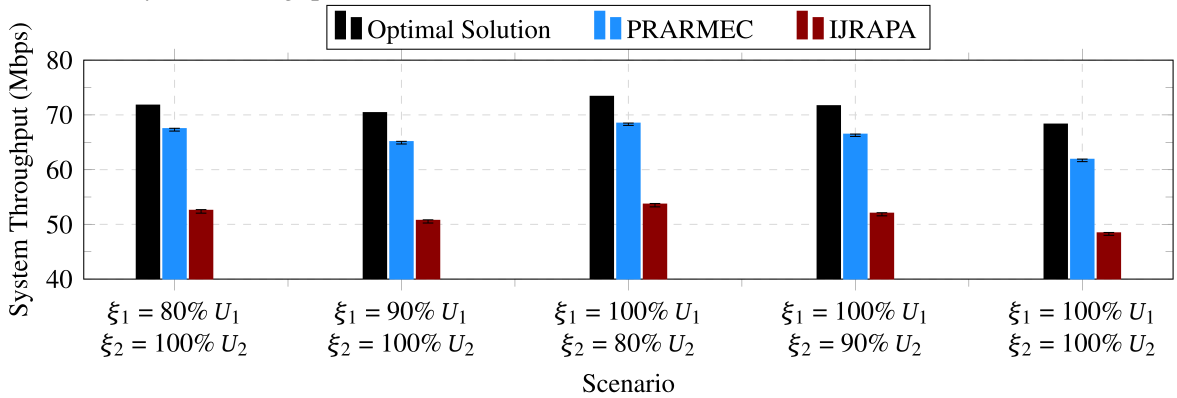
Observe that, in Fig. 5.4a, both PRARMEC and IJRAPA were capable of achieving results of average satisfaction rates almost equal to the minimum required by both services in all considered scenarios. In fact, in the most challenging scenario, where $\xi_1 = U_1$ and $\xi_2 = U_2$, the proposed algorithm achieves an average satisfaction rate of 99.93% and 99.83% for services 1 and 2, respectively. Meanwhile, the IJRAPA algorithm reaches an average satisfaction rate of 99.96% and 99.86% for services 1 and 2, respectively. As already explained in the first analyses of this section for the single service scenarios, since the heuristics in this chapter are capable of allocating jointly RB and power, they can compensate an unwise RB assignment with a proper power allocation. In other words, the joint RB and power RRA problem has a large feasible set,

Figure 5.4 – System performance considering $U = 30$ UEs and $S = 2$ service plans, where $U_1 = 20$ and $U_2 = 10$ UEs, $\Omega_1^{\text{target}} = 500$ kbps and $\Omega_2^{\text{target}} = 1.5$ Mbps.

(a) Average satisfaction rate.



(b) Overall system throughput.



Source: Created by the author.

which means that an allocation that meets the UEs' requirements is likely to be found without an elaborate RB assignment algorithm.

On the other hand, as also pointed out before, besides meeting the UEs requirements, the overall system throughput achieved by the RRA algorithm may be far from the optimal solution. In order to reach a high system throughput, the RRA algorithm must perform an elaborate RB and power allocation, which implies at a higher computational complexity. As a matter of fact, all the analyses performed in single service scenarios have shown that the additional computational complexity of the PRARMEC algorithm is justified by the substantial improvement at the overall system throughput when compared to the benchmark heuristic. The good performance of the PRARMEC algorithm can also be perceived in multi-service scenarios. Indeed, the proposed algorithm considerably outperforms the benchmark heuristic in terms of overall system throughput, as depicted in Fig. 5.4b.

In the first scenario, the system is required to satisfy at least $\xi_1 = 80\% \cdot U_1 = 16$ of service 1 and $\xi_2 = 100\% \cdot U_2 = 10$ of service 2. In other words, jointly, both services require that at least 26 out of 30 UEs meet their requirements. Moreover, the joint rate required by both services is equal to $\xi_1\Omega_1^{\text{target}} + \xi_2\Omega_2^{\text{target}} = 80\% \cdot 20 \cdot 500\ \text{kbps} + 100\% \cdot 10 \cdot 1.5\ \text{Mbps} = 23$ Mbps. In this scenario, the proposed algorithm achieves a throughput 6% smaller than the optimal solution, while the throughput loss reached by the benchmark algorithm is 26.84%.

When the minimum number of UEs that shall be satisfied in service 1 increases up to $\xi_1 = 90\% \cdot U_1 = 18$, the total required rate is equal to 24 Mbps. Compared to the first scenario, here the UE diversity decreases, i.e., the algorithms are free to neglect a smaller number of UEs. Since the system is free to disregard less UEs with poor channel conditions, the RRA problem becomes more challenging, which impacts at the overall system throughput achieved by the suboptimal heuristics when compared to the optimal solution. Indeed, the throughput losses of the PRARMEC and the IJRAPA algorithms compared to the optimal solution are 7.54% and 28.03%, respectively.

In the third scenario, the system is required to satisfy all UEs subscribing the service 1, i.e., $\xi_1 = U_1 = 20$, and at least $\xi_2 = 80\% \cdot U_2 = 8$ of service 2. Moreover, the sum of the throughput required by services 1 and 2 is equal to 22 Mbps. Observe that, in this scenario, the optimal solution, as well as the suboptimal heuristics achieve higher throughput values than those presented in the first scenario. This is a consequence of the lower required throughput considered in this scenario. On the other hand, when compared in relative terms to the optimal solution, the throughput loss achieved by both suboptimal heuristics slightly increases. In fact, the overall system throughput achieved by the PRARMEC and the IJRAPA algorithms are 6.7% and 26.9% smaller than the optimal solution, respectively. The reason for this higher throughput loss is due to the fact that although the system requires a smaller throughput to address the constraint of minimum number of satisfied UEs in both services, the UE diversity is greater in the first scenario. As matter of fact, in this scenario, two additional UEs are required to meet their requirements, which implies that the RRA algorithms shall address the requirements of more UEs with worse channel conditions.

In the fourth scenario, the minimum number of UEs subscribing the service 2 that are required to meet their requirements is at least $\xi_2 = 90\% \cdot U_2 = 9$ UEs, implying at a joint required throughput equal to 23.5 Mbps. The higher demanded rate together with the lower UE diversity cause to a higher throughput loss compared to the optimal solution, specifically 7.28% and 27.47% for the proposed and the benchmark algorithms, respectively. Furthermore, comparing the throughput achieved by the algorithms in this scenario with those reached in the second one, it is possible to infer the same analysis performed when compared the the first and the third scenarios.

Finally, in the last scenario, all UEs are required to meet their requirements, i.e., there is no UE diversity. In this scenario, the joint throughput required by all UEs is 25 Mbps. This is the most challenging scenario considered in Fig. 5.4b and, as expected from the previous analyses, in this scenario the algorithms deliver the lowest throughput values. Additionally, both suboptimal heuristics present a higher throughput loss when compared to the optimal solution. In fact, the PRARMEC and the IJRAPA algorithms achieve throughput values 9.44% and 29.16% smaller than the optimal solution.

For all considered scenarios, the proposed algorithm achieves throughput values at least 27.83% higher than the benchmark algorithm. Moreover, the analyses performed in Fig. 5.4

show that the proposed algorithm also yields good results in a multi-service environment. It is important to highlight that, although the scenarios analyzed in Fig. 5.4 are the same as the ones considered in Fig. 3.10, no direct comparison between the results obtained in Figs. 3.10 and 5.4 can be done. The reason is that in the results depicted in Fig. 3.10, only feasible instances of the RB allocation problem (3.1) are considered. In its turn, the results showed in Fig. 5.4 consider instances of the problem where (5.1) has a feasible solution. Recall that the optimization problem (3.1) considers that the power is equally divided among all RBs, while in (5.1), both power and RBs are jointly allocated. Therefore, the feasible set of (5.1) can be seen as a superset of the feasible set of (3.1). However, it is important to emphasize that even dealing with instances of the RRA problem with UEs in poorer channel conditions, the throughput achieved by (5.1) in Fig. 5.4 is considerably higher than the one reached by (3.1) in Fig. 3.10.

5.7 Chapter Summary

In this chapter, the problem of maximizing the overall system rate by jointly allocating both RBs and power has been studied, constrained by ensuring the QoS/QoE requirements of at least a minimum number of UEs per service. This problem is similar to the one addressed in Chapter 3, however, there the RBs are allocated to the UEs considering that the power is divided equally between all RBs.

The problem addressed in this chapter was rewritten as an ILP, which can be solved by standard methods, such as BB or BC. Nevertheless, due to the prohibitive computational complexity to obtain the optimal solution in real-time systems, a low-complexity suboptimal algorithm, called PRARMEC, was proposed. In addition to, it was also proposed an improvement over the state-of-the-art heuristic, which is referred in this thesis as JRAPA, and is proposed to solve the same problem addressed in this chapter, without increase the computational complexity. This improved version of the JRAPA algorithm is called IJRAPA.

During the analyses performed in Section 5.6, it is shown that the IJRAPA outperforms its predecessor, besides of solving two existing issues in the JRAPA algorithm: the improper priority at the UE removal and the fact that it does not provide a useful solution when it is not capable of meeting the UEs' QoS/QoE requirements. Besides that, it was shown that the IJRAPA algorithm is capable of often meeting the QoS/QoE requirements, hence achieving a high satisfaction rate. However, the IJRAPA algorithm is not capable of achieving high throughput values, mainly in challenging scenarios, where the UEs have poor channel conditions.

In this context, the PRARMEC algorithm substantially outperformed the state-of-the-art heuristic, as well as its improvement. The proposed algorithm is capable of achieving high throughput values, often close to the optimal solution, besides of properly meeting the RRA constraints. In addition to, the proposed algorithm is also more robust when dealing with instances of the problem where the constraints are impossible to be met.

It was verified that the PRARMEC algorithm inherits the "false feasible" fractional assignment issue from the solution framework employed also by the heuristics proposed in the previous chapters. Moreover, this problem appears more often in the PRARMEC than in the RMEC heuristic, proposed in Chapter 3. However, as shown in Section 5.6, even with this drawback, the proposed algorithm still yields better results than the JRAPA and the IJRAPA heuristics.

It is important to highlight that the better performance of the proposed algorithm comes with the cost of a higher computational complexity. Therefore, the IJRAPA algorithm stands as a good solution at non challenging scenarios, i.e., where the UEs have good channel conditions and low rate requirements. On the other hand, the PRARMEC algorithm is more suitable to challenging scenarios, mainly when a feasible solution is hard to find.

6 CONCLUSIONS AND FUTURE PERSPECTIVES

The work developed along this thesis has studied RRA methods aiming at maximizing the overall system throughput, constrained by guaranteeing a certain satisfaction rate per service in single and multi-service scenarios. It is important to emphasize that this type of RRA problem is very important to the mobile network operators, since it deals with different service requirements and seeks to provide a trade-off between a higher spectral efficiency and the UEs' satisfaction, which in turn can be adjusted by the system operator. Moreover, the RRA problems studied along this thesis have considered that the UEs' requirements are given in terms of both QoS and QoE.

In Chapter 3, the first RRA problem studied in this thesis is presented. In this chapter, it is considered that the total power available in the BS is equally divided among all RBs. Moreover, the UEs' requirements are addressed on a TTI basis. In other words, the RRA problem treated in this chapter must assign the available RBs in order to maximize the BS downlink throughput, ensuring that at least a minimum number of UEs subscribing each service plan has their requirements met on each TTI. The problem studied in this chapter was initially formulated as a non-convex and non-concave optimization problem and further rewritten as an ILP, which can be solved by standard algorithms from the literature, such as the BB algorithm. Even though, depending on the number of UEs and RBs in the system, solving a ILP may be prohibitive due to its exponential computational complexity. Therefore, a new low-complexity suboptimal heuristic, called RMEC, was proposed. The performance analyses have shown that the RMEC algorithm has a near optimal performance, in addition to a high scalability in terms of the problem's input size. The RMEC algorithm outperformed the state-of-the-art heuristic, namely RAISES, which intends to solve the same problem, but which was not designed to address QoE constraints. Furthermore, the results have shown that RMEC is also capable of providing near feasible solutions when the constraints of the problem are impossible to be met, i.e, the proposed algorithm reaches a solution that is close to the best one available. However, the results have also shown that in non challenging scenarios, i.e., where UEs have good channel conditions and low rate requirements, the RAISES algorithm is a suitable choice, due to its lower computational complexity.

In Chapter 4, a RRA problem similar to the one addressed in Chapter 3 is studied. However, in this chapter, the UEs requirements needed to be met in a given timespan. This RRA problem was formulated as an ILP, however, the addition of the time dimension to the optimization problem required a high computational effort, as well as the knowledge of the UEs' channel during the entire timespan. Therefore, a low complexity heuristic was proposed based on the results of Chapter 3. This new algorithm, called TRMEC, is designed to run on each time slot, considering the previous KPI values of the UEs to improve the scheduling on each TTI. Besides addressing the RRA problem considering the time dimension, the computational complexity of TRMEC on each TTI is the same as RMEC's one. The computational simulations have shown that the proposed heuristic considerably outperforms its predecessor, RMEC, and two benchmark algorithms, namely ASC and ATES. Both considered benchmark heuristics have the same constraints of satisfying a fraction of the UEs. However, both of them were designed to work in single service scenarios and to address only the UEs' minimum rate requirements. The results have shown that the TRMEC algorithm presents higher robustness and scalability to all analyzed parameters. Besides that, TRMEC also provides a near-feasible solution, i.e., when it does not address the problem constraints, it tries to satisfy as many UEs as it seems possible, in addition to achieve high throughput values.

Finally, the last contribution of this thesis is presented in Chapter 5. Here, the RRA problem is once again studied on a TTI basis, i.e., the UEs' QoS/QoE requirements are addressed during a single TTI, as done in Chapter 3. However, differently from Chapter 3, the total available power in the BS is considered as part of the RRA problem. The joint power and RB allocation problem was mathematically formulated as an optimization problem. In Chapter 2, it was stated that the system operates using discrete MCS levels, which means that the possible values of power needed by an UE to transmit in a given RB using a certain MCS can be prior calculated. This assumption allowed the RRA optimization problem to be restated as an ILP, in a similar manner as done in Chapter 3. However, due to the excessive computational complexity, a new low complexity suboptimal heuristic was proposed, called PRARMEC. The PRARMEC algorithm was conceived using the same solution framework as the other two heuristics proposed in Chapters 3 and 4, i.e., the RMEC and the TRMEC algorithms. Additionally, it was also proposed an improvement over the state-of-the-art heuristic in the literature, called JRAPA. This improved version of JRAPA was called IJRAPA. The computational simulations have shown that the IJRAPA algorithm in fact is a more suitable RRA method compared to the JRAPA. In fact, the results have shown that in general the IJRAPA presented lower outage rates and higher system throughputs. Moreover, the IJRAPA algorithm is capable of providing an useful solution when it is not feasible, in opposite to the JRAPA heuristic. On the other hand, the simulations results have also shown that besides the IJRAPA and the PRARMEC algorithms have similar outage rates, the latter one substantially outperforms the first in terms of system throughput, being often close to the optimal solution. Furthermore, the PRARMEC algorithm is also more robust in scenarios where no feasible solution exists, achieving satisfaction rates very close to the best solution, besides higher throughput values.

It is worth to highlight that the RMEC, TRMEC and PRARMEC algorithms were proposed over the same solution framework and all of them, in general, have yielded better results in all considered KPIs. However, the high scalability and robustness presented by the algorithms proposed along this thesis comes with a higher computational complexity compared to the existing heuristics used in the benchmarking. It was also observed that the three algorithms, specially the TRMEC and the PRARMEC algorithms, presented an issue, which was referred in this thesis as the "false feasible" fractional assignment problem, which may result in lower throughput values. Even though, recall from the performance analyses of the algorithms that this issue only happens in very challenging scenarios, where the benchmark algorithms were already significantly outperformed. A workaround to this issue would require an additional computations, which may not be worth the corresponding performance improvement.

Some future research perspectives that may be pursued starting from the work done in this thesis are:

- Extend the PRARMEC to consider the time scheduling: The first perspective left from the work done in this thesis is to study the RRA problem similar to the one addressed in Chapter 4, but considering that the power is also managed by the RRA algorithm.

- Multiuser MIMO: All the studies conducted in this thesis have considered that only one UE could be assigned to each RB. However, the evolution of the wireless communications points towards the employment of increasing number of antennas in both BS and UE. Due to this, an interesting future research would be to consider spatial multiplexing as an input, or more challengingly, as part of the RRA problem. This would significant improve the overall system rate.

- Multi Radio Access Technology (RAT): The most recent works in the literature show that the new generation of telecommunications, 5G, will work together with the current LTE. It means that, in a near future, the devices will be capable of connecting to one of the RATs or use both at the same time. Therefore, an interesting extension of the work done in this thesis would be to extend the RRA problems to cope with UEs connected to multiple RATs.

- Consider more realistic traffic models in temporal scheduling: In this thesis, all the computational simulations have considered that the UEs were demanding data from the BS uninterruptedly. However, in face of non-full buffer traffic models, the RRA algorithms shall cope with the limited transmit buffer, in order to not waste resources, scheduling them to UEs that have not enough data to receive. In future works, the RRA problem studied in Chapter 4 could be studied in scenarios with more realistic traffic models, such as the Constant Bit Rate (CBR), video and online gaming.

- Consider more realistic QoE metrics and better suited to the considered more realistic traffic models: The QoE to QoS mapping considered along this thesis has considered a straight mapping between rate and the MOS. In future works, MOS depending on other KPIs may be considered, such as delay, packet queue length, among others.

- Heuristics with lower complexity: The algorithms proposed in this thesis have polynomial-time worst-case computational complexity. However, as already explained, they are more complex than the benchmark algorithms. Thus, a study on how to reduce the complexity of the heuristics proposed in this thesis maintaining their good performance would be interesting.

- Uplink RRA algorithms: The RRA problems addressed in this thesis were proposed to work only over downlink transmissions. In future works, the heuristics proposed in this thesis could be adapted to uplink transmissions.

- Analyze the impact of channel hardening in the RB scheduling: With channel hardening, the scheduling process of the heuristics proposed in this thesis may be simplified. Therefore, it would be interesting to evaluate these scenarios and verify the impacts in the overall complexity of the scheduling algorithms.

- Study the fairness maximization problem with satisfaction constraints: Although the rate maximization is interesting from the spectral efficiency point of view, fairness maximization is also a relevant objective for network operators and, thus, can be a topic of research in the future.

REFERENCES

1 HASAN, M.; HOSSAIN, E. Toward 5G: Applications, requirements and candidate technologies. In: ed. by Rath Vannithamby and Shilpa Talwar. 1. ed. [S.l.]: John Wiley & Sons, Jan. 2017. Distributed Resource Allocation in 5G Cellular Networks, p. 129–161.

2 ITU. **Draft new Report ITU-R M.[IMT-2020.TECH PERF REQ] - Minimum requirements related to technical performance for IMT-2020 radio interface(s)**. [S.l.: s.n.], Feb. 2017. ITU-R SG05 Contribution 40.

3 CAPOZZI, F. et al. Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and a Survey. **IEEE Communications Surveys and Tutorials**, v. 15, n. 2, p. 678–700, May 2013. ISSN 1553-877X. DOI: `10.1109/SURV.2012.060912.00100`.

4 CASTANEDA, E. et al. An Overview on Resource Allocation Techniques for Multi-User MIMO Systems. **IEEE Communications Surveys and Tutorials**, v. 19, n. 1, p. 239–284, Firstquarter 2017. DOI: `10.1109/COMST.2016.2618870`.

5 AROUSSI, S.; MELLOUK, A. Survey on machine learning-based QoE-QoS correlation models. In: INTERNATIONAL CONFERENCE ON COMPUTING, MANAGEMENT AND TELECOMMUNICATIONS, April 2014, Da Nang, Vietnam. **Proceedings [...]** Da Nang, Vietnam: [s.n.], Apr. 2014. P. 200–204. DOI: `10.1109/ComManTel.2014.6825604`.

6 ITU. **Vocabulary for performance and quality of service. Amendment 2: New definitions for inclusion in Recommendation ITU-T P.10/G.100**. [S.l.], July 2008.

7 BARAKOVI, S.; SKORIN-KAPOV, L. Survey and Challenges of QoE Management Issues in Wireless Networks. **Journal of Computer Networks and Communications**, Hindawi Publishing Corporation, v. 2013, p. 1–28, 2013. DOI: `10.1155/2013/165146`.

8 FIEDLER, M.; HOSSFELD, T.; TRAN-GIA, P. A Generic Quantitative Relationship between Quality of Experience and Quality of Service. **IEEE Network**, v. 24, n. 2, p. 36–41, Mar. 2010. DOI: `10.1109/MNET.2010.5430142`.

9 REICHL, P.; TUFFIN, B.; SCHATZ, R. Economics of logarithmic Quality-of-Experience in communication networks. In: CONFERENCE ON TELECOMMUNICATIONS INTERNET AND MEDIA TECHNO ECONOMICS (CTTE), June 2010, Ghent, Belgium. **Proceedings [...]** Ghent, Belgium: IEEE, June 2010. P. 1–8. DOI: `10.1109/CTTE.2010.5557702`.

10 REICHL, P.; TUFFIN, B.; SCHATZ, R. Logarithmic laws in service quality perception: where microeconomics meets psychophysics and quality of experience. **Telecommunication Systems**, v. 52, n. 2, p. 587–600, 2013. DOI: `10.1007/s11235-011-9503-7`.

11 PONCELA, J. et al. Quality Assessment in 3G/4G Wireless Networks. **Wireless Personal Communications**, v. 76, n. 3, p. 363–377, Mar. 2014. DOI: `10.1007/s11277-014-1711-5`.

12 COLE, R. G.; ROSENBLUTH, J. H. Voice over IP Performance Monitoring. **ACM SIGCOMM Computer Communication Review**, v. 31, n. 2, p. 9–24, Apr. 2001. ISSN 0146-4833. DOI: `10.1145/505666.505669`.

13 MOK, R.; CHAN, E.; CHANG, R. Measuring the quality of experience of HTTP video streaming. In: IFIP/IEEE INTERNATIONAL SYMPOSIUM ON INTEGRATED NETWORK MANAGEMENT, May 2011, Dublin, Ireland. **Proceedings [...]** Dublin, Ireland: IEEE, May 2011. P. 485–492. DOI: `10.1109/INM.2011.5990550`.

14 POLITIS, I. et al. A model of network related QoE for 3D video. In: IEEE GLOBECOM WORKSHOPS, Dec 2012, Anaheim, CA, USA. **Proceedings [...]** Anaheim, CA, USA: IEEE, Dec. 2012. P. 1335–1340. DOI: `10.1109/GLOCOMW.2012.6477776`.

15 OLIVER-BALSALOBRE, P. et al. Self-tuning of scheduling parameters for balancing the quality of experience among services in LTE. **EURASIP Journal on Wireless Communications and Networking**, Springer, v. 2016, n. 1, p. 7, 2016.

16 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.

17 SOUSA, I.; QUELUZ, M. P.; RODRIGUES, A. A Survey on QoE-oriented Wireless Resources Scheduling. **arXiv preprint arXiv:1705.07839**, May 2017.

18 ZHANG, Y. J.; LETAIEF, K. B. Multiuser Adaptive Subcarrier-and-Bit Allocation with Adaptive Cell Selection for OFDM Systems. **IEEE Transactions on Communications**, v. 3, n. 5, p. 1566–1575, Sept. 2004. ISSN 1536-1276. DOI: `10.1109/TWC.2004.833501`.

19 LEE, T.-H.; HUANG, Y.-W. Resource Allocation Achieving High System Throughput with QoS Support in OFDMA-Based System. **IEEE Transactions on Communications**, v. 60, n. 3, p. 851–861, Mar. 2012. ISSN 0090-6778. DOI: `10.1109/TCOMM.2012.020912.100632`.

20 HE, T.; WANG, X.; NI, W. Optimal Chunk-Based Resource Allocation for OFDMA Systems With Multiple BER Requirements. **IEEE Transactions on Vehicular Technology**, v. 63, n. 9, p. 4292–4301, Nov. 2014. ISSN 0018-9545. DOI: `10.1109/TVT.2014.2313354`.

21 ZHAO, W.; WANG, S. Joint subchannel and power allocation in multiuser OFDM systems with minimal rate constraints. **International Journal of Communication Systems**, Wiley Online Library, v. 27, n. 1, p. 1–12, 2014. DOI: `10.1002/dac.2336`.

22 UMTS. **Selection procedures for the choice of radio transmission technologies of the UMTS**. Sophia Antipolis, France, Apr. 1998.

23 FURUSKÄR, A. **Radio Resource Sharing and Bearer Service Allocation for Multi-Bearer Service, Multi-Access Wireless Networks**. Apr. 2003. PhD thesis – Royal Institute of Technology (KTH), Radio Communication Systems.

24 LIMA, F. R. M. et al. Resource Assignment for Rate Maximization with QoS Guarantees in Multiservice Wireless Systems. **IEEE Transactions on Vehicular Technology**, v. 61, n. 3, p. 1318–1332, Mar. 2012. DOI: 10.1109/TVT.2012.2183905.

25 LIMA, F. R. M. et al. Improved Spectral Efficiency with Acceptable Service Provision in Multi-User MIMO Scenarios. **IEEE Transactions on Vehicular Technology**, v. 63, n. 6, p. 2697–2711, July 2014. DOI: 10.1109/TVT.2013.2293333.

26 MAURÍCIO, W. V. F. et al. Joint Resource Block Assignment and Power Allocation Problem for Rate Maximization With QoS Guarantees in Multiservice Wireless Systems. **Journal of Communication and Information Systems (JCIS)**, v. 31, n. 1, p. 211–223, 2016. ISSN 1980-6604. DOI: 10.14209/jcis.2016.19.

27 YAACOUB, E.; DAWY, Z. Fair optimization of video streaming quality of experience in LTE networks using distributed antenna systems and radio resource management. **Journal of Applied Mathematics**, Hindawi, v. 2014, 2014. DOI: 10.1155/2014/562079.

28 CHEN, F.; QIN, X.; WEI, G. QoE Optimized Resource Allocation in Multiuser OFDM Systems. **Przegląd Elektrochniczny (Electrical Review)**, v. 2012, 7b, p. 328–331, July 2012.

29 CHEN, X. et al. A Near Optimal QoE-Driven Power Allocation Scheme for Scalable Video Transmissions Over MIMO Systems. **IEEE Journal of Selected Topics in Signal Processing**, v. 9, n. 1, p. 76–88, Feb. 2015. ISSN 1932-4553. DOI: 10.1109/JSTSP.2014.2336603.

30 ZHENG, J. et al. Optimal Power Allocation and User Scheduling in Multicell Networks: Base Station Cooperation Using a Game-Theoretic Approach. **IEEE Transactions on Wireless Communications**, v. 13, n. 12, p. 6928–6942, Dec. 2014. ISSN 1536-1276. DOI: 10.1109/TWC.2014.2334673.

31 CHO, Y. H. et al. A QoE-Aware Proportional Fair Resource Allocation for Multi-Cell OFDMA Networks. **IEEE Communications Letters**, v. 19, n. 1, p. 82–85, Jan. 2015. ISSN 1089-7798. DOI: 10.1109/LCOMM.2014.2367115.

32 MONTEIRO, V. F. et al. Radio resource allocation framework for quality of experience optimization in wireless networks. **IEEE Network**, v. 29, n. 6, p. 33–39, Nov. 2015. ISSN 0890-8044. DOI: 10.1109/MNET.2015.7340422.

33 MONTEIRO, V. F. et al. A QoE-Aware Scheduler for OFDMA Networks. **Journal of Communication and Information Systems (JCIS)**, v. 31, n. 1, p. 41–48, Nov. 2016. ISSN 1980-6604. DOI: 10.14209/jcis.2016.3.

34 LIN, X.; WANG, S. QoE-oriented resource allocation in multiuser OFDM systems. In: 2016 IEEE INTERNATIONAL CONFERENCE ON COMMUNICATION SYSTEMS (ICCS), Dec. 2016, Shenzhen, China. **Proceedings [...]** Shenzhen, China: IEEE, Dec. 2016. P. 1–6. DOI: 10.1109/ICCS.2016.7833551.

35 RUGELJ, M. et al. Novel Cross-Layer QoE-Aware Radio Resource Allocation Algorithms in Multiuser OFDMA Systems. **IEEE Transactions on Communications**, v. 62, n. 9, p. 3196–3208, Sept. 2014. ISSN 0090-6778. DOI: 10.1109/TCOMM.2014.2347288.

36 SACCHI, C.; GRANELLI, F.; SCHLEGEL, C. A QoE-Oriented Strategy for OFDMA Radio Resource Allocation Based on Min-MOS Maximization. **IEEE Communications Letters**, v. 15, n. 5, p. 494–496, May 2011. ISSN 1089-7798. DOI: 10.1109/LCOMM.2011.031411.101672.

37 LEE, G. et al. QoE-Aware Scheduling for Sigmoid Optimization in Wireless Networks. **IEEE Communications Letters**, v. 18, n. 11, p. 1995–1998, Nov. 2014. ISSN 1089-7798. DOI: 10.1109/LCOMM.2014.2354409.

38 XIAO, X.; TAO, X.; LU, J. Energy-Efficient Resource Allocation in LTE-Based MIMO-OFDMA Systems With User Rate Constraints. **IEEE Transactions on Vehicular Technology**, v. 64, n. 1, p. 185–197, Jan. 2015. ISSN 0018-9545. DOI: 10.1109/TVT.2014.2319078.

39 3GPP. **Physical layer aspect for evolved Universal Terrestrial Radio Access (UTRA)**. [S.l.], Oct. 2006. Available from: <http://www.3gpp.org/ftp/Specs/html-info/25814.htm>.

40 GOLDSMITH, A. **Wireless Communications**. [S.l.]: Cambridge University Press, 2005. ISBN 9781139445849.

41 PÄTZOLD, M. **Mobile radio channels**. [S.l.]: John Wiley & Sons, 2011.

42 RAPPAPORT, T. S. et al. Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design. **IEEE Transactions on Communications**, v. 63, n. 9, p. 3029–3056, Sept. 2015.

43 3GPP. **Study on Channel Model for Frequencies from 0.5 to 100 GHz**. [S.l.], Sept. 2017. v.14.2.0. Available from: <http://www.3gpp.org/DynaReport/38901.htm>. Visited on: 26 Sept. 2017.

44 PAULRAJ, A.; NABAR, R.; GORE, D. **Introduction to Space-Time Wireless Communications**. 1st. [S.l.]: Cambridge University Press, 2003.

45 TSE, D.; VISWANATH, P. **Fundamentals of Wireless Communication**. [S.l.]: Cambridge University Press, 2005.

46 3GPP. **Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures**. [S.l.], Oct. 2010. Available from: <http://www.3gpp.org/ftp/Specs/html-info/36213.htm>.

47 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. DOI: `10.1109/MCOM.2014.6815890`.

48 BOUDREAU, G. et al. Interference coordination and cancellation for 4G networks. **IEEE Communications Magazine**, v. 47, n. 4, p. 74–81, Apr. 2009. ISSN 0163-6804. DOI: `10.1109/MCOM.2009.4907410`.

49 DEB, S. et al. Algorithms for Enhanced Inter-Cell Interference Coordination (eICIC) in LTE HetNets. **IEEE/ACM Transactions on Networking**, v. 22, n. 1, p. 137–150, Feb. 2014.

50 CADAMBE, V.; JAFAR, S. Interference Alignment and Spatial Degrees of Freedom for the K User Interference Channel. In: IEEE INTERNATIONAL CONFERENCE ON COMMUNICATIONS, May 2008, Beijing, China. **Proceedings [...]** Beijing, China: IEEE, May 2008. P. 971–975. DOI: `10.1109/ICC.2008.190`.

51 AZIZ, D.; MAZHAR, M.; WEBER, A. Multi User Inter Cell Interference Alignment in Heterogeneous Cellular Networks. In: IEEE VEHICULAR TECHNOLOGY CONFERENCE (VTC SPRING), May 2014, Seoul, Korea. **Proceedings [...]** Seoul, Korea: IEEE, May 2014. P. 1–5. DOI: `10.1109/VTCSpring.2014.7022994`.

52 SEOL, C.; CHEUN, K. A Statistical Inter-Cell Interference Model for Downlink Cellular OFDMA Networks under Log-Normal Shadowing and Multipath Rayleigh Fading. **IEEE Transactions on Communications**, IEEE Press, Piscataway, NJ, USA, v. 57, p. 3069–3077, 10 Oct. 2009. ISSN 0090-6778. DOI: `10.1109/TCOMM.2009.10.080152`.

53 MEHLFÜHRER, C. et al. Simulating the Long Term Evolution Physical Layer. In: PROCEEDINGS OF THE EUROPEAN SIGNAL PROCESSING CONFERENCE, Aug. 2009, Glasgow, Scotland. **Proceedings [...]** Glasgow, Scotland: [s.n.], Aug. 2009. P. 1471–1478.

54 BOYD, S.; VANDENBERGHE, L. **Convex optimization**. 1st. [S.l.]: Cambridge University Press, 2004.

55 NEMHAUSER, G.; WOSLEY, L. **Integer and Combinatorial Optimization**. [S.l.]: John Wiley & Sons, 1999.

56 EMBRECHTS, P.; HOFERT, M. A note on generalized inverses. **Mathematical Methods of Operations Research**, v. 77, n. 3, p. 423–432, 2013. ISSN 1432-5217. DOI: `10.1007/s00186-013-0436-7`.

57 KRUMKE, S. O.; THIELEN, C. The generalized assignment problem with minimum quantities. **European Journal of Operational Research**, v. 228, n. 1, p. 46–55, 2013. ISSN 0377-2217. DOI: `10.1016/j.ejor.2013.01.027`.

58 SHMOYS, D. B.; TARDOS, É. An Approximation Algorithm for the Generalized Assignment Problem. **Mathematical Programming**, Springer-Verlag New York, Inc., Secaucus, NJ, USA, v. 62, n. 3, p. 461–474, Dec. 1993. ISSN 0025-5610. DOI: `10.1007/BF01585178`.

59 MARTELLO, S.; TOTH, P. **Knapsack problems: algorithms and computer implementations**. [S.l.]: J. Wiley & Sons, 1990. (Wiley-Interscience series in discrete mathematics and optimization). ISBN 9780471924203.

60 CATTRYSSE, D. G.; WASSENHOVE, L. N. V. A survey of algorithms for the generalized assignment problem. **European Journal of Operational Research**, v. 60, n. 3, p. 260–272, 1992. ISSN 0377-2217. DOI: `10.1016/0377-2217(92)90077-M`.

61 KOSTER, A.; MUÑOZ, X. **Graphs and Algorithms in Communication Networks: Studies in Broadband, Optical, Wireless and Ad Hoc Networks**. [S.l.]: Springer, 2009. (Texts in Theoretical Computer Science. An EATCS Series). ISBN 9783642022500.

62 WOLSEY, L. **Integer Programming**. [S.l.]: Wiley, 1998. (Wiley Series in Discrete Mathematics and Optimization). ISBN 9780471283669.

63 KUHN, H. W. The Hungarian Method for the assignment problem. **Naval Research Logistics Quarterly**, v. 2, p. 83–97, 1955. DOI: `10.1002/nav.3800020109`.

64 KARMARKAR, N. A new polynomial-time algorithm for linear programming. **Combinatorica**, v. 4, n. 4, p. 373–395, Dec. 1984. ISSN 1439-6912. DOI: `10.1007/BF02579150`.

65 3GPP. **Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects**. [S.l.], Mar. 2010. Available from: <`http://www.3gpp.org/ftp/Specs/html-info/36814.htm`>.

66 ________. **Spatial channel model for Multiple Input Multiple Output (MIMO) simulations**. [S.l.], Dec. 2009. Available from: <`http://www.3gpp.org/ftp/Specs/html-info/25996.htm`>.

67 MICROSOFT. **How much bandwidth does Skype need?** Available from: <`https://support.skype.com/en/faq/FA1417/how-much-bandwidth-does-skype-need`>. Visited on: 12 Feb. 2018.

68 HAYKIN, S. O. **Adaptive Filter Theory**. [S.l.]: Pearson Education, 2013. ISBN 978-0132671453.

69 RODRIGUES, E. B. et al. Resource Allocation and MIMO for 4G and Beyond. In: ed. by Francisco Rodrigo Porto Cavalcanti. 1. ed. New York, USA: Springer, Oct. 2014. Capacity, Fairness, and QoS Trade-Offs in Wireless Networks with Applications to LTE, p. 63–104. ISBN 978-1-4614-8056-3. DOI: `10.1007/978-1-4614-8057-0_4`.

70 COSTA NETO, F. H. **Utility-Based Scheduling Algorithms to Enhance User Satisfaction in OFDMA Systems**. Feb. 2016. Master Thesis – Universidade Federal do Ceará (UFC).

71 MOGENSEN, P. et al. LTE Capacity Compared to the Shannon Bound. In: IEEE VEHICULAR TECHNOLOGY CONFERENCE (VTC - SPRING), Apr. 2007, Dublin, Ireland. **Proceedings [...]** Dublin, Ireland: IEEE, Apr. 2007. P. 1234–1238. DOI: `10.1109/VETECS.2007.260`.

72 HUGHES-HARTOGS, D. **Ensemble Modem Structure for Imperfect Transmission Media**. US, May 1989. US 4833706. Patent.

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