

Resource Management for Rate Maximization with QoE Provisioning in Wireless Networks

Diego A. Sousa, Victor F. Monteiro, Tarcisio F. Maciel,
Francisco Rafael M. Lima, and Francisco Rodrigo P. Cavalcanti

Abstract—In this work, we study the problem of allocating resources in a multi-service cellular network aiming at maximizing the total system rate while providing suitable Quality of Experience (QoE) to the network users. In our formulation, we try to satisfy at least a certain number of users per service plan, which is an important constraint from the mobile network operators' perspective. We manage to reformulate this nonlinear optimization problem as an Integer Linear Problem (ILP), that can be solved by standard methods. However, due to the exponentially high complexity to solve large instances of this problem, we propose and evaluate a suboptimal algorithm with a much lower complexity, called Rate Maximization under Experience Constraints (RMEC), whose main idea is to divide the problem into three smaller subproblems with reduced complexity. By means of computational simulations, we show that our proposed algorithm presents a near optimal performance and outperforms the state-of-art solution of the literature.

Index Terms—Quality of Experience, Radio Resource Allocation, rate maximization, multi-service

I. INTRODUCTION

The circuit-switched voice service was the dominant source of traffic on mobile networks during the first and second generations of cellular systems. The evolution of user devices and the advent of smartphones and tablets had an important influence on the rapid growth of the number of mobile networks users, leading operators to shift their focus to the many resource-demanding data-oriented services present nowadays in mobile networks. Following the expansion of the mobile communications, standards have been evolving to best meet users' needs. In this context, academy and industry are now discussing the 5th Generation (5G) of mobile networks, which shall be released by 2020 [1].

The future trends indicate that more and more equipments will be connected to the mobile networks increasing the system load. 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 [2].

In order to evaluate the Quality of Service (QoS), operators usually make use of objective parameters provided by the network, such as throughput, delay and jitter. However, due to the numerous service types provided by mobile networks

and the different requirements of each service, each user experiences a different level of satisfaction considering the same QoS parameters. Therefore, in order to better evaluate the specific user's requirements, Quality of Experience (QoE) models aim to provide a more subjective measurement of the services quality [3]. Accordingly to [4], the QoE is defined as the perception by the user of the acceptability of a service. The users' QoE is often evaluated in the literature in terms of Mean Opinion Scores (MOSs), which consist in measurements of the QoE subjectively perceived by the users [5].

In [6]–[8], generic models presenting a mathematical relationship between QoS and QoE are proposed. In [6], the IQX hypothesis (exponential interdependency of QoE and QoS) is presented, which consists in 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 [7], [8].

Other works also study and propose service-specific models to relate QoS and QoE, such as [9]–[12], which propose utility functions that map QoS metrics into QoE for web browsing, Voice over IP (VoIP), video streaming and 3D video traffic, respectively.

As aforementioned, higher system rates are highly desired in future mobile networks. 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 we assign the resources 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 [13].

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 [14]–[16]. In [14], the optimal solution is provided as an Integer Linear Problem (ILP) and a low complexity heuristic is proposed. In [15], 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 [16], 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.

The authors are with the Wireless Telecommunications Research Group (GTel) of the Teleinformatics Engineering Department, Federal University of Ceará (UFC), Fortaleza-CE, Brazil. Diego Aguiar Sousa is also with the Federal Institute of Education, Science, and Technology of Ceará, Acaraú-CE, Brazil. Francisco Rafael Marques Lima is also with the Computer Engineering Department, Sobral-CE, UFC.

Digital Object Identifier: 10.14209/jcis.2016.25

Notice that in [14]–[16], 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 [17], [18]. In this context, in [19] 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 [20].

Usually, the works about resource allocation considering QoE measurements intend to maximize the number of users having their minimum QoE requirements satisfied [21], maximize the minimum MOS of the system [22], [23], or maximize the energy efficiency of the system [24]. To the best of our knowledge, no previous works discuss the rate maximization problem subject to satisfaction constraints based on QoE measurements.

In this paper, we consider the problem of maximizing the overall system rate subject to having a minimum number per-service of users satisfied with their minimum QoE requirements. Note that our problem extends the problem studied in [19] since we consider a more holistic manner of measuring the user satisfaction, i.e., QoE instead of QoS. We summarize our main contributions as:

- We analyze 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. We highlight that to the best of our knowledge, this problem was not considered in the previous literature;
- We manage to reformulate the resource allocation problem as an ILP and solve it using standard algorithms;
- We propose a low-complexity suboptimal solution that has near optimal performance and presents high scalability in terms of the problem inputs. Moreover, differently of [19], our algorithm also treats unfeasible instances of the problem producing near feasible solutions.

The rest of this paper is organized as follows. In Section II, we present the system model considered in this work. In Sections III and IV, we describe the studied optimization problem and reformulate it as an ILP, respectively. In Section V, we propose a low complexity algorithm that provides a near optimal solution to our problem. In Sections VI, we evaluate the performance of our suboptimal algorithm and the optimum solution, as well as the performance of a state-of-art algorithm from the literature. In Section VII, we present our conclusions.

II. SYSTEM MODELING

We consider the downlink transmission of an Evolved Node B (eNB) located in the center of a sectorized cell in a Long

Term Evolution (LTE)-like wireless communication system. Each sector of the cell services a set $\mathcal{U} = \{1, 2, \dots, U\}$ of User Equipments (UEs) distributed on its coverage area.

The cell's eNB employs as multiple access scheme a combination of Orthogonal Frequency Division Multiple Access (OFDMA) and Time Division Multiple Access (TDMA), allocating a set $\mathcal{K} = \{1, 2, \dots, K\}$ of time-frequency Resource Blocks (RBs) to the UEs. Due to signaling constraints, an RB is the minimum allocable resource [25]. Each RB consists of a set of Q adjacent subcarriers spaced of Δf Hz in frequency-domain and a set of T consecutive symbols in time-domain, whose total duration corresponds to one Transmission Time Interval (TTI). In our analyses, each RB can be assigned to only one UE per TTI. We assume that the channel response of an RB k should remain approximately constant over its Q subcarriers during the period of one TTI. Thus, for a UE u , the channel response of RB k can be approximated by the complex channel coefficient $h_{u,k}$ of the first symbol transmitted to the UE over the middle subcarrier of the RB.

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. Due to the stochastic behavior of wireless data traffic, the interference modeling and its mitigation in packet-switched systems is a challenging issue. In the literature, some works propose methods to mitigate the interference, such as coordination among eNBs and interference alignment [26]–[30]. On the other hand, for the sake of simplicity, some papers treat the inter-cell interference as part of the Additive White Gaussian Noise (AWGN) added to the received signal in the UE, which is valid when the number of UEs and Base Stations (BSs) increases [31]. More detailed/sophisticated interference modeling and mitigation is out of the scope in this paper. Therefore, we perform our analyses considering a scenario composed by a single sector of a cell and adopt the simplifying assumption that interference is incorporated as part of the AWGN.

We consider that the eNB and the UE are equipped with single antennas, i.e., we consider a Single Input Single Output (SISO) scenario. We also consider that the signal transmitted from the eNB to a UE experiences long-term fading, in terms of average path loss and shadowing, and short-term fading.

In this paper, the Radio Resource Allocation (RRA) problem consists of a snapshot optimization problem which is solved on a TTI basis. Hence, without loss of generality and for simplicity of notation, we will omit any TTI index. We consider that the eNB has a total available power P_t , which is allocated to the RBs during the transmission. We also consider that Equal Power Allocation (EPA) is employed, so that the power p_k allocated to the k^{th} RB is given by $p_k = \frac{P_t}{K}$. The Signal to Noise Ratio (SNR) of UE u in the RB k is given by

$$\gamma_{u,k} = \frac{p_k |h_{u,k}|^2}{\sigma^2}, \quad (1)$$

where the σ^2 denotes the AWGN average power in the frequency band of an RB.

We consider that the eNB employs a link adaptation scheme that allows to transmit with different data rates depending on the estimated SNR $\gamma_{u,k}$ of the UE u at RB k . Therefore, the data rate allocated by the eNB to the UE u in RB k is $r_{u,k} = f(\gamma_{u,k})$, where $f(\cdot)$ represents the link adaptation function that maps SNR to rate values.

The total data rate allocated to a UE u is given by

$$R_u = \sum_{k \in \mathcal{K}_u} r_{u,k}, \quad (2)$$

where $\mathcal{K}_u \subset \mathcal{K}$ is the subset of RBs allocated to the UE.

The UEs' QoE measurements are given in terms of their MOSs, which depend on the multimedia applications used by the UEs. In our analyses, we consider that the MOS of a UE u can be obtained from the UE's data rate, i.e.,

$$\text{MOS}_u = \phi_u(R_u), \quad (3)$$

where $\phi_u(\cdot)$ is an increasing function that maps the achieved rate of an UE u into a MOS value.

Due to the diversity of applications with distinct requirements, we consider that the UEs are separated into different mobile service subscription plans. We define $\mathcal{S} = \{1, 2, \dots, S\}$ as the set of subscription plans and \mathcal{U}_s as the set of subscribers of the service plan $s \in \mathcal{S}$. Moreover, we consider that $\bigcup_{s \in \mathcal{S}} \mathcal{U}_s = \mathcal{U}$ and we also consider that $\bigcap_{s \in \mathcal{S}} \mathcal{U}_s = \emptyset$, i.e., that each user subscribes to only a single service plan.

Despite the fact that our modeling considers an OFDMA-based system, it is worth to note that the analysis proposed in this work could be performed using any multiple access scheme that ensures orthogonality among the resources guaranteeing no intra-cell interference. The models defined in this section are used to introduce the problem investigated in this paper, as well to evaluate their performance.

III. PROBLEM FORMULATION

In this work, we aim to allocate the available RBs in order to maximize the overall system rate, while ensuring that a minimum number m_s of UEs in each service plan s get their QoE requirements met.

We define \mathbf{X} as an $U \times K$ 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. Therefore, our problem can be written as an optimization problem as follows:

$$\max_{\mathbf{X}} \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} r_{u,k} x_{u,k}, \quad (4a)$$

$$\text{s.t.} \sum_{u \in \mathcal{U}} x_{u,k} = 1, \forall k \in \mathcal{K}, \quad (4b)$$

$$\sum_{u \in \mathcal{U}_s} u \left(\phi_u \left(\sum_{k \in \mathcal{K}} r_{u,k} x_{u,k} \right), \varphi_s \right) \geq m_s, \forall s \in \mathcal{S}, \quad (4c)$$

$$x_{u,k} \in \{0, 1\}, \forall u \in \mathcal{U} \text{ and } \forall k \in \mathcal{K}, \quad (4d)$$

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

The problem stated in (4) aims at finding the optimal resource assignment that maximizes the achievable total system rate in the objective function (4a). Constraints (4b) and (4d) guarantee that each RB is assigned to a single UE. Furthermore, (4c) requires that a minimum number m_s of UEs should be satisfied for each service plan s .

IV. OPTIMAL SOLUTION

Notice that (4) is a combinatorial optimization problem with a nonconvex constraint (4c); hence, it has a prohibitive computational complexity [32]. In this section we reformulate the problem stated in (4) into a more tractable form.

In order to convert the minimum MOS constraint into a minimum UE's rate requirement, we need to define a function which maps MOS into rate. Since we consider $\phi_u(\cdot)$ as an increasing function, we can not guarantee its invertibility, unless it is strictly increasing. In order to define a function $\phi^\dagger(\cdot)$ which maps MOS into rate, we adopt the concept of generalized inverse function for increasing functions stated in [33]. Thus, we define $\phi^\dagger(\cdot)$ as

$$\phi^\dagger(\varphi) = \inf\{R \in \mathbb{R} : \phi(R) \geq \varphi\}, \varphi \in \mathbb{R}, \quad (5)$$

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

Therefore, we can define the data rate (ψ_u) that UE u requires to achieve the MOS requirement φ_s of the service plan s as:

$$\psi_u = \phi_u^\dagger(\varphi_s), \quad u \in \mathcal{U}_s, \quad \forall s \in \mathcal{S}. \quad (6)$$

Therefore, we can rewrite the constraint (4c) as

$$\sum_{u \in \mathcal{U}_s} u \left(\sum_{k \in \mathcal{K}} r_{u,k} x_{u,k}, \psi_u \right) \geq m_s. \quad (7)$$

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 $\boldsymbol{\rho}$, we can rewrite (7) into two new constraints, (8c) and (8d), and restate problem (4) as follows:

$$\max_{\mathbf{X}, \boldsymbol{\rho}} \sum_{u \in \mathcal{U}} \sum_{k \in \mathcal{K}} r_{u,k} x_{u,k}, \quad (8a)$$

$$\text{s.t.} \sum_{u \in \mathcal{U}} x_{u,k} = 1, \forall k \in \mathcal{K}, \quad (8b)$$

$$\sum_{k \in \mathcal{K}} r_{u,k} x_{u,k} \geq \psi_u \rho_u, \quad \forall u \in \mathcal{U}, \quad (8c)$$

$$\sum_{u \in \mathcal{U}} p_{s,u} \rho_u \geq m_s, \forall s \in \mathcal{S}, \quad (8d)$$

$$x_{u,k} \in \{0, 1\}, \quad \forall u \in \mathcal{U} \text{ and } \forall k \in \mathcal{K}, \quad (8e)$$

$$\rho_u \in \{0, 1\}, \quad \forall u \in \mathcal{U}, \quad (8f)$$

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

It is noteworthy that in (4c), the functions $u(\cdot)$ and $\phi_u(\cdot)$ are applied over the optimization variables $x_{u,k}$. Since $u(\cdot)$ is neither convex nor concave, and $\phi_u(\cdot)$ is usually nonlinear, the optimal solution of (4) becomes harder to find. Differently,

the constraints (8c) and (8d) are linear, which simplifies the problem structure.

In fact, the problem stated in (8) can be rewritten in a compact form, where the variables are organized as matrices and vectors. Consider that the terms $r_{u,k}$ are organized into a matrix \mathbf{R} with dimensions $U \times K$. Also consider that the variable m_s , for $s \in \mathcal{S}$, is arranged into a column vector \mathbf{m} with length S and let $\boldsymbol{\psi} = [\psi_1 \ \psi_2 \ \dots \ \psi_U]^T$ be a column vector containing the rate requirements of all U UEs in the system. Moreover, we also group the terms $p_{s,u}$ into a matrix \mathbf{P} with dimension $S \times U$. Let us also denote by the operator \odot the Hadamard product, which consists in a element-wise matrix multiplication, and by \otimes the Kronecker product. We also represent by $*$ the Khatri-Rao product, which consists in a column-wise Kronecker product.

Using these definitions, we can rewrite the problem (8) as

$$\max_{\mathbf{X}, \boldsymbol{\rho}} \text{vec}^T\{\mathbf{R}\} \text{vec}\{\mathbf{X}\}, \quad (9a)$$

$$\text{s.t. } (\mathbf{I}_K \otimes \mathbf{1}_U^T) \text{vec}\{\mathbf{X}\} = \mathbf{1}_K, \quad (9b)$$

$$(\mathbf{R}^T * \mathbf{I}_U)^T \text{vec}\{\mathbf{X}\} \geq ((\boldsymbol{\psi} \otimes \mathbf{1}_U^T) \odot \mathbf{I}_U) \boldsymbol{\rho}, \quad (9c)$$

$$\mathbf{P}\boldsymbol{\rho} \geq \mathbf{m}, \quad (9d)$$

$$\mathbf{X} \text{ is a binary matrix and} \quad (9e)$$

$$\boldsymbol{\rho} \text{ is a binary vector,} \quad (9f)$$

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 $\text{vec}\{\mathbf{X}\} = [\mathbf{x}_1^T \ \mathbf{x}_2^T \ \dots \ \mathbf{x}_K^T]^T$, where \mathbf{x}_k denotes the k^{th} column of the matrix \mathbf{X} .

At this point, we can rearrange the optimization variables into a single vector $\mathbf{y} = [\text{vec}^T\{\mathbf{X}\} \ | \ \boldsymbol{\rho}^T]^T$. In order to obtain \mathbf{X} and $\boldsymbol{\rho}$ in terms of \mathbf{y} , we denote $\text{vec}\{\mathbf{X}\} = \mathbf{A}\mathbf{y}$ and $\boldsymbol{\rho} = \mathbf{B}\mathbf{y}$, where $\mathbf{A} = [\mathbf{I}_{UK} \ | \ \mathbf{0}_{UK \times U}]$, $\mathbf{B} = [\mathbf{0}_{0 \times UK} \ | \ \mathbf{I}_U]$, and $\mathbf{0}_{a \times b}$ is a matrix with dimensions $a \times b$ composed by zeros. Thus, we can rewrite (9) as

$$\max_{\mathbf{y}} \text{vec}\{\mathbf{R}\}^T \mathbf{A}\mathbf{y}, \quad (10a)$$

$$\text{s.t. } (\mathbf{I}_K \otimes \mathbf{1}_U^T) \mathbf{A}\mathbf{y} = \mathbf{1}_K, \quad (10b)$$

$$(\mathbf{R}^T * \mathbf{I}_U)^T \mathbf{A}\mathbf{y} \geq ((\boldsymbol{\psi} \otimes \mathbf{1}_U^T) \odot \mathbf{I}_U) \mathbf{B}\mathbf{y}, \quad (10c)$$

$$\mathbf{P}\mathbf{B}\mathbf{y} \geq \mathbf{m}, \quad (10d)$$

$$\mathbf{y} \text{ is a binary vector.} \quad (10e)$$

Notice that in (10), we reduce our optimization variables to a single vector \mathbf{y} , in opposite to (9), where our optimization variables are \mathbf{X} and $\boldsymbol{\rho}$. In order to further simplify the problem (10), we consider the following variables:

$$\mathbf{c} = \mathbf{A}^T \text{vec}\{\mathbf{R}\}, \quad (11a)$$

$$\mathbf{D} = \begin{bmatrix} ((\boldsymbol{\psi} \otimes \mathbf{1}_U^T) \odot \mathbf{I}_U) \mathbf{B} - (\mathbf{R}^T * \mathbf{I}_U)^T \mathbf{A} \\ -\mathbf{P}\mathbf{B} \end{bmatrix}, \quad (11b)$$

$$\mathbf{w} = [\mathbf{0}_{UK}^T \ | \ -\mathbf{m}^T]^T, \quad (11c)$$

and

$$\mathbf{F} = (\mathbf{I}_K \otimes \mathbf{1}_U^T) \mathbf{A}. \quad (11d)$$

Applying (11) into (10), we rewrite the problem (8) as

$$\max_{\mathbf{y}} \mathbf{c}^T \mathbf{y}, \quad (12a)$$

$$\text{s.t. } \mathbf{D}\mathbf{y} \leq \mathbf{w}, \quad (12b)$$

$$\mathbf{F}\mathbf{y} = \mathbf{1}_K, \quad (12c)$$

$$\mathbf{y} \text{ is a binary vector,} \quad (12d)$$

thus finally reformulating the initial optimization problem presented in (4) as the standard ILP in (12) to which standard methods to solve ILPs, such as Branch and Bound (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 [34]. However, the 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 (12) are highly desired.

V. SUBOPTIMAL SOLUTION

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

In order to develop the proposed heuristic, we adopt the divide-and-conquer strategy by dividing the problem into smaller subproblems. Our algorithm is divided into three stages: firstly, we disregard from the users' set \mathcal{U} a certain number of UEs which we choose not to satisfy; secondly, we calculate an initial assignment; finally, we reallocate the RBs between the users in order to ensure that the problem constraints are met. A general overview of our suboptimal algorithm is shown as a flowchart in Fig. 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. 1.

A. Step 1: User Selection

As mentioned before, one of our contributions is the possibility of separating the users in different sets and satisfying a fraction of these users in each set. This constraint is represented by (4c) in the optimization problem (4).

As presented in Section II, the rate of a user u in the RB k depends on its 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 QoE requirement of a UE is, the harder it is to satisfy it since the UE needs more RBs to get satisfied.

Our main goal is to maximize the transmit rate with the constraint of satisfying at least m_s users of each set \mathcal{S} , for all $s \in \mathcal{S}$. The idea of this phase of our suboptimal solution is to select exactly m_s users from each set \mathcal{U}_s . We remark that by selecting more than m_s users on each service, we will satisfy more users than the necessary. Moreover, the total transmit

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 [35], which means that finding a polynomial-time α -approximation algorithm with $\alpha < 2$ is NP-Hard [39].

Considering the same scenario 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 , where $q_j \in \mathbb{N}$ and $q_j \leq B_j$. For further details of GAP-MQ, see [35].

In our analysis, we consider a particular case of the GAP-MQ, where the number of bins is fixed, i.e., all the n_{bins} bins must be used. In [35], the author presents a (1,2)-approximation algorithm to solve this particular problem, where a (α, β) -approximation is defined as an algorithm that provides a solution at least a factor $1/\alpha$ of the optimal solution, while violating the constraints by at most a factor β .

We can relate this problem with (14), by considering that $n_{\text{bins}} = |\mathcal{L}|$, $n_{\text{items}} = |\mathcal{K}|$, $q_u = \psi_u$, and that there is no capacity limits, i.e, $B_u \rightarrow \infty$, for $u \in \mathcal{L}$. The profit and the size in our problem are equal to the user achievable rate in a RB, i.e, $g_{u,k} = s_{u,k} = r_{u,k}$, where $k \in \mathcal{K}$.

We adapt the algorithm presented in [35] in order to find an initial resource allocation. Initially, in block (2) of Fig. 1, we relax the optimization problem stated in (14), by modifying the constraint (14d) into

$$0 \leq \tilde{x}_{u,k} \leq 1, \quad (15)$$

and replacing all $x_{u,k}$ by $\tilde{x}_{u,k}$, for all $u \in \mathcal{L}$ and $k \in \mathcal{K}$. The problem composed by (14a), (14b), (14c) and (15) consists in a Linear Programming (LP), called $\mathbf{LP}(\mathbf{R})$, and provides a fractional assignment matrix $\tilde{\mathbf{X}}$ composed of the terms $\tilde{x}_{u,k}$. This problem can be efficiently solved by many algorithms proposed in the literature, such as the simplex method and interior point methods [32].

Notice that if the problem $\mathbf{LP}(\mathbf{R})$ is unfeasible, so the problem (14) is, which suggests that the original problem (4) has a high probability of being unfeasible since (14) is an approximation of (4). In order to deal with the unfeasibility, in blocks (3) and (4) of Fig. 1, if $\mathbf{LP}(\mathbf{R})$ presents no feasible solution, then we proposed to remove one user of the set \mathcal{L} using the same criterion (13) of the previous step of our suboptimal solution and we proposed to repeat this process until $\mathbf{LP}(\mathbf{R})$ presents a feasible solution. We highlight 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.

Next, in block (5) of Fig. 1, from $\tilde{\mathbf{X}}$ we can estimate the amount ν_u of resources that each user $u \in \mathcal{L}$ will obtain as

$$\nu_u = \left\lceil \sum_{k \in \mathcal{K}} \tilde{x}_{u,k} \right\rceil, \quad (16)$$

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

Therefore, in block (6) of Fig. 1, from $\tilde{\mathbf{X}}$ we create a bipartite graph $G(\mathcal{V}, \mathcal{K}, \mathcal{E})$. 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 $v_{u,n} \in \mathcal{V}$, where $n \in [1, \nu_u]$. The edges linking both sides $(v_{u,n}, k) \in \mathcal{E}$ are added accordingly to the following explanation.

As illustrated in block (7) of Fig. 1, in order to create $G(\mathcal{V}, \mathcal{K}, \mathcal{E})$, for each user $u \in \mathcal{L}$, we consider a counter $c = 0$ and initialize the index of the user's node $n = 1$. We also consider the indices of RBs sorted in non-increasing order of $r_{u,k}$, i.e., $r_{u,1} \geq r_{u,2} \geq \dots \geq r_{u,K}$. For each k and $\tilde{x}_{u,k} > 0$, we accumulate $\tilde{x}_{u,k}$ into the counter c and create a new edge $(v_{u,n}, k)$ with weight equal to $r_{u,k}$. When the value of c becomes greater or equal to one, which means that the user's node $v_{u,n}$ reaches its maximum capacity, we decrement the counter c 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, we add another edge $(v_{u,n}, k)$ with weight equal to $r_{u,k}$.

In our problem, obeying the 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 [35], we compute the minimum weighted matching. This modification in the algorithm proposed in [35] 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 our problem, we select the edges that produce the minimum sum rate 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 [40]. 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.

We denote by $\mathcal{M} \subset \mathcal{E}$ the subset of edges that composes the minimum weighed matching. In order to find \mathcal{M} , we apply the Hungarian algorithm [40] over G . Finally, we initialize \mathbf{X}_{sat} with all elements equal to zero, then for each edge $(v_{u,n}, k) \in \mathcal{M}$ we 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 our suboptimal solution is presented in Algorithm 2.

As an example, consider that three users that must be satisfied are competing for 5 RBs and all of them have the same MOS requirement that implies in a rate requirement of

Algorithm 2 Initial Resource Assignment

```

1: while  $\text{LP}(\mathbf{R})$  has no feasible solution do
2:    $u' \leftarrow \arg \min_{u \in \mathcal{L}} \left\{ \frac{\sum_{k \in \mathcal{K}} r_{u,k}}{\psi_u} \right\}$ 
3:    $\mathcal{L} \leftarrow \mathcal{L} \setminus \{u'\}$ 
4: end while
5: Solve  $\text{LP}(\mathbf{R})$  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:    $c \leftarrow 0$  ▷ Initialize a counter
9:    $n \leftarrow 1$ 
10:  for all  $k \in \mathcal{K}$  sorted in non-increasing order of  $r_{u,k}$  do
11:    if  $\tilde{x}_{u,k} > 0$  then
12:       $c \leftarrow c + \tilde{x}_{u,k}$ 
13:      Add edge  $(v_{u,n}, k) = r_{u,k}$  to  $G$ 
14:      if  $c \geq 1$  then
15:         $c \leftarrow c - 1$ 
16:         $n \leftarrow n + 1$ 
17:      end if
18:      if  $c > 0$  then
19:        Add edge  $(v_{u,n}, k) = r_{u,k}$  to  $G$ 
20:      end if
21:    end if
22:  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  $\mathcal{M}$ 
    
```

512 kbps. Consider also that the users achievable rates in the RBs are given by the 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 $\text{LP}(\mathbf{R})$ we obtain a fractional assignment 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 we construct the bipartite graph G . Notice that applying (16) on $\tilde{\mathbf{X}}$, we calculate the number of nodes that each user will hold in the bipartite graph, namely $\nu_1 = 2$, $\nu_2 = 1$ and $\nu_3 = 3$. Therefore, we create the bipartite graph depicted in Fig. 2 with 6 user nodes and 5 resource nodes. The edges are added to G accordingly to lines 7-23 of the Algorithm 2. For user 1, we initialize $c = 0$ and sort the indices of the RBs in non-increasing order of achievable rate, thus the indices are sorted as [1 2 3 5 4]. Taking the first RB in the sorted list, we get $\tilde{x}_{1,1} = 0.4029 > 0$, so we add an edge $(v_{1,1}, 1) = r_{1,1} = 655$ and add $\tilde{x}_{1,1}$ to c , resulting in $c = 0.4029$. The next resource on the list has $\tilde{x}_{1,2} = 0$, thus we go further to the next RB, 3, which has $\tilde{x}_{1,3} = 1$. We add a new edge $(v_{1,1}, 3) = r_{1,2} = 248$ to \mathcal{E} and add $\tilde{x}_{1,3}$ to c , resulting in $c = 1.4029$. As $c > 1$, it means that the node $v_{1,1}$ reaches its maximum capacity, so we go to the next node, $v_{1,2}$, and decrement the value of c by one, i.e., $c = c - 1 = 0.4029$. As $c > 0$, we add a new edge $(v_{1,2}, 3) = r_{1,3} = 248$ to \mathcal{E} . The next RBs in the sorted list has no assigned portion, thus we pass to the next user. The edge addition to the users 2 and 3 follows the same idea as the user 1. After the creation of the bipartite graph, we compute the edges that compose the minimum weighted matching using

the Hungarian algorithm, denoted by the solid edges in Fig. 2.

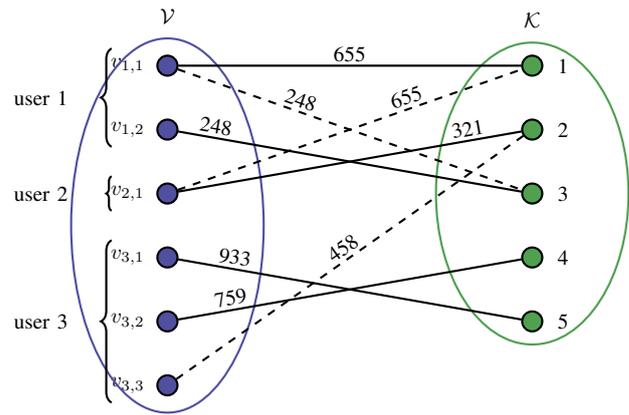


Figure 2: Bipartite Graph with solid edges denoting the maximum matching

Neglecting the edges $(v_{u,n}, k) \notin \mathcal{M}$, we notice that all RB nodes are linked to only one user node, meeting the constraint (8b). From Fig. 2, the initial allocation is given by

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

C. Step 3: Reallocation

In this step of our algorithm, we propose a resource reallocation in order to respect the constraint (14c) rather than achieve higher overall system rates.

As depicted in block (8) of Fig. 1, if all UEs in \mathcal{L} are already satisfied, then the initial solution provided in the previous step is used. Otherwise, we start by removing from \mathcal{L} all users that have already achieved their QoE, then we choose the UE $u \in \mathcal{L}$ that needs more resources to achieve its requirement. Considering the user u , we define an auxiliary set \mathcal{J} that will contain all $k \in \mathcal{K} \setminus \{x_{u,k} = 1\}$, i.e., all resources indices that are not allocated to the user u , as shown in block (9) of Fig. 1. After that, in block (10) of Fig. 1, we create a priority vector $\mathbf{w} = [w_1 w_2 \dots w_K]^T$, 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. 1, we get the index of the RB $k \in \mathcal{J}$ with the highest value w_k and remove it from the auxiliary set \mathcal{J} . We denote by $i \in \mathcal{L}$ the user that owns resource k . In blocks (12) and (13) of Fig. 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 u . If the user u gets satisfied with the received resource, then we go further to the next unsatisfied user, if any. Otherwise, if the user u do not get satisfied, then we continue the process until removing all resources from \mathcal{J} . These steps are depicted in blocks (14) and (15) of Fig. 1. This step of our algorithm finishes when all unsatisfied users are parsed. The algorithm description is presented in Algorithm 3.

VI. PERFORMANCE ANALYSIS

In this section, we evaluate the performance of our algorithm proposed in Section V by comparing it to the

Algorithm 3 Reallocation

```

1: while  $|\mathcal{L}| > 0$  do
2:    $u \leftarrow \arg \min_{u \in \mathcal{L}} \{\psi_u - \sum_{k \in \mathcal{K}} r_{u,k} x_{u,k}\}$ 
3:    $\mathcal{L} \leftarrow \mathcal{L} \setminus \{u\}$ 
4:   if  $\sum_{k \in \mathcal{K}} r_{u,k} x_{u,k} < \psi_u$  then ▷ If user  $u$  is not satisfied
5:      $\mathcal{J} \leftarrow \mathcal{K} \setminus \{x_{u,k} = 1\}$ 
6:      $w_k \leftarrow \frac{r_{u,k}}{\sum_{i \in \mathcal{U}} r_{i,k} x_{i,k}} \quad \forall k \in \mathcal{K}$ 
7:     while  $|\mathcal{J}| > 0$  do
8:        $k \leftarrow$  the resource index  $k \in \mathcal{J}$  with highest  $w_k$ 
9:        $\mathcal{J} \leftarrow \mathcal{J} \setminus \{k\}$ 
10:       $i \leftarrow$  the user that owns the resource  $j$ 
11:      if  $\sum_{j \in \mathcal{K}} r_{i,j} x_{i,j} - r_{i,k} > \psi_i$  then
12:         $x_{i,k} \leftarrow 0$ 
13:         $x_{u,k} \leftarrow 1$ 
14:        if  $\sum_{j \in \mathcal{K}} r_{u,j} x_{u,j} > \psi_u$  then
15:          break
16:        end if
17:      end if
18:    end while
19:  end if
20: end while

```

optimal solution (12) provided in Section IV and to the algorithm presented in [19].

A. Scenario Modeling

The simulations were done in a snapshot-based system-level network simulator and the adopted system model is aligned with 3rd Generation Partnership Project (3GPP) LTE standards [25], [41]. We consider a three-sector hexagonal cell with 1 km radius and an eNB located on its center. We consider that the system operates at a frequency of 2 GHz with a downlink bandwidth of 10 MHz. The system considers $K = 50$ RBs, each composed by $Q = 12$ adjacent subcarriers spaced of $\Delta f = 15$ kHz and by $T = 14$ consecutive symbols. The channel modeling considers path loss, shadowing and Independent and Identically Distributed (IID) small-scale fading (i.e, the IID fading is modeled as a random variable with Zero Mean Circularly Symmetric Complex Gaussian (ZMCSG) distribution). In order to reduce inter-cell interference, we also consider a downtilted antenna that provides horizontal and vertical gain, as in [42]. In Table I, we present the main adopted system parameters.

In all simulations, we consider that the services are web browsing based [9], [45], with MOS given by

$$\phi_u(R_u) = 5 - \frac{578}{1 + \left(\frac{R_u + 541.1}{45.98}\right)^2}, \quad (17)$$

where R_u is given in kbps.

In the following analysis, we compare our proposed algorithm (RMEC) with the optimal solution, obtained by solving (12), and with the RAISES heuristic provided by [19]. The comparisons are performed in terms of outage and total system rate. We consider that an outage event occurs when the constraint of minimum number of satisfied UEs is violated.

Notice that RAISES intends to solve a similar problem to the one treated in this paper, however considering QoS instead of QoE constraints. In order to compare the results of our

Table I: Simulation parameters.

Parameter	Value
Maximum eNB transmit power (P_t)	46 dBm [41]
eNB antenna radiation pattern	Three-sector [41]
Cell radius	1 km
UE speed	3 km/h [25]
Carrier frequency	2 GHz [25]
System bandwidth	10 MHz [41]
Subcarrier bandwidth (Δf)	15 kHz
Number of RBs (K)	50
Number of subcarriers per RB (Q)	12
Number of symbols per RB (T)	14
Path loss ^a	$34.5 + 35 \log_{10}(d)$ [43]
Antenna gain ^b	$G_h(\theta_h) + G_v(\theta_v)$ [42]
Downtilt angle	8 degrees
Log-normal shadowing standard deviation	8 dB [25]
Small-scale fading	IID
AWGN power per sub-carrier	-123.24 dBm
Noise figure	9 dB
Link adaptation	Link level curves from [44]
Traffic model	Full buffer
Transmission Time Interval	1 ms
Number of snapshots	10500

^a d is the distance from the eNB to the UE in meters.

^b θ_h and θ_v represents the horizontal and vertical angles related to the eNB, respectively.

algorithm with those provided in [19], we convert the QoS into QoE metrics, using the function $\phi_u(\cdot)$.

B. Results

In Fig. 3, we present the outage rate as function of the required MOS considering three different number of served UEs, $U = 10, 20$ and 30 . In this analysis, we consider that all users subscribe a single service and that the satisfaction requirement (m_1) is equal to 90% of the total number of UEs.

In Fig. 3a, for $U = 10$ UEs, the outage slightly varies with the required MOS. Moreover, the performance of both RMEC and RAISES algorithm regarding the required MOS remains near optimal. In Fig. 3b, for $U = 20$ UEs, both algorithms presents near optimal outage rates for low MOS values. However, with the increase of the MOS requirement, we notice that the outage rate of the RAISES algorithm increases faster, oppositely to RMEC that continues showing near optimal results. Observe that, for a required MOS of 4.4, the outage rate of RAISES is approximately 3.6 times greater than for RMEC, although the outage still remains on acceptable rates. In Fig. 3c, for $U = 30$ UEs, the outage rate of both algorithms continues presenting near optimal results for low MOS requirements. However, notice that with

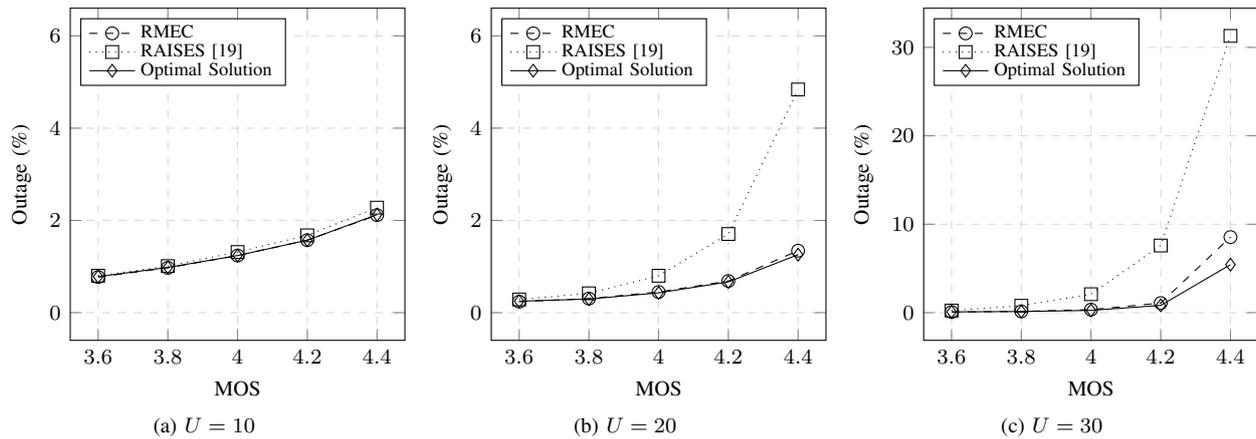


Figure 3: Impact of the number of UEs in the outage.

the increasing MOS requirement, the outage rate of RAISES increases even faster, reaching an unacceptable outage rate of 31.29% for a MOS requirement of 4.4, while RMEC presents 8.54% of outage. Comparing Figs. 3a, 3b and 3c, we observe that for low MOS requirements, the outage rate is slightly smaller when we consider a greater number of served UEs in the system. This happens due to the increase of the UE diversity, i.e., when the number of UEs increases, the chances of choosing the 90% of the UEs with good channel conditions also increase.

In Fig. 4 we analyze the total system rate versus the MOS required by the users' service, considering the same setup of the outage analysis. In order to perform a fair comparison, we consider only the cases where the optimal solution is feasible.

In Fig. 4a, for $U = 10$ UEs, the total system rate of both RAISES and RMEC algorithms are near optimal. Nevertheless, it is worth noting that for a MOS requirement of 4.4, the RMEC algorithm provides a rate 1.3 Mbps higher than RAISES. In Fig. 4b, for $U = 20$ UEs, the performance of RAISES degrades with the increasing required MOS. Comparing this result with the outage rate presented in Fig. 3b, we notice that the RAISES performance decreases with its outage increasing, which indicates that the RAISES algorithm presents a scalability issue with the increasing number of UEs. In Fig. 4c, for $U = 30$ UEs, the performance of RMEC continues presenting near optimal results. Meanwhile, the difference between the total system rate achieved by RMEC and RAISES reaches approximately 17% for a required MOS of 4.4. Comparing this result with Fig. 3c, the scalability issue appears more evidently. We remark that in both outage and total system rate analysis, our proposed algorithm presented better results than RAISES. Furthermore, the RMEC always showed near optimal performance and did not register any scalability issue with the increasing number of UEs.

This difference on the algorithms' performance relies mainly in 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. 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 increases, the capability of the reallocation process of finding a feasible solution decreases. This explains why there is a lack of scalability for RAISES algorithm. On the other hand, our heuristic considers as initial allocation a graph-based rounding of 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 RAISES one. Thus, in the reallocation process, our algorithm finds a near-optimal feasible solution in most of the cases where it exists.

In Fig. 5 we depict the outage rate as function of the required MOS, but in this analysis we evaluate the impact of the minimum number of satisfied UEs, m_1 . We consider three different values for m_1 , namely, 80, 90 and 100% of the total number of UEs. Moreover, we also consider that the eNB serves 20 users subscribed in the same service.

In Fig. 5a, for a requirement of 80% of the UEs, both algorithms follow the optimal solution performance, by presenting almost no outage for all considered MOS requirements. It means that in this scenario, the system has a high probability of serving at least 16 of the 20 UEs with good channel conditions, even when the UEs have high QoE requirements. When the minimum number of satisfied UEs is 90% (18 UEs), as in Fig. 5b, the outage rate increases, but remains in acceptable levels. Notice that with the increasing MOS requirement, while the RMEC algorithm presents a near optimal outage rate, the difference between our proposed algorithm and RAISES grows. In Fig. 5c, we try to satisfy all 20 UEs. In this case, if at least one UE has poor channel conditions, the optimal solution of the allocation problem becomes unfeasible. This lack of UE diversity increases significantly the outage rate. Notice that in this case the RAISES algorithm comes to present an outage rate 14% higher than our proposed algorithm.

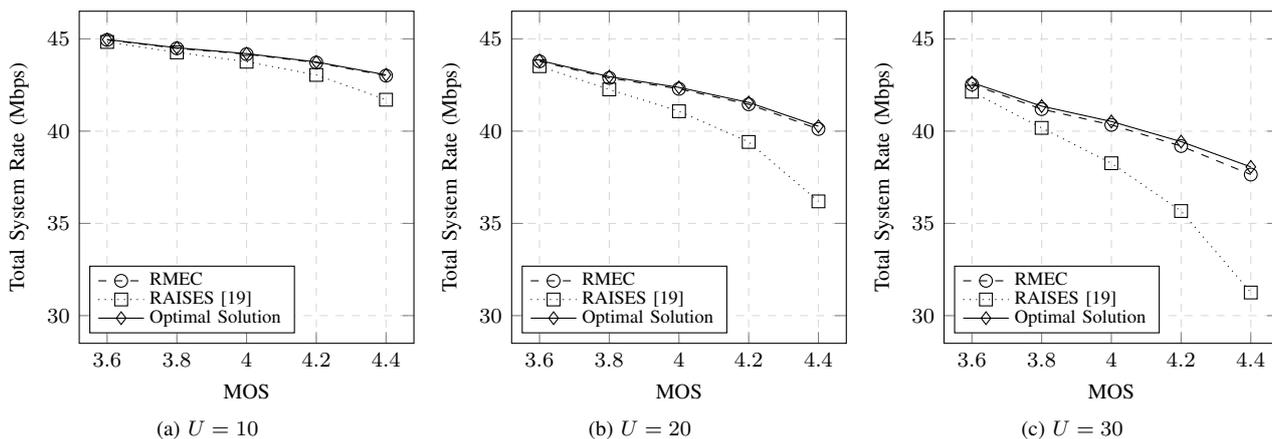


Figure 4: Impact of the number of UEs in the system rate.

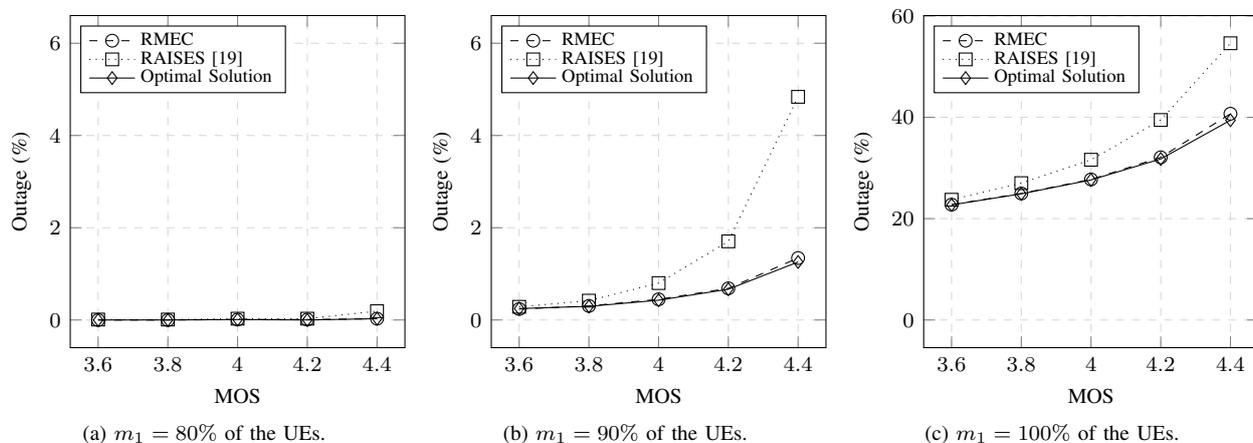


Figure 5: Impact of the satisfaction requirement m_1 in the outage.

In Fig. 6 we analyze the total rate achieved by the system considering the same setup of Fig. 5. In this analysis, we also considered only cases where (12) presented a feasible solution.

In Fig. 6a we notice that for $m_1 = 80\%$ of the number of UEs, both algorithms presented a good performance. However, notice that our proposed algorithm achieves a total system rate closer to the optimal solution than RAISES, specially for high MOS requirements, as we can observe for a MOS requirement of 4.4, where our proposed solution presented a rate approximately 1.6 Mbps higher. By comparing this result with the outage in Fig. 5a, we notice that our proposed solution presents a better result even in scenarios where there is no resource scarcity. In Fig. 6b, for $m_1 = 90\%$ of the number of UEs, our proposed solution presents a near optimal performance. On the other hand, the RAISES algorithm achieves a total system rate approximately optimal for low MOS requirement; however, its performance deteriorates with the increasing of the required MOS. For instance, for a requirement of 4.4, the RAISES algorithm reaches a total system rate approximately 9.8% lower than that of our proposed algorithm. We highlight that by comparing this result with those in Fig. 5b, we notice that the performance degradation is also linked with the outage rate increasing,

which indicates that RAISES also presents a lack of scalability regarding to the minimum number of satisfied users, m_1 . In Fig. 6c, for $m_1 = 100\%$ of the UEs, the total achieved rate of our proposed algorithm also presents a near optimal performance for all considered MOS requirements. In this analysis, we notice that the lack of scalability of RAISES regarding m_1 becomes more evident. Notice that in this case, even for a low MOS requirement of 3.6, the rate achieved by RAISES is 3.8% lower than that of RMEC. For a requirement of 4.4, this difference increases and RAISES achieves a rate approximately 21.1% lower than the one achieved by our proposed algorithm.

Observe that besides of the higher scalability of our algorithm with the number of users in the system, RMEC has also higher scalability regarding to m_1 . Notice that independently of the satisfaction target, m_1 , our algorithm presents results close to the optimal solution. On the other hand, RAISES performance deteriorates with the increase of m_1 . This scalability issue of RAISES regarding to m_1 is also explained by the selection of the initial solution.

In order to verify the near feasibility property of our proposed algorithm, in Fig. 7, we present the average of UEs that achieved their QoE requirements when the optimal

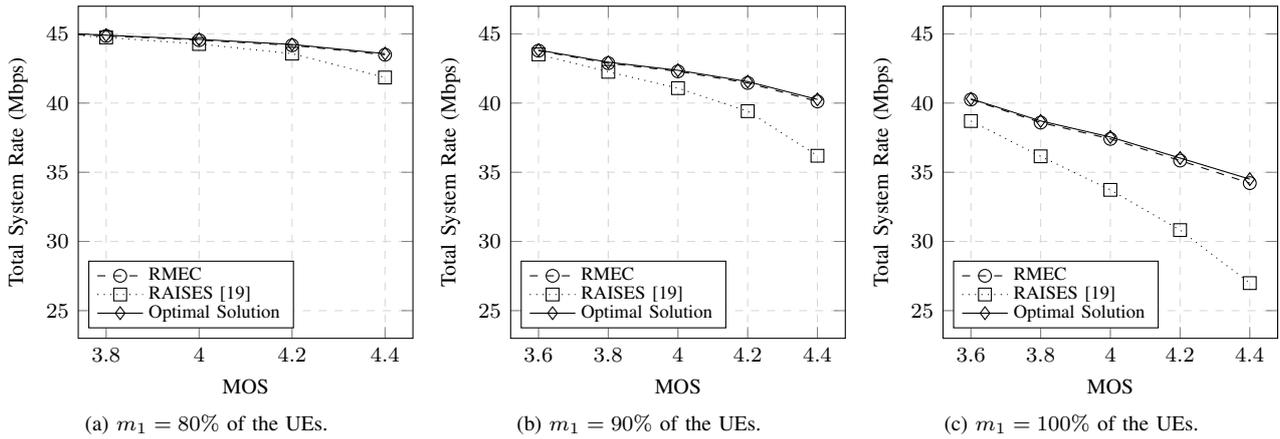


Figure 6: Impact of the satisfaction requirement m_1 in the system rate.

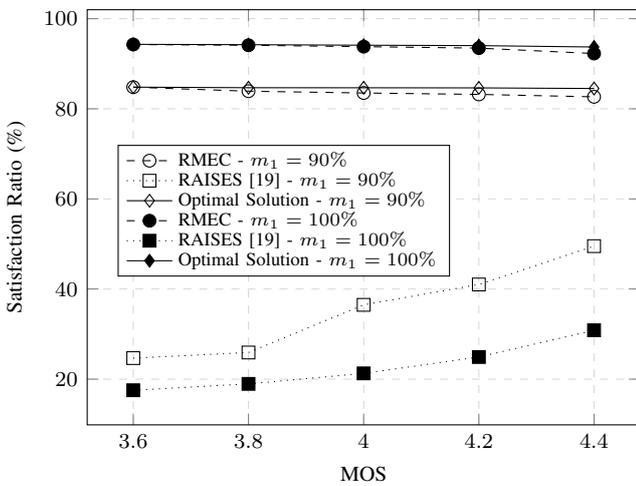


Figure 7: Percentage of satisfied UEs when the optimal solution is not feasible.

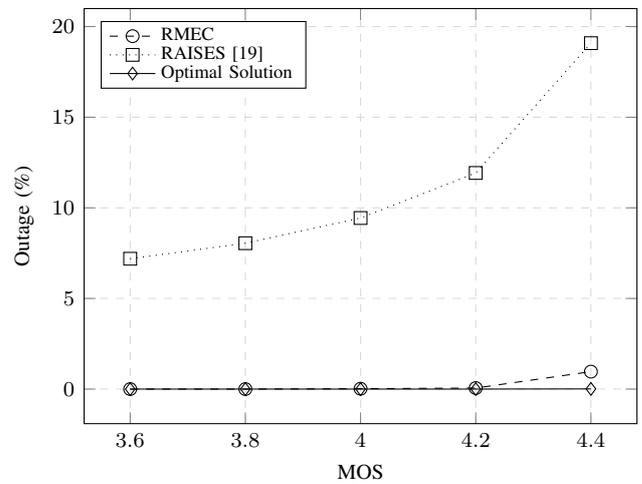


Figure 8: Outage Probability considering 30 UEs and 3 services.

solution is unfeasible. We consider $U = 20$ and $m_1 = 90\%$ and 100% of the UEs. In order to obtain an optimal number of users that must be satisfied, we reduce the requirement m_1 by one and solve (12), until a feasible solution is found.

Notice that our proposed algorithm achieves a solution close to the optimal solution (12). Therefore, our proposed algorithm produces near feasible solutions in unfeasible instances of the allocation problem, unlike RAISES. In situations where the eNB has not enough resources to satisfy the QoE constraints of the m_1 UEs, we remark that the near feasibility property of RMEC is very useful since the heuristic tries to satisfy the higher number of UEs as possible, providing a good resource allocation.

Until now, we evaluated the performance of our algorithm varying the number of users, the minimum number of users that should be satisfied and the minimum user's MOS requirement in a single service scenario. These analyses showed that our algorithm provides a good solution with reduced complexity independently of the magnitude of the evaluated parameters. On the other hand, the benchmarking algorithm, RAISES, presents an scalability issue and presented inefficiency in dealing with users with the poorest channel

conditions, even when a feasible solution exists.

In Figs. 8 and 9, we analyze the outage probability and the total system rate in a scenario with 30 UEs and 3 services. We consider that 5 UEs subscribe the service plan 1, 10 UEs subscribe the service 2 and 15 subscribe the service 3. We also consider that the services 1, 2 and 3 have different targets of minimum number of satisfied UEs, namely, $m_1 = 100\% = 5$ UEs, $m_2 = 90\% = 9$ UEs and $m_3 = 80\% = 12$ UEs. In this analysis, we consider that all services have the same MOS requirement. As in the previous results, in the total system rate analysis we consider only cases where (12) has feasible solution.

In this multi-service scenario, the performance of our algorithm presents much better performance than the RAISES algorithm. Notice that, even for high MOS requirements, the outage of our proposed algorithm remains with low values, unlike RAISES that presents high outage probability even for low MOS requirements. Besides that, our proposed algorithm also provides near optimal total system rate for all analyzed MOS requirements. As we can observe, for a MOS equal to 4.4, our proposed algorithm achieves a total system rate 9.8 Mbps higher than RAISES and an outage 18.1% lower.

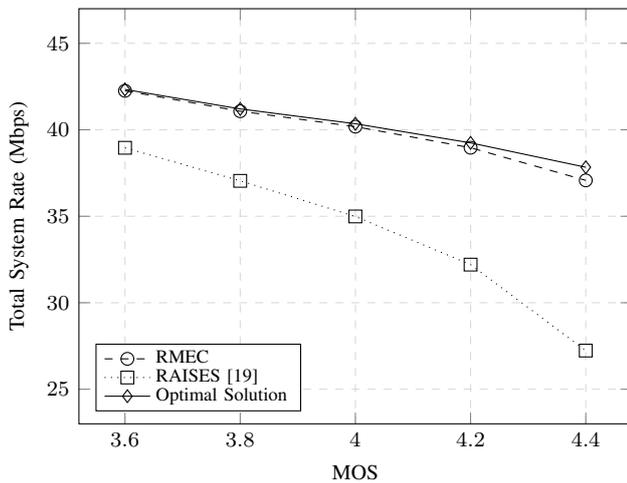


Figure 9: Total System Rate considering 30 UEs and 3 services.

Once more, despite of the increasing number of services, our algorithm provides near-optimal solutions, while RAISES suffers a performance deterioration. Notice that from all analyzed parameters, the number of services in the system was the most harmful to RAISES.

VII. CONCLUSION

In this work, we have studied the problem of maximizing the total system rate, subject to meet the Quality of Experience (QoE) requirements of at least a minimum number of users per service in a multi-service cellular scenario. We remark that this constraint is very important to the mobile network operators.

In this paper, we reformulate our problem as an Integer Linear Problem (ILP), which can be solved by standard methods, like Branch and Bound (BB) or Branch and Cut (BC). However, as the computational complexity to obtain the optimal solution is prohibitive in real-time systems, we proposed a low-complexity suboptimal algorithm, called Rate Maximization under Experience Constraints (RMEC).

In our analysis, the RMEC algorithm presented a near optimal behavior in terms of both total system rate and outage. Furthermore, we have shown that our suboptimal algorithm also provides near feasible solutions to unfeasible instances of the resource allocation problem. We also showed that RMEC outperforms the state-of-art heuristic that intends to solve the same problem, but considering Quality of Service (QoS) constraints instead QoE.

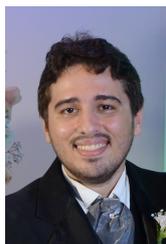
ACKNOWLEDGMENT

The authors acknowledge the technical and financial support from Ericsson Research, Wireless Access Network Department, Sweden and from the Ericsson Innovation Center, Brazil, under EDB/UFC.40 Technical Cooperation Contract. Victor F. Monteiro and F. Rafael M. Lima would like to acknowledge FUNCAP for its scholarship support.

REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. E. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?", *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 1065–1082, 2014. DOI: 10.1109/JSAC.2014.2328098.
- [2] Y. Wang, J. Xu, and L. Jiang, "Challenges of system-level simulations and performance evaluation for 5g wireless networks", *IEEE Access*, vol. 2, pp. 1553–1561, 2014, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2014.2383833.
- [3] S. Aroussi and A. Mellouk, "Survey on machine learning-based qoe-qos correlation models", in *International Conference on Computing, Management and Telecommunications*, 2014, pp. 200–204. DOI: 10.1109/ComManTel.2014.6825604.
- [4] ITU, "Vocabulary for performance and quality of service. amendment 2: new definitions for inclusion in recommendation itu-t p.10/g.100", International Telecommunication Union, Tech. Rep. ITU-T P.10/G.100, Jul. 2008.
- [5] S. Barakovi and L. Skorin-Kapov, "Survey and challenges of qoe management issues in wireless networks", *Journal of Computer Networks and Communications*, vol. 2013, pp. 1–28, 2013. DOI: 10.1155/2013/165146.
- [6] M. Fiedler, T. Hossfeld, and P. Tran-Gia, "A generic quantitative relationship between quality of experience and quality of service", *IEEE Network*, vol. 24, no. 2, pp. 36–41, Mar. 2010. DOI: 10.1109/MNET.2010.5430142.
- [7] P. Reichl, B. Tuffin, and R. Schatz, "Economics of logarithmic quality-of-experience in communication networks", in *Conference on Telecommunications Internet and Media Techno Economics (CTTE)*, Jun. 2010, pp. 1–8. DOI: 10.1109/CTTE.2010.5557702.
- [8] P. Reichl, B. Tuffin, and R. Schatz, "Logarithmic laws in service quality perception: Where microeconomics meets psychophysics and quality of experience", *Telecommunication Systems*, vol. 52, no. 2, pp. 587–600, 2013. DOI: 10.1007/s11235-011-9503-7.
- [9] J. Poncela, G. Gomez, A. Hierrezuelo, F. Lopez-Martinez, and M. Aamir, "Quality assessment in 3g/4g wireless networks", *Wireless Personal Communications*, vol. 76, no. 3, pp. 363–377, Mar. 2014. DOI: 10.1007/s11277-014-1711-5.
- [10] R. G. Cole and J. H. Rosenbluth, "Voice over ip performance monitoring", *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 2, pp. 9–24, 2001, ISSN: 0146-4833. DOI: 10.1145/505666.505669.
- [11] R. Mok, E. Chan, and R. Chang, "Measuring the quality of experience of http video streaming", in *IFIP/IEEE International Symposium on Integrated Network Management*, 2011, pp. 485–492. DOI: 10.1109/INM.2011.5990550.
- [12] I. Politis, L. Dounis, C. Tselios, A. Kordelas, T. Dagiuklas, and A. Papadakis, "A model of network related qoe for 3d video", in *IEEE Globecom Workshops*, 2012, pp. 1335–1340. DOI: 10.1109/GLOCOMW.2012.6477776.
- [13] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in lte cellular networks: key design issues and a survey", *IEEE Communications Surveys and Tutorials*, vol. 15, no. 2, pp. 678–700, 2013, ISSN: 1553-877X. DOI: 10.1109/SURV.2012.060912.00100.
- [14] Y. J. Zhang and K. B. Letaief, "Multiuser Adaptive Subcarrier-and-Bit Allocation with Adaptive Cell Selection for OFDM Systems", *IEEE Transactions on Communications*, vol. 3, no. 5, pp. 1566–1575, Sep. 2004, ISSN: 1536-1276. DOI: 10.1109/TWC.2004.833501.
- [15] T.-H. Lee and Y.-W. Huang, "Resource allocation achieving high system throughput with qos support in ofdma-based system", *IEEE Transactions on Communications*, vol. 60, no. 3, pp. 851–861, 2012, ISSN: 0090-6778. DOI: 10.1109/TCOMM.2012.020912.100632.
- [16] T. He, X. Wang, and W. Ni, "Optimal chunk-based resource allocation for ofdma systems with multiple ber requirements", *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4292–4301, 2014, ISSN: 0018-9545. DOI: 10.1109/TVT.2014.2313354.
- [17] UMTS, "Selection procedures for the choice of radio transmission technologies of the UMTS", Universal Mobile Telecommunications System (UMTS), Sophia Antipolis, France, Tech. Rep. TR 101 112 V3.2.0 - UMTS 30.03, Apr. 1998.
- [18] A. Furuskär, "Radio Resource Sharing and Bearer Service Allocation for Multi-Bearer Service, Multi-Access Wireless Networks", PhD thesis, Royal Institute of Technology (KTH), Radio Communication Systems, Apr. 2003.
- [19] F. R. M. Lima, T. F. Maciel, W. C. Freitas, and F. R. P. Cavalcanti, "Resource Assignment for Rate Maximization with QoS Guarantees

- in Multiservice Wireless Systems”, *IEEE Transactions on Vehicular Technology*, vol. 61, no. 3, pp. 1318–1332, Mar. 2012, ISSN: 0018-9545. DOI: 10.1109/TVT.2012.2183905.
- [20] F. R. M. Lima, T. F. Maciel, W. da Cruz Freitas Jr, and F. R. P. Cavalcanti, “Improved Spectral Efficiency with Acceptable Service Provision in Multi-User MIMO Scenarios”, *IEEE Transactions on Vehicular Technology*, vol. 63, no. 6, pp. 2697–2711, Jul. 2014. DOI: 10.1109/TVT.2013.2293333.
- [21] G. Lee, H. Kim, Y. Cho, and S.-H. Lee, “QoE-aware scheduling for sigmoid optimization in wireless networks”, *IEEE Communications Letters*, vol. 18, no. 11, pp. 1995–1998, 2014, ISSN: 1089-7798. DOI: 10.1109/LCOMM.2014.2354409.
- [22] M. Rugelj, U. Sedlar, M. Volk, J. Sterle, M. Hajdinjak, and A. Kos, “Novel cross-layer qoe-aware radio resource allocation algorithms in multiuser ofdma systems”, *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3196–3208, 2014, ISSN: 0090-6778. DOI: 10.1109/TCOMM.2014.2347288.
- [23] C. Sacchi, F. Granelli, and C. Schlegel, “A qoe-oriented strategy for ofdma radio resource allocation based on min-mos maximization”, *IEEE Communications Letters*, vol. 15, no. 5, pp. 494–496, May 2011, ISSN: 1089-7798. DOI: 10.1109/LCOMM.2011.031411.101672.
- [24] X. Xiao, X. Tao, and J. Lu, “Energy-efficient resource allocation in lte-based mimo-ofdma systems with user rate constraints”, *IEEE Transactions on Vehicular Technology*, vol. 64, no. 1, pp. 185–197, 2015, ISSN: 0018-9545. DOI: 10.1109/TVT.2014.2319078.
- [25] 3GPP, “Physical layer aspects for evolved universal terrestrial radio access (UTRA)”, Third Generation Partnership Project, TR 25.814 V7.1.0, Sep. 2006.
- [26] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, “Scenarios for 5G mobile and wireless communications: The vision of the METIS project”, *IEEE Communications Magazine*, vol. 52, no. 5, pp. 26–35, May 2014, ISSN: 0163-6804. DOI: 10.1109/MCOM.2014.6815890.
- [27] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, “Interference coordination and cancellation for 4g networks”, *IEEE Communications Magazine*, vol. 47, no. 4, pp. 74–81, 2009, ISSN: 0163-6804. DOI: 10.1109/MCOM.2009.4907410.
- [28] S. Deb, P. Monogioudis, J. Miernik, and J. Seymour, “Algorithms for enhanced inter-cell interference coordination (eicc) in lte hetnets”, *IEEE/ACM Transactions on Networking*, vol. 22, no. 1, pp. 137–150, 2014.
- [29] V. Cadambe and S. Jafar, “Interference alignment and spatial degrees of freedom for the k user interference channel”, in *IEEE International Conference on Communications*, 2008, pp. 971–975. DOI: 10.1109/ICC.2008.190.
- [30] D. Aziz, M. Mazhar, and A. Weber, “Multi user inter cell interference alignment in heterogeneous cellular networks”, in *IEEE Vehicular Technology Conference (VTC Spring)*, 2014, pp. 1–5. DOI: 10.1109/VTCSpring.2014.7022994.
- [31] C. Seol and K. Cheun, “A Statistical Inter-Cell Interference Model for Downlink Cellular OFDMA Networks under Log-Normal Shadowing and Multipath Rayleigh Fading”, *IEEE Transactions on Communications*, vol. 57, pp. 3069–3077, 10 Oct. 2009, ISSN: 0090-6778. DOI: 10.1109/TCOMM.2009.10.080152.
- [32] S. Boyd and L. Vandenberghe, *Convex optimization*, 1st. Cambridge University Press, 2004.
- [33] P. Embrechts and M. Hofert, “A note on generalized inverses”, *Mathematical Methods of Operations Research*, vol. 77, no. 3, pp. 423–432, 2013, ISSN: 1432-5217. DOI: 10.1007/s00186-013-0436-7.
- [34] G. Nemhauser and L. Wolsey, *Integer and Combinatorial Optimization*. John Wiley & Sons, 1999.
- [35] S. O. Krumke and C. Thielen, “The generalized assignment problem with minimum quantities”, *European Journal of Operational Research*, vol. 228, no. 1, pp. 46–55, 2013, ISSN: 0377-2217. DOI: 10.1016/j.ejor.2013.01.027.
- [36] D. B. Shmoys and E. Tardos, “An approximation algorithm for the generalized assignment problem”, *Mathematical Programming*, vol. 62, no. 3, pp. 461–474, 1993, ISSN: 0025-5610. DOI: 10.1007/BF01585178.
- [37] S. Martello and P. Toth, *Knapsack problems: Algorithms and computer implementations*, ser. Wiley-Interscience series in discrete mathematics and optimization. J. Wiley & Sons, 1990, ISBN: 9780471924203.
- [38] D. G. Cattrysse and L. N. V. Wassenhove, “A survey of algorithms for the generalized assignment problem”, *European Journal of Operational Research*, vol. 60, no. 3, pp. 260–272, 1992, ISSN: 0377-2217. DOI: 10.1016/0377-2217(92)90077-M.
- [39] A. Koster and X. Muñoz, *Graphs and Algorithms in Communication Networks: Studies in Broadband, Optical, Wireless and Ad Hoc Networks*, ser. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2009, ISBN: 9783642022500.
- [40] H. W. Kuhn, “The hungarian method for the assignment problem”, *Naval Research Logistics Quarterly*, vol. 2, pp. 83–97, 1955. DOI: 10.1002/nav.3800020109.
- [41] 3GPP, “Evolved universal terrestrial radio access (E-UTRA); further advancements for E-UTRA physical layer aspects”, 3rd Generation Partnership Project (3GPP), TR 36.814, Mar. 2010. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/36814.htm>.
- [42] F. Gunnarsson, M. Johansson, A. Furuskär, M. Lundevall, A. Simonsson, C. Tidestav, and M. Blomgren, “Downtilted base station antennas - a simulation model proposal and impact on HSPA and LTE performance”, in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Sep. 2008, pp. 1–5. DOI: 10.1109/VETECF.2008.49.
- [43] 3GPP, “Spatial channel model for multiple input multiple output (MIMO) simulations”, Third Generation Partnership Project, TR 25.996 V7.0.0, Jun. 2007.
- [44] C. Mehlführer, M. Wrulich, J. C. Ikuno, D. Bosanska, and M. Rupp, “Simulating the long term evolution physical layer”, in *Proceedings of the European Signal Processing Conference*, Glasgow, Scotland, Aug. 2009, pp. 1471–1478.
- [45] J. Navarro-Ortiz, J. M. Lopez-Soler, and G. Stea, “Quality of experience based resource sharing in ieee 802.11e hcca”, in *European Wireless Conference (EW)*, Apr. 2010, pp. 454–461. DOI: 10.1109/EW.2010.5483493.



Diego Aguiar Sousa received the B.Sc. degree in Computer Engineering in University of Ceará (UFC), Sobral, Brazil, in 2011. In 2013, he received the M.Sc. degree in Telecommunications Engineering from the UFC, Fortaleza, Brazil. Since 2013, he has been a Ph.D. student and researcher at the Wireless Telecom Research Group (GTEL), UFC, participating of projects in a technical and scientific cooperation with Ericsson Research. Also, since 2013, he has been a Professor of the Federal Institute of Education, Science, and Technology of Ceará (IFCE), Acaraú, Brazil. His research interests include numerical optimization, 5G wireless communication networks, coordinated scheduling, radio resource allocation for QoS/QoE provisioning.



Victor Farias Monteiro received the double B.Sc. degree in General Engineering, from the École Centrale Lyon, France, and in Telecommunications Engineering (magna cum laude), from the Federal University of Ceará (UFC), Fortaleza, Brazil, in 2013. In 2015, he received the M.Sc. degree in Telecommunications Engineering from UFC. He is currently a Ph.D. student and researcher at the Wireless Telecom Research Group (GTEL), UFC, where he works in projects in cooperation with Ericsson Research. In 2016, he was in an internship at Ericsson Research in Luleå, Sweden, where he studied challenges involving Radio Access Technologies (RAT) scheduling in a 5G multi-RAT scenario. From 2010 to 2012, he took part, in France, of the Eiffel Excellence Scholarship Programme, established by the French Ministry of Foreign Affairs. In 2012, he was intern at Gemalto, Paris, France, working on Machine-to-Machine technology for photovoltaic systems. His research interests include numerical optimization, 5G wireless communication networks and radio resource allocation algorithms for QoS/QoE provisioning.



Tarcisio Ferreira Maciel received his B.Sc. and M.Sc. degrees in Electrical Engineering from the Federal University of Ceará (UFC) in 2002 and 2004, respectively, and his Dr.-Ing. degree from the Technische Universität Darmstadt (TUD), Germany, in 2008, also in Electrical Engineering. Since 2001, he has actively participated in several projects in a technical and scientific cooperation between Wireless Telecom Research Group (GTEL), UFC, and Ericsson Research. From 2005 to 2008, he was a research assistant with the Communications Engineering Laboratory, TUD. Since 2008, he has been a member of the Post-Graduation Program in Teleinformatics Engineering, UFC. In 2009, he was a Professor of computer engineering with UFC-Sobral and since 2010, he has been a Professor with the Center of Technology, UFC. His research interests include radio resource management, numerical optimization, and multiuser/multiantenna communications.



Francisco Rafael Marques Lima received the B.Sc. degree with honors in Electrical Engineering in 2005, and M.Sc. and D.Sc. degrees in Telecommunications Engineering from the Federal University of Ceará, Fortaleza, Brazil, in 2008 and 2012, respectively. In 2008, he has been in an internship at Ericsson Research in Luleå, Sweden, where he studied scheduling algorithms for LTE system. Since 2010, he has been a Professor of Computer Engineering Department of Federal University of Ceará, Sobral, Brazil. Prof. Lima is also a researcher at the Wireless Telecom Research Group (GTEL), Fortaleza, Brazil, where he works in projects in cooperation with Ericsson Research. He has published several conference and journal articles as well as patents in the wireless telecommunications field. His research interests include radio resource allocation algorithms for QoS guarantees in scenarios with multiple services, resources, antennas and users.



Francisco Rodrigo Porto Cavalcanti received the B.Sc. and M.Sc. degrees in Electrical Engineering from Federal University of Ceará (UFC), Fortaleza, Brazil, in 1994 and 1996, respectively, and the D.Sc. degree in Electrical Engineering from the State University of Campinas, São Paulo, Brazil, in 1999. Upon graduation, he joined the UFC, where he is currently an Associate Professor and holds the Wireless Communications Chair with the Department of Teleinformatics Engineering. In 2000, he founded and, since then has directed the Wireless Telecom Research Group (GTEL), which is a research laboratory based on Fortaleza, which focuses on the advancement of wireless telecommunications technologies. At GTEL, he manages a program of research projects in wireless communications sponsored by the Ericsson Innovation Center in Brazil and Ericsson Research in Sweden. Prof. Cavalcanti has produced a varied body of work including two edited books, conference and journal papers, international patents and computer software dealing with subjects such as radio resource allocation, cross-layer algorithms, quality of service provisioning, radio transceiver architectures, signal processing and project management. Prof. Cavalcanti is a distinguished researcher of the Brazilian Scientific and Technological Development Council for his technology development and innovation record. He also holds a Leadership and Management professional certificate from the Massachusetts Institute of Technology, Cambridge.