

# Optimization of Multitenant Radio Admission Control through a Semi-Markov Decision Process

Jordi Pérez-Romero, Oriol Sallent

**Abstract**—Network slicing in future 5G systems enables the provision of multitenant networks in which a network infrastructure owned by an operator is shared among different tenants, such as mobile virtual operators, over-the-top providers or vertical market players. The support of network slicing within the radio access network requires the introduction of appropriate radio resource management functions to ensure that each tenant gets the required radio resources in accordance with the expected service level agreement (SLA). This paper addresses radio admission control (RAC) functionality in multiservice and multitenant scenarios as a mechanism for regulating the acceptance of new guaranteed bit rate service requests of different tenants. This paper proposes an optimization framework that models the RAC as a semi-Markov decision process and, as a result, derives an optimal decision-making policy that maximizes an average long-term function representing the desired optimization target. A reward function is proposed to capture the degree of tenant satisfaction with the received service in relation to the expected SLA, accounting for both the provision of excess capacity beyond the SLA and the cost associated with sporadic SLA breaches. The proposed approach is evaluated by means of simulations, and its superiority to other reference schemes in terms of reward and other key performance indicators is analyzed.

**Index Terms**—5G, multitenancy, radio admission control, RAN slicing, semi-Markov decision process.

## I. INTRODUCTION

NETWORK slicing, which allows the formation of multiple dedicated logical networks with specific functionality running on top of a common infrastructure, is a key capability of fifth generation (5G) systems [1] for enabling the simultaneous support of a wide range of application scenarios (e.g., the automotive industry, utilities, smart cities, and high-tech manufacturing) and for consolidating more flexible and cost-efficient service delivery models (e.g., neutral host network providers, Network as a

Service, and private cellular networks). Each of these logical networks is referred to as a *network slice* and can be tailored to provide a particular system behavior to best support specific service/application domains. Therefore, network slicing enables the provision of multitenant networks in which a network infrastructure owned and deployed by an operator is shared among different organizations or business entities such as mobile virtual operators, over-the-top providers or vertical market players, referred to in general as *tenants*, who are entitled to use a network slice with a given level of guaranteed network resources.

In the most general case, the realization of a network slice involves the support of specific features and resources both in the 5G core network part and in the next generation radio access network (NG-RAN) part. The latter is referred to as a radio access network (RAN) slice. The issue of how the pool of radio resources available to one gNB (i.e., a NG-RAN node) can be configured and operated to simultaneously deliver multiple and diverse behaviors [2] and achieve isolation between slices (meaning that the traffic in one network slice should not negatively affect the performance seen by other slices) makes the realization of RAN slices particularly challenging. This isolation in the NG-RAN may be achieved by means of implementation-dependent radio resource management (RRM) policies [3], and there exists a general trade-off between the level of isolation and the efficiency of the radio resource utilization [4].

The split of the available radio resources among RAN slices in multitenant scenarios has received attention in recent years in the literature. This split can be performed at various levels with the support of different RRM functions [4]. A significant number of studies have focused on packet scheduling (PS) to carry out the dynamic assignment of radio resources to the users of each RAN slice. The network virtualization substrate (NVS) concept, designed as a hierarchical scheduler composed of a slice scheduler to define the resource allocation to a slice and a flow scheduler to decide the assignment of resources to the different users of the slice, is used in [5][6][7][8]. The CellSlice approach is presented in [9] and acts as a gateway-level solution that constrains the uplink scheduling decisions and uses NVS for the downlink. In [10], an application-oriented framework is proposed to optimize the allocation of resources to the different operators sharing the RAN. Recently, the Hap-SliceR framework has been proposed in [11] to optimize the allocation of radio resources to slices using a Q-learning algorithm, while a game theory-based

Copyright (c) 2019 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org)

Date of submission: March, 22nd, 2019. Revisions: July, 17th, 2019 and September, 13th, 2019. Accepted October 27th, 2019.

This work has been supported by the EU funded H2020 5G-PPP project 5G ESSENCE under the grant agreement 761592 and by the Spanish Research Council and FEDER funds under SONAR 5G grant (ref. TEC2017-82651-R).

The authors are with the Department of Signal Theory and Communications at Universitat Politècnica de Catalunya (UPC), Barcelona, Spain (e-mail: [jorperez,sallent]@tsc.upc.edu).

approach is used in [12][13] to model the RAN slicing problem and to determine the allocation of resources to the users of each slice.

Some other studies have considered the allocation of resources to RAN slices rather than to individual users of each slice. For example, a heuristic algorithm integrating spectrum allocation, admission control for slice requests and spatial multiplexing is proposed in [14] to allocate resources to each slice. Similarly, [15] proposes the 5G network slice broker concept as a mechanism that facilitates on-demand resource allocation and performs slice admission control. This concept was further developed in [16], which introduces a slice admission control algorithm with traffic forecasting capabilities, as well as in [17] and [18], which study the optimization of admission control for slice requests that need to support a given number of users for a known period of time.

The consideration of radio admission control (RAC) in multitenant scenarios as a mechanism to regulate the number of admitted users of guaranteed bit rate (GBR) services for each slice or tenant has received less attention in the literature. For example, in [8], a tenant-specific admission control is included in the NVS framework of [5]. It is based on comparing the GBR load of a tenant to a threshold that accounts for some resources reserved for that tenant. In turn, the game theory-based network slicing approach of [12] considers two possible admission control strategies, a worst-case admission control and a load-driven admission control, that aim to ensure that the subsequent resource allocation algorithm will be able to find a solution that meets the rate requirements of GBR services. Similarly, a multitenant admission control for cellular networks that exploits traffic multiplexing principles at both intra- and multicellular levels was proposed in [19]. It is based on a heuristic algorithm that compares the GBR load of each tenant against a tenant-specific threshold that is dynamically adjusted based on the loads of the tenants in the different cells.

In this context, this paper focuses on the optimization of the RAC decision-making policy for admitting users in a multitenant and multiservice 5G scenario. For this purpose, RAC is modeled as a semi-Markov decision process (SMDP). The use of the SMDP as an optimization tool has been studied in different contexts in the literature. For example, SMDPs were used in the past to derive optimal circuit access policies for Integrated Service Digital Networks in [20]. Later, it was used in [21] to optimize the air interface selection in heterogeneous networks. More recently, the SMDP has been used in the network slicing domain. In particular, [11] uses the SMDP to decide the allocation of resource units to slices in different cells, while [17][18] have used SMDPs to optimize the admission of slice requests. However, none of these studies considers RAC decision making in multitenant networks, which deals with the dynamics associated with user-level traffic generation processes and the congestion that can appear due to the varying number of users and the corresponding propagation conditions.

The contributions of this paper can be summarized as follows:

- This paper proposes an SMDP model for optimizing the RAC decision making policy in a multitenant and multiservice 5G scenario, which is something that, to the authors' best knowledge, has not been addressed in the literature.
- The model is defined in accordance with the network slicing capabilities of 5G, which allow the assignment of a different RAN slice to each tenant and the enforcement of different policies on different slices when admitting the users of a tenant according to the tenant's service level agreement (SLA).
- As an optimization target, the paper proposes a reward function that captures the degree of satisfaction of the tenants with the received service in relation to their expected SLA. The proposed reward function includes different parameters to account for the provision of excess capacity beyond the SLA and for the cost associated with sporadic SLA breaches. Then, through the setting of these parameters, it is possible to balance the trade-off between a more conservative RAC policy that targets a strict fulfillment of the SLA and a more flexible RAC policy that targets a more efficient utilization of the radio resources.
- The proposed SMDP-based RAC policy is evaluated by means of simulations and compared against other reference schemes in order to analyze its behavior in terms of obtained reward and in terms of different key performance indicators such as blocking probability, throughput or congestion probability.

The rest of the paper is organized as follows. Section II introduces the considered system model to characterize the radio resources and the system dynamics in the multitenant and multiservice 5G scenario. Based on this, Section III presents the formulation of the RAC as an SMDP problem, the algorithmic solution to this problem based on the value iteration method and the formulation of the proposed reward function to reflect the degree of fulfillment of the SLA. A performance evaluation of the proposed approach by means of system-level simulations is presented in Section IV. Finally, concluding remarks and future work directions are highlighted in Section V.

## II. SYSTEM MODEL

Let us assume a 5G network operated by an infrastructure provider and configured to support  $Q$  different network slices. Each slice is associated with a different tenant, and therefore, the terms slice and tenant will be used interchangeably throughout this paper. Each slice is identified through the Single Network Slice Selection Assistance Information (S-NSSAI) [1] and provides end-to-end connectivity between the user equipment (UE) of a tenant and a data network.

A network slice is composed of one network slice subnet instance that includes the NG-RAN functions, referred to as the RAN slice, and another one with the 5G core network functions, referred to as the 5G core slice. Following the Third Generation Partnership Project (3GPP) management model [22], the creation and operation of the end-to-end slice is

carried out by a network slice management function (NSMF) residing at the 3GPP management system, as illustrated in Fig. 1. In turn, the NSMF relies on one network slice subnet management function (NSSMF) for creating and operating the RAN slice, referred to here as the RAN slicing management function (RSMF), and on another one for the core network, referred to as the 5G core slicing management function.

The focus of this paper is on the RAN slice. For this reason, let us assume a 5G new radio (NR) gNB of the NG-RAN with a cell that supports the  $Q$  slices. The cell includes a total of  $P$  resource blocks (RBs), each with bandwidth  $B$ .

Tenant  $q$  uses its associated slice to provide  $S_q$  services numbered  $s=1\dots S_q$  to its users. Following the quality of service (QoS) model of the 5G system [1], which defines the QoS flow as the finest granularity of QoS differentiation in the 5G system architecture, and assuming GBR services, the requirements of service  $s$  of tenant  $q$  are specified in terms of the guaranteed flow bit rate (GFBR), denoted as  $GFBR_{q,s}$ .

The contractual relationship between a tenant and the infrastructure provider is defined in terms of an SLA that records a common understanding about the service offered by the infrastructure provider, together with the measurable target values characterizing the level of the offered service. The SLA can be specified in terms of different parameters depending on the services involved. This paper assumes that the SLA of a tenant is defined in terms of the contractual aggregate GFBR to be guaranteed for all the QoS flows of the tenant in the cell, referred to as the slice aggregated guaranteed bit rate (SAGBR). The corresponding value for tenant  $q$  is denoted  $SAGBR_q$ .

To configure the operation of the RAN slice in the cell according to the SLA, and following the Network Resource Model specified in [23], this paper assumes that the RSMF uses the so-called *RRMPolicy* attribute, which is implementation-dependent, to configure the  $SAGBR_q$  value for the RAN slice of tenant  $q$  in the gNB. This is done through the management interface between RSMF and gNB shown in Fig. 1.

To model the dynamic traffic generation process, it is assumed that the UEs of the  $s$ -th service of the  $q$ -th tenant establish QoS flows following a Poisson arrival process with rate  $\lambda_{q,s}$  requests/s and that the QoS flow durations are exponentially distributed with average  $(1/\mu_{q,s})$  s.

Fig. 1 illustrates the main steps of the UE-initiated QoS flow establishment and the architectural entities of the 5G system architecture involved based on [24]. This process is executed as part of the protocol data unit (PDU) session establishment procedure, in which an IP connection is created between the UE and a data network including one QoS flow. In step number 1 of Fig. 1, the UE sends a *PDU session establishment request* message. This message belongs to the non-access stratum (NAS) control protocol defined between the UE and the access and mobility management function (AMF), which is the network function of the 5G core network in charge of handling the control signaling with the UEs. It is sent to the gNB encapsulated in a radio resource control (RRC) message and transferred through the 5G NR interface

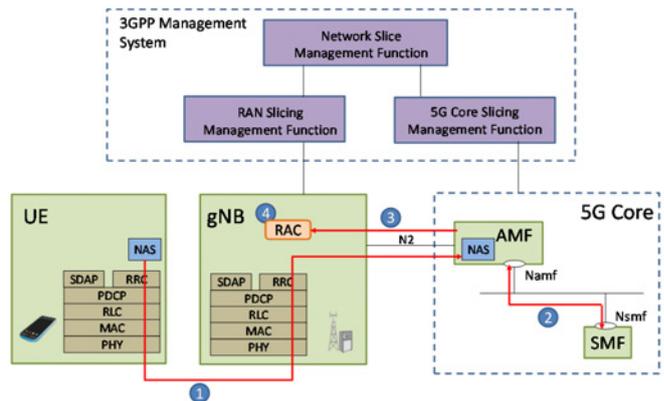


Fig. 1. Architectural model and interactions between involved entities.

using the different layers of the control plane protocol stack, namely, the packet data convergence protocol (PDCP), radio link control (RLC), medium access control (MAC) and physical (PHY) layers. The gNB forwards this message to the AMF through the N2 interface. In the *PDU session establishment request* message, the UE identifies the associated slice through the S-NSSAI. In step number 2 of Fig. 1, the AMF interacts with the session management function (SMF) of the 5G core network in order to create the session management context with the new QoS flow. At this point, the SMF determines the GFBR of the QoS flow in accordance with the user subscription profile. For this purpose, it may interact with other network functions of the 5G core that are not shown here for the sake of simplicity. Interactions between the AMF and the SMF are performed through their service-based interfaces Namf and Nsmf, respectively.

To set up resources for the new QoS flow in the radio interface, in step number 3 shown in Fig. 1, the AMF sends a *PDU session resource setup request* message to the gNB through the N2 interface, indicating the S-NSSAI and the GFBR of the QoS flow being established. In this way, the gNB is aware of the network slice associated with the new QoS flow and can apply a differentiated treatment depending on its slice and the associated SLA.

At the gNB, the new service data adaptation protocol (SDAP) introduced in 5G NR maps the QoS flow to a data radio bearer (DRB) that enables data transfer through the radio interface between the UE and the gNB according to the expected QoS. Finally, in step number 4 of Fig. 1, the RAC function is triggered at the gNB to determine if the DRB associated with the new QoS flow can be admitted or not. The proposed RAC will act based on the SAGBR configured through the *RRMPolicy* attribute to enforce different policies on each slice in accordance with the SLA.

In general terms, the RAC decision should take into account the QoS requirements of the different services and the number of QoS flows that have been already admitted to the cell for each service and tenant. This information is captured in the *state* of the cell at a certain time, which is defined by the matrix  $\mathbf{X}=\{n_{q,s}\}$ , where  $n_{q,s}$  is the number of admitted QoS flows of the  $s$ -th service of the  $q$ -th tenant.  $\mathbf{X}$  is a  $Q \times S$  matrix

with  $S = \max(S_1, \dots, S_Q)$ . To indicate that a tenant can have a number of services  $S_q$  lower than  $S$ , it is assumed that  $n_{q,s} = 0$  and  $\lambda_{q,s} = 0$  for  $s > S_q$ . Therefore, the state space  $\mathcal{X}$  is defined as:

$$\mathcal{X} = \left\{ \mathbf{X} = \{n_{q,s}\} \in \mathbb{N}^{Q \times S} : 0 \leq n_{q,s} \leq N_{\max,q,s} \right\} \quad (1)$$

where  $N_{\max,q,s}$  is the maximum possible number of QoS flows for the  $s$ -th service of the  $q$ -th tenant, defined as:

$$N_{\max,q,s} = \begin{cases} \left\lfloor \frac{P \eta_{\max} \cdot B}{GFBR_{q,s}} \right\rfloor & 1 \leq s \leq S_q \\ 0 & S_q < s \leq S \end{cases} \quad (2)$$

where  $\eta_{\max}$  (b/s/Hz) is the maximum spectral efficiency that can be achieved in the radio interface. It is worth mentioning that  $N_{\max,q,s}$  is an upper bound for the number of QoS flows that can be admitted for each service and tenant, while the RAC function can introduce stricter limitations on the actual number of admissible users depending on the considered RAC policy.

### III. FORMULATION OF THE RADIO ADMISSION CONTROL AS A SEMI-MARKOV DECISION PROCESS

#### A. Problem formulation

According to the traffic generation process defined in Section II, the dynamic evolution of the system follows a Markov chain in which state transitions occur due to either the generation of new QoS flows that are admitted by the RAC function or the finalization of existing QoS flows. The RAC function makes an acceptance or rejection decision each time that a new QoS flow is generated. Therefore, the time between RAC decisions is a continuous random variable related to the Poisson QoS flow generation process. Taking into consideration these characteristics, i.e., the Markovian dynamics and the random time between decisions, the RAC can be formulated as an SMDP [25].

An SMDP can be seen as an extension of a Markov decision process (MDP), which is a versatile and powerful tool for analyzing probabilistic sequential decision processes with an infinite planning horizon. MDPs have become a fundamental formalism for decision-theoretic planning, reinforcement learning and other learning problems in stochastic domains. They are useful for modeling dynamic systems evolving over time where the probabilistic law of motion can be controlled by making decisions (actions) that lead to state changes and earning rewards (or incurring costs) as a consequence of these decisions [25]. SMDPs extend MDPs to deal with situations in which the times between decision instants are not constant but random, as occurs with the RAC problem considered here.

In general, an SMDP model is defined through the following components: (i) the state space  $\mathcal{X}$ ; (ii) the action space  $\mathcal{A}_X$ , which includes the set of possible actions that can be taken in each state  $\mathbf{X}$ ; (iii) the sojourn time in state  $\mathbf{X}$  under action  $\mathbf{A}$ ,

denoted  $\tau(\mathbf{X}, \mathbf{A})$ , which corresponds to the average time that the modeled system will remain in state  $\mathbf{X}$  before changing to a new state when  $\mathbf{A}$  is the selected action; (iv) the state transition probability from state  $\mathbf{X}$  to another state  $\mathbf{Y}$  assuming that  $\mathbf{A}$  is the action selected in state  $\mathbf{X}$ , denoted  $P_{\mathbf{X}\mathbf{Y}}^{\mathbf{A}}$ ; and (v) the reward rate obtained until the next decision time if action  $\mathbf{A}$  is selected in state  $\mathbf{X}$ , which is denoted  $r(\mathbf{X}, \mathbf{A})$ .

In the following, the abovementioned components of the SMDP model are particularized for the RAC decision problem considered in this paper. The associated notation is summarized in Table I.

TABLE I.  
NOTATION USED IN RAC FORMULATION

Notation	Definition
$q$	Tenant identifier. Range $q=1, \dots, Q$
$s$	Service identifier. Range $s=1, \dots, S_q$
$Q$	Number of tenants
$S_q$	Number of services of tenant $q$
$\lambda_{q,s}$	QoS flow generation rate of the $s$ -th service of the $q$ -th tenant
$(1/\mu_{q,s})$	Average duration of the $s$ -th service of the $q$ -th tenant
$GFBR_{q,s}$	Guaranteed flow bit rate of the $s$ -th service of the $q$ -th tenant
$SAGBR_q$	Aggregate GFBR to be guaranteed for all the QoS flows of the $q$ -th tenant in the cell
$P$	Number of RBs in the cell
$B$	Bandwidth of an RB
$\eta_{\max}$	Maximum spectral efficiency achievable in the radio interface
$n_{q,s}$	Number of admitted QoS flows of the $s$ -th service of the $q$ -th tenant
$\mathbf{X} = \{n_{q,s}\}$	State with the number of admitted QoS flows for all tenants and services in a $Q \times S$ matrix
$\mathcal{X}$	State space
$N_{\max,q,s}$	Upper bound for the number of admissible QoS flows of the $s$ -th service of the $q$ -th tenant
$a_{q,s}$	Admission decision for a new QoS flow arrival of the $s$ -th service of the $q$ -th tenant (0 for rejection, 1 for acceptance)
$\mathbf{A}(\mathbf{X}) = \{a_{q,s}\}$	Action to be made in state $\mathbf{X}$ . It includes the admission decisions $a_{q,s}$ for each service and tenant in a $Q \times S$ matrix
$\mathcal{A}_X$	Action space in state $\mathbf{X}$ .
$\mathbf{e}_{q,s}$	$Q \times S$ matrix with all elements equal to 0 except the element in the $q$ -th row and $s$ -th column, which equals 1.
$r(\mathbf{X}, \mathbf{A})$	Reward rate function while the system is in state $\mathbf{X}$ and action $\mathbf{A}$ is chosen.
$f(\mathbf{X}, a_{q,s})$	Reward obtained as a result of the RAC decision $a_{q,s}$ made upon the arrival of a new QoS flow of the $s$ -th service of the $q$ -th tenant while being in state $\mathbf{X}$
$\tau(\mathbf{X}, \mathbf{A})$	Sojourn time in state $\mathbf{X}$ under action $\mathbf{A}$
$\tau$	Minimum sojourn time among all states and actions
$P_{\mathbf{X}\mathbf{Y}}^{\mathbf{A}}$	State transition probability from state $\mathbf{X}$ to state $\mathbf{Y}$ assuming that $\mathbf{A}$ is the action selected in state $\mathbf{X}$
$V_m(\mathbf{X})$	Value function of state $\mathbf{X}$ at iteration $m$
$L_m$	Maximum difference between the value of any state at iteration $m$ and at iteration $m-1$ .
$l_m$	Minimum difference between the value of any state at iteration $m$ and at iteration $m-1$ .
$\varepsilon$	Parameter that determines the stop condition of the value iteration algorithm
$R_q(\mathbf{X})$	Total GFBR of all the admitted QoS flows of the $q$ -th tenant when the cell is in state $\mathbf{X}$
$\Delta C_{q,s}$	Cost of an SLA breach
$\Delta I_{q,s}$	Extra reward for admitting a QoS flow exceeding the SLA limit
$P_c(\mathbf{X})$	Congestion probability when the cell is in state $\mathbf{X}$ .

For any state of the cell, the RAC function should decide whether to accept or reject any new QoS flow request of any service and tenant. Then, the RAC operation can be formalized through an RAC policy, defined as a stationary rule that associates each state  $\mathbf{X} \in \mathcal{X}$  with an action  $\mathbf{A}(\mathbf{X}) = \{a_{q,s}\}$ , represented as a  $Q \times S$  matrix, each of whose components  $a_{q,s}$  denote the decision to be made for a new QoS flow arrival of the  $s$ -th service of the  $q$ -th tenant when the cell is in state  $\mathbf{X}$ , where  $a_{q,s}=1$  indicates acceptance and  $a_{q,s}=0$  indicates rejection.

For a given state  $\mathbf{X} \in \mathcal{X}$ , the action space  $\mathcal{A}_X$  defines the set of eligible actions in this state and is composed of those actions that lead to another state of the state space  $\mathcal{X}$ . This is formally defined as:

$$\mathcal{A}_X = \left\{ \mathbf{A} = \{a_{q,s}\} \in \mathbb{N}^{Q \times S} : a_{q,s} \in \{0,1\}, a_{q,s} = 0 \text{ if } \mathbf{X} + \mathbf{e}_{q,s} \notin \mathcal{X} \right\} \quad (3)$$

where  $\mathbf{e}_{q,s}$  is a  $Q \times S$  matrix with all elements equal to 0 except the element in the  $q$ -th row and  $s$ -th column, which equals 1. Therefore,  $\mathbf{X} + \mathbf{e}_{q,s}$  denotes the new state that will be reached if a new QoS flow of the  $s$ -th service of the  $q$ -th tenant is admitted while being in state  $\mathbf{X}$ .

The RAC policy has an impact on the system dynamics and, consequently, on the obtained performance, which can be characterized through different metrics. This paper considers a general reward rate function  $r(\mathbf{X}, \mathbf{A})$  that defines the performance obtained while the system is in state  $\mathbf{X}$  and action  $\mathbf{A} \in \mathcal{A}_X$  is chosen in this state. A specific formulation of this function will be presented in Section III.C.

The SMDP problem considered here consists of finding the optimum RAC policy, i.e., the optimum action  $\mathbf{A}(\mathbf{X}) \in \mathcal{A}_X$  to be selected in every state  $\mathbf{X} \in \mathcal{X}$  in order to maximize the average long-term reward obtained in the cell as a result of its dynamics.

### B. Algorithmic solution to the SMDP problem

In general, different algorithms for finding an optimal decision policy in an SMDP exist, such as linear programming, policy iteration or value iteration (see ch. 7 of [25]). This paper makes use of the value iteration algorithm for solving the RAC function problem formulated in Section III.A, since it has been empirically demonstrated that this algorithm is able to find adequate solutions with low computational time. Value iteration has also been used in other references that have addressed SMDP, such as [20] and [21].

The value iteration algorithm for solving an SMDP is defined in [25] by means of a data transformation that converts the SMDP into an equivalent discrete time Markov decision process. In this way, the algorithm can recursively compute a sequence of value functions that approximate the maximum average reward per time unit. Specifically, the formulation of the value iteration algorithm in [20][25] is particularized to the problem of Section III.A as follows.

The value functions are initialized to arbitrary values  $V_0(\mathbf{X})$  for each state  $\mathbf{X} \in \mathcal{X}$  at  $m=0$ . Then, at iteration  $m$ , the new value functions are obtained as:

$$V_m(\mathbf{X}) = \max_{\mathbf{A} \in \mathcal{A}_X} \left\{ r(\mathbf{X}, \mathbf{A}) + \sum_{\mathbf{Y} \in \mathcal{X}} \frac{\tau}{\tau(\mathbf{X}, \mathbf{A})} P_{\mathbf{XAY}} V_{m-1}(\mathbf{Y}) + \left( 1 - \frac{\tau}{\tau(\mathbf{X}, \mathbf{A})} \right) V_{m-1}(\mathbf{X}) \right\}, \quad \mathbf{X} \in \mathcal{X} \quad (4)$$

where:

- $\tau(\mathbf{X}, \mathbf{A})$  is the sojourn time in state  $\mathbf{X}$  under action  $\mathbf{A}$  and is given by:

$$\tau(\mathbf{X}, \mathbf{A}) = \frac{1}{\sum_{q=1}^Q \sum_{s=1}^{S_q} n_{q,s} \mu_{q,s} + \sum_{q=1}^Q \sum_{s=1}^{S_q} \lambda_{q,s} a_{q,s}} \quad (5)$$

- $\tau$  is the minimum sojourn time among all states and actions, i.e.

$$\tau = \min_{\mathbf{X}, \mathbf{A}} \tau(\mathbf{X}, \mathbf{A}) \quad (6)$$

- $P_{\mathbf{XAY}}$  is the state transition probability from state  $\mathbf{X}$  to another state  $\mathbf{Y}$  assuming that  $\mathbf{A}$  is the action selected in state  $\mathbf{X}$  and is defined as:

$$P_{\mathbf{XAY}} = \begin{cases} \lambda_{q,s} a_{q,s} \tau(\mathbf{X}, \mathbf{A}) & \mathbf{Y} = \mathbf{X} + \mathbf{e}_{q,s}, \quad s=1, \dots, S_q, \quad q=1, \dots, Q \\ n_{q,s} \mu_{q,s} \tau(\mathbf{X}, \mathbf{A}) & \mathbf{Y} = \mathbf{X} - \mathbf{e}_{q,s}, \quad s=1, \dots, S_q, \quad q=1, \dots, Q \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

The iterative process stops at the first iteration  $m$  that fulfills the following criterion:

$$0 \leq L_m - l_m \leq \varepsilon l_m \quad (8)$$

where  $\varepsilon$  is a parameter and  $L_m, l_m$  are defined as:

$$L_m = \max_{\mathbf{X} \in \mathcal{X}} (V_m(\mathbf{X}) - V_{m-1}(\mathbf{X})) \quad (9)$$

$$l_m = \min_{\mathbf{X} \in \mathcal{X}} (V_m(\mathbf{X}) - V_{m-1}(\mathbf{X})) \quad (10)$$

After stopping at iteration  $m$ , the obtained RAC policy is given by the following actions for each state  $\mathbf{X} \in \mathcal{X}$ :

$$\mathbf{A}(\mathbf{X}) = \arg \max_{\mathbf{A} \in \mathcal{A}_X} \left\{ r(\mathbf{X}, \mathbf{A}) + \sum_{\mathbf{Y} \in \mathcal{X}} \frac{\tau}{\tau(\mathbf{X}, \mathbf{A})} P_{\mathbf{XAY}} V_{m-1}(\mathbf{Y}) + \left( 1 - \frac{\tau}{\tau(\mathbf{X}, \mathbf{A})} \right) V_{m-1}(\mathbf{X}) \right\}, \quad \mathbf{X} \in \mathcal{X} \quad (11)$$

This policy provides an average reward within  $\varepsilon$  of the optimum one.

### C. Reward formulation

The formulation of the RAC as an SMDP presented in Sections III.A and III.B is general and can be used for

optimizing the RAC function based on different targets, depending on how the reward rate function  $r(\mathbf{X}, \mathbf{A})$  is defined. This paper targets the optimization of the RAC from the perspective of fulfillment of the SLA (i.e., the SAGBR value) for the different tenants. For this reason, a reward function is chosen that captures the satisfaction of each tenant with the received service in relation to the expected SLA.

Specifically, when the cell is in state  $\mathbf{X}=\{n_{q,s}\}$  and the action  $\mathbf{A}=\{a_{q,s}\}$  is selected, the reward rate  $r(\mathbf{X}, \mathbf{A})$  is defined as:

$$r(\mathbf{X}, \mathbf{A}) = \sum_{q=1}^Q \sum_{s=1}^{S_q} \lambda_{q,s} \cdot f(\mathbf{X}, a_{q,s}) \quad (12)$$

where  $f(\mathbf{X}, a_{q,s})$  is defined in (13) at the bottom of the page and represents the reward obtained as a result of the RAC decision  $a_{q,s}$  made upon the arrival of a new QoS flow of the  $s$ -th service of the  $q$ -th tenant in accordance with the selected action  $\mathbf{A}$ . The formulation of (13) takes into consideration the requested  $GFBR_{q,s}$  of the new QoS flow and the total GFBR of all the admitted QoS flows of the  $q$ -th tenant when the cell is in state  $\mathbf{X}$ , which is denoted  $R_q(\mathbf{X})$  and given by:

$$R_q(\mathbf{X}) = \sum_{s=1}^{S_q} n_{q,s} GFBR_{q,s} \quad (14)$$

Then, the different conditions in (13) are explained as follows:

- If the total aggregate GFBR of all the QoS flows of the  $q$ -th tenant is still below the contractual limit  $SAGBR_q$ , i.e.,  $R_q(\mathbf{X})+GFBR_{q,s} \leq SAGBR_q$ , and the RAC decision is to reject the new QoS flow, i.e.,  $a_{q,s}=0$ , then an SLA breach occurs. Therefore, the obtained reward is negative and corresponds to a cost of  $\Delta C_{q,s}$  units. This condition intends to drive the RAC towards the fulfillment of the SLA terms.
- In the case in which the RAC decision is to admit the new QoS flow, i.e.,  $a_{q,s}=1$ , and this involves exceeding the contractual SLA limit, i.e.,  $R_q(\mathbf{X})+GFBR_{q,s} > SAGBR_q$ , this means that the RAC is indeed granting excess capacity to the  $q$ -th tenant. This will be beneficial for both the infrastructure provider and the tenant: the infrastructure provider will achieve a better utilization of radio resources, and the tenant will sporadically experience better service than expected. However, this excess capacity will only be effectively provided if no congestion occurs, i.e., if the cell in the new state  $\mathbf{X}+\mathbf{e}_{q,s}$  has sufficient RBs to ensure the GFBR requirements of all the admitted QoS flows. Otherwise, the admitted QoS flows will be negatively affected. For this reason, the reward includes a positive

term  $\Delta I_{q,s} \cdot (1 - P_c(\mathbf{X} + \mathbf{e}_{q,s}))$  to reflect the extra capacity if there is no congestion (where  $P_c(\mathbf{X} + \mathbf{e}_{q,s})$  denotes the congestion probability in the new state) and a negative term  $\Delta C_{q,s} \cdot P_c(\mathbf{X} + \mathbf{e}_{q,s})$  for each admitted QoS flow to account for the degradation if there is congestion.

- All the other cases, i.e., when  $R_q(\mathbf{X})+GFBR_{q,s} \leq SAGBR_q$  and the decision is to admit the new QoS flow or when  $R_q(\mathbf{X})+GFBR_{q,s} > SAGBR_q$  and the decision is to reject the new QoS flow, reflect the normal operation of the RAC that would be expected by the  $q$ -th tenant and by the infrastructure provider. Consequently, they have neither positive nor negative influence, so the obtained reward is 0.

The congestion probability  $P_c(\mathbf{X})$  when the cell is in a given state  $\mathbf{X}$  is the probability that the cell will not have sufficient RBs to serve the admitted QoS flows with their required GFBRs. The detailed computation of  $P_c(\mathbf{X})$  is given in the Appendix and assumes that the mapping between GFBR requirements and the number of required RBs is stochastic and depends on the propagation conditions that determine the spectral efficiency achievable by each UE.

The setting of the parameters  $\Delta I_{q,s}$  and  $\Delta C_{q,s}$  in the reward function allows control of the trade-off between the level of isolation among the slices of each tenant and resource utilization efficiency. In particular, if a large value is set for  $\Delta C_{q,s}$  in relation to  $\Delta I_{q,s}$ , the reward function will drive the decision making toward a strict adherence to the SLA and thus a high isolation between slices. In contrast, increasing the value of  $\Delta I_{q,s}$  will drive the decision making toward making it easier for a tenant to obtain a higher bit rate than initially considered in its SLA, thus facilitating more efficient resource utilization but at the expense of a lower degree of isolation.

#### IV. PERFORMANCE EVALUATION

This section presents the performance evaluation of the proposed SMDP-based RAC by means of system-level simulations and compares it with different reference schemes used for benchmarking purposes in order to quantify the benefits that can be obtained. The details of the scenario considered in the simulations are described in Section IV.A, while the benchmarking approaches are presented in Section IV.B. Section IV.C includes some illustrative results to analyze the behavior of the SMDP-based RAC process. Sections IV.D to IV.F discuss the comparison with the reference schemes, focusing on different metrics, namely, reward, blocking probability, throughput and congestion probability. Finally, Section V.G discusses the issues involved in the practical implementation of the proposed approach.

$$f(\mathbf{X}, a_{q,s}) = \begin{cases} -\Delta C_{q,s} & \text{if } (a_{q,s} = 0) \text{ AND } (R_q(\mathbf{X}) + GFBR_{q,s} \leq SAGBR_q) \\ \Delta I_{q,s} \cdot (1 - P_c(\mathbf{X} + \mathbf{e}_{q,s})) - \left( \Delta C_{q,s} + \sum_{q'=1}^Q \sum_{s'=1}^{S_{q'}} n_{q',s'} \Delta C_{q',s'} \right) P_c(\mathbf{X} + \mathbf{e}_{q,s}) & \text{if } (a_{q,s} = 1) \text{ AND } (R_q(\mathbf{X}) + GFBR_{q,s} > SAGBR_q) \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

### A. Scenario description

The performance of the proposed SMDP-based RAC is evaluated by means of system-level simulations. The scenario is composed of one gNB with one omnidirectional cell that includes  $P=51$  RBs with bandwidth  $B=360$  kHz, which corresponds to a subcarrier spacing of  $\Delta f=30$  kHz according to the numerologies given in [3]. Table II presents the parameters considered in the simulation.

The cell is configured to support  $Q=2$  tenants, and each is associated with a different RAN slice. Each tenant provides two different GBR services. The GFBR, QoS flow generation rate and average duration of each service are detailed in Table III. The table also includes the contractual SLA values per tenant  $SAGBR_q$ . The average offered load of each tenant, defined as the aggregate of  $\lambda_{q,s} \cdot (1/\mu_{q,s}) \cdot GFBR_{q,s}$  for all the services of the tenant, is varied in the different simulations by changing the total generation rate per tenant based on the values indicated in Table III. Only traffic in the downlink direction is considered.

### B. Benchmarking approaches

For benchmarking purposes, the proposed SMDP-based RAC will be compared with the following reference schemes:

- **Myopic policy:** This strategy only considers the immediate reward when making RAC decisions. Thus, when the cell is in a given state  $\mathbf{X}$ , it accepts a new QoS flow of the  $s$ -th service of the  $q$ -th tenant if the reward in case of acceptance  $f(\mathbf{X}, a_{q,s}=1)$  is higher than the reward in case of rejection  $f(\mathbf{X}, a_{q,s}=0)$ . Otherwise, the new request is rejected.
- **Non-slice-aware policy:** This strategy considers a non-slice-aware RAC that makes decisions based only on the planned capacity  $C$  of the cell. Thus, when the cell is in state  $\mathbf{X}$ , a new QoS flow of the  $s$ -th service of the  $q$ -th tenant is admitted if the following condition is fulfilled:

$$\sum_{q=1}^Q R_q(\mathbf{X}) + GFBR_{q,s} \leq C \quad (15)$$

where the planned cell capacity  $C$  equals the aggregate SAGBR of all the tenants, that is:

$$C = \sum_{q=1}^Q SAGBR_q \quad (16)$$

- **Strict slicing policy:** This RAC strategy decides on the admission or rejection of a new QoS flow of the  $q$ -th tenant based on the agreed  $SAGBR_q$  of this tenant and considers a strict slicing that does not allow deviations exceeding this value. Therefore, a new QoS flow of the  $s$ -th service of the  $q$ -th tenant is admitted if the following condition holds:

$$R_q(\mathbf{X}) + GFBR_{q,s} \leq SAGBR_q \quad (17)$$

TABLE II.  
SIMULATION PARAMETERS

Parameter	Value
Cell radius	115 m
Path loss and shadowing model	Urban microcell model with hexagonal layout (see [26]) with a gNB antenna height of 10 m, a UE height of 1.5 m and a minimum distance of 10 m between UE and gNB.
Shadowing standard deviation	3 dB in line of sight (LOS) and 4 dB in non-line of sight (NLOS) (see [26])
Base station antenna gain	5 dB
Frequency	3.6 GHz
Transmitted power per RB	24 dBm
Number of RBs ( $P$ )	51
UE noise figure	9 dB
Link-level model to map the signal-to-interference-and-noise ratio and bit rate	Model in section A.1 of [27] with maximum spectral efficiency $\eta_{\max}=8.8$ b/s/Hz.
Simulation duration	100000 s
$\Delta I_{q,s}, \Delta C_{q,s}$	Varied in the simulations
Parameter $\varepsilon$ of the value iteration algorithm	$10^{-3}$
Initial values of the value iteration algorithm	$V_0(\mathbf{X}) = \max_{\mathbf{A}} \tau(\mathbf{X}, \mathbf{A})$

TABLE III.  
CHARACTERISTICS OF THE TENANTS AND THE SERVICES

Tenant	$q=1$		$q=2$	
Generation rate per tenant ( $\lambda_q$ )	$\lambda_1$ : from $3.125 \cdot 10^{-3}$ to $9.1 \cdot 10^{-2}$ requests/s		$\lambda_2$ : from $2.31 \cdot 10^{-3}$ to 0.278 requests/s	
Offered load per tenant	From 1 Mb/s to 120 Mb/s		From 1 Mb/s to 120 Mb/s	
$SAGBR_q$	60 Mb/s		40 Mb/s	
Service	$s=1$	$s=2$	$s=1$	$s=2$
$GFBR_{q,s}$	10 Mb/s	20 Mb/s	2 Mb/s	10 Mb/s
Generation rate per service ( $\lambda_{q,s}$ )	$0.9 \cdot \lambda_1$ requests/s	$0.1 \cdot \lambda_1$ requests/s	$0.8 \cdot \lambda_2$ requests/s	$0.2 \cdot \lambda_2$ requests/s
QoS flow duration ( $1/\mu_{q,s}$ )	120s	120s	120s	120s

### C. Behavior analysis of the SMDP

To gain insight into decision-making mechanisms, this section analyzes the actions  $\mathbf{A}(\mathbf{X})$  that the value iteration algorithm of Section III.B found to be optimal for the SMDP problem for an illustrative state  $\mathbf{X}$ . Each action specifies whether an acceptance or rejection decision has to be made when the system is in state  $\mathbf{X}$  and a new QoS flow of the  $s$ -th service of the  $q$ -th tenant arrives. Therefore, the actions specify the allowed transitions between the different states. A specific example of these transitions is shown in Fig. 2. It has been obtained by assuming the simulation parameters of Table II and the service characteristics of Table III with an average offered load of 60 Mb/s for tenant  $q=1$  and 60 Mb/s for tenant  $q=2$  and with  $\Delta I_{q,s}=100$  and  $\Delta C_{q,s}=10$ . The figure depicts the transitions when the cell is in state  $[n_{1,1}=6, n_{1,2}=0, n_{2,1}=21, n_{2,2}=0]$  for all the possible QoS flow arrivals in this state; it

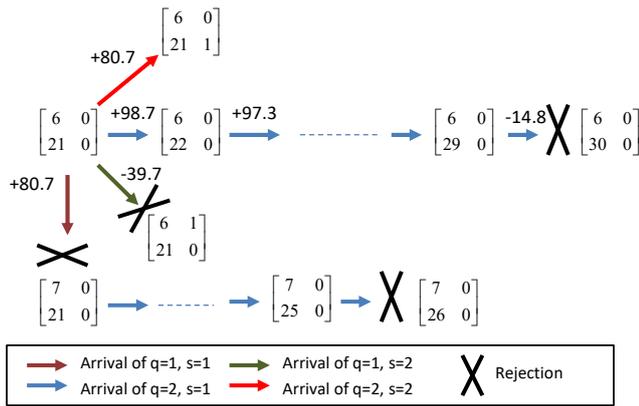


Fig. 2. Example of the allowed state transitions from state  $[n_{1,1}=6, n_{1,2}=0, n_{2,1}=21, n_{2,2}=0]$  for the different service arrivals.

represents the admission acceptance decisions ( $a_{q,s}=1$ ) with an arrow and represents the admission rejection decisions ( $a_{q,s}=0$ ) with an arrow followed by a cross. The number above each arrow indicates the obtained reward  $f(\mathbf{X}, a_{q,s}=1)$  in case of acceptance. In state  $[n_{1,1}=6, n_{1,2}=0, n_{2,1}=21, n_{2,2}=0]$ , the aggregate GFR of tenant 1 is 60 Mb/s, which equals the SAGBR value of this tenant, while the aggregate GFR of tenant 2 is 42 Mb/s, which is above its  $SAGBR_2=40$  Mb/s. Therefore, the reward in the case of rejection  $f(\mathbf{X}, a_{q,s}=0)$  is 0 for all possible arrivals.

Fig. 2 shows that in the case of a new arrival of service  $s=2$  from tenant  $q=1$ , the decision made by the value iteration algorithm is to reject the new request. This is indeed a logical decision since the actual reward that could be obtained in case of acceptance is negative (-39.7) due to the high GFR of this service, which would lead to a state with high congestion probability (37%) in case of acceptance. In contrast, the decision is made to admit the new arrivals of the services  $s=1$  and  $s=2$  from tenant  $q=2$ . In both cases, the reward associated with acceptance is positive (+98.7 and +80.7, respectively) because the aggregate GFR of tenant  $q=2$  is above its SAGBR value.

Remarkably, the long-term perspective that the algorithm embraces to gain insight into the future can be observed in the case of a new arrival of service  $s=1$  from tenant  $q=1$ , where the decision made by the algorithm is to reject the new request, even though the immediate reward that can be obtained in case of acceptance is quite positive (+80.7). The reason is that the algorithm detects that a higher long-term reward can be obtained by rejecting the new request and waiting for the arrival of new QoS flows of other services such as  $s=1, q=2$ . Note that, as shown in Fig. 2, by admitting new requests of this service, the cell can progressively accumulate rewards until reaching a maximum of 29 users. In contrast, if the decision had been made to accept the new arrival of service  $s=1, q=1$  and thus move to state  $[n_{1,1}=7, n_{1,2}=0, n_{2,1}=21, n_{2,2}=0]$ , the maximum number of new requests of service  $s=1, q=2$  that could have been admitted would be 25 users.

As a result of this long-term perspective when solving the SMDP problem, the cell will be able to achieve a higher

reward than that obtained with other strategies. To illustrate this, Table IV presents the average reward obtained along a simulation when the RAC follows the admission policy used above to solve the SMDP problem and compares it with the myopic RAC strategy that only considers the immediate reward. The simulations use the parameters of Tables II and III with  $\Delta I_{q,s}=100$  and  $\Delta C_{q,s}=10$  and an offered load of 60 Mb/s for each tenant. Table IV shows that the resulting reward with the SMDP-based approach is approximately 30% higher than that obtained with the myopic approach.

Another characteristic of the long-term perspective inherent in SMDP-based decision making is an improved ability to avoid states with high congestion probability. This ability is demonstrated in Table IV, which includes the percentage of the total simulation time in which the cell has remained in states  $\mathbf{X}$  with congestion probability  $P_c(\mathbf{X})$  higher than 5%. Clearly, the myopic approach spends much more time in these states than the SMDP-based policy, and as a result, the overall congestion probability experienced along the simulation is much higher, as can also be seen in Table IV.

To provide further insight into this behavior, Table V presents a detailed analysis of the congestion in four selected states, in which  $n_{1,1}=4, n_{1,2}=1, n_{2,1}=13$  and the value of  $n_{2,2}$  varies. For each state, the table presents the congestion probability  $P_c(\mathbf{X})$  computed (as explained in the Appendix) from the probability density function (pdf) of the number of RBs required by the QoS flows of the state, which are shown for illustration purposes in Fig. 3. On the other hand, the table includes the amount of simulation time (out of a total time of 100000 s) that the cell has remained in each of the four states under each RAC policy. It is observed that when the RAC operates according to the myopic policy, the cell spends a nonnegligible amount of time in states with high congestion probability, whereas when the RAC follows the SMDP-based approach, these states are not observed during the simulation.

TABLE IV. COMPARISON BETWEEN THE SMDP AND THE MYOPIC APPROACH

	Average reward	% of time in states with $P_c(\mathbf{X}) > 5\%$	Congestion probability along the simulation
SMDP approach	11.27	0.47%	0.08%
Myopic approach	8.63	25.5%	4.98%

TABLE V. CONGESTION PROBABILITIES FOR DIFFERENT STATES AND AMOUNT OF TIME IN THESE STATES WITH THE SMDP AND MYOPIC APPROACHES

State $\mathbf{X}$	$P_c(\mathbf{X})$	Time Myopic	Time SMDP
$n_{1,1}=4, n_{1,2}=1, n_{2,1}=13, n_{2,2}=0$	0.0057%	27 s	0 s
$n_{1,1}=4, n_{1,2}=1, n_{2,1}=13, n_{2,2}=1$	0.126%	96 s	0 s
$n_{1,1}=4, n_{1,2}=1, n_{2,1}=13, n_{2,2}=2$	2.21%	194 s	0 s
$n_{1,1}=4, n_{1,2}=1, n_{2,1}=13, n_{2,2}=3$	16.9%	208 s	0 s

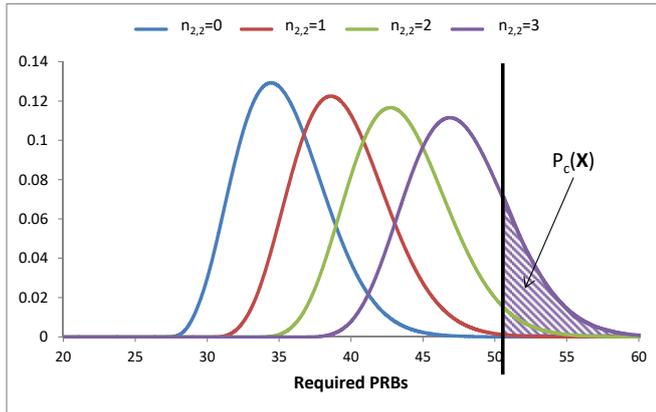


Fig. 3. Probability density functions of the number of required RBs for states with  $n_{1,1}=4$ ,  $n_{1,2}=1$ ,  $n_{2,1}=13$  and different values of  $n_{2,2}$ .

#### D. Comparison in terms of reward

Fig. 4 plots the rewards obtained with the different strategies as a function of the offered load of tenant 1 when the offered load of tenant 2 is 20 Mb/s, which is much lower than its  $SAGBR_2$  value. Then, Fig. 5 depicts the corresponding result for the case in which the offered load of tenant 2 is 60 Mb/s, i.e., higher than  $SAGBR_2$ . Both figures assume a reward function specified by  $\Delta I_{q,s}=100$  and  $\Delta C_{q,s}=10$  for all services. It can be observed in both figures that the reward obtained by the SMDP-based approach outperforms the rewards of all the other strategies thanks to its ability to optimize the RAC decision-making policy to maximize the long-term reward. Among the benchmarking strategies, the best reward is achieved by the myopic policy because this policy also includes the reward function in the decision-making process. However, due to its lack of insight into the future, this policy becomes suboptimal from the perspective of long-term reward, leading to substantial differences from the SMDP approach. These differences are particularly high when the offered load of at least one of the two tenants is higher than its SAGBR (e.g., the cases in Fig. 4 in which the offered load of tenant 1 is above 60 Mb/s and the cases in Fig. 5), which is when the SMDP-based decision-making policy can provide extra capacity to the tenants and thus increase the reward through the term  $\Delta I_{q,s}$ . The non-slice-aware strategy is also able to achieve a certain reward; it is able to admit requests of a tenant even when the aggregate GFBR of the tenant's admitted flows is higher than the SAGBR because it does not impose tenant-specific limits on the admission of new requests. However, since this policy does not explicitly incorporate the reward into its formulation, it yields a lower reward than both the SMDP and the myopic policies. Finally, the reward of the strict slicing policy is 0 in all cases since, by definition, this policy does not allow admissions when the aggregate GFBR of the admitted flows exceeds the SAGBR level, so it can never achieve either a positive or a negative reward.

To analyze the impact of the parameters  $\Delta I_{q,s}$  and  $\Delta C_{q,s}$  of the reward function, Fig. 6 and Fig. 7 show the reward obtained with the different strategies and an offered load of

tenant 2 equal to 60 Mb/s in two cases: the case in which the reward achieved by extra capacity equals the cost of blocking a request that is below the contractual SAGBR limit (i.e.,  $\Delta I_{q,s}=100$ ,  $\Delta C_{q,s}=100$  in Fig. 6); and the case in which this cost is much higher than the reward due to extra capacity (i.e.,  $\Delta I_{q,s}=100$ ,  $\Delta C_{q,s}=200$  in Fig. 7). The situation in which the reward due to extra capacity is much higher than the cost has been already discussed in Fig. 5 (i.e.,  $\Delta I_{q,s}=100$ ,  $\Delta C_{q,s}=10$ ).

The results shown in Fig. 6 and 7 reveal a similar behavior as in Fig. 5 in the sense that the SMDP-based approach outperforms all the other reference schemes. However, it can be observed that the actual reward decreases when the value of  $\Delta C_{q,s}$  increases, and the superiority of the SMDP-based approach to the myopic policy also decreases. This is due to the higher cost associated with rejections when a tenant is below its SAGBR level.

The percentage gain in reward achieved by the SMDP-based approach over the myopic policy for various offered traffic loads of the two tenants and for the three configurations of  $\Delta I_{q,s}$  and  $\Delta C_{q,s}$  considered above is presented in the 3-dimensional graphs of Fig. 8. It can be seen that the highest gains are achieved for high loads of both tenants and that the gain decreases when the cost term  $\Delta C_{q,s}$  increases. Specifically, the highest gain obtained for the configuration  $\Delta I_{q,s}=100$ ,  $\Delta C_{q,s}=10$  is 110%, which is reduced to 76% for the

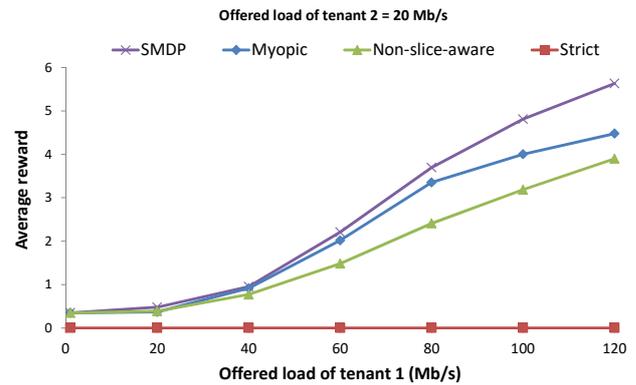


Fig. 4. Reward obtained with the different strategies as a function of the offered load of tenant 1 when the offered load of tenant 2 is 20 Mb/s, for the case  $\Delta I_{q,s}=100$ ,  $\Delta C_{q,s}=10$ .

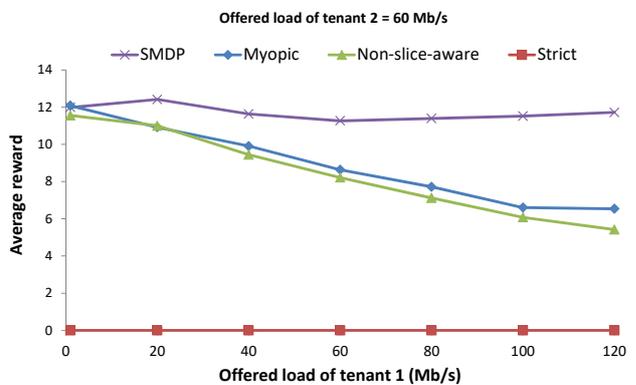


Fig. 5. Reward obtained with the different strategies as a function of the offered load of tenant 1 when the offered load of tenant 2 is 60 Mb/s, for the case  $\Delta I_{q,s}=100$ ,  $\Delta C_{q,s}=10$ .

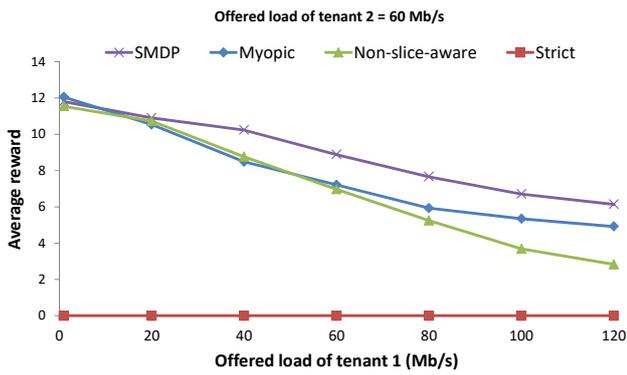


Fig. 6. Reward obtained with the different strategies as a function of the offered load of tenant 1 when the offered load of tenant 2 is 60 Mb/s, for the case  $\Delta I_{q,s}=100, \Delta C_{q,s}=100$ .

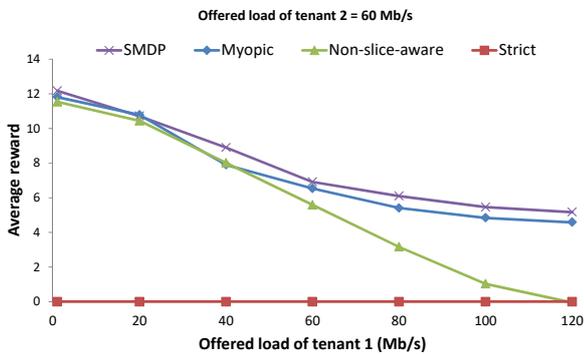


Fig. 7. Reward obtained with the different strategies as a function of the offered load of tenant 1 when the offered load of tenant 2 is 60 Mb/s, for the case  $\Delta I_{q,s}=100, \Delta C_{q,s}=200$ .

configuration  $\Delta I_{q,s}=100, \Delta C_{q,s}=100$  and reduced even further to 50% for the configuration  $\Delta I_{q,s}=100, \Delta C_{q,s}=200$ .

A relevant characteristic of the proposed SMDP-based framework is its flexibility: it can be configured to achieve different behaviors by properly tuning the values of the parameters  $\Delta I_{q,s}$  and  $\Delta C_{q,s}$  in the reward function. To illustrate this, Fig. 9 presents the reward obtained for the configuration  $\Delta I_{q,s}=0, \Delta C_{q,s}=200$ . This configuration reflects a conservative case in which the RAC should strictly adhere to SLA

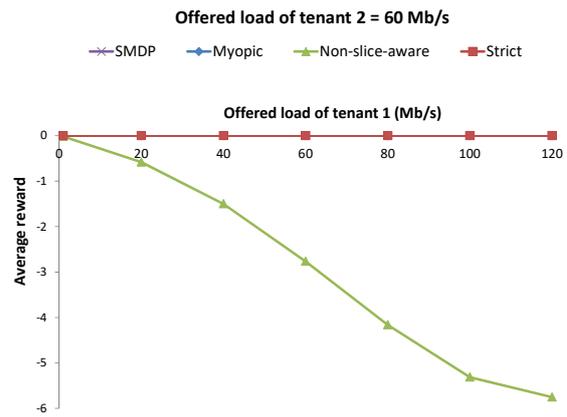


Fig. 9. Reward obtained with the different strategies as a function of the offered load of tenant 1 when the offered load of tenant 2 is 60 Mb/s, for the case  $\Delta I_{q,s}=0, \Delta C_{q,s}=200$ .

fulfillment because there are high penalties for rejecting new requests that are under the SAGBR limit but no reward for providing extra capacity beyond the SAGBR limit. As a result, it can be seen in Fig. 9 that the reward obtained by the SMDP-based policy is 0 in all cases, reflecting the fact that the decision-making policy avoids the possibility of incurring costs. This behavior ensures a very strong isolation between slices, since each tenant will obtain resources in accordance with its SLA, regardless of the traffic of the other tenants. Indeed, Fig. 9 shows that the SMDP-based approach has exactly the same behavior as the strict slicing policy and the myopic policy. In contrast, the non-slice-aware policy leads to negative rewards in this case due to the indiscriminate rejections that do not take into account the SLA of each slice.

#### E. Comparison in terms of blocking probability and throughput

Going beyond analyzing performance in terms of reward, this section compares the behavior obtained with the different approaches from the perspective of different key performance indicators (KPIs).

Fig. 10 plots the blocking probability experienced by each tenant, defined as the percentage of rejections out of the total

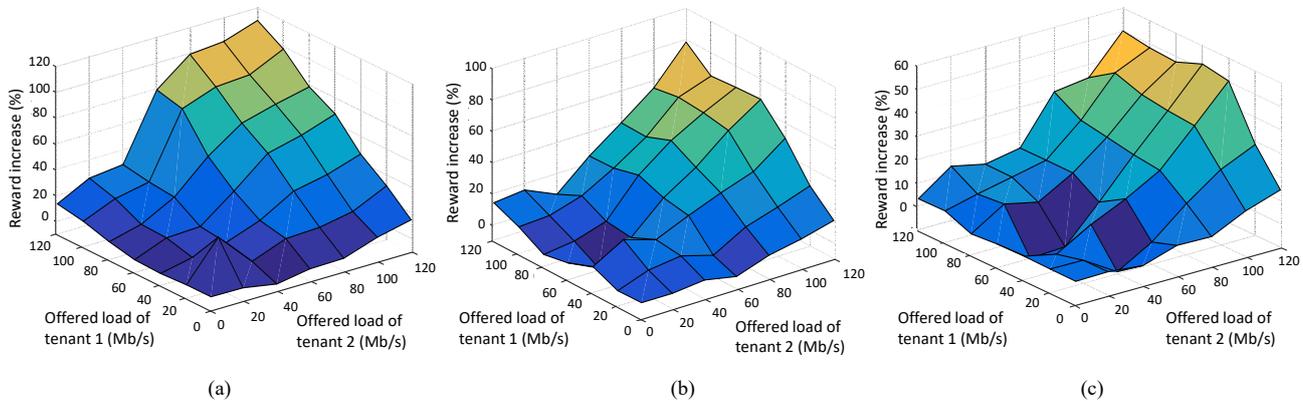


Fig. 8. Reward increase achieved by the SMDP-based approach over the myopic approach for (a)  $\Delta I_{q,s}=100, \Delta C_{q,s}=10$ ; (b)  $\Delta I_{q,s}=100, \Delta C_{q,s}=100$ ; and (c)  $\Delta I_{q,s}=100, \Delta C_{q,s}=200$ .

number of admission attempts. The results are presented for two different values of the cost  $\Delta C_{q,s}$ , assuming  $\Delta I_{q,s}$  is set to 100, and for two different traffic mixes, the first characterized by an offered load of 40 Mb/s for tenant 1 and 20 Mb/s for tenant 2 and the second by an offered load of 60 Mb/s for tenant 1 and 20 Mb/s for tenant 2.

For the case in which the cost  $\Delta C_{q,s}$  is set to 200, i.e., a large value in relation to  $\Delta I_{q,s}$ , Fig. 10 shows that for both traffic mixes, the SMDP and myopic approaches provide a very similar blocking probability that is lower than the probability for the non-slice-aware and strict slicing approaches. It is also observed that tenant 1 experiences higher blocking than tenant 2 because in both traffic mixes, the offered load of tenant 2 in relation to its  $SAGBR_2$  is lower than the offered load of tenant 1 in relation to its  $SAGBR_1$ .

In contrast, when the cost  $\Delta C_{q,s}$  is set to 10, i.e., much lower than  $\Delta I_{q,s}$ , Fig. 10 reveals some changes in the blocking probability obtained with the SMDP-based approach. In particular, note that the blocking probability of tenant 1 is lower than that in the case  $\Delta C_{q,s}=200$ . The reason is that, with the lower cost associated with rejections, the RAC allows the admission of more requests of tenant 1 exceeding the  $SAGBR$  limit, even if these extra admissions could represent some blockings of tenant 2. Then, as a result of this behavior, the blocking probability of tenant 2 experiences some increase, particularly when the offered load of tenant 1 is 60 Mb/s, since for this load level, which is equal to  $SAGBR_1$ , the RAC will have more opportunities to obtain a positive reward by providing extra capacity to this tenant.

Fig. 11 depicts the aggregate throughput obtained by each tenant with the different RAC strategies. The results are presented for the same conditions as in Fig. 10. Indeed, similar observations can be made for the throughput as for the blocking probability. Specifically, both the SMDP and the myopic approaches provide similar throughput levels that are higher than those obtained with the non-slice-aware and strict slicing strategies. Similarly, reducing the cost  $\Delta C_{q,s}$  leads to an increase in the throughput of the tenant with a higher offered load (i.e., tenant 1 in this case) due to the reward associated with the extra capacity that can be provided to this tenant.

The blocking probability per service is depicted in Fig. 12, assuming an offered load of 40 Mb/s for tenant 1 and 20 Mb/s for tenant 2 and reward parameters  $\Delta I_{q,s}=100$ ,  $\Delta C_{q,s}=200$ . The results show that the worst performance is achieved by the strict RAC, while the SMDP and myopic approaches reduce the blocking probability, with a slightly better performance achieved by the SMDP approach. In turn, it can be seen that service 2 of tenant 1, which requires the highest GFBR (see Table III), also experiences the highest blocking probability, while the lowest blocking probability is obtained by service 1 of tenant 2, which requires the lowest GFBR.

#### F. Comparison in terms of congestion probability

Fig. 13 depicts the congestion probability, i.e., the probability that the cell does not have enough RBs to fulfill the requirements of all the admitted QoS flows at a certain point in time. It can be seen that the SMDP-based approach,

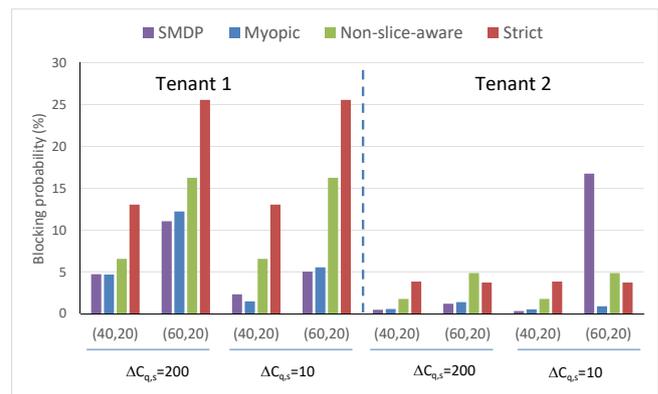


Fig. 10. Blocking probability of tenant 1 and tenant 2 for different values of  $\Delta C_{q,s}$  and two different traffic mixes. The numbers in parentheses indicate the offered loads of tenant 1 and tenant 2 in Mb/s, respectively.

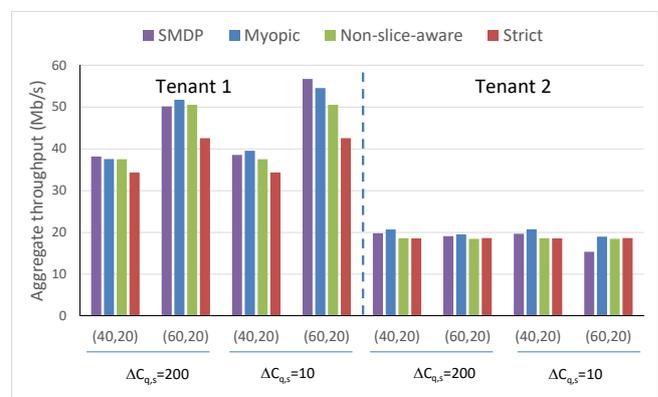


Fig. 11. Aggregate throughput of tenant 1 and tenant 2 for different values of  $\Delta C_{q,s}$  and two different traffic mixes. The numbers in parentheses indicate the offered loads of tenant 1 and tenant 2 in Mb/s, respectively.

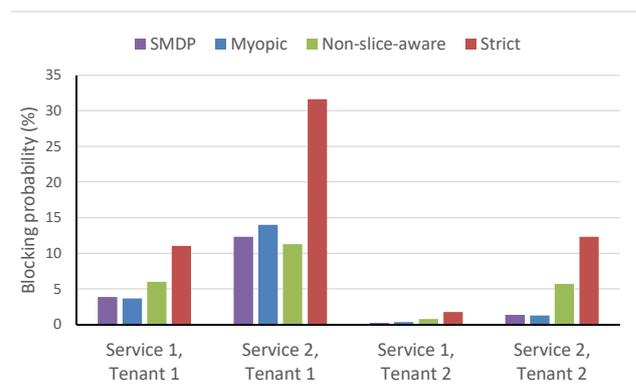


Fig. 12. Blocking probability per service.

the non-slice-aware approach and the strict slicing approach maintain very reduced values of the congestion probability. For the non-slice-aware and strict slicing approaches, the reason is that both approaches impose a limitation on the total number of admitted QoS flows that is determined by the planned cell capacity  $C$ , whose value provides a sufficient margin to avoid congestion in most situations. In contrast, both the SMDP and the myopic approach allow this limit to be exceeded in certain situations, i.e., when both tenants obtain extra capacity above their  $SAGBR$  value. However, the long-

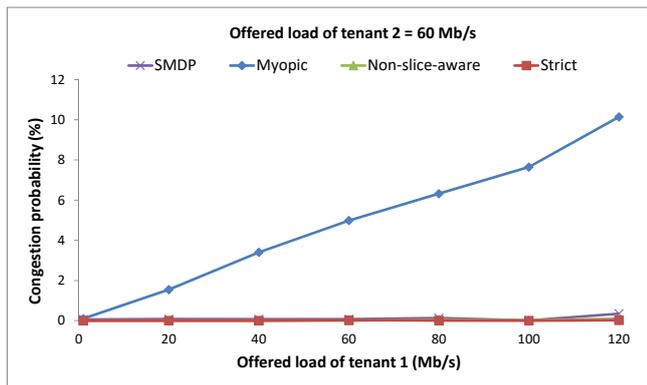


Fig. 13. Congestion probability as a function of the offered load of tenant 1 when the offered load of tenant 2 is 60 Mb/s, for the case  $\Delta I_{q,s}=100$ ,  $\Delta C_{q,s}=10$ .

term reward perspective achieved by the SMDP-based strategy gives this strategy better control of the congestion probability than the myopic policy. Consequently, the myopic approach is the one that experiences the worst congestion probability.

### G. Practicality considerations

Although a detailed practical implementation of the proposed SMDP-based algorithm is beyond the scope of this paper, this subsection intends to provide a discussion of the practical aspects that should be taken into account for such an implementation.

#### a) Configuration of the RAC algorithm

As discussed in Section II, the configuration of the RAC on a per-slice basis is performed by the RSMF through the implementation-dependent *RRMPolicy* attribute defined in the gNB network resource model [23]. Specifically, this attribute can be specified by means of different fields, namely, the  $SAGBR_q$  values, as previously explained in Section II, as well as the  $\Delta I_{q,s}$  and  $\Delta C_{q,s}$  values.

#### b) Acquisition of the required input parameters for executing the optimization model

The optimization process to determine the RAC policy described in Section III.B requires knowledge about the service characterization in terms of the QoS flow generation rates ( $\lambda_{q,s}$ ) and the average service duration ( $1/\mu_{q,s}$ ). In practice, these could be obtained from measurements collected by the network regarding service duration and time between arrivals observed from multiple QoS flows in each cell. These measurements should be statistically processed to obtain the average values and can be defined depending on the specific period of the day (e.g., every hour) so that different RAC policies can be applied in different periods. In the context of the 5G system architecture, the network data analytics function (NWDAF) [28], which can perform data analytics on UE communication patterns, is envisaged here to support the collection of statistics regarding the establishment and duration of the different services. Furthermore, the optimization process also requires a characterization of the congestion probability in the different states, following the formulation specified in the Appendix. This formulation

depends on the spectral efficiency distribution of the users in the cell, which determines the bit rate per RB and correspondingly affects the number of RBs required to achieve a certain GFBR. The distribution of the spectral efficiency can be obtained from the wideband channel quality indicator (CQI) distribution [29], which is computed by the gNB and delivered to the management system, e.g., to a management data analytics function (MDAF) [30], according to a certain periodicity (e.g., 15 min). Each CQI is an integer index that indicates the modulation and coding scheme that a UE can use in accordance with its experienced propagation and interference conditions and is mapped to a spectral efficiency value according to [31]. The MDAF gathers samples of the wideband CQI distribution and averages them for a longer time period in order to obtain adequate statistical validity to be representative of the cell conditions. Then, the averaged distribution of the CQI indices is directly mapped to the distribution of the spectral efficiency values experienced by the users in the cell. Using this distribution, the computations presented in the Appendix can be performed to obtain the congestion probability of each state to be delivered to the RAC function. Note that the congestion probability reflects the long-term behavior of the cell, so its computation will be performed on a long-term time scale; thus, computational complexity is not a critical issue.

#### c) Computation of the optimal RAC policy

The determination of the optimal RAC policy involves the execution of the value iteration algorithm presented in Section III.B. The optimal RAC policy should specify the operation of the RAC algorithm for a period of time in which the abovementioned input parameters (i.e., QoS flow generation rates, durations and congestion probabilities) are valid. The variation in these inputs is expected to occur on a relatively long-term basis, e.g., over the course of hours to capture different traffic conditions during a day. For this reason, the value iteration algorithm can be executed off-line. The result can be stored in a database and remain valid as long as the traffic conditions do not change significantly. Because of this off-line execution, the computational complexity of the value iteration algorithm does not place significant constraints on the operation of the RAC. To illustrate this computational complexity, note that the duration of the algorithm execution for the different simulations presented in the previous sections has been 126 s on average, on an Intel(R) core CPU at 3.4 GHz with 16 GB of RAM.

## V. CONCLUSIONS AND FUTURE WORK

This paper has proposed an optimization framework for the radio admission control function in multitenant and multiservice 5G scenarios based on the formulation of a semi-Markov decision process. The general formulation has then been particularized to include a specific reward function that captures the degree of fulfillment of the service level agreement from the perspective of the involved tenants. The reward function includes different parameters to account for the possibility of providing excess capacity beyond the SLA, thus achieving a more efficient usage of the radio resources,

and for the possibility of SLA breaches.

The proposed SMDP-based RAC policy has been evaluated by means of system-level simulations in order to assess its performance and compare it with other reference approaches. The behavior analysis of the SMDP approach has allowed identification of how the decision-making policy obtained incorporates the long-term reward perspective and differs from a myopic approach that only considers the immediate reward of an RAC decision. Thanks to this perspective, the SMDP-based approach is able to increase the obtained reward and reduce the time spent in states with high congestion probability. Specifically, the analysis of different traffic loads and different parameters of the reward function reflects reward improvements of up to 110% achieved by the SMDP-based policy over the myopic policy. The results have also shown that among all the configurations studied, the SMDP-based policy always outperforms the other reference strategies, namely, the myopic policy, the non-slice-aware policy and the strict slicing policy, with respect to the reward obtained.

Another characteristic of the proposed approach is its flexibility: it can be configured to achieve different behaviors by properly tuning the parameters of the reward function, thus achieving a trade-off between a more conservative policy that targets strict fulfillment of the SLA and a more flexible policy that achieves higher efficiency in the utilization of radio resources by providing excess capacity beyond the SLA.

Finally, the behavior of the RAC policies being considered has been analyzed in terms of different performance indicators such as blocking probability, throughput and congestion probability. In general, it has been observed that the SMDP-based and myopic policies exhibit small differences in terms of blocking and throughput and that both of them outperform the non-slice-aware and strict slicing policies. In turn, the SMDP-based approach clearly outperforms the myopic policy in terms of congestion probability since it is able to maintain a negligible congestion probability, while the myopic policy exhibits congestion probability values of up to 10%.

Based on these results, it is concluded that the proposed optimization framework establishes a solid basis for configuring the operation of the RAC function in multitenant and multiservice scenarios. Future work will consider the more practical aspects related to the implementation of the proposed approach, such as determining the conditions under which the optimized policy would need to be recomputed to address variations in the stationarity conditions of the scenario. Similarly, the framework presented here can also be studied with other optimization targets leading to different reward functions. For example, a possible extension in this direction would be the consideration of delay guarantees for handling the admission of ultra-reliable and low-latency communication (URLLC) services.

#### APPENDIX: COMPUTATION OF CONGESTION PROBABILITY

Congestion occurs whenever the number of RBs required to satisfy the GFBR requirements of all admitted QoS flows in the cell, denoted  $K_R$ , is higher than the number  $P$  of available RBs. This occurs stochastically depending on the variations in

the propagation conditions associated with the positions of the different UEs and the presence of shadowing losses. These effects impact the spectral efficiency achieved by each UE and, therefore, the number of required RBs.

Specifically, the number of RBs required by one QoS flow with requirement GFBR is  $K=GFBR/R$ , where  $R$  is the bit rate per RB and is related to the actual spectral efficiency  $S$  by  $R=S \cdot B$ ; it is a random variable that depends on the propagation conditions experienced by the UE of the QoS flow. Then, the proposed methodology assumes, based on measurements collected from the different UEs in terms of the wideband CQI distribution [29] that is directly mapped to the spectral efficiency distribution, that it is possible to derive the probability density function (pdf) of the random variable  $Y=1/R$ , denoted  $f_Y(y)$ . Consequently, the pdf of the number  $K$  of required RBs by one QoS flow with requirement GFBR is:

$$f_K(k) = f_Y\left(\frac{k}{GFBR}\right) \frac{1}{GFBR} \quad (18)$$

Then, assuming that the cell is in state  $\mathbf{X}=\{n_{q,s}\}$  and that each QoS flow experiences independent propagation conditions, the aggregate number  $K_R$  of required RBs by all the QoS flows is another random variable whose pdf is obtained by the convolution of (18) as many times as the number of QoS flows:

$$f_{K_R|\mathbf{X}}(k) = \left(\frac{1}{GFBR_{1,1}}\right)^{n_{1,1}} \cdots \left(\frac{1}{GFBR_{Q,S_Q}}\right)^{n_{Q,S_Q}} \left[ f_Y\left(\frac{k}{GFBR_{1,1}}\right) * \cdots * f_Y\left(\frac{k}{GFBR_{1,1}}\right) \right] * \cdots \left[ f_Y\left(\frac{k}{GFBR_{Q,S_Q}}\right) * \cdots * f_Y\left(\frac{k}{GFBR_{Q,S_Q}}\right) \right] \quad (19)$$

The congestion probability in state  $\mathbf{X}$  is thus given by:

$$P_c(\mathbf{X}) = \Pr(K_R > P | \mathbf{x}) = \int_P^\infty f_{K_R|\mathbf{X}}(k) dk \quad (20)$$

#### REFERENCES

- [1] System Architecture for the 5G System; Stage 2 (Release 15), 3GPP TS 23.501 v15.2.0, June, 2018
- [2] R. Ferrús, O. Sallent, J. Pérez-Romero, R. Agustí, "On 5G Radio Access Network Slicing: Radio Interface Protocol Features and Configuration", *IEEE Communications Magazine*, May, 2018, pp.184-192.
- [3] NR and NG-RAN Overall Description; Stage 2 (Release 15), 3GPP TS 38.300 v15.2.0, June, 2018.
- [4] O. Sallent, J. Pérez-Romero, J. Pérez-Romero, R. Ferrús, R. Agustí, "On Radio Access Network Slicing from a Radio Resource Management Perspective", *IEEE Wireless Communications*, October, 2017, pp. 166-174.
- [5] R. Kokku, R. Mahindra, H. Zhang, S. Rangarajan, "NVS: A substrate for Virtualizing Wireless Resources in Cellular Networks", *IEEE/ACM Transactions on Networking*, Vol. 20, No. 5, October, 2012.
- [6] R. Mahindra, M. Khojastepour, H. Zhang, S. Rangarajan, "Radio Access Networks Sharing in Cellular Networks", *21st IEEE Int. Conference on Network Protocols (ICNP)*, Göttingen, Germany, October, 2013

[7] X. Costa-Perez, J. Swetina, T. Guo, R. Mahindra, "Radio Access Network Virtualization for Future Mobile Carrier Networks", *IEEE Communications Magazine*, July, 2013.

[8] T. Guo, R. Arnott, "Active LTE RAN Sharing with Partial Resource Reservation", *IEEE 78th Vehicular Technology Conference (VTC Fall)*, Las Vegas, NV, USA, September, 2013.

[9] R. Kokku, R. Mahindra, H. Zhang, S. Rangajaran, "CellSlice: Cellular Wireless Resource Slicing for Active RAN Sharing", *5th International Conference on Communication Systems and Networks*, 2013.

[10] J. He, W. Song, "AppRAN: Application-Oriented Radio Access Network Sharing in Mobile Networks", *IEEE International Conference on Communications (ICC)*, 2015.

[11] A. Aijaz, "Hap-SliceR: A Radio Resource Slicing Framework for 5G Networks with Haptic Communications", *IEEE Systems Journal*, Vol. 12, No. 3, September, 2018, pp. 2285-2296.

[12] P. Caballero, A. Banchs, G. de Veciana, X. Costa-Pérez, A. Azcorra, "Network Slicing for Guaranteed Rate Services: Admission Control and Resource Allocation Games", *IEEE Transactions on Wireless Communications*, Vol. 17, No. 10, October, 2018.

[13] P. Caballero, A. Banchs, G. de Veciana, X. Costa-Pérez, "Network Slicing Games: Enabling Customization in Multi-Tenant Mobile Networks", *IEEE/ACM Transactions on Networking*, February, 2019.

[14] H. M. Soliman, A. Leon-Garcia, "QoS-Aware Frequency-Space Network Slicing and Admission Control for Virtual Wireless Networks", *IEEE GLOBECOM*, 2016.

[15] K. Samdanis, X. Costa-Perez, V. Sciancalepore, "From Network Slicing to Multi-Tenancy: The 5G Network Slice Broker", *IEEE Communications Magazine*, July, 2016.

[16] V. Sciancalepore, K. Samdanis, X. Costa-Perez, D. Bega, M. Gramaglia, A. Banchs, "Mobile Traffic Forecasting for Maximizing 5G Network Slicing Resource Utilization", *IEEE INFOCOM*, 2017.

[17] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, K. Samdanis, X. Costa-Perez, "Optimising 5G infrastructure markets: The Business of Network Slicing", *IEEE INFOCOM*, 2017.

[18] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, K. Samdanis, X. Costa-Perez, "A Machine Learning approach to 5G Infrastructure Market optimization", *IEEE Transactions on Mobile Computing*, February, 2019.

[19] J. Pérez-Romero, O. Sallent, R. Ferrus, R. Agustí, "Admission Control for Multi-tenant Radio Access Networks", *IEEE Int. Conference on Communications (ICC) - workshops*, Paris, France, May, 2017.

[20] K. W. Ross, D. H. K. Tsang, "Optimal Circuit Access Policies in an ISDN Environment: A Markov Decision Approach", *IEEE Transactions on Communications*, Vol. 37, No. 9, September, 1989, pp. 934-939.

[21] J. Bühler, G. Wunder, "Traffic-Aware Optimization of Heterogeneous Access Management", *IEEE Transactions on Communications*, Vol. 58, No. 6, June, 2010, pp. 1737-1747

[22] Management and orchestration; Provisioning (Release 16), 3GPP TS 28.531 v16.2.0, June, 2019.

[23] Management and orchestration; 5G Network Resource Model (NRM); Stage 2 and stage 3 (Release 16), 3GPP TS 28.541 v16.1.0, June, 2019.

[24] Procedures for the 5G System; Stage 2 (Release 16), 3GPP TS 23.502 v16.1.1, June, 2019.

[25] H.C. Tijms, *A First Course in Stochastic Models*, John Wiley & Sons, 2013.

[26] E-UTRA; Further advancements for E-UTRA physical layer aspects (Release 9), 3GPP TR 36.814 v9.2.0, March, 2017.

[27] Radio Frequency (RF) system scenarios (Release 15), 3GPP TR 36.942 v15.0.0, June, 2018

[28] Architecture enhancements for 5G System (5GS) to support network data analytics services (Release 16), 3GPP TS 23.288 v16.0.0, June, 2019.

[29] Management and Orchestration; 5G Performance Measurements (Release 16), 3GPP TS 28.552 v 16.2.0, June, 2019

[30] Management and orchestration; Architecture framework (Release 16), 3GPP TS 28.533 v16.0.0, June, 2019.

[31] NR; Physical layer procedures for data (Release 15), 3GPP TS 38.214 v15.6.0, June, 2019.



**Jordi Pérez-Romero** is a professor in the Dept. of Signal Theory and Communications of the Universitat Politècnica de Catalunya (UPC) in Barcelona, Spain, where he received a degree in telecommunications engineering in 1997 and the Ph.D. degree in 2001.

He has been working in the field of wireless communication systems, with a

particular focus on radio resource management, cognitive radio networks and network optimization. He has been involved in different European projects with different responsibilities, such as researcher, work package leader, and Project Responsible. He has also participated in different projects for private companies and has contributed to the 3GPP and ETSI standardization bodies. He is the author of more than 200 papers in international journals and conferences, a coauthor of three books and a contributor of seven book chapters.

Prof. Pérez-Romero serves as an Associate Editor for two international journals (*IEEE Vehicular Technology Magazine* and the *EURASIP Journal on Wireless Communications Networks*). Moreover, he has participated in the organization of several workshops, special sessions, and special issues in international conferences and journals.



**Oriol Sallent** is a professor at the Universitat Politècnica de Catalunya (UPC). He has participated in a wide range of European projects with diverse responsibilities such as work package leader and coordinator partner and has contributed to standardization bodies such as 3GPP, IEEE and ETSI. He has published more than 200 papers, mostly in

IEEE journals and conferences. His research interests include cognitive management in cognitive radio networks, self-organizing networks, radio network optimization and QoS provisioning in heterogeneous wireless networks.