IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 7, 2018, accepted December 7, 2018, date of publication December 12, 2018,
date of current version January 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2886447

Spectrum Sharing in Multi-Tenant 5G Cellular Networks: Modeling and Planning

OBADA AL-KHATIB[1], (Member, IEEE), WIBOWO HARDJAWANA[2], (Member, IEEE), AND BRANKA VUCETIC[2], (Fellow, IEEE)

[1]Faculty of Engineering and Information Sciences, University of Wollongong in Dubai, Dubai 20183, UAE
[2]School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia

Corresponding author: Obada Al-Khatib (obadaalkhatib@uowdubai.ac.ae)

The work of W. Hardjawana was supported by the Australian Research Council Discovery Early Career Research Award under Grant DE150101704.

ABSTRACT A multi-tenant cellular network is a paradigm where the physical infrastructure of the network is leased by various big industries, e.g., power utilities and transportation. Hence, a major challenge in a multi-tenant cellular network is the efficient allocation of the physical spectrum to various tenants with broadly distinct quality-of-service (QoS) requirements and communications traffic characteristics. In this paper, we approach this issue by presenting a versatile spectrum sharing scheme, which may be deployed to model any spectrum sharing strategy between various tenants in a multi-tenant cellular network. The proposed spectrum sharing scheme is based upon a queuing system that considers the various communications traffic characteristics of the tenants. In addition, by using the developed queuing system, mathematical expressions for the blocking probability and spectrum utilization are derived. We then propose an optimal spectrum planning scheme, referred to as reservation-based sharing (RBS) policy that maximizes the spectrum utilization by allocating the spectrum resources to various tenants according to their traffic loads. The computational complexity of the optimal RBS policy is reduced by developing a learning automata technique, referred to as pursuit learning-based RBS policy. By using real traffic parameters for various tenants, the results show that the simulation and analytical results match well, ensuring the accuracy of the proposed analytical model. Moreover, the results indicate that the proposed pursuit learning-based RBS policy firmly matches the optimal solution and delivers a higher spectrum utilization that increases linearly with the number of tenants.

INDEX TERMS Network virtualization, pursuit learning, queuing systems, spectrum sharing.

I. INTRODUCTION

A multi-tenant cellular network allows all available spectrum resources to be shared by multiple virtual networks, referred to as tenants. The tenants span a wide range of industries, such as eHealth, power utilities, and transportation, each with different traffic load requirements. Each tenant represents a large wireless network consisting of enormous number of wireless nodes, with different traffic load characteristics, behavior and requirements [1]. These tenants will be connected via 5G to provide various 5G services, such as ultra-reliable low latency communications (URLLC), enhanced mobile broadband (eMBB), and machine type communications (MTC) [2]. For example, millions of sensors will be installed to manage the energy distribution and usage in smart grid networks [3] and to control the congestion and road traffic lights in wireless vehicular networks [4].

In a multi-tenant cellular network, each physical Base station (BS) is partitioned into multiple virtual BSs, where each tenant will lease a single virtual BS. In other word, the traffic generated by those tenants is multiplexed over the same physical Network Infrastructure (NI) [5], [6]. Therefore, a major issue in a multi-tenant cellular network is the efficient allocation of the physical spectrum to various tenants with broadly distinct Quality-of-Service (QoS) requirements and communications traffic characteristics. By allocating less spectrum resources than what a tenant requires will lead to a deterioration in the QoS offered by the tenant to its users, in terms of the achieved data rates, transmission reliability, and latency requirement. As a result, there is a new challenge in planning the amount of spectrum resources that should be allocated to each tenant according to its traffic load, such that each tenant can guarantee achieving its QoS goals. The ''one-fit-all'' architecture of the 4G cannot guarantee the various communication requirements of the different tenants. Therefore, 5G will adopt a software-based architecture to consider the diverse demands of various tenants [7].

In this paper, we present a traffic load-based spectrum sharing scheme that maximizes the spectrum utilization, which is achieved by allowing the tenants to share all available Resource Blocks (RBs) of the 5G Network Infrastructure (NI). Here, the available RBs at each physical BS will be shared between the tenants. The proposed model is a long-standing spectrum planning since it takes into account the long-term traffic characteristics of each tenant. Hence, the computations of the RBs allocation may now be performed at longer intervals, e.g., daily, weekly, etc. We first develop a versatile spectrum sharing scheme, which may be deployed to model any spectrum sharing strategy between various tenants in a multi-tenant cellular network. The proposed spectrum sharing scheme is based upon a queuing system that considers the various communications traffic characteristics of the tenants. To guarantee that the transmission quality meets the stipulated latency and reliability requirements, a request for spectrum resources by a certain tenant will be blocked if the number of required RBs is larger than the number of free RBs in the NI. Based on the proposed analytical framework, mathematical formulae for the blocking probability and spectrum utilization in a multi-tenant cellular network are obtained. The blocking probability is an important metric that directly affects the QoS requirements, such as data rate and latency as explained earlier. By using these expressions, we formulate an optimisation problem, referred to Reservation-Based Sharing (RBS), to obtain the optimal spectrum allocation for each tenant in a multi-tenant cellular network, while considering the historical traffic load distribution and subject to the blocking probability requirements for the tenants. A low-complexity, iterative, and probabilistic pursuit learning algorithm is then proposed to solve this optimisation problem [8]. The simulation results show that the pursuit learning based RBS policy closely matches the solution obtained by exhaustive search and outperforms other known policies by at least 15% in terms of spectrum utilization. It also delivers a gain in the spectrum utilization over a network with an exclusive spectrum for each tenant, where the spectrum utilization is proportional to the number of tenants. Moreover, the results show that the number of needed iterations for the proposed pursuit learning algorithm to converge decreases as the number of tenants increases.

A. RELATED WORK

The existing research on spectrum sharing in multi-tenant networks may be classified into two broad categories. The first category is referred to as instantaneous spectrum sharing, in which the allocation of RBs for every tenant is performed instantaneously at each time slot, without taking into account the traffic load statistics [9]–[15]. Hence, the computational complexity of the resource allocation at each scheduling interval will be prohibitively high for this category.

Conversely, the second category is referred to as long-term spectrum sharing, which considers traffic load statistics to achieve long-term spectrum planning, since the RBs allocation is now performed at longer intervals, e.g., daily, weekly, etc. This will reduce the computational complexity for this resource allocation. Similar models have been considered in [16]–[25]. Conversely, such models have considered Poisson processes for the incoming traffic for various tenants [16]–[18], a single tenant [19]–[23], or a fixed sharing policy for every tenant [24], [25]. Therefore, these models cannot capture arbitrary traffic characteristics as well as traffic periodicity in some tenant applications, such as smart metering. In [26] and [27], we developed spectrum sharing models that deal with a non-Poisson arrival process. However, the proposed models cannot be directly used for multi-tenant wireless networks as they either consider a single tenant and/or have a very high computational complexity. Lastly, in [16]–[25], the number of allocated RBs to a tenant is constant and does not consider the channel fluctuations. Thus, the reliability of the transmission is degraded if the channel quality is poor.

With regards to the spectrum utilization, none of the aforementioned models has considered maximizing the spectrum utilization. The work presented in [28] developed a traffic load-based spectrum sharing model that can be applied to describe analytically various spectrum sharing schemes. However, the presented work in [28] has not proposed a solution to obtain the optimal spectrum sharing policy that maximizes the spectrum utilization. Paper [28] has presented only an analytical model to evaluate various spectrum sharing policies without any capability to find the optimal sharing policy, which means it cannot be used for spectrum planning. In addition, the analytical model in [28] has assumed Poisson processes for the incoming and outgoing traffic for all tenants sharing the spectrum, which is not realistic as mentioned earlier. The optimisation of the spectrum allocation has been extensively studied in [14], [15], and [29]–[32]; however, the interaction between the spectrum utilization and the traffic characteristics for various tenants in multi-tenant cellular networks has not been considered yet. For this complex optimisation problem, a specific type of Learning Automata (LA) technique, referred to as pursuit learning [8], is shown to be effective in reducing the computational complexity of such complex optimisation problems [8], [33]–[35]. To date, pursuit learning technique has not been applied to spectrum sharing in multi-tenant cellular networks while considering traffic statistics [26], [27]. Most papers are utilizing the stochastic learning [14], [36]–[38] that converges to a solution very slowly, which is not suitable for many practical applications. For example, the delay allowance for high-speed protection information in smart grid communication networks should be less than 10 ms [39].

B. CONTRIBUTIONS

In this paper, we propose a versatile spectrum sharing framework for a multi-tenant cellular network taking into account the traffic load statistics. The main contributions of this paper may be summarized as:

- We propose a general traffic model to evaluate various schemes for the spectrum sharing among

multiple tenants. The traffic model is based on a queuing system and may be applied to any arrival process and any distribution of the transmission times. That is unlike [16]–[25] and [28], where the arrival process is always assumed to be Poisson, the distribution of the transmission times is exponential, or a single traffic distribution is considered. In addition, the proposed traffic model considers multiple tenants, where each tenant has its own traffic distribution. This is also unlike [26] and [27], where only a single tenant has been assumed.

- We propose an optimal spectrum sharing policy, referred to as the RBS policy, that takes into account the traffic statistics of all tenants when allocating the RBs to maximize the spectrum utilization. To enforce the transmission quality, we introduce constraints on the blocking rates of various tenants. This is unlike [9]–[15], where the traffic load statistics are not taken into account when allocating the RBs to the tenants, and unlike [26] and [27], where no measures were taken to guarantee the transmission quality.
- Furthermore, by deploying the proposed framework, expressions for the blocking probability and spectrum utilization in a multi-tenant wireless network are obtained. These expressions are utilized to formulate an optimisation problem, referred to as the Reservation-Based Sharing (RBS) policy, that maximizes the spectrum utilization with the number of allocated RBs for each tenant as the optimisation variable, subject to blocking rate requirement for each tenant. A low-complexity optimisation solver based on pursuit learning [8] is then developed. This is unlike [28], where a solution to obtain the optimal spectrum sharing policy that maximizes the spectrum utilization has not been proposed.
- Given the amount of information involved (including channel state information and dynamic changes in tenant's demands), instantaneous approaches for spectrum optimisation in multi-tenant cellular networks [9]–[15] will result in significant inter-tenant signaling overheads and require performing the spectrum optimisation at every time transmission interval (TTI) in 5G cellular networks [40]. Thus, in this paper, we abstract the above information by constructing the probability distribution functions of the incoming and outgoing traffic for each tenant by using historical data captured by each tenant. We then use this data to plan the spectrum for multi-tenant wireless networks.

The organization of this paper is as follows. The analytical framework, that is formulated as a queuing system, is presented in Section II. The steady-state equations, which describe the system behavior when it becomes independent of time, are obtained followed by deriving mathematical expressions for the performance metrics. The formulation of the RBS policy as an optimisation problem and its solution, which is based on pursuit learning, are given in Section III. Section V discusses the simulation results and followed by conclusions in Section VI.

II. SYSTEM MODEL

A multi-tenant cellular wireless network is considered. The NI operator owns and manages the physical infrastructure and radio spectrum of this network. The set of the physical BSs in the network is denoted by $\mathcal{C} \triangleq \{1, \cdots, c, \cdots, C\}$, such that the target cell is in the center. We assume a frequency reuse of 1, i.e. every physical BS has a spectrum bandwidth of BW, which is orthogonally divided into N RBs, such that $\mathcal{N} \triangleq \{1, \cdots, n, \cdots, N\}$. An RB is the smallest resource unit that can be allocated in 5G with 180-KHz bandwidth and 0.5-ms duration. Each physical BS may be leased by K industries, where $\mathcal{K} \triangleq \{1, \cdots, k, \cdots, K\}$. Alternatively, each physical BS is partitioned into K virtual BSs, where each virtual BS is leased by an industry.

The NI operator is responsible for managing the scheduler, which is placed at the physical BS. Specifically, the scheduler divides the N available RBs at each physical BS orthogonally between the K tenants, in compliance with predefined spectrum sharing policies. The allocation policy is determined based on a Service Level Agreement (SLA) between the NI operator and the tenants. In Section III, we provide an optimal spectrum sharing policy that maximizes the spectrum utilization while guaranteeing the SLA of each tenant, expressed in terms of the maximum permissible blocking rate. After the NI operator assigns the available RBs to the K tenants, each tenant deploys its own scheduling policy to assign its RBs to its users/devices. The spectrum allocation among the K tenants sharing the same physical BS may be implemented distinctly in various cells. The partitioning of the spectrum bandwidth of the NI operator among various tenants is known as wireless network virtualization (WNV). In WNV, every tenant builds its own virtual network by utilizing its own virtual resources, allocated by the NI operator. Fig. 1 exemplifies a simple virtualised wireless network.

A. TRAFFIC DISTRIBUTIONS

We describe a queuing system to model the spectrum sharing in a multi-tenant cellular network. In this system, tenant k may request a number of RBs to provide services to its users/devices. When the number of requested RBs by tenant k is larger than the number of free RBs, this request will not be granted, i.e. blocked. The arrival process for tenant k, that defines the number of requests for RBs initiated by tenant k, follows a general probability distribution function (pdf), denoted by A_k, with λ_k is the average arrival rate. Alternatively, tenant k generates an average of λ_k requests per time unit. The number of requested RBs by tenant k in every initiated request for RBs is a discrete random variable $\tilde{\mathcal{F}}_k$, that may be described by any discrete pdf $\mathcal{F}_k$, e.g., geometric, uniform, or constant. Note that A_k is the pdf of the number of requests for RBs, while $\mathcal{F}_k$ is the pdf of the number of RBs per request. By denoting $l_{k,j} = Pr(\tilde{\mathcal{F}}_k = j)$ as the probability of requesting j RBs by all users of tenant k for a given

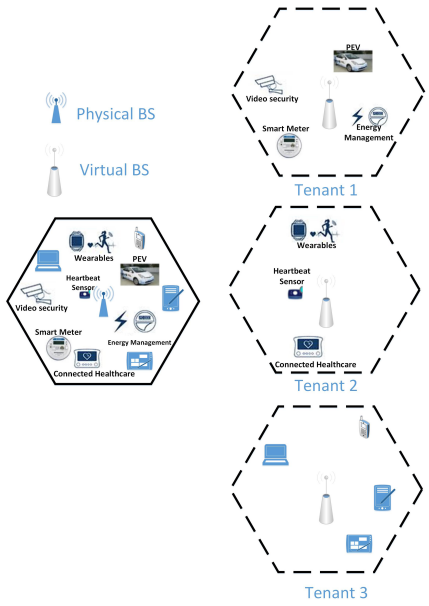

FIGURE 1. Multi-tenant Wireless Network.

RB request, the average number of requested RBs by tenant k per request event is given as

$$\mathbb{E}[\tilde{\mathcal{F}}_k] = \sum_{j=1}^{d_k} j \frac{l_{k,j}}{z_l}, \quad k \in \mathcal{K} \tag{1}$$

such that $z_l = \sum_{j=1}^{d_k} l_{k,j} = 1$ and d_k represents the largest number of RBs that may be assigned to tenant k. The departure process of tenant k defines the number of releases of RBs by tenant k. This process follows an arbitrary general distribution with pdf T_k and an average departure rate μ_k. Alternatively, μ_k represents the average number of releases for RBs per unit time. The number of RBs released by tenant k per release event depends on the number of transmitted symbols, users' locations, and its type of users. In addition, the pdf T_k is dependent on the tenant-specific scheduling policy. That is, different scheduling policies give different distributions for T_k [41]. The number of released RBs by tenant k in every release event is a discrete random variable $\tilde{\mathcal{W}}_k$, following any discrete pdf $\mathcal{W}_k$. By denoting $b_{k,j} = Pr(\tilde{\mathcal{W}}_k = j)$ as the probability of releasing j RBs by tenant k for a given RB release, the average number of released RBs by tenant k per release event is given as

$$\mathbb{E}[\tilde{\mathcal{W}}_k] = \sum_{j=1}^{d_k} j \frac{b_{k,j}}{z_b}, \quad k \in \mathcal{K} \tag{2}$$

such that $z_b = \sum_{j=1}^{d_k} b_{k,j} = 1$ and d_k is the largest number of RBs that can be released by tenant k. In real networks, the pdfs of the random variables $\tilde{\mathcal{F}}_k$ and $\tilde{\mathcal{W}}_k$, denoted by $\mathcal{F}_k$ and $\mathcal{W}_k$, respectively, may be determined by tracing the numbers of requested and released RBs by the users of tenant k over a long period of time, respectively, and then using the distribution fitting approach [42] to obtain those pdfs. Assuming different pdfs for $\tilde{\mathcal{F}}_k$ and $\tilde{\mathcal{W}}_k$ is due to the blocking that might happen when the number of free RBs is less than the number of requested RBs. The two pdfs $\mathcal{F}_k$ and $\mathcal{W}_k$ would be similar in case there is no blocking in the network. Specifically, when $\tilde{\mathcal{F}}_k$ is larger than the number of free RBs, the request for the RBs is blocked. Therefore, $\tilde{\mathcal{W}}_k = \tilde{\mathcal{F}}_k$ only if $\tilde{\mathcal{F}}_k$ is smaller than or equal to the number of free RBs, which indicates there is no blocking.

B. TRAFFIC BASED SPECTRUM SHARING MODEL

At any time instant, the state of the queuing system may be defined by the numbers of RBs assigned to each tenant. Therefore, at time t, the state space of the queuing system is described by the stochastic process $\Omega(t) = \mathbf{r}$, such that $\mathbf{r} = [v_1, \cdots, v_k, \cdots, v_K]$, where v_k is the number of RBs allocated to tenant k and $\sum_{i=1}^{K} v_i \leq N$, since the same RB cannot be allocated to different tenants. This stochastic process may not be Markovian as the pdfs A_k and T_k can take any form. Consequently, the next request or release of RBs is dependent on the time of the previous request or release of RBs, respectively [41]. Therefore, the process $\Omega(t)$ does not define the complete past history of the queuing system at time t since it does not know the time of the previous request or release of RBs. Note that for $\Omega(t)$ to be a Markovian process, the time of the next request or release of RBs should not depend on the time of the previous request or release, respectively. This is only achievable if the pdfs A_k and $T_k, \forall k \in \mathcal{K}$, are modeled as Poisson or exponential distribution [41], respectively.

To analyze the proposed queuing system, the first step is to convert it into a Markovian process. Thus, we define the joint process $[\Omega(t), \tilde{A}_1(t), \cdots, \tilde{A}_K(t), \tilde{T}_1(t), \cdots, \tilde{T}_K(t)]$, such that $\tilde{A}_k(t)$ and $\tilde{T}_k(t)$ are random variables following the pdfs A_k and T_k, respectively, and describing the remaining times until the next request or release of RBs by tenant k at time t, respectively. Based on this joint process, we define the probability of the system being in state $\mathbf{r}$ at time t, where the remaining times until the next request or release of RBs by tenant k are denoted by x_k and y_k, respectively, as

$$\begin{aligned} &P_{\mathbf{r}}(t, x_1, \cdots, x_K, y_1, \cdots, y_K) \\ &\quad \equiv P[\Omega(t) = \mathbf{r}, \tilde{A}_1(t) = x_1, \cdots, \tilde{A}_K(t) = x_K, \\ &\qquad \tilde{T}_1(t) = y_1, \cdots, \tilde{T}_K(t) = y_K]. \end{aligned} \tag{3}$$

The above joint process is Markovian because the complete past history is abstracted by the random variables $\tilde{A}_k(t)$ and $\tilde{T}_k(t), \forall k \in \mathcal{K}$. This method of transforming a non-Markovian process to a Markovian process is known as the supplementary variable method [43]. To ease the presentation, we define the state $\mathbf{r}_b^a$, which is identical to the state $\mathbf{r}$ except for its a^{th} entry, which is given by $\mathbf{r}(a) + b$, such that $\mathbf{r}(a)$ is the a^{th} entry of $\mathbf{r}$ and b is an integer.

C. STATE TRANSITIONS AND STEADY-STATE ANALYSIS

A new request or release of RBs by any tenant will cause the queuing system to move from one state to another. Fig. 2 shows all possible transitions into and out of state $\mathbf{r}$, such

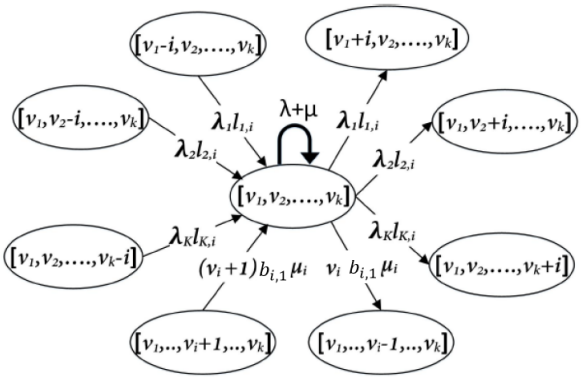

FIGURE 2. Possible transitions into and out of state r, where $i = 1, \cdots, N$.

that the numbers of RBs assigned to the tenants are given by $\mathbf{r} = [v_1, v_2, \cdots, v_K]$. For instance, a new request for RBs by tenant 1 requesting 3 RBs implies that the transition to the state $\mathbf{r} = [v_1, v_2, \cdots, v_K]$ happens from the state $\mathbf{r}^1_{-3} = [v_1 - 3, v_2, \cdots, v_K]$. By using Fig. 2, (3), and defining the remaining times until the next request or release of RBs in all tenants as $\mathcal{V} = \{x_1, \cdots, x_K, y_1, \cdots, y_K\}$, we write (4), shown at the bottom of the this page, for the probability of state $\mathbf{r}$ at time t when the remaining times until the next request or release of RBs in all tenants are given by $\mathcal{V}$. In (4), h is the length of an infinitesimal interval $[t-h, t]$, $\mathcal{V}^i_x$ is equivalent to $\mathcal{V}$, except for the i-th entry of the x variables, which equals to $x_i + h$. Similarly, $\mathcal{V}^i_y$ is equivalent to $\mathcal{V}$, except for the i-th entry of the y variables, which equals to $y_i + h$.

The first, third, and fifth terms in (4) give the probability that the queuing system is in state $\mathbf{r}$ at time $t-h$ when there are no requests or releases of RBs. The second and fourth terms give the probability that the queuing system is in state $\mathbf{r}$ at time $t-h$ when the remaining time until the next request of RBs by tenant i at time $t-h$ is $x_i + h$ and the remaining time until the next release of RBs by tenant i at time $t-h$ is $y_i + h$, respectively. The sixth term corresponds to a new request of RBs by tenant i, requesting j RBs, $1 \le j \le d_k$. The seventh term represents a release of RBs by tenant i. Thus, the sixth and seventh terms represent the total rate at which the traffic model moves into state $\mathbf{r}$ from other states during the interval $[t-h, t]$. Finally, the eighth and ninth terms on the right-hand side of (4) show the total rate at which the traffic model moves out of state $\mathbf{r}$ during the interval $[t-h, t]$. We highlight here that the proposed model is valid for any number of tenants, K. Increasing the number of tenants will only increase the dimension of the state space, represented by $\mathbf{r} = [v_1, \cdots, v_k, \cdots, v_K]$, which increases the number of terms in (4) and does not change the approach for finding the state probabilities.

The steady-state behavior of the proposed queuing system is achieved when its state probabilities do not change with time t, referred to as steady-state probabilities. To derive the steady-state probabilities, we take the limit as h goes to zero for each term on both sides of (4) and set the rate of change in the state probabilities with respect to the time to zero to obtain the steady-state equation for state $\mathbf{r}$ as

$$-\frac{\partial}{\partial \mathcal{V}} P_{\mathbf{r}}(\mathcal{V}) = \sum_{i=1}^{K}\sum_{j=1}^{N_i} \lambda_i l_{i,j} A_i(x_i) P_{\mathbf{r}^i_{-j}}(\mathcal{V}) + \sum_{i=1}^{K} \mu_i T_i(y_i) P_{\mathbf{r}^i_1}(\mathcal{V}) - (\lambda+\mu) P_{\mathbf{r}}(\mathcal{V}) \quad (5)$$

where $\lambda = \sum_{i=1}^{K}\sum_{j=1}^{d_i} \lambda_i l_{i,j}$, $\mu = \sum_{i=1}^{K} \mu_i$, $A_i(x_i)$ and $T_i(y_i)$ are the probabilities that the remaining times until the next request or release of RBs by tenant i is given by x_i and y_i, respectively. The first and second terms on the right-hand side of (5) describe all possible transitions into and out of state $\mathbf{r}$, respectively. The third term on the right-hand side of (5) corresponds to the case of remaining in state $\mathbf{r}$, i.e. no requests or releases of RBs.

The first step in solving the steady-state equations in (5) is to write them as linear difference equations by applying Laplace transform to obtain

$$-sP_{\mathbf{r}}(s) + P_{\mathbf{r}}(0) = \sum_{i=1}^{K}\sum_{j=1}^{N} \lambda_i l_{i,j} A_i(x_i) P_{\mathbf{r}^i_{-j}}(s) + \sum_{i=1}^{K} \mu_i T_i(y_i) P_{\mathbf{r}^i_1}(s) - (\lambda+\mu) P_{\mathbf{r}}(s) \quad (6)$$

$$\begin{aligned}\frac{P_{\mathbf{r}}(t,\mathcal{V})}{h} = {} & \underbrace{\frac{P_{\mathbf{r}}(t-h,\mathcal{V})}{h}}_{1} + \sum_{i=1}^{K}\Bigg\{\underbrace{\frac{P_{\mathbf{r}}(t-h,\mathcal{V}^i_x)}{h}}_{2} - \underbrace{\frac{P_{\mathbf{r}}(t-h,\mathcal{V})}{h}}_{3} + \underbrace{\frac{P_{\mathbf{r}}(t-h,\mathcal{V}^i_y)}{h}}_{4} - \underbrace{\frac{P_{\mathbf{r}}(t-h,\mathcal{V})}{h}}_{5}\Bigg\} \\ & + \underbrace{\sum_{i=1}^{K}\sum_{j=1}^{N_i} \lambda_i l_{i,j} A_i(x_i+h) P_{\mathbf{r}^i_{-j}}(t-h,\mathcal{V}^i_x)}_{6} + \underbrace{\sum_{i=1}^{K}\sum_{j=1}^{d_i} \mu_i b_{i,j} T_i(y_i+h) P_{\mathbf{r}^i_1}(t-h,\mathcal{V}^i_y)}_{7} \\ & - \underbrace{\sum_{i=1}^{K}\sum_{j=1}^{d_i} \lambda_i l_{i,j} P_{\mathbf{r}}(t-h,\mathcal{V}^i_x)}_{8} - \underbrace{\sum_{i=1}^{K}\sum_{j=1}^{d_i} \mu_i b_{i,j} P_{\mathbf{r}}(t-h,\mathcal{V}^i_y)}_{9} \end{aligned} \quad (4)$$

where

$$P_{\mathbf{r}}(s) = \int_0^\infty e^{-s\mathcal{V}} P_{\mathbf{r}}(\mathcal{V})\, \mathrm{d}\mathcal{V},$$

The second step involves solving the system of linear equations in (6) besides the normalization constraint written as

$$\sum_{\mathbf{r}\in\mathbf{R}_\varphi} P_{\mathbf{r}}(s) = 1, \tag{7}$$

where $\mathbf{R}_\varphi$ is the set of all possible states. $\mathbf{R}_\varphi$ is dependent on the deployed spectrum sharing policy $\varphi \in \{PS, NS, CS\}$, where PS refers to Partial Sharing, NS refers to No Sharing, and CS refers to Complete Sharing. The case of $\varphi = PS$ corresponds to all sharing policies, where the total number of available RBs N is larger than the total number of RBs assigned to all tenants, ($\sum_{k=1}^{K} N_k < N, \forall k \in \mathcal{K}$). The unallocated RBs represent a pool of RBs, that may be assigned to any tenant according to a first-come-first-serve policy. Conversely, the case of $\varphi = NS$ corresponds to all sharing policies, in which all available RBs are assigned to the various tenants, $\sum_{k=1}^{K} N_k = N, \forall k \in \mathcal{K}$. Thus, every tenant owns a fixed spectrum for its private usage and there is no pool of unallocated RBs. Finally, when $\varphi = CS$, the complete spectrum is shared by all tenants, i.e, $\sum_{k=1}^{K} N_k = 0, \forall k \in \mathcal{K}$. In other words, the number of RBs in the common pool in N.

Mathematically, the state space for PS case, $\mathbf{R}_{PS}$, is given by

$$\mathbf{R}_{PS} = \Bigg\{(v_1, v_2, \cdots, v_K) \mid 0 \le v_k \le N - \sum_{j=1, j\neq k}^{K} N_j, \sum_{k=1}^{K} N_k < N, \forall k \in \mathcal{K}\Bigg\} \tag{8}$$

such that N_k is the number of allocated RBs to tenant k. Therefore, the proposed framework may be deployed to compare different spectrum sharing policies under PS via controlling the number of assigned RBs for each tenant k, denoted by N_k. For the NS policy, the state space $\mathbf{R}_{NS}$ is described by

$$\mathbf{R}_{NS} = \Bigg\{(v_1, v_2, \cdots, v_K) \mid 0 \le v_k \le N - \sum_{j=1, j\neq k}^{K} N_j, \sum_{k=1}^{K} N_k = N, \forall k \in \mathcal{K}\Bigg\} \tag{9}$$

To obtain the steady-state probabilities of the proposed queuing system under sharing policy φ, the system of linear equations, given by (6) and (7), should be solved via a numerical method, e.g., the successive over relaxation method or the iterative power method [44]. To apply the iterative power method, (6) is written as a matrix equation given by $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$, such that $\boldsymbol{\pi}$ is a row vector, whose entries are the steady-state probabilities given by $P_{\mathbf{r}}(s)$ and $P_{\mathbf{r}}(0)$, $\forall \mathbf{r} \in \mathbf{R}_\varphi$, and $\mathbf{Q}$ is the transition rate matrix, whose entries are the coefficients of $P_{\mathbf{r}}(s)$ and $P_{\mathbf{r}}(0)$ that represent the transition rates between the different states. By assuming that there are no users in the system at $t = 0$, the probabilities $P_{\mathbf{r}}(0)$ are set to zero. We now solve this matrix equation iteratively as $\boldsymbol{\pi}(i+1) = \boldsymbol{\pi}(i)\mathbf{V}$, where $\boldsymbol{\pi}(i)$ is the i-th iteration and $\mathbf{V} = \beta\mathbf{Q} + \mathbf{I}$, such that $\mathbf{I}$ is the identity matrix and $\beta \in [0, 1]$. At each iteration i, $c(i) = \frac{\partial}{\partial s} P_{\mathbf{r}}(s)|_{s=0}$, $\forall \mathbf{r} \in \mathbf{R}_\varphi$, is computed, which represents the mean value of the steady state probability of state $\mathbf{r}$. The iterations will stop when the absolute difference between $c(i)$ at iteration i and $c(i-1)$ at iteration $i-1$, $\forall \mathbf{r} \in \mathbf{R}_\varphi$, is smaller than the threshold value of $\epsilon \in [0, 1]$. When a state is not feasible under sharing policy φ, then its probability equals to zero.

D. PERFORMANCE METRICS

By utilizing the proposed queuing system, we derive two performance metrics under sharing policy φ: 1) blocking probability of tenant k, denoted by $B_{k,\varphi}$, and 2) the spectrum utilization, denoted by Ψ_φ. The spectrum utilization is the ratio of the average number of allocated RBs to the total number of available RBs N. The average blocking probability for tenant k may be obtained as the first order Moment Generating Function (MGF) of the Laplace transform steady-state probabilities, given by

$$B_{k,\varphi} = \frac{\partial}{\partial s}\Bigg\{\sum_{\mathbf{r}\in\mathbf{R}_\varphi} P_{\mathbf{r}}(s) \sum_{j=\Gamma_k(\mathbf{r})+1}^{N_k} l_{k,j}\Bigg\}|_{s=0}, \tag{10}$$

where $\Gamma_k(\mathbf{r})$ is the number of available RBs for tenant k under sharing policy φ, when the system is in state $\mathbf{r}$. $\Gamma_k(\mathbf{r})$ is described by

$$\Gamma_k(\mathbf{r}) = (N_k - v_k)\delta_{Nv} + N - \sum_{k=1}^{K} N_k - \sum_{k=1}^{K} (v_k - N_k)\delta_{vN}, \tag{11}$$

where δ_{Nv} is 1 if $N_k \ge v_k$ and 0 otherwise, while δ_{vN} is 1 if $v_k \ge N_k$ and 0 otherwise. The average spectrum utilization is calculated by

$$\Psi_\varphi = \frac{\partial}{\partial s}\Bigg\{\frac{\sum_{\mathbf{r}\in\mathbf{R}_\varphi}(v_1 + v_2 + \cdots + v_K)P_{\mathbf{r}}(s)}{N}\Bigg\}|_{s=0} \tag{12}$$

which is the first order MGF of the Laplace transform steady-state probabilities. In reference to (10) and (12), the effective spectrum utilization, that considers the blocking probability, under sharing policy φ is defined as

$$\Psi_\varphi^{eff} = (1 - B_\varphi)\Psi_\varphi, \tag{13}$$

where $B_\varphi = \sum_{k=1}^{K} B_{k,\varphi}/K$ is the average blocking probability of all tenants under sharing policy φ. The efficiency gain Υ is defined as the ratio of the effective spectrum utilization under the PS sharing policy and the NS sharing policy, which is given by

$$\Upsilon = \frac{(1 - B_{PS})\Psi_{PS}^{eff}}{(1 - B_{NS})\Psi_{NS}^{eff}}. \tag{14}$$

The efficiency gain in (14) corresponds to the gain in the effective spectrum utilization when deploying the *PS* spectrum sharing policy, in comparison to the *NS* spectrum sharing policy. As Υ in (14) increases, the effective spectrum utilization under *PS* policy increases, leading to smaller blocking probabilities and accordingly larger number of devices admitted into the network, in comparison to the *NS* policy.

III. SPECTRUM UTILIZATION OPTIMISATION

In this section, we propose an optimum spectrum sharing scheme, referred to as the Reservation-Based Sharing (RBS) policy. Specifically, the RBS policy aims at maximizing the spectrum utilization in (12) under given blocking probability threshold for each tenant, with the numbers of RBs allocated to each tenant $N_k, \forall k \in \mathcal{K}$, as the optimisation variables. Thus, the RBS policy may be described by the following optimisation problem:

$$\max_{N_k, k\in\mathcal{K}} \Psi_\varphi \sum_{k\in\mathcal{K}} w_k \mathrm{sgn}(B_k^{th} - B_{k,\varphi}) \tag{15}$$

where B_k^{th} is the blocking probability threshold for tenant k and $\mathrm{sgn}(a-b) = 1$ if $a \geq b$, and 0 otherwise, and $w_k \in (0, 1]$ is the priority weight for tenant k such that 0 is the lowest priority and 1 is the highest priority. Note that the value of the blocking probability threshold for tenant k, B_k^{th}, is specified in the SLA between tenant k and the NI operator. Note that in the event that the thresholds for the blocking probabilities are set too low, the value of (15) will be zero, implying no feasible solution for (15). In this case, the NI operator in agreement with the tenants will increase the blocking probabilities thresholds or reduce the number of tenants allowed to share its spectrum, such that the spectrum utilization is optimized. The global optimal solution for (15) may be found by using an exhaustive search over all the possibilities of $\{N_1, \cdots, N_k, \cdots, N_K\}, \forall k \in \mathcal{K}$, such that $\sum_{k=1}^{K} N_k \leq N$, where N_k is the number of allocated RBs for tenant k, and then selecting the possibility that gives the maximum of (15). However, the exhaustive search requires a high computational complexity so that the optimal solution is hard to obtain in a reasonable time.

A. PURSUIT LEARNING BASED RBS POLICY

Learning Automata (LA) has been shown to effectively reduce the computational complexity of complex optimisation problems, such as (15), in [8], [34], and [35]. The LA is an adaptive and iterative machine learning algorithm that finds the best action from a finite set of potential actions based on the rewards received from an unknown environment at every iteration. Specifically, an action is selected randomly at each iteration according to an action probability vector. The environment then provides a reward to the selected action and the action probability vector is then updated according to this reward. This process is repeated until the learning automaton finds the optimal action from the action set that yields the highest reward. In this paper, we apply a specific learning automata technique, referred to as pursuit learning. The main idea behind the pursuit learning scheme is to update the action probability vector at every iteration to "pursue" the action that results in the highest average reward.

To apply the pursuit learning to (15), we first regard each tenant as a learning automaton with its tuple $\{\mathbf{a}, \mathbf{p}, R, \mathbf{L}\}$ defined as follows.

- Action set ($\mathbf{a}$): An action for tenant k is to select the number of required RBs, N_k, such that $0 \leq N_k \leq N$. Thus, the action set for each tenant is given by $\mathbf{a} = \{a_0, a_1, \cdots, a_N\}$, such that $a_i, 0 \leq i \leq N$, is the action of requesting i RBs.
- Action probability vector ($\mathbf{p}$): The probability of selecting action a_i from the action set $\mathbf{a}$ by tenant k at iteration f is denoted by $p_{ik}(f)$. Thus, the action probability vector for tenant k may be defined as $\mathbf{p}_k(f) = [p_{0k}(f), p_{1k}(f), \cdots, p_{Nk}(f)]$.
- Reward (R): Based on the optimisation problem in (15), the instantaneous reward received by each tenant k at iteration f is the same, referred to as common payoff, and is given by

$$R(f) = \Psi_\varphi \prod_{k\in\mathcal{K}} \mathrm{sgn}(B_k^{th} - B_{k,\varphi}) \tag{16}$$

  Here, we define $\mathbf{q}_k(f) = \{q_{0k}(f), \cdots, q_{ik}(f), \cdots, q_{Nk}(f)\}$ as the average reward vector for tenant k at iteration f, where $q_{ik}(f)$ is the average reward received by tenant k for taking action a_i up to and including iteration f, which is updated at iteration f as

$$q_{ik}(f) = \begin{cases} q_{ik}(f-1) + \dfrac{R(f) - q_{ik}(f-1)}{W_{ik}(f)}, \\ \qquad\qquad\qquad \text{if } a_k(f) = a_i, \\ q_{ik}(f-1), \qquad \text{if } a_k(f) \neq a_i \end{cases} \tag{17}$$

  where

$$W_{ik}(f) = \begin{cases} W_{ik}(f-1) + 1, & \text{if } a_k(f) = a_i \\ W_{ik}(f-1), & \text{if } a_k(f) \neq a_i \end{cases} \tag{18}$$

  which denotes the number of times action a_i has been selected by tenant k up to and including iteration f. In (17), $a_k(f) = a_i$ indicates that action a_i has been selected by tenant k at iteration f. As seen from (17), for tenant k, the average reward of taking action a_i will be increased if it has been selected at iteration f and the average rewards of all other actions will remain unchanged, i.e. equal to their average rewards at iteration $f-1$.
- Learning scheme ($\mathbf{L}$): After updating the average reward vector $\mathbf{q}_k(f), \forall k \in \mathcal{K}$, the action probability vector for all tenants will be updated accordingly as

$$\mathbf{p}_k(f) = (1-\theta)\mathbf{p}_k(f-1) + \theta \mathbf{e}_{m_k(f)} \tag{19}$$

  where $0 < \theta < 1$ is the learning rate, $\mathbf{e}_i$ is an $(N+1)$-vector whose i^{th} entry is one and the others are zeros, and $m_k(f)$ is the index of the action at iteration f that incurs

the highest average reward among all other actions for tenant k, i.e. the highest average spectrum utilization. Mathematically,

$$m_k(f) = \text{argmax}\ \{q_{0k}(f), \cdots, q_{Nk}(f)\} \tag{20}$$

Based on the above mathematical development, the proposed pursuit learning based RBS policy is presented in Algorithm 1 and can be explained as follows. At iteration f, each tenant $k, \forall k \in \mathcal{K}$, will take an action a_i according to its action probability vector $\mathbf{p}_k(f)$. The a_i action indicates that i RBs will be reserved for tenant k for its exclusive use. Then, based on the actions taken by all tenants, each tenant will receive the same reward $R(f)$, which is calculated by using (16). Then, the entries of the average reward vector $\mathbf{q}_k$ for each tenant k will be updated by using (17), in which the average reward for the selected action at step f will be increased while keeping the average rewards of all other actions unchanged. Then, we find the index of the action that incurs the highest average reward for each tenant k by using (20). Next, by using (19), the action probability vector $\mathbf{p}_k(f)$ for each tenant will be updated. In (19), the probability of the action that resulted in the highest average reward will be increased, while the probabilities of all other actions will be decreased. Finally, Algorithm 1 will start a new iteration until we have an action that has a probability greater than 0.99 for each tenant $k, \forall k \in K$.

The aforementioned optimisation process will be valid for long periods of time unless major system parameters change, for example, the spectrum bandwidth or the number of tenants allocated on the same physical BS, which do not change unless the operator decides to buy more spectrum or to introduce a new tenant into its network, respectively. The optimisation process is to be done as well in case the long-term traffic load characteristics of any tenant has changed, for example, more sensors are being installed by a utility in its smart grid virtual network.

B. CONVERGENCE BEHAVIOR

The pursuit learning scheme is considered the best among all learning automata schemes as it is shown to be ϵ-optimal with fast convergence [33]. The pursuit learning scheme is ϵ-optimal if, for any $\epsilon > 0$, there exists a fixed $\theta > 0$ such that $\lim\inf_{f\to\infty} \mathbf{p}_k(f) > 1 - \epsilon$ with probability one [45], $\forall i \in \{0, \cdots, K\}$, $\forall k \in \{1, \cdots, N\}$. It also has been shown in [33] that the pursuit learning algorithm, such as the one in Algorithm 1, will eventually select a local optimal action for each tenant k [33]. In other words, the number of allocated RBs for each tenant will eventually give the maximum spectrum utilization. Thus, if $q_{ik}(f)$ in Algorithm 1 is the largest average reward in the vector $\mathbf{q}_k = [q_{1k}, \cdots, q_{Nk}]$, then the probability of action a_i gets close to 1 as $f \to \infty$. This indicates that the probability of the optimum action increases at each iteration until it gets close to 1. Note that the learning rate θ in (19) is to increase the probability of the action that leads to the maximum average reward estimate at the current iteration, while decreasing the probabilities of all other actions. Note that the big O notation of this algorithm is given by $O(KN)$. The convergence and the optimality of the pursuit learning algorithm are discussed in details in [8] and [33].

Algorithm 1 Pursuit Learning Based RBS Policy for Tenant $k, k = 1, \ldots, K$

Inputs: λ_k, μ_k, T_k, A_k, $\mathcal{F}_k$, $\mathcal{W}_k$, and d_k.

1. Initialisation:
   - Set $p_{ik}(0) = 1/(N+1), \forall k \in \mathcal{K}, 0 \le i \le N$.
   - $W_{ik}(0) = 0, 0 \le i \le N, \forall k \in \mathcal{K}$.
   - $q_{ik}(0)$ is initialized by selecting action a_i several times and recording the proportion of rewards.
   - Set $f = 1$.
2. At step f, each tenant k selects an action $a_k(f)$ based on its probability action vector $\mathbf{p}_k(f)$ and then all tenants receive the same reward $R(f)$ calculated by using (16).
3. If $R(f) \neq 0$, update $W_{ik}(f)$ by using (18).
4. Compute the average reward obtained by taking action a_i by tenant k up to and including step f by using (17).
5. $\forall k \in \mathcal{K}$, compute the index of the action that results in the highest average reward by using (20).
6. $\forall k \in \mathcal{K}$, update the action probability vector by using (19).
7. If, $\forall k \in \mathcal{K}$, there is a probability $p_{ik}(f), 0 \le i \le N$, which is larger than 0.99, then stop. Otherwise, set $f \leftarrow f + 1$ and go to step 2.

Outputs: The number of RBs allocated for each tenant k, $(N_k, k = 1, \cdots, K)$

IV. MODEL VALIDATION IN WIRELESS CELLULAR NETWORKS

To prove the accuracy of the proposed analytical framework, we develop a system-level event-based simulator built by using MATLAB [27], [28]. In this simulator, an event represents a request for RBs from a tenant or a release of RBs from a tenant. While the inputs for this simulator will be abstracted from the network, in reality, those inputs may be abstracted from the MAC layer. The proposed simulator is based on 5G wireless cellular network having C cells with the target cell being in the center. The available bandwidth at every cell, BW, is divided to N RBs. The simulator considers the contiguous allocation constraints required by the 3GPP standards on the allocation of RBs to the users in 5G uplink [46]. Those constraints ensure that the same RB cannot be allocated to more than one user, and all RBs allocated to the same user are adjacent and have the same Modulation and Coding Scheme (MCS).

The distance between two adjacent BSs, referred to as inter-site distance, is assumed to be r meters. The distribution of the users is taken to be uniform. As in [47], the channel gain between user u of tenant k, which is assigned RB n, and BS c, is given by

$$G_{u,n}^{c,k} = H_{u,n}^{c,k} \Upsilon_{u,n}^{c,k} \tag{21}$$

where $H_{u,n}^{c,k}$ is the mutlipath Rayleigh fading or small scale fading, and $\Upsilon_{u,n}^{c,k}$ is the large scale fading, which includes the path loss and the log normal shadowing. Since the signal envelope of the small scale fading is described by a Rayleigh distribution, we model $H_{u,n}^{c,k}$ as statistically independent random variables following an exponential distribution with a mean of zero a variance of one. This is possible because $H_{u,n}^{c,k}$ represent geographically separated wireless channels, that show independent multipath fading characteristics [47]. The large scale fading $\Upsilon_{u,n}^{c,k}$ is constant during the optimisation process. The path loss in dB is given by $128.1+37.6\log_{10}(d)$, such that d is the distance in kilometers between a user and its respective BS. Accordingly, $G_{u,n}^{c,k}$ are independent random variables following the exponential distribution with parameters $\chi_{u,n}^{c,k}$, such that

$$\chi_{u,n}^{c,k} = \frac{1}{\mathbb{E}[G_{u,n}^{c,k}]} = \frac{1}{\Upsilon_{u,n}^{c,k}} \tag{22}$$

where $\mathbb{E}[.]$ is the expected value operator. Since $H_{u,n}^{c,k}$ fluctuates so fast, the BS may average it out on its channel measurements. Therefore, BS c estimates $\Upsilon_{u,n}^{c,k}$ (and equivalently $\chi_{u,n}^{c,k}$) for all of its users. In such a setting, the long term uplink Signal-to-Interference-plus-Noise Ratio (SINR) for user u of tenant k in cell c, when assigned RB n, is calculated as

$$\gamma_{u,n}^{c,k}(t) = \frac{\xi_{u,n}\Upsilon_{n,u}^{c,k}}{\sum_{i\in\mathcal{I}_c}\xi_{i,n}\Upsilon_{u,n}^{c} + \sigma^2}, \tag{23}$$

where σ^2 is the noise power, $\mathcal{I}_c$ represents the set of users outside cell c and interfering with user u in cell c, and $\xi_{u,n}$ is the transmit power of user u on RB n. The transmit power is calculated by using an open-loop power control scheme given in [48] as

$$\xi_{u,n}[\text{dBm}] = \min\{\xi_{\max}, \xi_0 + 10\log_{10}|\mathcal{D}_u^k| + \alpha\text{PL}\} \tag{24}$$

where $\xi_{\max}$ is the user maximum transmit power, ξ_0 is a reference power, $\mathcal{D}_u^k$ is the set of RBs assigned to user u of tenant k and $|\mathcal{D}_u^k|$ is its cardinality, α is a constant path loss factor, and PL is the path loss between a user and its serving BS. Therefore, the long term capacity for user u of tenant k in cell c, when assigned $|\mathcal{D}_u^k|$ RBs, is given by

$$R_u^{c,k}(t) = \sum_{n\in\mathcal{D}_u^k} R_{u,n}^{c,k}(t) = \sum_{n\in\mathcal{D}_u^k} \frac{BW}{N} log_2(1+\gamma_{u,n}^{c,k}), \tag{25}$$

and the long term transmission time of tenant k is calculated as

$$\tilde{T}_k^c = \sum_{u\in\mathcal{U}_k} S_u^{c,k}(t) = \sum_{u\in\mathcal{U}_k} \frac{Z_u^{c,k}}{\min_{n\in\mathcal{D}_u^k} R_{u,n}^{c,k}(t) \times |\mathcal{D}_u^k|}, \tag{26}$$

where $Z_u^{c,k}$ is a random variable representing the number of transmitted bits by user u of tenant k in cell c. Thus, the pdf of $Z_u^{c,k}$ is written as

$$P(Z_u^k = x) = \lambda_z e^{-\lambda_z x} \tag{27}$$

where λ_z is the average number of transmitted bits per unit time. This assumption of exponentially-distributed random variables $Z_u^{c,k}$ is widely used in the literature in order to keep the analytical model tractable [49]. The denominator of (26) shows that the capacity for each of the adjacent RBs in $\mathcal{D}_u^k$ allocated to user u is the same capacity of the RB with the worst channel conditions. This is known as the robust Modulation and Coding Scheme (MCS) mode, that ensures the same Block-Error Rate (BLER) performance for each RB by allocating the same MCS for all RBs assigned to the same user [50]. Given (25), the capacity for tenant k is calculated by

$$R^{c,k}(t) = \sum_{u\in\mathcal{U}_k} R_u^{c,k}(t) \tag{28}$$

where $\mathcal{U}_k \equiv \{1,\cdots,U^k\}$ and U^k is the number of users of tenant k. The transmission time in (26) indicates the time duration over which a tenant will keep the RBs for its exclusive use.

In the uplink transmission, the number of bits each user has in its buffer will be sent to the BS by using the Buffer Status Report (BSR) [51]. Therefore, in the proposed simulator, each user u sends a BSR to its serving virtual BS. Then, tenant k uses this BSR to find the number of RBs needed to serve this user by

$$\begin{aligned}\tilde{\mathcal{F}}_k &= \sum_{u\in\mathcal{U}_k} \min\{d_k, |\mathcal{D}_u^k|\} \\ &= \sum_{u\in\mathcal{U}_k} \min\left\{d_k, \left\lfloor \frac{Q_u^k}{\min_{n\in\mathcal{N}}\{R_{u,n}^{c,k}\} \times T_s} \right\rfloor\right\}\end{aligned} \tag{29}$$

where Q_u^k is the buffer length in bits of user u of tenant k, where $u \in \mathcal{U}^k$. The buffer length is conveyed to the BS by using BSR. T_s is the minimum scheduling interval in 5G, normally 1 ms, $\lfloor x \rfloor$ is the maximum integer less than x. The constant d_k represents the maximum number of RBs which may be assigned to or released by tenant k. It is specified in the SLA between the tenant and the NI operator. Based on (29), each tenant estimates the number of RBs that should be requested from the NI operator to serve its users. By investigating (29), we can see that the number of RBs allocated to a user is proportional to the number of bits in its buffer.

To find the pdfs of the transmission times in (26) for tenant k, denoted by $T_k(y_k)$, and its mean service time, denoted by μ_k, we exploit the distribution fitting approach. Firstly, the simulator generates a set of the transmission times by using (26). This set consists of 10000 values for the transmission times of the users belonging to tenant k. Secondly, we use the distribution fitting approach to obtain the best distribution that fits this set of data. The steps of the distribution fitting approach may be summarized in three steps as follows: 1) select the pdfs that may fairly describe the perceived empirical data; 2) calculate the parameters for each possible distribution by applying the maximum likelihood estimation method; and 3) verify the accuracy of the fit by the Kolmogorov-Smirnov (K-S) test. More details on the distribution fitting can be found in [42]. By following the above steps, we calculate the pdf of the transmission times for

TABLE 1. Simulation parameters.

Parameter	Setting
Operating Frequency	2 GHz
System bandwidth (BW)	20 MHz
No. of cells C	19
inter-site distance r	500 m
Number of RBs (N)	100
Channel model	Typical urban [57]
Distance-dependent path loss	$128.1 + 37.6\log_{10}(R)$, R in km
Shadowing Standard deviation	8 dB
Power control	α=0.8, $\xi_0 = -60$ dBm
Max. transmission power ξ_{max}	24 dBm
Thermal noise spectral density	-174 dBm/Hz

each tenant. Once all these pdfs are obtained, we may then apply the analytical models presented in Section II to find the blocking probability and spectrum utilization. We also apply the above approach to determine the pdf of the random variable $\tilde{\mathcal{F}}_k$, $\forall k \in \mathcal{K}$, that represents the number of RBs requested by tenant k.

V. SIMULATION RESULTS

We show the accuracy of the proposed analytical framework by using the simulator presented in Section IV. We assume four tenants, $K = 4$, which are smart metering (MTC), transportation (URLLC), eHealth (MTC), and connected consumer (eMBB). The processes describing the number of requests for RBs, i.e., $A_i(x_i)$, $\forall i \in \{1, 2, 3, 4\}$, is assumed to be periodic for smart metering and transportation tenants [3], [4], [52] and to be Poisson for eHealth and connected consumer tenants [53]–[55]. As shown in [3], [4], and [52]–[55], these distributions are shown to be the best distributions to describe the traffic generated by the services offered by the aforementioned tenants. The value of d_k is set to be 5 RBs for all tenants. As in [56], the packet size for all users follows an exponential distribution with a mean size of 200 bytes, since this is the typical packet size in M2M communications. In the figures, 'Ana' and 'Sim' refer to the analytical and simulation results produced from the analytical model and the simulator, respectively. The simulation parameters are given in Table 1.

To determine the pdf T_k of the transmission times for tenant k, the pdf $\mathcal{W}_k$ of the number of RBs released by tenant k per release event, and the pdf $\mathcal{F}_k$ of the number of requested RBs by tenant k per request event, we generate 10000 simulation values for the tenants by using (26) and (29) and obtain their means, μ_k, $\mathbb{E}[\tilde{W}_k]$, and $\mathbb{E}[\tilde{F}_k]$, respectively. Then, we use the distribution fitting method to obtain the best distribution that fits this set of data as explained in previous section. The pdf of the arrival process for tenant k, denoted by A_k, depends on the application that tenant k is serving (examples are given in the first paragraph of Section V). Once all these pdfs are obtained, we can then use the analytical framework presented in Section II to find the blocking probability and spectrum utilization.

A. MODEL VALIDATION

We validate the accuracy of the proposed analytical framework and the expressions in Section II-B by plotting the blocking probability and spectrum utilization as functions of the total arrival rate $\lambda = \sum_{i=1}^{4} \lambda_i$ and the total number of RBs allocated to all tenants $N_T = \sum_{k=1}^{4} N_k$. In addition, we put $\lambda/N_T = \lambda_i/N_i, \forall i \in \{1, 2, 3, 4\}$, to assure fairness among the tenants and to guarantee that the number of RBs reserved by each tenant is proportional to its arrival rate. For example, if $\lambda = 100$ RBs/sec and $N_T = 40$ RBs, then we may have $\lambda_1 = 10$ RBs/sec, $\lambda_2 = 20$ RBs/sec, $\lambda_3 = 30$ RBs/sec, $\lambda_4 = 40$ RBs/sec, $N_1 = 4, N_2 = 8, N_3 = 12$, and $N_4 = 16$. Alternatively, for the same λ and N_T, we may have $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 25$ RBs/sec, $N_1 = N_2 = N_3 = N_4 = 10$.

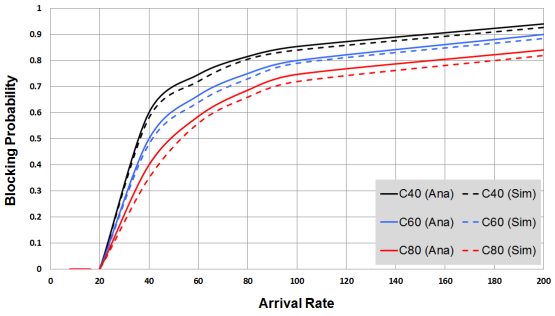

FIGURE 3. Blocking probability vs. the arrival rate.

In Fig. 3, the blocking probability as a function of the arrival rate λ is shown. Here, the total number of allocated RBs, N_T, is taken to be 40, 60, and 80, denoted by scenarios C40, C60, and C80, respectively. Firstly, it is clear that the blocking probabilities calculated by using the analytical model and the simulation match very well. This validates the accuracy and universality of the proposed analytical framework in modeling different sharing policies. Secondly, as expected, the blocking probability increases as the arrival rate increases in all of the three scenarios. A higher arrival rate indicates more users arrive to the system leading to more requests and consequently a higher competition for the limited network resources. As a result, the blocking probability increases. A higher blocking probability means more users will be rejected and will not get any of the network resources, which eventually degrades the QoS offered to the users. Finally, we notice that the blocking probability decreases as the total number of allocated RBs to all tenants, N_T increases. By having more assigned RBs for its exclusive use, each tenant can serve more of its incoming users, leading to smaller blocking probability. However, when the number of assigned RBs decreases, a smaller number of users can be served and more users will be competing for the common pool of RBs shared by all tenants. This leads to a higher blocking probability.

For the same aforementioned scenarios, C40, C60, and C80, Fig. 4 presents the spectrum utilization as a function of the arrival rate λ. Firstly, similar to the blocking probability, the analytical and simulation results for the spectrum utilization match very well. Secondly, as the number of total number of reserved RBs N_T decreases, a higher spectrum utilization

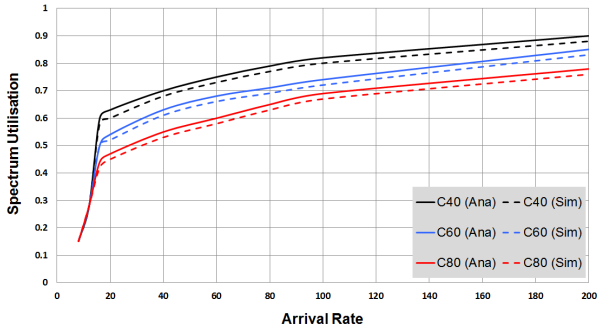

FIGURE 4. Spectrum utilization vs. the arrival rate.

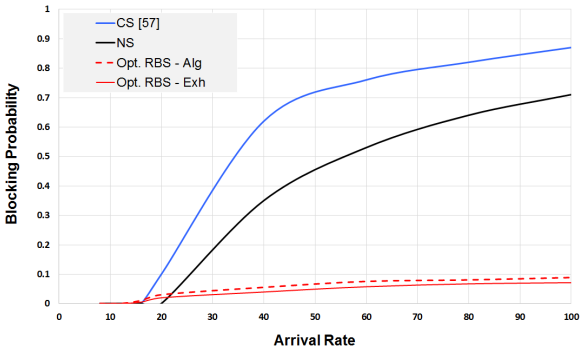

FIGURE 5. Blocking probability vs. the arrival rate for various sharing policies.

is achieved. A smaller number for N_T means that there are more RBs available in the common pool of resources shared by all tenants, leading to a higher spectrum utilization. This is because a tenant with a larger number of users may get more RBs from the common pool to serve them, leading to more RBs being allocated and consequently a higher spectrum utilization. On the contrary, when N_T increases, the number of RBs in the common pool decreases and more RBs are assigned for the exclusive use of each tenant. Therefore, in this case, a tenant with small number of users will have most of its assigned RBs not being used, leading to smaller spectrum utilization. From Figs. 3 and 4, we deduce that N_T is inversely proportional to both blocking probability and spectrum utilization. Therefore, it is vital to determine the optimal number of assigned RBs, N_T, that guarantees the blocking rate requirements of each tenant, whilst increasing the spectrum utilization.

B. PERFORMANCE OF THE PURSUIT LEARNING BASED RBS POLICY

Here, we evaluate the performance of the pursuit learning based RBS policy with respect to its blocking probability and spectrum utilization. The blocking probability threshold in (15) for all considered tenants is taken to be $B_1^{th} = B_2^{th} = B_3^{th} = B_4^{th} = 0.1$. For simplicity, the priority weight is also taken to be the same for the four tenants, $w_k = 1$, $\forall k \in \{1, 2, 3, 4\}$. We compare the pursuit learning based RBS policy with the *NS* policy, where $\sum_{k=1}^{K} N_k = N$, and with the *CS* policy [58], where $N_k = 0, \forall k \in \{1, 2, 3, 4\}$. Fig. 5 presents the blocking probability for the *CS*, *NS*, and the learning based RBS sharing policies as a function of the total arrival rate λ. The learning based RBS policy is obtained by solving (15) using the exhaustive search and Algorithm 1. We notice that Algorithm 1 provides a solution for problem (15), which closely matches the solution obtained by exhaustive search. This is because, at each iteration of Algorithm 1, the probability of selecting the solution that maximizes the spectrum utilization in (15) is increased while decreasing the probabilities of selecting all other solutions, as shown from (19). In other words, the probability of the optimal solution increases at each iteration until it becomes almost 1.

We can also notice that the blocking probability performance for the learning based RBS policy is less than the threshold value of 0.1 for the whole selected range of the arrival rates. This confirms that the proposed learning based RBS policy outperforms the *CS* and *NS* policies and guarantees the blocking probability requirements for all tenants. In addition, it can be seen that the blocking probability performance for the *NS* policy is better than the blocking probability performance for the *CS* policy. This could be explained by the fact that in the *NS* policy, each tenant will have a dedicated exclusive number of RBs. Thus, each tenant guarantees to serve its users by using its own fixed number of RBs. That is, the RBs allocated to a tenant cannot be accessed by other tenants. However, in the *CS* policy, all tenants share all available RBs. Therefore, a tenant with high traffic intensity may then hold a large number of RBs in order to meet its high traffic demands, leaving other tenants with only few RBs and consequently a higher blocking probability. By investigating Fig. 5, we notice that the *CS* and learning based RBS policies represent the upper and lower limits for the blocking probability performance, respectively. Note that, in the learning based RBS policy, the blocking probability cannot exceed the blocking probability threshold for each tenant, as shown in (15).

Fig. 6 shows the spectrum utilization under the *CS*, *NS*, and learning based RBS policies, where the learning based RBS policy is obtained by solving (15) using the exhaustive search and Algorithm 1. We notice that the learning based RBS policy provides the highest spectrum utilization, compared to the *CS* and *NS* policies, since it assigns the RBs to the tenants dynamically, according to their loads, in order to meet the blocking probability thresholds. We also notice that Algorithm 1 provides a solution for problem (15), which closely matches the solution obtained by exhaustive search. The figure also shows that the *CS* policy offers a better spectrum utilization compared to the *NS* policy, because under *CS* policy, the complete spectrum is shared by all tenants. Thus, any tenant may seize as many RBs as it requests, leading to a higher spectrum utilization. Conversely, for the *NS* policy, each tenant may obtain only a predefined number of RBs all the time. Thus, in case the tenants are

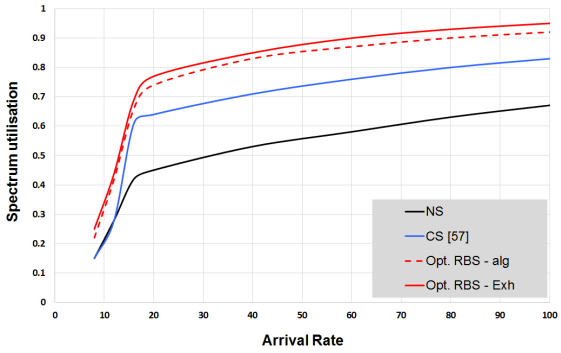

FIGURE 6. Spectrum utilization vs. the arrival rate for various sharing policies.

TABLE 2. Traffic characteristics for various tenants.

Tenant Index	Tenant Application	Traffic Models
1	Smart Metering [3], [52], [56]	Periodic, 5 mins periodicity [56]
2	eHealth [53], [54]	Poisson, λ = 55 RBs/sec [54]
3	Transportation [4], [56]	Periodic, 10 s periodicity [56]
4	Connected Consumer [55]	Poisson, λ = 200 RBs/sec
5	Tracking and Tracing [59]	Periodic, 5 sec periodicity
6	Payment [60]	Poisson, λ = 100 RBs/sec
7	Remote Maintenance and Control [59]	Poisson, λ = 1000 RBs/sec
8	Environment Monitoring (Earthquake) [56]	Uniform distribution over a period of 1 s for 126 sensors/cell
9	Public Services [61]	Poisson, λ = 300 RBs/sec

lightly loaded, the utilization of their portions of the spectrum is small, resulting in a smaller overall spectrum utilization. By investigating Fig. 6, we notice that the *NS* and learning RBS policies represent the lower and upper limits for the spectrum utilization, respectively. We can also conclude that the *CS* policy is preferred for a heavily-loaded tenant while the *NS* policy is preferred for a lightly-loaded tenant.

C. EFFICIENCY GAIN

In this section, we show how the number of tenants sharing the spectrum may affect efficiency gain in (14), Υ. We assume nine tenants, i.e. $K = 9$, as illustrated in Table 2. For the *NS* policy, we consider that every tenant owns a physical spectrum based on a 10-MHz 5G system having 50 RBs. For the case of *PS* sharing policy, all tenants share their spectrums and the RBs are assigned according to the pursuit learning based RBS policy in Algorithm 1. Fig. 7 shows the efficiency gain in (14) as a percentage on the right y-axis, and the effective spectrum utilization in (13) for the *PS* and *NS* policies on the left y-axis. First, we notice that the *PS* policy leads to a higher effective spectrum utilization. This is because the RBs are allocated to the tenants based on their traffic loads under the *PS* policy. This ensures that the spectrum is not wasted by assigning a quite large number of RBs to

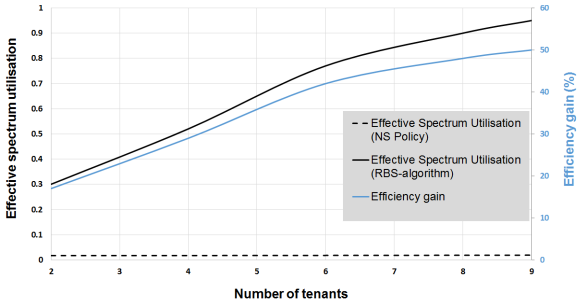

FIGURE 7. Effective Spectrum utilization and efficiency gain as a function of the number of tenants.

TABLE 3. Convergence of algorithm 1 and its utilization for various numbers of tenants and actions.

# Tenants	Env.	# actions	# iterations	Utilisation
2	E_5	25	1157	0.85
	E_{10}	50	1823	0.79
	E_{20}	100	2814	0.76
3	E_5	25	712	0.91
	E_{10}	50	1356	0.87
	E_{20}	100	1983	0.83
4	E_5	25	578	0.98
	E_{10}	50	994	0.95
	E_{20}	100	1479	0.93

the tenants with small traffic loads. An important observation here is that the efficiency gain is almost proportional to the number of tenants sharing the spectrum. In other words, for the *PS* sharing policy, which is solved by Algorithm 1, we may achieve a higher spectrum utilization in addition to a smaller blocking probability. This is because, under the *PS* policy, more RBs are available for sharing, in comparison to the *NS* sharing policy, where no RBs are shared at all. For instance, two tenants sharing the spectrum achieve an efficiency gain of only 15%. In other words, 15% more users may be admitted to the network in contrast to the *NS* policy. On the contrary, the case of 8 tenants achieves an efficiency gain of 48%.

D. CONVERGENCE OF RBS ALGORITHM

To investigate the convergence speed of Algorithm 1, we consider three environments with different numbers of actions, i.e., different numbers of available RBs. These environments, referred to as E_5, E_{10} and E_{20}, represent the cases of 5-MHz, 10-MHz, and 20-MHz 5G physical spectrums with 25, 50, and 100 RBs, and consequently 25, 50, and 100 actions, respectively. Table 3 illustrates the number of iterations required until Algorithm 1 of the learning based RBS policy converges for each of the above environments for various number of tenants, when the stopping criterion is taken to be 0.99. First, for a given number of tenants, it is noticed that the number of iterations increases nearly in proportion to the number of actions. In other words, as the set of possible actions increases, Algorithm 1 needs more time to find the optimum action from this set. Second, for a given environment, the number of required iterations until

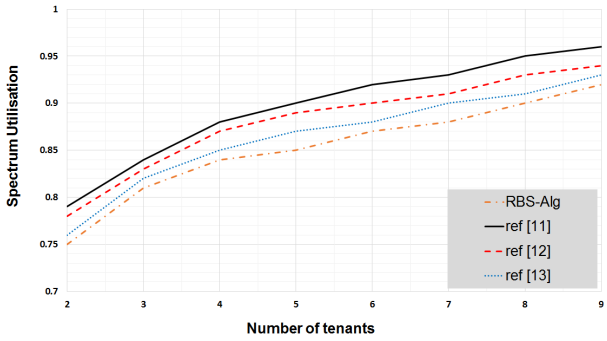

FIGURE 8. Spectrum utilization vs. the number of tenants for various spectrum sharing models.

Algorithm 1 converges decreases as the number of tenants increases. In particular, we notice that the convergence speed of the algorithm increases by approximately n-fold as the number of tenants increases by n-fold. For example, for E_{10}, the number of required iterations for 2 tenants is 1823, while it is 994 for 4 tenants. A similar finding is obtained in [62]. This finding indicates that the proposed pursuit learning-based RBS scheme is scalable and can be implemented for spectrum sharing involving a large number of tenants. Table 3 also shows the utilization achieved for the considered environments for various number of tenants. It is clear that, for a given environment, as the number of tenants increases, we achieve a better spectrum utilization.

E. COMPARISON OF PURSUIT LEARNING BASED RBS POLICY WITH INSTANTANEOUS LOAD-BASED SPECTRUM SHARING SCHEMES

We compare the performance and complexity of the proposed pursuit learning based RBS sharing policy that takes into account the traffic load statistics, with the instantaneous load-based sharing schemes proposed in [11]–[13]. For these simulations, we consider the 20-MHz 5G and the traffic characteristics of the tenants are shown in Table 2. In Fig. 8, the spectrum utilization of the instantaneous load-based sharing schemes in [11]–[13] is shown to be higher than the spectrum utilization of the proposed pursuit learning based RBS policy. This is because the spectrum allocation in these instantaneous schemes is changed every 1 ms, equivalent to one transmission time interval (TTI) in 5G, to adapt to the changes in instantaneous loads, since the tenant's traffic statistics are assumed unknown. This results in a heavy computation. The spectrum allocation of the pursuit learning based RBS policy exploits the knowledge of the traffic statistics of all tenants to fix the spectrum allocation for longer periods of time until the traffic statistics change. In other words, as long as the traffic distributions remain the same, the optimisation may be done only once and the allocation parameters can be monitored on a long-term basis.

Table 4 compares the complexity of the instantaneous load-based sharing schemes in [11]–[13] with the proposed algorithm in terms of the percentage of the algorithm running time. It is the ratio of the total time required for the algorithm to run to the total simulation time. For example, if the simulation time is 1000 seconds and the total time taken for the algorithm to find a solution in theses 1000 seconds is 200 seconds, then the percentage of the algorithm running time is 200/1000 = 20%. From Table 4, it is clear that the proposed pursuit learning based RBS algorithm needs much less running time compared to the instantaneous load-based sharing schemes, since the pursuit learning deploys the average reward, i.e. average spectrum utilization, in allocating the resources instead of the instantaneous reward. From Fig. 8 and Table 4, we see that the proposed algorithm provides a comparable spectrum utilization compared to the instantaneous load-based sharing schemes, while requiring much less time to be executed completely.

TABLE 4. Percentage of the algorithm running time as a function of the number of tenants for various spectrum sharing schemes.

	No. of tenants							
	2	3	4	5	6	7	8	9
RBS-Alg	3	5	7	10	12	15	17	19
ref [11]	14	21	25	28	31	32	33	34
ref [12]	11	14	18	20	24	26	28	29
ref [13]	13	16	21	24	27	28	30	32

VI. CONCLUSION

In this paper, we derived an analytical framework to investigate various spectrum sharing strategies in a multi-tenant wireless cellular network, which is based on a queuing system. The proposed framework is used to derive analytical expressions for various performance metrics of each tenant under a given sharing policy. We also proposed an optimum sharing policy, which is formulated as an optimisation problem, with the number of RBs reserved for each tenant as the optimisation variables. This optimisation problem is solved by an algorithm based on the pursuit learning scheme. The results show that the framework provides an accurate and effective tool to obtain the optimal spectrum sharing, which simultaneously minimizes the blocking probability and maximizes the spectrum utilization, in proportion to the number of tenants.

REFERENCES

[1] P. Rost *et al.*, "Mobile network architecture evolution toward 5G," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 84–91, May. 2016.
[2] A. Prasad, A. Benjebbour, O. Bulakci, K. I. Pedersen, N. K. Pratas, and M. Mezzavilla, "Agile radio resource management techniques for 5G new radio," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 62–63, Jun. 2017.
[3] *Information Technology—Internet of Things (IoT) Use Cases*, Standard ISO/IEC TR 22417:2017, Danish Standards Foundation (DS), Dec. 2017.
[4] *Machine to Machine Communications (M2M); Use cases of Automotive Applications in M2M Capable Networks*, document ETSI TR 102 898 V1.1.1, European Telecommunications Standards Institute (ETSI), 2013.
[5] C. Sexton, N. J. Kaminski, J. M. Marquez-Barja, N. Marchetti, and L. A. DaSilva, "5G: Adaptable networks enabled by versatile radio access technologies," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 688–720, 2nd Quart., 2017.
[6] *5G Network Architecture—A High Level Perspective*, Huawei, Shenzhen, China, Jul. 2016.

[7] I. Afolabi, T. Taleb, K. Samdanis, A. Ksentini, and H. Flinck, "Network slicing and softwarization: A survey on principles, enabling technologies, and solutions," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2429–2453, 3rd Quart., 2018.

[8] A. Rezvanian, A. M. Saghiri, S. M. Vahidipour, M. Esnaashari, and M. R. Meybodi, *Recent Advances in Learning Automata* (Studies in Computational Intelligence). Cham, Switzerland: Springer, 2018.

[9] P. Rost *et al.*, "Benefits and challenges of virtualization in 5G radio access networks," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 75–82, Dec. 2015.

[10] M. Richart, J. Baliosian, J. Serrat, and J.-L. Gorricho, "Resource slicing in virtual wireless networks: A survey," *IEEE Trans. Netw. Service Manag.*, vol. 13, no. 3, pp. 462–476, Sep. 2016.

[11] K. Zhu and E. Hossain, "Virtualization of 5G cellular networks as a hierarchical combinatorial auction," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2640–2654, Oct. 2016.

[12] S. M. A. Kazmi, N. H. Tran, T. M. Ho, and C. S. Hong, "Hierarchical matching game for service selection and resource purchasing in wireless network virtualization," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 121–124, Jan. 2018.

[13] M. Dighriri, A. S. D. Alfoudi, G. M. Lee, T. Baker, and R. Pereira, "Resource allocation scheme in 5G network slices," in *Proc. 32nd Int. Conf. Adv. Inf. Netw. Appl. Workshops (WAINA)*, May 2018, pp. 275–280.

[14] H. Cao and J. Cai, "Distributed opportunistic spectrum access in an unknown and dynamic environment: A stochastic learning approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4454–4465, May 2018.

[15] M. Fahimi and A. Ghasemi, "A distributed learning automata scheme for spectrum management in self-organized cognitive radio network," *IEEE Trans. Mobile Comput.*, vol. 16, no. 6, pp. 1490–1501, Jun. 2017.

[16] O. Jouini and A. Roubos, "On multiple priority multi-server queues with impatience," *J. Oper. Res. Soc.*, vol. 65, no. 5, pp. 616–632, 2014.

[17] A. Sleptchenko, J. Selen, I. Adan, and G.-J. van Houtum, "Joint queue length distribution of multi-class, single-server queues with preemptive priorities," *Queueing Syst.*, vol. 81, no. 4, pp. 379–395, Dec. 2015.

[18] N. Li and D. A. Stanford, "Multi-server accumulating priority queues with heterogeneous servers," *Eur. J. Oper. Res.*, vol. 252, no. 3, pp. 866–878, 2016.

[19] M. Yu and A. S. Alfa, "A simple method to obtain the stochastic decomposition structure of the busy period in *Geo/Geo/1/N* vacation queue," *4OR*, vol. 13, no. 4, pp. 361–380, Dec. 2015.

[20] A. Brandwajn and T. Begin, "Breaking the dimensionality curse in multi-server queues," *Comput. Oper. Res.*, vol. 73, pp. 141–149, Sep. 2016.

[21] X. Zhao, J. Hou, and K. Gilbert, "Measuring the variance of customer waiting time in service operations," *Manage. Decis.*, vol. 52, no. 2, pp. 296–312, 2014.

[22] T. Van Do, "A closed-form solution for a tollbooth tandem queue with two heterogeneous servers and exponential service times," *Eur. J. Oper. Res.*, vol. 247, no. 2, pp. 672–675, 2015.

[23] U. N. Bhat, "The general queue *G/G/1* and approximations," in *An Introduction to Queueing Theory*. Boston, MA, USA: Birkhäuser, 2015, pp. 201–214.

[24] T. D. Tran and L. B. Le, "Stackelberg game approach for wireless virtualization design in wireless networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[25] W. Mélange, H. Bruneel, B. Steyaert, D. Claeys, and J. Walraevens, "A continuous-time queueing model with class clustering and global FCFS service discipline," *J. Ind. Manage. Optim.*, vol. 10, no. 1, pp. 193–206, Jan. 2014.

[26] O. Al-Khatib, W. Hardjawana, and B. Vucetic, "Traffic modeling and optimization in public and private wireless access networks for smart grids," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1949–1960, Jul. 2014.

[27] O. Al-Khatib, W. Hardjawana, and B. Vucetic, "Wireless networks virtualisation: Traffic modeling and spectrum sharing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 5859–5864.

[28] O. Al-Khatib, W. Hardjawana, and B. Vucetic, "Traffic load-based spectrum sharing for multi-tenant cellular networks for IoT services," in *Proc. IEEE Int. Conf. Commun.(ICC)*, May 2018, pp. 1–6.

[29] D. B. Rawat, "A novel approach for shared resource allocation with wireless network virtualization," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2017, pp. 665–669.

[30] N. Zhang, S. Zhang, J. Zheng, X. Fang, J. W. Mark, and X. Shen, "QoE driven decentralized spectrum sharing in 5G networks: Potential game approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 7797–7808, Sep. 2017.

[31] R. Han, Y. Gao, C. Wu, and D. Lu, "An effective multi-objective optimization algorithm for spectrum allocations in the cognitive-radio-based Internet of Things," *IEEE Access*, vol. 6, pp. 12858–12867, 2018.

[32] B. Zhuang, D. Guo, E. Wei, and M. L. Honig, "Scalable spectrum allocation for large networks based on sparse optimization," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 2518–2522.

[33] R. Martin and O. Tilak, "On ϵ-optimality of the pursuit learning algorithm," *J. Appl. Probab.*, vol. 49, no. 3, pp. 795–805, Sep. 2011.

[34] T. Sanguanpuak, S. Guruacharya, N. Rajatheva, M. Bennis, and M. Latva-Aho, "Multi-operator spectrum sharing for small cell networks: A matching game perspective," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3761–3774, Jun. 2017.

[35] A. Asheralieva and Y. Miyanaga, "An autonomous learning-based algorithm for joint channel and power level selection by D2D pairs in heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3996–4012, Sep. 2016.

[36] X. Chen, Z. Han, H. Zhang, G. Xue, Y. Xiao, and M. Bennis, "Wireless resource scheduling in virtualized radio access networks using stochastic learning," *IEEE Trans. Mobile Comput.*, vol. 17, no. 4, pp. 961–974, Apr. 2018.

[37] M. Zandi, M. Dong, and A. Grami, "Distributed stochastic learning and adaptation to primary traffic for dynamic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1675–1688, Mar. 2016.

[38] Y. Gai and B. Krishnamachari, "Distributed stochastic online learning policies for opportunistic spectrum access," *IEEE Trans. Signal Process.*, vol. 62, no. 23, pp. 6184–6193, Dec. 2014.

[39] K. C. Budka, J. G. Deshpande, and M. Thottan, *Communication Networks for Smart Grids: Making Smart Grid Real*. London, U.K.: Springer-Verlag, 2014.

[40] *Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN), Overall Description, Stage 2 (Release 15)*, document 3GPP ETSI TS 36.300 V15.3.0, Third Generation Partnership Project (3GPP), Oct. 2018.

[41] U. N. Bhat, *An Introduction to Queueing Theory: Modeling and Analysis in Applications* (Statistics for Industry and Technology). Boston, MA, USA: Birkhäuser, 2015.

[42] I. Burr and J. Schmidt, *Applied Statistical Methods* (Operations Research and Industrial Engineering). Amsterdam, The Netherlands: Elsevier, 2014.

[43] M. H. A. Davis, *Markov Models & Optimization*. Boca Raton, FL, USA: CRC Press, 2018.

[44] *Numerical Methods in Markov Chain Modeling*. Nat. Aeronaut. Space Admin., Washington, DC, USA, 2018.

[45] J. Zhang, C. Wang, and M. C. Zhou, "Fast and epsilon-optimal discretized pursuit learning automata," *IEEE Trans. Cybern.*, vol. 45, no. 10, pp. 2089–2099, Oct. 2015.

[46] *Evolved Universal Terrestrial Radio Access (E-SUTRA); Physical Channles and Modulation*, document 3GPP ETSI TS 36.211 v. 15.0.0, Third Generation Partnership Project (3GPP), 2018.

[47] H. Boostanimehr and V. K. Bhargava, "Unified and distributed QoS-driven cell association algorithms in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1650–1662, Mar. 2015.

[48] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures*, document 3GPP ETSI TS 36.213 v. 15.0.0, Third Generation Partnership Project (3GPP), 2018.

[49] X. Chen, C. Wu, Y. Zhou, and H. Zhang, "A learning approach for traffic offloading in stochastic heterogeneous cellular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 3347–3351.

[50] A. Kumar, A. Sengupta, R. Tandon, and T. C. Clancy, "Dynamic resource allocation for cooperative spectrum sharing in LTE networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5232–5245, Nov. 2015.

[51] *Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) Protocol Specification; (Release 15)*, document 3GPP ETSI TS 36.321 V15.3.0, Third Generation Partnership Project (3GPP), Oct. 2018.

[52] Ausgrid. *Smart Grid Smart City*. [Online]. Available: http://www.smartgridsmartcity.com.au/

[53] *Machine-to-Machine Communications (M2M); Use Cases of M2M Applications for eHealth*, document ETSI TR 102 732 V1.1.1, European Telecommunications Standards Institute (ETSI), Sep. 2013.

[54] S. M. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K.-S. Kwak, "The Internet of Things for health care: A comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015.

[55] *Machine-to-Machine communications (M2M); Use Cases of M2M Applications for Connected Consumer*, document ETSI TR 102 857 V1.1.1, European Telecommunications Standards Institute (ETSI), Aug. 2013.
[56] *Technical Specification Group Services and System Aspects; Service Requirements for Machine—Type Communications (MTC); (Release 14)*, document 3GPP TR 22.368 V14.0.1, Third Generation Partnership Project (3GPP), Aug. 2017.
[57] *Technical Specification Group Radio Access Network; Deployment Aspects (Release 15)*, document 3GPP Std. TR 25.943 v. V15.0.0, Third Generation Partnership Project (3GPP), Jul. 2018.
[58] M. Vincenzi, A. Antonopoulos, E. Kartsakli, J. Vardakas, L. Alonso, and C. Verikoukis, "Multi-tenant slicing for spectrum management on the road to 5G," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 118–125, Oct. 2017.
[59] S. Ramnath, A. Javali, B. Narang, P. Mishra, and S. K. Routray, "IoT based localization and tracking," in *Proc. Int. Conf. IoT Appl. (ICIOT)*, May 2017, pp. 1–4.
[60] *Machine-to-Machine Communications (M2M); M2M Service Requirements*, document ETSI TS 102 689 V2.1.1, European Telecommunications Standards Institute (ETSI), Jul. 2013.
[61] *ONEM2M: Standards for M2M and the Internet of Things*, Standard TR-0001-V2.4.2, Mar. 2018.
[62] H. Ge, Y. Guo, and S. Li, "An efficient parallel pursuit algorithm," in *Proc. 8th Int. Conf. Intell. Hum.-Mach. Syst. Cybern. (IHMSC)*, vol. 1, Aug. 2016, pp. 587–591.

OBADA AL-KHATIB (M'17) received the B.Sc. degree (Hons.) in electrical engineering from Qatar University, Qatar, in 2006, the M.Eng. degree (Hons.) in communication and computer from the National University of Malaysia, Bangi, Malaysia, in 2010, and the Ph.D. degree in electrical and information engineering from The University of Sydney, Sydney, Australia, in 2015. From 2006 to 2009, he was an Electrical Engineer with Consolidated Contractors International Company, Qatar. In 2015, he joined the Centre for IoT and Telecommunications, The University of Sydney, as a Research Associate. Since 2016, he has been an Assistant Professor with the Faculty of Engineering and Information Sciences, University of Wollongong in Dubai, UAE. His current research interests include the areas of smart grid communication, wireless resource allocation and management, cooperative communications, and wireless network virtualization.

WIBOWO HARDJAWANA (M'09) received the Ph.D. degree in electrical engineering from The University of Sydney, Australia, in 2009. From 1999 to 2004, he was with Singapore Telecom Ltd. He is currently an Australian Research Council Discovery Early Career Research Award Fellow with the School of Electrical and Information Engineering, The University of Sydney. His current research interests include wireless communications, with a focus on multiple-input-multiple-output, cloud, virtualization, multiple access, cooperative communications, and coding techniques.

BRANKA VUCETIC (SM'00–F'03) held research and academic positions in Yugoslavia, Australia, U.K., and China. She is currently an ARC Laureate Fellow, a Professor of telecommunications, and the Director of the Centre of Excellence in Telecommunications, The University of Sydney. She has co-authored four books and more than 400 papers in telecommunications journals and conference proceedings. Her research interests include coding, communication theory, and signal processing and their applications in wireless networks and industrial internet of things. She is a Fellow of the Australian Academy of Technological Sciences and Engineering.