Economics of 5G Network Slicing: Optimal and Revenue-based allocation of radio and core resources in 5G

Menglan Jiang, Massimo Condoluci, Toktam Mahmoodi, Luis Guijarro

Abstract

Network slicing in 5G is considered mainly to address the diverse need of vertical industries. However, network slicing will open new avenues for operators business models, and opportunity for such new business models will be the significance of slicing. On one hand, vertical industries as users of 5G have very different needs in terms of data rate, latency, reliability and consistency of the offered service, and level of control over their network slice. Nevertheless, given the diverse business nature of these verticals, the price they might agree to pay for the given slice might vary significantly, depending on the revenue of the vertical business. Hence, there is no clear value for a given network slice. To this end, we propose an auction business model for the operators. We also define a network slice as an aggregation of resources in the central cloud, in the operators cloud, and in the radio access and further discuss how radio network, computation and storage resources can be amalgamated to shape an end-to-end slice. The novel resource allocation and slice allocation we propose have been examined via extensive simulation results assessing the introduced enhancements in providing requirements of the slice owner, network utilization, operators' revenue.

Index Terms

5G, Network slicing, Auctioning, Bertrand equilibrium.

M. Jiang, M. Condoluci, and T. Mahmoodi are with the Department of Informatics, King's College London, UK. e-mail: {menglan.jiang, massimo.condoluci, toktam.mahmoodi}@kcl.ac.uk.

L. Guijarro is with Institute ITACA, Universitat Politecnica de Valencia, Spain. e-mail: lguijar@upv.es.

I. INTRODUCTION

The next-to-come fifth generation (5G) mobile network is expected to open unprecedented business opportunities to telco operators by increasing their market to the business owners and providing not only business-to-customer service (B2C) services but also business-to-business (B2B) and and business to business to consumer (B2B2C) services. Large number of vertical industries are foreseen as natural users of 5G beyond the mobile broadband services, including healthcare automotive, smart cities, and industry automation [1]. In order to deliver services to such wide range of industries with diverse requirements, *network slicing* has been introduced in 5G networks, which can be then foreseen as a composition of multiple slices, each one designed with a set of functionalities tailored to serve a specific business [2].

From the network performance point of view, slicing implies that each 5G slice needs to have its own set of allocated resources and this aspect introduces a novelty in the management of network resources in mobile systems. Indeed, in the previous generations of mobile networks, the resources to be assigned to each application were mainly radio resources [3], while in the 5G network, resources represent both radio and core network [4], by means of computational and storage capabilities in addition to the over-the-air data rate. From the business point of view, however, the issue of pricing a given slice is similar to pricing given spectrum, i.e., there is no clear way to value a given created slice by the operator. Therefore, there is a need for novel business model by the operators.

In order to consider the fact that slice allocation in 5G means allocating resources throughout the network, we model 5G network resources as multiple *chunks*, each one with a different capacity, spread across the whole physical network. This allows to take into consideration the management of resources in the core network in addition to the resources in the radio access. In order to optimize the network revenue by considering network utilization aspects, we focus on a resource management strategy based on a novel competitive auction mechanism combined with an optimization algorithm for network resources allocation. Therefore, objectives are increasing network utilization, enhancing satisfaction of requirements of network slices and increasing the incentive for operators by maximizing their revenue.

The pricing mechanism considered in this paper is built based on the assumption that sellers (i.e., network chunks) are assumed to have the knowledge of products (i.e., network resources) for potential buyers (i.e., network slices). Using the Bertrand model for pricing [5], [6], network chunks receive the information from network slices through the auction mechanism while network slices receives their optimized price values. To formulate the problem of price decision, we propose a game theory based auction mechanism as an effective method for analysing interactive decision making [7] in order to consider the interaction between network chunks and network slices. For maximizing the network revenue, we propose an auction mechanism which is derived from the Vickrey-Clarke-Groves (VCG) [8], [9] approach. Therefore, compared with previous researches which mainly focused on price definition, we design our own auction mechanism by considering the amount of network resources the network slices are requesting to the network. Concerning the resource allocation, unlike game theory based resource management mechanisms in the more traditional way, we take into account the efficiency of resource utilization with the constraints of maximizing the network revenue. Therefore, the Lagrangian based optimization mechanism has been designed for the purpose of optimizing the allocated network resources to network slices without changing their cost values.

The key contributions of this paper can be summarized as follows:

1) Enabling new business model for offering 5G network slices to wide range of vertical industries, through novel pricing model and auctioning mechanism.
2) Presenting a novel augmentation of end-to-end network slices based on both radio and core network resources.
3) Formulating a novel resource management model considering points 1 and 2.
4) Investigating performance of the discussed models thoroughly, and through simulations.

The remainder of this paper is organized as follows. Section II focuses on the business models for 5G network slicing and also review the state of art in resource allocation strategies for network slicing and auction mechanisms. After elaborating our system model in Section III, the problem formulations, our proposed resource allocation mechanism and pricing mechanism will be described in Section V. Section VI provides numerical results to verify our design objectives and analyses performance observations. Finally, the summary of our work is detailed in Section VII.

II. RELATED WORKS

Network slicing drives the business models behind 5G ecosystem by providing an effective way to delivery heterogeneous services of interest for different verticals [10]. As previously discussed, there are three different of business models for network slicing: B2B, B2C, and B2B2C [11]. In B2B model, operators sell customized network resources to enterprises and release full control of end consumers to enterprises. In B2C model, customers purchases customized network resource based on their requirements without considering which operator provides the requested resources. In B2B2C model, operators just provide customized network resources to a broker, and the broker gets more control of the network and engages with end consumer directly. It can be thus noticed that above mentioned business models deal with the allocation of network resources; consequently, the effective revenue as well as the network performance directly depend on the way the operator manages such resources. Nevertheless, resource allocation for network slicing and auction mechanisms to improve network utilization or revenue have been addressed in literature as two separated contexts. In the remainder of this Section, we provide an overview of the related work from both resource allocation and auction points of view and we then describe the novelties of our paper with respect to the existing works in literature.

From a resource allocation point of view, network slicing is strictly related to virtualization [12], enabling the management (deployment, placement, moving, etc.) of network functionalities across the network. Several solutions for efficiently supporting network resource virtualization [13] and resource allocation by using auction approaches [8] have already been proposed. They have been designed to improve the quality of experience (QoE) of mobile users (thus focusing on an end-user point of view and not from a slice point of view) and network utilization. In this section, we briefly overview the most relevant studies. Focusing from a network resource slicing point of view, a resource allocation strategy of virtualized resources for Long Term Evolution (LTE) networks has been proposed in [14]. This work proposed a slicing scheme to allocate virtualized radio resources, i.e., resource blocks, to different service providers (SPs) in order to maximize the radio utilization. The proposed scheme was dynamic and flexible for addressing arbitrary fairness requirements of different SPs. Similarly, [15] proposed a framework of wireless resource virtualization in LTE networks to allow radio network resources to be shared among mobile network operators. An iterative algorithm has been proposed to solve the Binary Integer

Programming (BIP) problem with less computational overhead. Finally, in our previous work [16], we focused on the topic of slice association and resource allocation for mobile users with the aim of increasing the QoE the users.

From an auction point of view, game theory based resource auction mechanism have been widely investigated in existing works [17]. Nash equilibrium was considered as the solution for solving the problem of spectrum sharing in cognitive radio networks in [18]. Stackelberg game mechanism has been formulated for power control in wireless networks to maximize the capacity [19] [20], while a non-cooperative game based power control algorithm has been proposed in [21] together with a base station association scheme for heterogeneous networks. In [22], an auction mechanism has been proposed to maximize the expected revenue of sellers. An auction-based scheme which can be used to develop a synchronous algorithm for solving the optimization problem of resource allocation has been proposed in [23]. The game theory based network virtualization framework has been described in [24] and an auction mechanism has been used for pricing the instantaneous rate consumption. Compared with the other game theory mechanisms, the auction mechanism was widely applied [23] [25] in the situation of competitive resource allocation. Meanwhile, effective allocation of network resources can improve the revenue of both users and networks. In [5], a competitive pricing model has been formulated in a dynamic spectrum access where a few primary services offer spectrum access opportunities to a secondary service.

The above considered resource allocation mechanisms were mainly considering the radio segment (e.g., power allocation, radio resources) without taking into account the other resources in the mobile core networks such as storage and computational resources. In addition, the competitive price mechanism in the existing research did not consider the global resource utilization as well as priority values (i.e., network resources spread across the edge of the network are scarcer than the ones in a central cloud and are more subject to overload) of different network resources. From the discussion above, the relationship between resource allocation and operator revenue becomes one of the key aspect when designing a business-driven network slice allocation. With this aim, our paper extends the approaches in literature by proposing an auction business model for the operators which takes into account the availability of different types of network resources across the network (i.e., both radio and core network segments). The proposed auction mechanism is modelled by considering the total revenue for the operator and the total amount of assigned resources.

III. SYSTEM MODEL

We divide the network architecture into two parts, i.e., the radio access layer (handling over-the-air resources) and the mobile core layer (handling computational and storage resources and consisting of two parts as central cloud and an operator cloud located at the edge). The system model can be seen in Fig 1 and further elaboration are different part of the model are given in the following subsections.

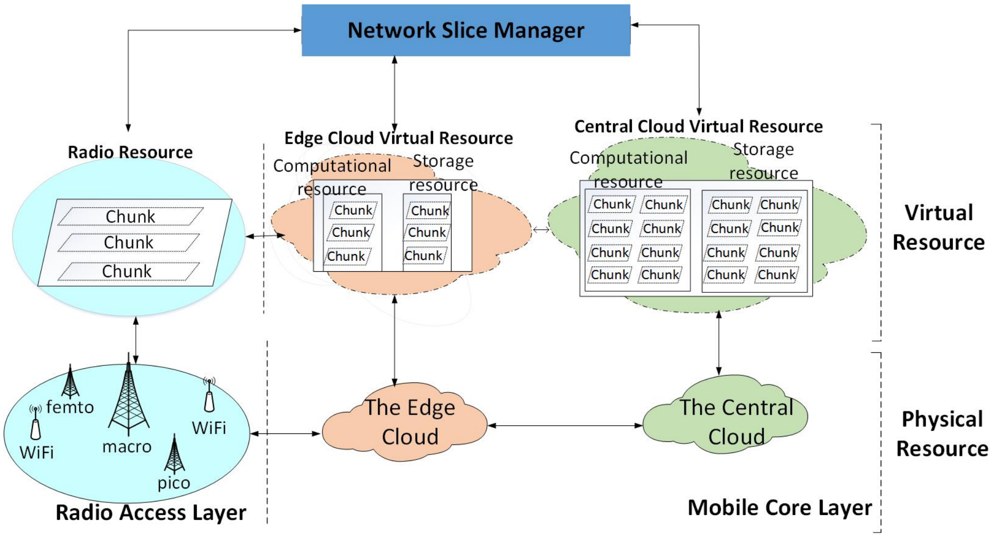


Fig. 1. Our proposed system model for the network architecture with network resources represented as chunks spread across the whole network.

A. Network Slices

The set of network slices requesting is denoted by $\mathbb{K} = \{1, 2, ..., K\}$. We introduce the concept of slice priority in order to consider the fact that some slices could have different importance from an operator point of view (e.g., a slice related to emergency services or a slice of a vertical which is a premium business partner of the operator). For each slice k, the priority has been defined in the range of $\alpha_k = \{1, 2, ..., A\}$, where A indicates the maximum priority level in our system model. The higher α_k, the higher the priority for the given slice, k.

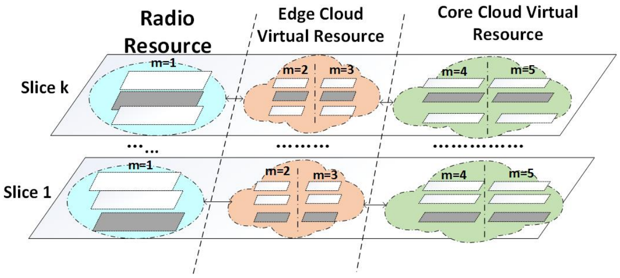


Fig. 2. An example of allocation of resources to different slices over different chunks of the network.

B. Network Resources

In our system model, as shown in Fig 1, the available network resources are modelled as diverse types of network chunks, and the total number of available network chunks is denoted by $\mathbb{M} = \{1, 2, ..., M\}$. From a radio access point of view, different chunks can be seen as different radio access technologies (RATs) providing wireless resource to the network slices. From a core network point of view, the chunks offer resources in terms of computational capability (i.e., CPU) and resource storage capacity (i.e., RAM) in two different locations, i.e., a central cloud and an operator cloud close to the edge.

The resources of each network chunk m are allocated to different slices, as in Fig 2. The capacity of each network chunk m is denoted by η_m, where $m \in \mathbb{M}$. The amount of resources assigned to slice k from resource chunk m is denoted by $\sigma_{k,m}$, where $m \in \mathbb{M}$ and $k \in \mathbb{K}$, respectively. Therefore, the ratio of network resource allocation that network slice k receives from the network chunk m can be explained as $\frac{\sigma_{k,m}}{\eta_m}$. Moreover, in the real network environment, different network chunks will have different amount of network resources which have been described as “weight values” denoted by β_m. This choice allows us to abstract the effective capacity of each network chunk and to focus only on the portion of resources each slice requires from each chunk.

The network slice manager, in charge of running the auction process, price definition and allocating network resources to slices, is assumed to be an impartial entity (i.e. third party). In this case, cost values of different network chunks’ per unit have been described as $\Omega_m = \{\omega_m | m \in \mathbb{M}\}$, and $C_m = \{c_m | m \in \mathbb{M}\}$ indicates the selling price per unit of the network chunk m. The network slice manager will make auction-based resource allocation decisions for network slicing. It is assumed that each network slice is able to provide the slice manager with a request in terms of needed resources. Therefore, network chunks with the best pay-off values generated based on Equation (9), will be allocated to the slices (details can be seen in Fig 3).

C. Slice Requirements

As shown in Fig 2, each network slice in the architecture requires resources in different types of network chunks, which are five in our model: over-the-air resources, the computational and the storage resources from the operator cloud at the edge, and the computational and the storage resources from the central cloud. The radio resource required by the network slice k from the network chunk m is denoted by $\mathbb{R}_{k,m}$. The value of $r_{k,m}$ represents the minimum resource requirements by network slice k from the network chunk m. Value of $R_{k,m}$ represents the maximum resource requirements by network slice k from the network chunk m; such values are assumed to be an indication of the needs of the slices.

IV. OPERATORS REVENUE AND AUCTION MODEL

In the above explained system model, competitive resource allocation mechanism involves three aspects: *a)* the price competition model; *b)* an auction based resource allocation mechanism for network slicing; *c)* an optimized resource allocation for all required slices.

Before determining resource allocation for maximizing the network revenue and optimizing the amount of resources allocated to network slices, the network slice manager should determine the selling price c_m of the network chunk m by using auction mechanism. Then, competitive auction mechanism of network slicing and optimization algorithm for allocating resource to all slices are performed.

A. Price Competition: Bertrand model

As mentioned earlier, the price competition among different network chunks is defined based on Bertrand model, where limited network chunks can compete with each other in order to achieve the optimal profit by controlling the price of each network chunk. The assumption is that both competing parties have the same constant unit cost of production (i.e. unit cost of network chunks), so that marginal and average costs are the same and equal to the competitive price.

Such price competition can be applied to analyze and obtain equilibrium price strategies over the 5G mobile networks. The system model contains several network chunks, and these chunks can be sold to network slices. The price strategies are determined by the network slice manager based on utility value and capacity of network chunks as well as the requested resource by network slices.

The profit maximization problem of the network chunks while serving all slices can be formulated as follows,

P1: $maximize\ \Theta(c_m, \sigma_{k,m})$

The profit function of $\Theta(c_m, \sigma_{k,m})$ is defined in Equation (1).

$$\Theta(c_m, \sigma_{k,m}) = \sum_{k=1}^{K} \sigma_{k,m} c_m - \omega_m \eta_m, \tag{1}$$

where ω_m is the cost value per unit of the virtual network chunk m. The value of ω_m has been defined as $\omega_m = \frac{\eta_m}{\sum_{k=1}^{K} \sigma_{k,m}}$, while the value of $\omega_m \eta_m$ is fixed. The maximum value of Θ is equal to zero.

The profit value per unit of the network chunk m is given by,

$$\theta_m(c_m, \sigma_{k,m}) = \sum_{k=1}^{K} \sigma_{k,m} c_m - \omega_m. \tag{2}$$

The partial differential of the price of virtual network chunk unit is given by,

$$\begin{aligned} \frac{\partial \theta_m(\eta_m)}{\partial c_m} &= \frac{\partial}{\partial c_m} (\sum_{k=1}^{K} \sigma_{k,m} c_m - \omega_m) \\ &= \sum_{k=1}^{K} \sigma_{k,m} + c_m \frac{\partial \sum_{k=1}^{K} \sigma_{k,m}}{\partial c_m} - \omega_m \frac{\partial \eta_m}{\partial c_m}. \end{aligned} \tag{3}$$

In order to maximize total profit for each network chunk, profit per unit of network chunk is maximized. Therefore, $\frac{\partial \theta_m(\eta_m)}{\partial c_m} = 0$, hence Equation (3) is transformed to Equations 4, and 5.

$$\sum_{k=1}^{K} \sigma_{k,m} + c_m \frac{\partial \sum_{k=1}^{K} \sigma_{k,m}}{\partial c_m} = \omega_m \frac{\partial \eta_m}{\partial c_m}, \tag{4}$$

$$\frac{\sum_{k=1}^{K} \sigma_{k,m}}{\frac{\partial \eta_m}{\partial c_m}} + c_m \frac{\partial \sum_{k=1}^{K} \sigma_{k,m}}{\partial \eta_m} = \omega_m. \tag{5}$$

In Equations (4) and (5), $\sum_{k=1}^{K} \sigma_{k,m} = \eta_m$, assuming resource of the network chunk m has been fully allocated to network slices. Therefore,

$$\frac{c_m}{\frac{\partial \eta_m}{\partial c_m}\frac{c_m}{\eta_m}} + c_m = \omega_m \tag{6}$$

$$c_m\left(\frac{1}{\frac{\partial \eta_m}{\partial c_m}\frac{c_m}{\eta_m}} + 1\right) = \omega_m \tag{7}$$

Let $\varepsilon = \frac{\partial \eta_m}{\partial c_m}\frac{c_m}{\eta_m}$, where ε is a small positive value. Therefore, the value of c_m based on the Bertrand Nash equilibrium can be given by $c_m = \frac{\varepsilon}{\varepsilon+1}\omega_m$. Based on the result of Bertrand Nash Equilibrium mechanism, we define a proposition which can be described as follows:

Proposition 1: Based on Bertrand model, equilibrium price values of network chunks per unit is given by $c_m = \frac{\varepsilon}{\varepsilon+1}\omega_m$. This is the selling price per unit of the network chunk m that will be charged to the network slice k while no unused resource in the edge cloud is generated.

Proof 1: If $c_m = \frac{\varepsilon}{\varepsilon+1}\omega_m$, the value of Θ is given by Equation (8), where ε is a small positive value (explained earlier) and $\sum_{k=1}^{K}\sigma_{k,m} = \eta_m$, hence Θ equals to 0.

$$\Theta(C_m, \sigma_{k,m}) = \sum_{k=1}^{K}\sigma_{k,m}\frac{\varepsilon\omega_m(\eta_m)}{\varepsilon+1} - \omega_m\eta_m \tag{8}$$

In Equation (8), the total profit value of allocated resources is equal to the profit value of the total network resource capacity. In this case, there is no incentive for competitors to deviate from their optimized cost values. That is because if one of the competitors chooses to improve the selling price, the selling price will be increased as well by allocating its resource. If one of the competitors chooses to decrease the selling price, it will not have enough resource to meet the increased demands. Therefore, the selling price generated from the competitive mechanism is the optimal value for maximizing the function Θ.

Given c_m is the price per unit of the network chunk m, we can also define the network slice paid value according to their priority levels as,

$$p_{k,m} = \alpha_k c_m \beta_m \tag{9}$$

where $p_{k,m}$ is the price paid by network slice k to the network chunk m, and $c_m\beta_m$ indicates the price of network chunk m.

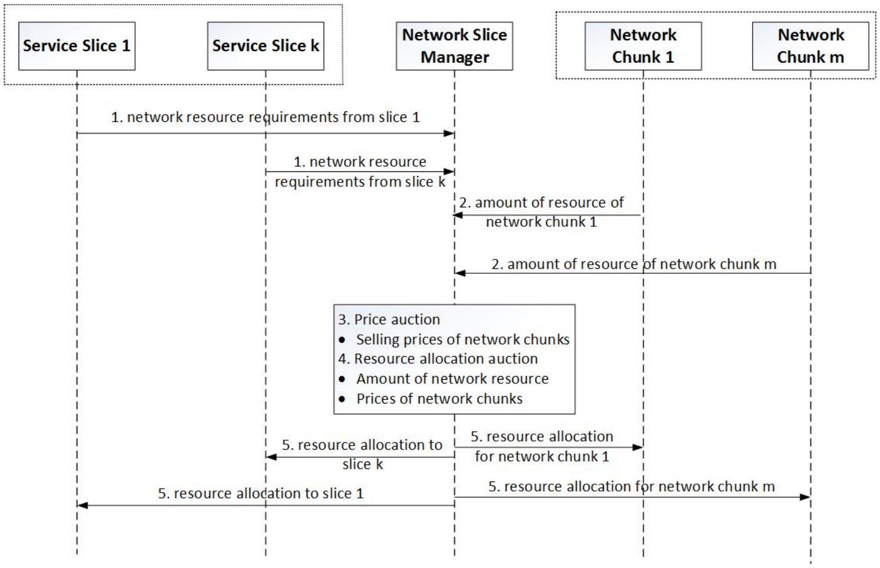


Fig. 3. Auction Flow Chart. Steps of our proposed auction model can be mainly described as: Network slices send a resource request to the slice manager which runs the auction mechanism considering the revenue and the network utilization of all the chunks of the network. Afterwards, slices will run auction mechanism in order to access their optimized resources of all network chunks.

B. *Auction Mechanism for Network Slicing*

In this section, properties of the auction model is detailed. The proposed auction mechanism is based on the VCG mechanism, which has been widely used for pricing in order to improve the efficiency of mobile/wireless resource allocation [23], [25]. According to the system model used in this paper, network slices are players and the aim of auction mechanism is to maximize network revenue by making strategic decisions for all network chunks, i.e., by looking at the auction mechanism from an overall point of view instead of focusing on only one network chunk. Furthermore, the auction mechanism satisfies the social efficiency (social efficiency is guaranteed of all participants' requests are optimized over all feasible allocation mechanisms.).

The network resource of each chunk allocated to service slices are determined by their own resource requirements and the impact on other network slices. Therefore, our auction based network slicing mechanism can be depicted as follows.

- In the first step, the optimal amount of network chunks for network slice k is chosen based on its resource requirements. The generated network revenue because of network slice k is denoted by $U_k(\sigma_{k,m}, p_{k,m}, \omega_m)$ as in equation (10).

$$U_k = \sum_{m=1}^{M} p_{k,m}\sigma_{k,m} - \sum_{m=1}^{M} \omega_m \tag{10}$$

- In the next step, the network revenue is calculated after removing slice k from the network. It can be described as $U_{\mathbb{K}/\{k\}}(\sigma_{\mathbb{K}/\{k\},m}, p_{\mathbb{K}/\{k\},m}, \omega_m)$. In addition to allocating network resources to slice k based on its resource requirements, the VCG mechanism force the network slice manager to consider profit values of other $K-1$ network slices. Given the objective of maximizing total network revenue, function of $U_{\mathbb{K}/\{k\}}$ has been defined in equation (11).

$$U_{\mathbb{K}/\{k\}} = \sum_{\mathbb{K}/\{k\}} \sum_{m=1}^{M} p_{\mathbb{K}/\{k\},m}\sigma_{\mathbb{K}/\{k\},m} - \sum_{m=1}^{M} \omega_m \tag{11}$$

- Finally, the total network revenue is maximized by computing $\sum_{k=1}^{K}(U_k - U_{\mathbb{K}/\{k\}})$.
- The revenue generated by slice k based on auction mechanism Γ, and the total revenue can be expressed as equations (12) and (13) consecutively.

$$U_k^{\Gamma} = U_k - (\max U_{\mathbb{K}/\{k\}}) \tag{12}$$

$$U^{\Gamma} = \sum_{k=1}^{K} U_k^{\Gamma} \tag{13}$$

- Hence, the optimized network resources allocated to slice k based on the auction mechanism can be expressed as equation (14).

$$\sigma_{k,m}^{*} = \arg U_k^{\Gamma} \tag{14}$$

The idea behind our proposal aiming at maximizing the revenue coming while keeping into consideration the amount of resources allocated to network slices, can be expressed with the following proposition.

Proposition 2: In order to allow network slice manager to allocate available network resources using competitive auction mechanisms, while maximizing profit of each network slice k, and maximizing total network revenue, all slices should provide their minimum and maximum requirements.

Proof: Based on the above described auction mechanism, from the network slice side, the optimal received resources of each network chunk m can be expressed as $\sigma^*_{k,m}$, and profit of the network is explained in equation (15).

$$q^*_k = \frac{\sum_{m=1}^{M} p_{k,m}\sigma_{k,m}}{\sum_{m=1}^{M} c_m\eta_m} \tag{15}$$

In order to maximize U^{AUC}, the $p_{k,m}\sigma_{k,m}$ should also be maximized; this shows the relationship between the network revenue and the amount of allocated resources, i.e., $\sigma_{k,m}$.

If all network slices provide the network with their requirements in terms of upper and lower amount of expected resources, all profit values q^*_k can be optimized according to our proposed auction model. So doing, we can guarantee a minimum amount of resources assigned to each slice (this means assigning at least the minimum amount of requested resources to each slice) and we can assign the remaining amount of resources with the aim to maximize the total network revenue. Otherwise, if network slices do not provide their resource requirements in terms of upper/lower bounds, the resource allocation task is not able to properly assign the available network resources to the slices; as noted in (12) and (13), this would decrease the total network revenue.

V. PROBLEM FORMULATION

In this section,formulations of the problem of interest is detailed. From economic point of view, the total network revenue is maximized(i.e., the total revenue generated from the allocation of resources in all network chunks). From the demand point of view, we take the allocated network resources throughout the network into account with the aim of keeping the level of network load, low. Hence, the trade-off between the maximization of operator's revenue and the minimization of the network load measured in terms of allocated resources is expressed in problem P2.

P2: $\underset{\sigma^*_{k,m}}{\arg\max}\ \frac{U^{\Gamma}}{\sum_{k=1}^{K}[\sum_{m=1}^{M}\sigma^*_{k,m}]^{\alpha_k}}$

$$\text{s.t. } \sum_{k=1}^{K}\sigma^*_{k,m} \leq \eta_m \tag{16a}$$

$$r_{k,m} \leq \sigma^*_{k,m} \leq R_{k,m},\ \forall k \in K,\ \forall m \in M \tag{16b}$$

In Equation (16), U^{Γ} is the total network revenue gained from network slice k, using our auction mechanism; $\sigma^*_{k,m}$ is the allocated resources to slice k from the network chunk m and

α_k is the normalized priority level for slice k. Constraint (16a) indicates that the amount of allocated network resources cannot exceed the maximum available resources for each network chunk, while constraint (16b) indicates that the received network resources of slice k should satisfy its own resource requirements.

We also use Jain's fairness index, denoted by ξ, to analyze distribution of network resources among slices. The Jain's fairness index is widely used in literatures to evaluate the level of fairness achieved by resource allocation algorithm [26], [27]. The Jains fairness index corresponds to all network resources, $T-\xi$, is given in Equation (18). The fairness index among slices in accessing resources of each network chunk, $C-\xi$, can be defined as in Equation (17).

$$C-\xi_m = \frac{(\sum_{k=1}^{K} \sigma_{k,m}^*)^2}{K(\sum_{k=1}^{K} \sigma_{k,m}^*)^2} \tag{17}$$

$$T-\xi = \frac{(\sum_{k=1}^{K} \sum_{m=1}^{M} \sigma_{k,m}^*)^2}{KM \sum_{k=1}^{K} \sum_{m=1}^{M} (\sigma_{k,m}^*)^2} \tag{18}$$

A. *Resource Slicing Mechanism*

1) Price Auction Mechanism: In this paper, we use Bertrand auctioning model for each set of network chunks in order to receive the selling price of network chunks per unit. The **players** in each step are sets of network chunks. The **strategy** of each player is the price per unit of network resources which are non-negative. The **solution** of this mechanism is based on Bertrand auction mechanism.

According to the network resources required by slice k, the network revenue can be expressed as in Equation (13). The price auction mechanism is a strategy where network chunks cannot increase their profit by choosing a different action without affecting others [5]. In this case, the optimized result is obtained by defining strategy for all participants from the overall view of the auction model. Selling price for each network chunk is computed according to the of Bertrand auction mechanism in Section IV-A.

2) Resource Auction Mechanism: The auction-based resource allocation mechanism aims to maximize the network revenue and has been described in Algorithm 1. In the proposed mechanism, the benefit of using our auction mechanism is that it can guarantee to all the participants their QoS requirements, i.e., the assignment of, at least, the minimum amount of requested resources for each chunk. Moreover, while satisfying allocated network resources $\sigma_{k,m}$ to all participants, network revenue can be improved by our proposed auction mechanism as well.

Algorithm 1: TWO-STEP AUCTION MECHANISM FOR RESOURCE ALLOCATION

```
$k$: number of required network slices;
$\eta_m^{rest}$: the rest resource of chunk $m$;

for $k := 1$ to $K$ do
    if $R_{k,m} < \eta_m^{rest}$ then
        Calculate the network revenue because of slice $k$ based on the Equation 10 where
          the value of received resource $\sigma_{k,m}$ equals to the value of their maximized
          requirements $R_{k,m}$;
        Then, calculate the network revenue value of the rest slices $K - 1$ according to the
          Equation (11) where the value of received resource $\sigma_{K/\{k\},m}$ equals to the value of
          their maximized requirements $R_{K/\{k\},m}$;
    end
    if $r_{k,m} < \eta_m^{rest} < R_{k,m}$ then
        Calculate the network revenue because of slice $k$ based on the Equation (10) ;
        Then, calculate the network revenue value of the rest slices $K - 1$ according to the
          Equation (11);
    end
    Calculate the revenue value of slice $k$ based on Equation (13).
end
```

3) Optimization: Since the optimization problem P2 needs to be solved only when network slices are allocated, real-time solution is not required. Therefore, we solve P2 using Lagrangian multipliers, as described as in Equation (19).

$$\begin{aligned} L(\sigma, \lambda, \mu, \phi) = & \frac{U^{\Gamma}}{\sum_{k=1}^{K}[\sum_{m=1}^{M}\sigma_{k,m}^{*}]^{\alpha_k}} \\ & + \sum_{m \in \mathbb{M}} \lambda_m (\eta_m - \sum_{k=1}^{K} \sigma_{k,m}^{*}) \\ & + \sum_{k \in \mathbb{K}} \mu_k (\sum_{m \in \mathbb{M}} R_{k,m} - \sum_{m \in \mathbb{M}} \sigma_{k,m}^{*}) \\ & + \sum_{k \in \mathbb{K}} \phi_k (\sum_{m=1}^{M} r_{k,m} - \sum_{m=1}^{M} \sigma_{k,m}^{*}). \end{aligned} \tag{19}$$

Based on Equation (19), one slice can receive the optimized resource value $\sigma_{k,m}^{*}$, while the following holds:

$$\begin{cases} \partial L_{\sigma_{k,m}}(\sigma_{k,m}^{*}, \lambda_m, \mu_k, \phi_k) = 0 \\ \partial L_{\lambda_m}(\sigma_{k,m}^{*}, \lambda_m, \mu_k, \phi_k) = 0 \\ \partial L_{\mu_k}(\sigma_{k,m}^{*}, \lambda_m, \mu_k, \phi_k) = 0 \\ \partial L_{\phi_k}(\sigma_{k,m}^{*}, \lambda_m, \mu_k, \phi_k) = 0 \end{cases} \tag{20}$$

The solution is summarized in Algorithm 2, which explains how allocated resources to slices are optimized according to their requests while network revenue is maximized simultaneously.

VI. PERFORMANCE INVESTIGATION

This section provides performance investigations of our proposed network slicing. In the simulations, we consider five different types of network resources including the radio network chunk, the edge network and the core network chunks. The radio network chunk $(m = 1)$ provides over-the-air resources, while the edge and the core network chunks provide storage and computational resources to network slices ($m = 2, 3, 4$, and 5.

We assume the amount of network resources in each network chunk is 100% and the resource required by network slices from each network chunk are a portion of that, i.e. for each chunk, the amount of allocated network resources to each slice is $\sum_{k=1}^{K} \sigma_{k,m}^{*} \leq 100\%$. We assume computational and storage resources in the core segment of the network are split between the edge and the central cloud such that 30% of these resources are allocated to the edge cloud while the remaining 70% are located at the central cloud of the core network. In other words, the radio chunk $(m = 1)$ contains all the over-the-air network resources, computational and

Algorithm 2: OPTIMIZED THREE-STEP AUCTION MECHANISM FOR RESOURCE ALLOCATION

```
if $R_{k,m} < \eta_m^{rest}$ then
    Calculate $\sigma_{k,m}$ based on Equation (14) and (20) separately, which have been denoted
      by $\sigma^*_{k,m}$ and $\sigma_{k,m}$, respectively;
    if $\sigma^*_{k,m} \geq \sigma_{k,m}$ then
        Calculate the network revenue value of slice $k$ and the network revenue of the rest
          of slices $K-1$ as Algorithm 1;
        else
            Calculate the network revenue of the slice $k$ and the network revenue of the
              rest of slice $K-1$ as Algorithm 1 with received value from Equation (20).
        end
    end
end
Calculate the revenue value of slice $k$ based on Equation (13).
```

storage capabilities for chunks $m = 2$ and $m = 3$ (i.e., the edge cloud) are the 30% of overall network capabilities, and chunks $m = 4$ and $m = 5$ (i.e., the central cloud) contain the 70% of the overall computational and storage capabilities, respectively, of the core network.

A. Simulation Scenarios

We consider three different strategies for network slicing, as described below.

1) priced-based network slicing (PB-NS): We use priced-based network slicing (PB-NS) scenario as our benchmark, and it is implemented by considering the minimization of allocated resources across the network.

2) Two-step auction mechanism for network slicing (TA-NS): The second scenario is based on our proposed two-step auction mechanism for network slicing (i.e. TA-NS), aimed at maximizing the network revenue. This is based on solving problem P2, and the solution is detailed in Algorithm 1, in section V-A2.

3) Three-step combination auction mechanism for network slicing (TC-NS): The third scenario is based on our proposed three step combination auction mechanism for network slicing (i.e.

TC-NS). In this scenario, we still focus on problem P2, but also additionally maximizing slices satisfaction, as detailed in Algorithm 2.

B. Results Analysis

Network Slices with the same Requests: In the first step, we assume all network slices have the same priority value ($\alpha_k = 5$) and request the same amount of resources (equals to $\frac{\frac{R_{k,m}}{\eta_m}-\frac{r_{k,m}}{\eta_m}}{2}$, where $\frac{R_{k,m}}{\eta_m} = 0.05$, $\frac{r_{k,m}}{\eta_m} = 0.01$) from each network chunk.

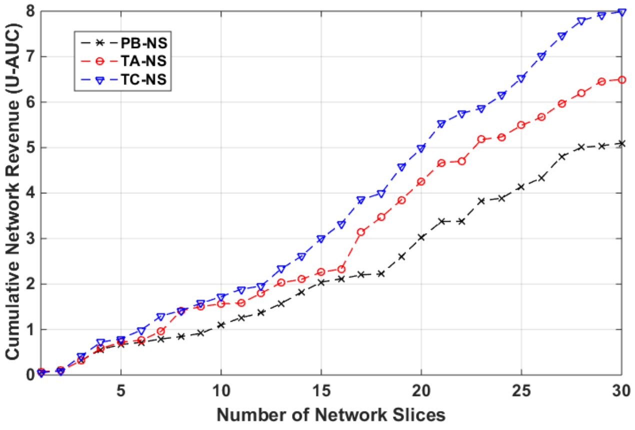
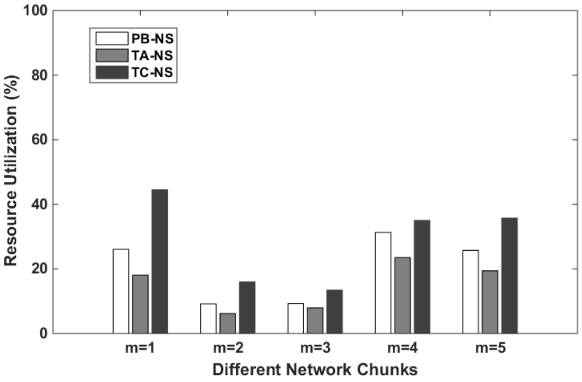


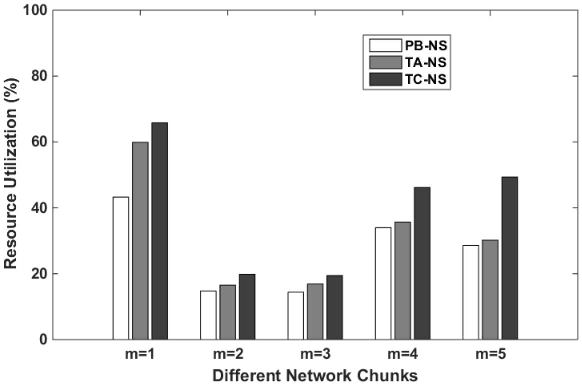
Fig. 4. Cumulative Network Revenue

Fig. 4 indicates the cumulative network revenue versus total number of network slices. It can be observed that the highest aggregated network revenue is obtained with the TC-NS strategy. This result is expected based on the definition of the TC-NS scenario.

Fig. 5 depicts resource utilization of different network chunks. It can be seen that the TC-NS scenario allocates the highest amount of network resources to the slices, i.e. has the highest resource utilization. This is mainly due to the fact that goal of TC-NS is maximizing allocated network resources to the slices. In the Focusing on TA-NS, the lowest resource utilization can be observed in the case with 10 slices, since the aim of TA-NS is mainly maximizing the network revenue. Hence, utilization in PB-NS is higher than the TA-NS, as shown in Fig. 5(a). However, in Fig. 5(b), slightly higher utilization can be observed in TA-NS comparing with PB-NS. This is mainly because TA-NS can support higher number of network slices more efficiently compared with that in PB-NS. In PB-NS, network resources have been allocated only based on price



(a) 10 Network slices



(b) 30 Network slices

Fig. 5. Analysis of Resource Utilization

auction mechanism without considering their capacity requirements. Further observation from the two plots of Fig. 5 shows that increasing number of slices, can also improve efficiency of resource allocation.

Fig. 6 shows the average satisfaction for network slices in different types of network chunks, when there are 30 network slices. The average is calculated as follows,

$$f_{k,m} = \frac{1}{k} \sum_{k=1}^{K} \frac{\sigma^{*}_{k,m} - r_{k,m}}{R_{k,m} - r_{k,m}} \tag{21}$$

From this figure, it can be seen that the TC-NS provides up to 50% higher slice satisfaction compared to TA-NS and PB-NS. The observation here confirms better performance of the strategy

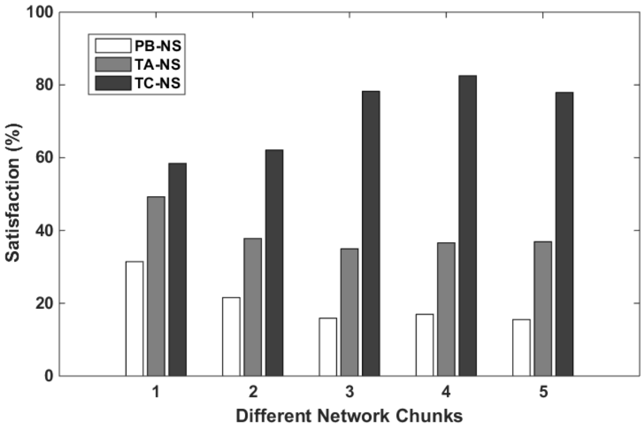


Fig. 6. Average level of satisfaction from each network chunks

TABLE I
JAIN'S FAIRNESS INDEX OF EACH NETWORK CHUNK

Fairness Index	PB-NS	TA-NS	TC-NS
$C - \xi_1$	0.5858	0.7172	0.8190
$C - \xi_2$	0.5068	0.6193	0.7683
$C - \xi_3$	0.4566	0.6189	0.7444
$C - \xi_4$	0.4092	0.7201	0.8047
$C - \xi_5$	0.3931	0.6109	0.7016
$T - \xi$	0.4911	0.6785	0.8543

in TC-NS that combines TA-NS revenue maximization strategy with optimal resource allocation. In PB-NS, network resources have been allocated to slices based on their own requirements without considering global optimization of resource allocations. Moreover, PB-NS does not consider effective resource allocation, hence, limited network resources will not satisfy the requirements of a portion of slices and this causes decrease in the overall satisfaction.

Table I lists the Jain's fairness index of allocated resources from each network chunk in different simulation scenarios based on the fairness, expressed in Equation (17). From this table, it can be noticed that the Jain's fairness index is the highest, for each chunk, in strategy TC-NS.

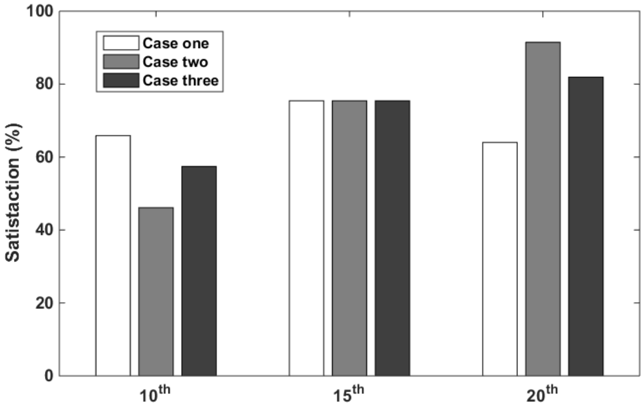


Fig. 7. Different Requests and Different Priority for network slices for the strategy TC-NS. In case two, the 10^{th} network slice requests 30% more network resources, while the 20^{th} network slice requests 30% less network resources. In case three, the 10^{th} network slice holds 30% higher priority value in case three, while the 20^{th} network slice holds 30% less priority value. The request/priority of the 15^{th} network slice in three cases are the same.)

It means that the each network chunk has fairly allocated its network resources according to resource requirements from each network slices. It is also worth noticing that the fairness of PB-NS decreases from 0.58 (radio chunk) down to 0.39 (central cloud chunk), while TA-NS and TC-NS decrease from 0.71 down to 0.61 and from 0.81 down to 0.7, respectively; there results underline how TA-NS and TC-NS can effectively provide better distribution of fairness among the available network chunks. The Jain's fairness index of the total allocated resource in different simulation strategies based on the fairness Equation (18), is also listed in the last row of Table I. It can be seen that fairness index has the highest value in the third scenario, i.e. TC-NS. This underlines that TC-NS is able to increase the network revenue while considering fairness in the allocation of network resources.

Network Slices with different Requests: In the next step, we focus our attention on the TC-NS strategy and we assume network slices can have different requests in terms of resources and priority. Three different cases are studied here:

1) All network slices have the same priority value and request the same amount of resources from the network chunks as in the previous simulation results;
2) All network slices have the same priority but have different resource requirements. In this case, 30% more network resources have been requested from slices $1^{st} - 14^{th}$, while 30%

less network resources have been requested from slices $16^{th} - 30^{th}$. The slice 15^{th} has been considered to have the same amount of network resource requirements as in case one.

3) All network slices have the same amount of resource requirements but with different priority. In this case, higher priority ($\alpha_k = 7$) is considered for slices $1^{st} - 14^{th}$, while lower priority ($\alpha_k = 3$) is considered for slices $16^{th} - 30^{th}$. The slice 15^{th} has the same priority as in case one.

Fig. 7 shows satisfaction values for network slices according to their different resource requirements. Observing from Fig. 7, the average satisfaction of slice 15^{th} with TC-NS is around 70% (the same in all three cases): this underlines that the performance of our proposed TC-NS strategy in satisfying one slice is not affected by changes in the requests or priorities of other slices. Network slice 10^{th} and 20^{th} are selected as representatives higher and lower resource requests and priorities, respectively. Compared with slice 15^{th}, it can be seen that, if the network slice with higher resource requirements (i.e., slice 10^{th} in case two) receives the same amount of network resource, it will not receive similar average slice satisfaction. Therefore, the network slice with higher resource requirements (i.e., 10^{th} in case two) will increase their payment in order to receive its required network resource and average satisfaction.

Further examination from this simulation shows network slices with higher priority value (e.g., 10^{th} in case three) will receive higher paying price from network chunks, computed based on Equation (9) (the revenue values are summarized in Table II). Therefore, the allocated network resources to the service slice k will be less, and it will generate lower satisfaction values, as observed in Fig. 7. From this point, the network slice with higher priority value will pay more in order to receive their required network resources and achieve their average satisfaction values.

VII. CONCLUSION

In this paper, we study economics of network slicing, how resources can be allocated to the slices competitively and how mobile operators can define revenue model for slicing their network resources for different vertical industries. Our approach is an auction-based optimal mechanism that takes into consideration both radio network resources and the storage and computational resource for performing resource allocation. Accordingly, our proposed mechanism can be used to maximize the total network revenue and optimize the resource allocation by considering the

TABLE II

REVENUES OF SELECTED SLICES OF THREE CASES IN TC-NS

	Case One	Case Two	Case Three
$10^{th} Slice$	0.20	0.35	0.25
$15^{th} Slice$	0.30	0.30	0.30
$20^{th} Slice$	0.27	0.13	0.21

limited network virtual resources. Extensive simulations show increase in total network revenue and extra efficiency in allocating network resources to the slices, as well as enhancement in satisfying slices requirements. We further study fairness in allocating resources to the network slices and show how fairness can be improved.

ACKNOWLEDGMENT

This work has been supported in part by the 5GPP VirtuWind (Virtual and programmable industrial network prototype deployed in operational wind park) Project.

REFERENCES

[1] "View on 5G Architecture." 5GPPP Architecture Working Group, White Paper, July 2016.

[2] H. J. Einsiedler, A. Gavras, P. Sellstedt, R. Aguiar, R. Trivisonno, and D. Lavaux, "System design for 5g converged networks," in *Networks and Communications (EuCNC), 2015 European Conference on*, pp. 391–396, June 2015.

[3] Y. Li, L. Liu, H. Li, J. Zhang, and Y. Yi, "Resource allocation for delay-sensitive traffic over lte-advanced relay networks," *IEEE Transactions on Wireless Communications*, vol. 14, pp. 4291–4303, Aug 2015.

[4] P. K. Agyapong, M. Iwamura, D. Staehle, W. Kiess, and A. Benjebbour, "Design considerations for a 5g network architecture," *IEEE Communications Magazine*, vol. 52, pp. 65–75, Nov 2014.

[5] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion," *Selected Areas in Communications, IEEE Journal on*, vol. 26, pp. 192–202, Jan 2008.

[6] S. Y. Jung, S. M. Yu, and S. L. KIM, "Optimization of spectrum allocation and subsidization in mobile communication services," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2015.

[7] J. Martyna, "Oligopoly bertrand model for price competition in cognitive radio networks," in *Communication Systems, Networks Digital Signal Processing (CSNDSP), 2014 9th International Symposium on*, pp. 227–231, July 2014.

[8] Y. Zhang, C. Lee, D. Niyato, and P. Wang, "Auction approaches for resource allocation in wireless systems: A survey," *IEEE Communications Surveys Tutorials*, vol. 15, pp. 1020–1041, Third 2013.

[9] P. Lin, X. Feng, and Q. Zhang, "Flexauc: Serving dynamic demands in spectrum trading markets with flexible auction," in *INFOCOM, 2014 Proceedings IEEE*, pp. 2265–2273, April 2014.

[10] J. F. Gomes, P. Ahokangas, and S. Moqaddamerad, "Business modeling options for distributed network functions virtualization: Operator perspective," in *European Wireless 2016; 22th European Wireless Conference*, pp. 1–6, May 2016.

[11] X. Zhou, R. Li, T. Chen, and H. Zhang, "Network slicing as a service: enabling enterprises' own software-defined cellular networks," *IEEE Communications Magazine*, vol. 54, pp. 146–153, July 2016.

[12] S. Abdelwahab, B. Hamdaoui, M. Guizani, and T. Znati, "Network function virtualization in 5g," *IEEE Communications Magazine*, vol. 54, pp. 84–91, April 2016.

[13] K. Zhu and E. Hossain, "Virtualization of 5g cellular networks as a hierarchical combinatorial auction," *CoRR*, vol. abs/1511.08256, 2015.

[14] M. Kamel, L. B. Le, and A. Girard, "Lte wireless network virtualization: Dynamic slicing via flexible scheduling," in *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*, pp. 1–5, Sept 2014.

[15] M. Kalil, A. Shami, and Y. Ye, "Wireless resources virtualization in lte systems," in *Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on*, pp. 363–368, April 2014.

[16] M. Jiang, M. Condoluci, and T. Mahmoodi, "Network slicing management amp; prioritization in 5g mobile systems," in *European Wireless 2016; 22th European Wireless Conference*, pp. 1–6, May 2016.

[17] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wang, and T. Q. S. Quek, "Resource allocation for cognitive small cell networks: A cooperative bargaining game theoretic approach," *IEEE Transactions on Wireless Communications*, vol. 14, pp. 3481–3493, June 2015.

[18] D. Niyato and E. Hossain, "Competitive spectrum sharing in cognitive radio networks: a dynamic game approach," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 2651–2660, July 2008.

[19] X. Kang, R. Zhang, and M. Motani, "Price-based resource allocation for spectrum-sharing femtocell networks: A stackelberg game approach," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, pp. 1–5, Dec 2011.

[20] S. Guruacharya, D. Niyato, D. I. Kim, and E. Hossain, "Hierarchical competition for downlink power allocation in ofdma femtocell networks," *IEEE Transactions on Wireless Communications*, vol. 12, pp. 1543–1553, April 2013.

[21] V. N. Ha and L. B. Le, "Distributed base station association and power control for heterogeneous cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 63, pp. 282–296, Jan 2014.

[22] J. Jia, Q. Zhang, Q. Zhang, and M. Liu, "Revenue generation for truthful spectrum auction in dynamic spectrum access," in *Proceedings of the Tenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, MobiHoc '09, (New York, NY, USA), pp. 3–12, ACM, 2009.

[23] Z. Suli, M. Zhongjing, and L. Xiangdong, "Efficient distributed resource allocation under synchronous auction-based algorithm," in *Control Conference (CCC), 2015 34th Chinese*, pp. 2720–2725, July 2015.

[24] F. Fu and U. C. Kozat, "Stochastic game for wireless network virtualization," *IEEE/ACM Transactions on Networking*, vol. 21, pp. 84–97, Feb 2013.

[25] P. Samadi, H. Mohsenian-Rad, R. Schober, and V. W. S. Wong, "Advanced demand side management for the future smart grid using mechanism design," *IEEE Transactions on Smart Grid*, vol. 3, pp. 1170–1180, Sept 2012.

[26] T. Mahmoodi, V. Friderikos, O. Holland, and H. Aghvami, "Cross-layer optimization to maximize fairness among tcp flows of different tcp flavors," in *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1–6, Nov 2008.

[27] T. Mahmoodi, V. Friderikos, O. Holland, and A. H. Aghvami, "Optimal design of forward error correction for fairness

maximisation among transmission control protocol flavours over wireless networks," *IET Communications*, vol. 4, pp. 1196–1206, July 2010.