Learning in SDN-Based Multi-Tenant Cellular Networks: A Game-Theoretic Perspective

Omer Narmanlioglu
P.I. Works,
Istanbul, Turkey 34912.
E-mail: omer.narmanlioglu@piworks.net

Engin Zeydan
Türk Telekom Labs,
Istanbul, Turkey 34770.
E-mail: engin.zeydan@turktelekom.com.tr

***Abstract*—In order to cope with the challenges of increasing user bandwidth demands as well as create new revenues by offering innovative services and applications, Mobile Network Operators (MNOs) are willing to increase their networks' capabilities by making it more flexible, programmable and agile. MNOs are also seeking new technologies to benefit from recent advances in cloud for rapid deployments and elastically scaling services that cloud providers are mostly benefiting today. On one hand, Software-Defined Networking (SDN) concept can be helpful for enabling network infrastructure sharing/slicing and elasticity for "softwarization" of network elements. On the other hand, machine learning and game-theoretical concepts can also be utilized to address network management and orchestration needs of services and applications and improve network infrastructure's operational needs. In that regard, joint utilization of machine learning, game theoretical approaches and SDN concepts for network slicing can be beneficial to MNOs as well as infrastructure providers. In this paper, we utilize regret-matching based learning approach for efficient Radio Remote Head (RRH) assignments among MNOs in software-defined based cloud radio access network (C-RAN). Using game-theoretical approach, we demonstrate convergence of RRH allocations to mixed strategy Nash equilibrium and present significant performance improvements compared to traditional assignment approach.**

***Keywords*—Software-Defined Networking; Game Theory, Machine Learning; C-RAN; Network Slicing.**

I. Introduction

Network slicing exploiting the principles of Software-Defined Networking (SDN) and Network Functions Virtualization (NFV) is an important concept which is expected to evolve as a fundamental feature for the next generation cellular network systems [1]. Network slicing can provide several benefits including dynamic multi-service support, multi-tenancy and integration with vertical players [2]. Therefore, Mobile Network Operators (MNOs) are willing to exploit network slicing for reducing capital expenditure (CapEx) and operating expenditure (OpEx), allowing programmable networks in order to offer enriched business services, and shared resources such as Radio Access Network (RAN) elements, spectrum or transport/core network equipment. At the same time, both creating and managing network slicing are challenging technological tasks where appropriate resource scheduling/assignments, creation and orchestration of new service instantiations, need to be solved.

Cloud RAN (C-RAN) is a promising technology that provides virtualization in the RAN aspect of the cellular networks. In C-RAN, digital signal processing functionality of Base Stations (BSs) is performed in cloud, e.g. in the baseband units (BBUs) pool. C-RAN's virtualized RAN also includes remote radio heads (RRHs) for the transmissions of radio signals to user equipments (UEs) depending on the baseband signals obtained from the cloud. In order to meet diverse demands of users and applications in the future, current RAN architectures need to evolve integrated with virtualization concept and SDN technologies. There exists different developments utilizing recent open source and standards for network slicing as well as cloud utilization in the telco domain [1], [3], [4]. In [1], the authors have provided an overview of network slicing concept from the 3GPP standardization point of view. A recent collaborative project between USA Tier-1 operator and ON Lab called central Office Rearchitected as a data center (CORD) project aims to transform legacy central offices in the telecommunication network by utilizing the elasticity and agility of cloud, SDN and NFV concepts [3].

Machine learning and game theory concept are also another two promising approaches for improving the operation and management of networks and engineering applications in the future cellular networks. These approaches will be more helpful to MNOs for improving the decision making process. Machine learning algorithms applied in the context of mobile networks can offer plentiful opportunities [5] to improve capacity planning [6], [7], enhance mobile user experience [8], generate new insight and foresight [9], and increase revenue [10]. A game theoretical aspect for network sharing that ensures resource sharing among multiple tenants is provided in [11] for mobile networks.

Network slicing utilizing scheduling algorithms in the context of SDN has also recently been studied in our recent paper [4] where the trade-offs between the resource allocations and quality-of-service (QoS) requirements of competing multiple MNOs operating in a shared mobile cellular network are investigated. However, it should be noted that applying machine learning and game theoretic computing paradigms towards optimization of network behaviour has not been adequately investigated from the perspective of SDN-based networks utilizing multi-tenancy in above papers. In this paper, we demonstrate the applicability of machine learning based computing approaches for usefulness of network and service management in SDN-based network sharing for multi-operators. One of the design goals of our approach is to assign multiple RRHs to different MNOs while learning from the previous assignment strategies in order to maximize received signal strength (RSS) levels of UEs in the network. This will not only benefit to UEs associated with different MNOs, but also to infrastructure providers as well as MNOs (or Mobile Virtual Network Operators (MVNOs)). In order to achieve this goal, we have employed a *regret*-matching based MNO selection algorithm for the problem of RRH assignment to multiple MNOs based on dynamic network behaviour of the infrastructure using a game-theoretic approach. Performance of proposed architecture is evaluated through Monte-Carlo simulation results where traditional homogeneous RRH assignment is considered as a benchmark. The simulation results reveal significant RSS level improvements with the use of *regret*-matching based MNO selection algorithm.

The rest of this paper is organized as follows: In Section II, we introduce system model of shared networks that is applicable for RRH assignments. In Section III, we propose a *regret*-matching based MNO selection mechanism. Section IV demonstrates the performance of the proposed mechanism and we finally conclude the paper in Section V.

II. System Model and Multi-Tenant Network Architecture

Fig. 1 illustrates an SDN-based C-RAN architecture where RAN slicing can be performed using C-RAN controller. In this SDN-enabled shared mobile architecture with K RRHs and M MNOs, let $\mathcal{M} = \{1, 2, \ldots, M\}$ denote the MNO set and $\mathcal{K} = \{1, 2, \ldots, K\}$ denote the RRH set. UEs associated with m-th MNO can be chosen from the set $\mathcal{N}_m = \{1, 2, \ldots, N_m\}$, thereby, total number of UEs in the given network architecture can be defined as $N = \sum_{m=1}^{M} \mathcal{N}_m$. A binary variable $q_{m,k}$ can be introduced to indicate whether RRH $k \in \mathcal{K}$ is assigned to MNO $m \in \mathcal{M}$ or not (i.e., if k-th RRH is assigned to m-th MNO then, $q_{m,k} = 1$ else $q_{m,k} = 0$). One of the main constraints is the fact that each RRH $k \in \mathcal{K}$ can be assigned to only one MNO during a certain time interval,

$$\sum_{m \in \mathcal{M}} q_{m,k} = 1. \tag{1}$$

Let $\mathbf{\Delta}_{K \times M} := [\mathbf{\Delta}_1 \;\; \mathbf{\Delta}_2 \;\; \ldots \;\; \mathbf{\Delta}_M] = (\mathbf{\Delta}_m, \mathbf{\Delta}_{-m})$ or alternatively $\mathbf{\Delta}_{K \times M} = [\mathbf{\Psi}_1 \;\; \mathbf{\Psi}_2 \;\; \ldots \;\; \mathbf{\Psi}_K]^T = (\mathbf{\Psi}_k, \mathbf{\Psi}_{-k})$ as the $K \times M$ RRH assignment matrix of all MNO. Here, $\mathbf{\Delta}_m = [q_{m,1} \;\; q_{m,2} \;\; \ldots \;\; q_{m,K}]^T$ is a $K \times 1$ is RRH assignment vector of m-th MNO and Δ_{-m} is the assignment vector of all MNOs other than the m-th MNO. Moreover, $\mathbf{\Psi}_k = [q_{1,k} \;\; q_{2,k} \;\; \ldots \;\; q_{M,k}]^T$ denotes k-th RRH's $M \times 1$ MNO assignment vector and $\mathbf{\Psi}_{-k}$ as the assignment vector of all RRHs other than the k-th RRH where $\mathbf{\Psi}_k \in \mathcal{I}_k$ and $\mathcal{I}_k$ denotes the set of all possible MNO assignments for k-th RRH and assume that $\mathcal{I} = \mathcal{I}_k = \{\mathbf{\Psi}_k^1, \mathbf{\Psi}_k^2, \ldots, \mathbf{\Psi}_k^M\}$ where each $\mathbf{\Psi}_k^m$ is $M \times 1$ orthogonal identity vector $\mathbf{I}_{M \times 1}$. For an RRH assignment profile $(\mathbf{\Psi}_k, \mathbf{\Psi}_{-k})$, denote the set of users of m-th MNO as $u \in \mathcal{N}_m$ choosing RRH $k \in \mathcal{K}$ as $\mathcal{C}_{k,m}$, i.e. $\mathcal{C}_{k,m} = \{u \in \mathcal{N}_m : q_{m,k} = 1, \; \forall k \in \mathcal{K}\}$, then the total number of users connected to k-th RRH can be expressed as $\Upsilon_{m,k} = |\mathcal{C}_{k,m}|$.

Other binary variable, $\vartheta_{k,i}^m$, can be introduced to indicate whether RRH $k \in \mathcal{K}$ is in the range of the user $i \in \mathcal{N}_m$ or not. It should be noted that UEs receive signals from multiple RRHs in a particular region, however, a finite number channel measurements can be reported due to capabilities of UEs, which is the second constraint. In respect to this, the maximum

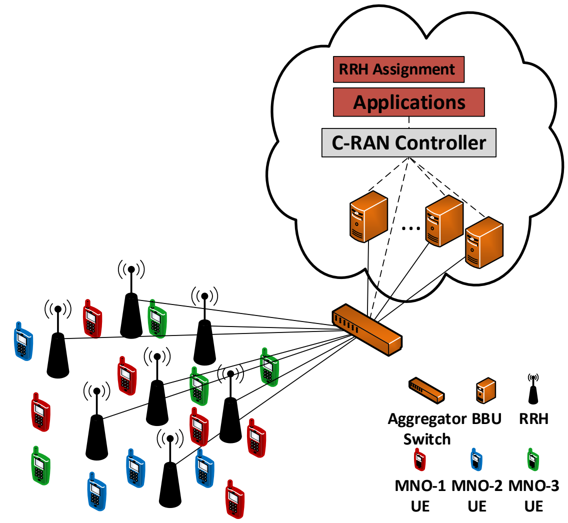

Fig. 1: SDN-based multi-tenant C-RAN architecture with three MNOs and associated UEs.

number of measured and estimated channels that are related to different RRHs are identified by an integer value of α. For this reason, each user can be connected to at most α different number of RRHs, i.e.,

$$\sum_{k \in \mathcal{K}} \vartheta_{k,i}^m \leq \alpha. \tag{2}$$

We define $\mathbf{\Theta}_{m_i} = [\vartheta_{1,i}^m, \vartheta_{2,i}^m, \ldots, \vartheta_{K,i}^m]^T$ as a $K \times 1$ vector associated with i-th user of m-th MNO and $\mathbf{\Omega}_{m_i} = [w_{1,i}^m, w_{2,i}^m, \ldots, w_{K,i}^m]^T$ as $K \times 1$ vector of the measured channel quality indicator (CQI) values from K different RRHs which is called as *channel measurement report* from i-th user of m-th MNO. Note that the values inside $\mathbf{\Omega}_{m_i}$ vector can have at most α non-zero values due to (2).

After introducing above parameters, the problem definition can be described as follows: Given a network state $\mathbf{S} = (\mathbf{\Psi}_k, \mathbf{\Psi}_{-k})$ where $(\mathbf{\Psi}_k, \mathbf{\Psi}_{-k})$ is a combination of each MNO assignments in the set of MNOs $\mathcal{M}$ to each RRH in the set $\mathcal{K}$, we look for the optimal values of assignments to minimize a cost function,

$$f(\mathbf{\Psi}_k, \mathbf{\Psi}_{-k}) = -\sum_{k \in \mathcal{K}} U_k', \tag{3}$$

where U_k is the utility of the k-th RRH. In order to accomplish this, each RRH's utility needs to be maximized by choosing appropriate MNO assignments. Using obtained CQI, average RSS of UEs is set as the maximization parameter. Then, the utility function of k-th RRH is expressed as

$$U_k' = \sum_{m \in \mathcal{M}} \sum_{i=1}^{N_m} \left(q_{m,k} \times \omega_{k,i}^m \right), \tag{4}$$

where the term $q_{m,k} \times \omega_{k,i}^m$ is the obtained CQI value of the UE $i \in \mathcal{N}_m$ that is attached to k-th RRH. Then, the optimization problem can be described as follows: Our goal is to maximize the sum of RSSs of all UEs (which also maximizes the values of CQIs) with the decision variables: (i) *Assignment problem:*

the assignment of RRHs to each MNO is represented by the variables $q_{m,k}$. (ii) *Connected users problem:* the successful assignment of all UEs of each MNO to various RRHs is specified by the multiplication of variables $\mathbf{\Theta}_{m_i}^T \mathbf{\Delta}_m$. For the shared mobile architecture, we use the following formulation for our optimization problem:

$$\underset{\mathbf{\Delta}}{\text{minimize}} \quad f(\mathbf{\Psi}_k, \mathbf{\Psi}_{-k}) \tag{5}$$

$$\text{subject to} \quad \mathbf{\Theta}_{m_i}^T \mathbf{\Delta}_m > 0, \quad \forall i \in \mathcal{N}_m, \forall m \in \mathcal{M}, \tag{5a}$$

$$0 < \Upsilon_{m,k} \leq N_m, \quad \forall k \in \mathcal{K}, \forall m \in \mathcal{M}, \tag{5b}$$

$$\sum_{k \in \mathcal{K}} q_{m,k} = 1, \ \ \forall m \in \mathcal{M}, \tag{5c}$$

$$0 < \sum_{k \in \mathcal{K}} \vartheta_{k,i}^m \leq \alpha, \ \ \forall m \in \mathcal{M}, \forall i \in \mathcal{N}_m, \tag{5d}$$

$$\{q_{m,k}, \vartheta_{k,i}^m\} \quad \in \{0,1\}, \ \ \forall m \in \mathcal{M}, \forall k \in \mathcal{K}, \forall i \in \mathcal{N}_m. \tag{5e}$$

In particular, the constraint (5a) tackles the case when there should not be any unconnected UEs in RRH assignments. The constraint in (5b) represents the fact that there should be nonzero number of UE connections to each RRH for all users of each MNO. The constraint in (5c) enforces each RRH be assigned to only one MNO, (5d) ensures each user be in the range of RRHs and (5e) denotes the binary decision variables of assignment and channel measurement reports.

Game Theoretic Interpretation: We consider the problem of (5) as a normal form game Π which can be mathematically defined as triplet $\Pi =< \mathcal{K}, \mathcal{W}, \{U_k'\}_{k=1}^K >$ where $\mathcal{K} = \{1, 2, \ldots, K\}$ is the finite set of players of the game, $\mathcal{W} = \Psi_1 \times \Psi_2 \times \ldots \Psi_K$ represents the set of all available actions for all the players and $\{U_k'\}_{k=1}^K : \mathcal{W} \to \mathbb{R}$ is the set of utility functions that the players associate with their strategies. The actions $\Psi_k \in \mathcal{I}_k$ for player k are the set of MNO selections $\Psi_k^m \in \mathcal{I}_k, \ \ \forall m \in \mathcal{M}$. Players select actions to maximize their utility functions. One of the questions that arise is if there exists a convergence point, a set of strategies, in our case a set of MNO selections $\Psi_k^m \in \mathcal{I}_k, \ \ \forall m \in \mathcal{M}$ from which no player would deviate. In game theory such a set of strategies is called a Nash Equilibrium (NE). A NE for a game is a set of strategy profiles $\mathbf{\Psi} = [\Psi_1, \Psi_2, \ldots, \Psi_K]$ from which no player can increase his utility by unilateral deviations. A strategy profile (Ψ_k, Ψ_{-k}) is a NE iff

$$U_k'((\Psi_k, \Psi_{-k})) \geq U_k'((\Psi_k', \Psi_{-k})), \ \ \forall k \in \mathcal{K}, \ \ \Psi_k, \Psi_k' \in \mathcal{I}_k \tag{6}$$

where (Ψ_k', Ψ_{-k}) refers to the strategy profile in which the action of UE k is changed from Ψ_k to Ψ_k' while the actions of all the other players in the game remain the same.

In order to tackle the above problem, the RRH assignment matrix $\mathbf{\Delta}$ needs to be optimized considering the constraints of (5a)-(5e). However, solving (5) problem is challenging due to coupling behaviour between the RRHs assignments and connected users problem. A naive approach for solving the problem (5) is to all RRH assignment vector profiles of each MNO exhaustively and pick the assignment profile with the maximum utility that gives successful assignments of all UEs of MNOs to RRHs as well. In order to compute (5), the centralized agent calculates the total CQI values for M^K possible RRH assignment vector combinations. For example, for a network topology with 160 RRHs, where infrastructure provider need to assign 3 MNO, the search space is 3^{160} assignment profiles. Therefore, finding the centralized MNO selections for all RRHs is cumbersome in large-scale wireless network. To alleviate the complexity problem, while maintaining good performance results, in the next section, we propose a *regret*-matching based algorithm, including capability of global and local view of the network, using centralized techniques aided with C-RAN controller.

III. Regret-matching-based MNO selection algorithm (*REGRET*)

In this section, our goal is to obtain a distributed learning algorithm for the joint connected user and RRH assignment problem of (5) of SDN-based shared mobile network architecture that requires only local information for updates. We derive a distributed mechanism of the problem of (5) named as *regret* which can work in C-RAN controller aided mode where each BBU connected to RRHs can solve the problem with low computation complexity and with limited effort. We will use the utility function defined in (4) for constructing a new utility function of (7) for non-cooperative RRHs with the C-RAN controller aided decision making:

$$U_k(\mathbf{\Psi}_k, \mathbf{\Psi}_{-k}) = \begin{cases} f(\mathbf{\Psi}_k, \mathbf{\Psi}_{-k}) & \text{if } \mathbf{\Theta}_{m_i}^T \mathbf{\Delta}_m > 0, \ \forall m \in \mathcal{M} \ \ \forall i \in \mathcal{N}_m, \\ -\infty & \text{if } \mathbf{\Theta}_{m_i}^T \mathbf{\Delta}_m = 0, \ \ \forall m \in \mathcal{M} \ \ \forall i \in \mathcal{N}_m. \end{cases} \tag{7}$$

Note that the interaction among K "selfish" RRHs can be defined as non-cooperative RSS maximization game where each RRH is attempting to find their own MNO selection vectors to maximize their corresponding total RSS. In the non-cooperative joint RRH assignment and connected user game, the K RRHs care only about their own RSS maximization exclusively, rather than accounting for the overall network RSS. Each player's utility function depends on the choice of the MNO selections, as well as on the other users' selections for MNO selections via the successful assignments of all UEs.

We study a C-RAN-aided learning algorithm called the *regret*-matching adaptive algorithm from [12], in which the players choose their actions based on their regret for not choosing particular actions in the past. The steady-state solution of the *regret*-matching-based learning algorithm exhibits "no regret" and the probability of choosing a strategy is proportional to the player's "regret" for not having chosen other strategies.

Let $\bar{\mathcal{I}}_k$ denote the vector of all strategies or actions for RRH $k \in \mathcal{K}$, i.e. $\bar{\mathcal{I}}_k = \{\mathbf{\Psi}_k^1, \mathbf{\Psi}_k^2, \ldots, \mathbf{\Psi}_k^M\}$ and $\mathbf{\Psi}_k(n)$ denote the MNO assignment vector selected by the k-th RRH at iteration n. Define the average regret vector $\mathbf{R}_k^{\bar{\mathcal{I}}_k}(k)$ of RRH k for an action vector $\bar{\mathcal{I}}_k$ at iteration (or time) i as

$$\mathbf{R}_k^{\bar{\mathcal{I}}_k}(i) = \frac{1}{i-1} \sum_{n=1}^{k-1} (U_k(\bar{\mathcal{I}}_k, \mathbf{\Psi}_k(-n)) - U_k(\mathbf{\Psi}_k(n))). \tag{8}$$

In the *regret*-matching-based MNO selection game algorithm (*regret*), each RRH k computes $\mathbf{R}_k^{\bar{\mathcal{I}}_k}$ for every action $\mathbf{\Psi}_k \in \mathcal{I}_k$ in all past steps when all other player's actions remain unchanged. Each player k updates its regret $\mathbf{R}_k^{\bar{\mathcal{I}}_k}(i)$ for every set of actions $\bar{\mathcal{I}}_k$ based on the following recursion formula:

$$\mathbf{R}_k^{\bar{\mathcal{I}}_k}(i+1) = \tfrac{i-1}{i} \mathbf{R}_k^{\bar{\mathcal{I}}_k}(i) + \tfrac{1}{i} (U_k(\bar{\mathcal{I}}_k, \mathbf{\Psi}_{-k}(n)) - U_k(\mathbf{\Psi}_k(n))). \tag{9}$$

At every step $i > 1$, each RRH k updates its own average regret vector $\mathbf{R}_k^{\bar{\mathcal{I}}_k}(i)$ for every strategy in $\bar{\mathcal{I}}_k$. In *regret* matching, after computing the average regret vector, $\mathbf{R}_k^{\bar{\mathcal{I}}_k}(i)$, each RRH k chooses an action or strategy $\mathbf{\Psi}_k(i)$, $i > 1$, according to probability distribution $\varrho_k^{\bar{\mathcal{I}}_k}(i)$ defined as,

$$\varrho_k^{\bar{\mathcal{I}}_k}(i+1) = \text{Prob}(\mathbf{\Psi}_k = \bar{\mathcal{I}}_k) = \frac{[\mathbf{R}_k^{\bar{\mathcal{I}}_k}(i)]^+}{\sum_{\bar{\mathcal{I}}_k \in \mathcal{I}_k} \mathbf{R}_k^{\bar{\mathcal{I}}_k}(i)]^+}, \quad (10)$$

where $[x]^+$ equals x when x is positive and zero otherwise. Notice that in the *regret*-matching game, each RRH k chooses a strategy $\mathbf{\Psi}_k \in \mathcal{I}_k$ at any step with a probability proportional to the average regret for not choosing that strategy $\mathbf{\Psi}_k \in \mathcal{I}_k$ in the past steps. The detailed summary of RMSG using Gauss-Seidel updating scheme [13] is given in Algorithm 1 where κ is the predefined number of iterations.

Algorithm 1 REGRET algorithm

1: **Initialization:** Set $\mathbf{R}_k^{\bar{\mathcal{I}}_k}(1) = 0, \ \forall k \in \mathcal{K}$ and select the initial MNO assignments with probabilities $1/M$
2: **for** iter $i = 1, 2, \ldots \kappa$ **do**
3:   **for** $k = 1, 2, \ldots K$ **do**
4:     Update $\mathbf{R}_k^{\bar{\mathcal{I}}_k}(i)$ using (9)
5:     Update probability distribution $\varrho_k^{\bar{\mathcal{I}}_k}(i)$ using (10)
6:     Select MNO selection vector $\mathbf{\Psi}_k(i)$ based on
7:     updated $\varrho_k^{\bar{\mathcal{I}}_k}(i)$.
8:   **end for**
9: **end for**
10: Return $\mathbf{\Delta}$

Every finite strategy game has a mixed strategy NE [14]. Therefore, using a good learning algorithm, any finite game can be shown to converge to a mixed strategy NE. Our proposed *regret*-matching-based selection method is distributed and requires limited information exchange between the RRHs if the utility function is properly selected. The time-averaged behavior of *regret*-matching game converges almost surely (with probability one) to the set of coarse-correlated equilibrium [15]. Therefore, the MNO selections of each RRH converge to a mixed strategy equilibrium solution. In fact, in our regret-based MNO selection game, the average regret of a RRH using regret matching becomes asymptotically zero, which is confirmed by our simulations. The utility function of non-cooperative or "selfish" RRHs for the MNO selection game at iteration n is given by (4). Note that by using this utility function, each RRH selects a MNO $\mathbf{\Psi}_k \in \mathcal{I}_k$ that maximizes its own RSS (or value of CQI). Moreover, the average regret in the recursion formula (9) is being updated locally as the best MNO is being selected.

IV. Performance Evaluation

In this section, we present the benefits of proposed RRH assignment mechanism on shared network region including 180 RRHs associated with 3 MNOs that is depicted in Fig. 2. This network region is generated on Matlab. The RRHs have omni-directional antennas and each MNO has 60 RRHs which are homogeneously distributed in the given region and very close to each other in order to serve the same coverage region. The distance between adjacent RRHs of each MNO is set to 5 km and the distance between adjacent MNOs associated with different MNOs is set to 0.4 km. The performance improvements by our proposed model are shown through Monte-Carlo simulations with the use of defined parameters in Table I under the consideration of 10 MHz system bandwidth and antenna diversity. Based on High Speed Downlink Shared Channel (HS-DSCH) power, RRH transmitter antenna gain and cable loss, the output power of RRH becomes 62 dBm. Additionally, based on UE noise figure, thermal noise (calculated by (Boltzmann constant $\times$ Temperature (290K) $\times$ Bandwidth)) and signal-to-interference-plus-noise ratio (SINR) [16], receiver sensitivity becomes -107 dBm. When the size of *channel measurement report* is set to $\alpha = 9$, each UE terminal forwards its report including the highest 9 channel measurement associated with RRHs whose associated RSS is higher than -107 dBm.

An optimization problem with feasible points exceeding 2^N when $N > 30$ is very difficult to find [17]. The centralized approach is no longer feasible in this scenario due to the enormous strategy space of 3^{60} profiles.

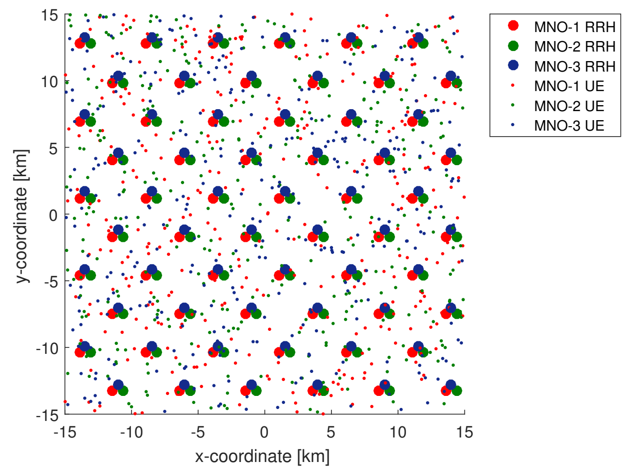

Fig. 2: Homogeneously distributed RRHs associated with different MNOs.

TABLE I: Downlink channel simulation parameters [16].

HS − DSCH power	46 dBm
RRH transmitter antenna gain	18 dBi
Cable loss	2 dB
UE noise figure	7 dB
Thermal noise	−104 dBm
SINR	−10 dB
Height of RRH antenna	80 m
Height of UE antenna	1.5 m

We consider urban environment Okumura − Hata path loss model [16] which can be written as

$$\begin{aligned}\text{Path Loss} =& 69.55 + 26.16 \log(f) - 13.82 \log(h_B) \\ & - C_H + (44.9 - 6.55 \log(h_B)) \log(d) \text{ dB},\end{aligned} \quad (11)$$

where d is the UE distance to RRH in km and C_H is antenna height correction factor and for small and medium-sized cities, it is calculated by

$$C_H = 0.8 + (1.1 \log(f) - 0.7) h_M - 1.56 \log(f), \quad (12)$$

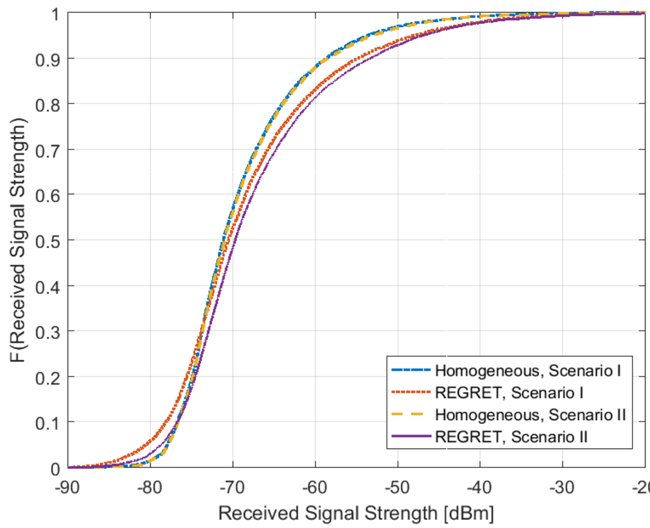

Fig. 3: Average RSSs of UEs with respect to homogeneous distribution and proposed method.

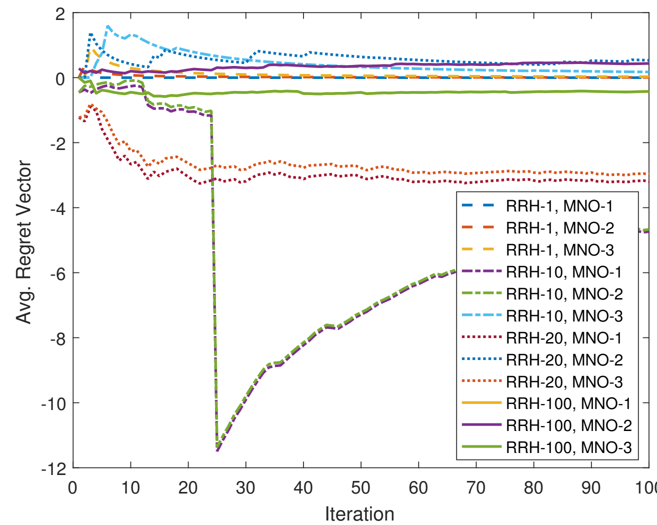

Fig. 4: Average values of regret vector with respect to different RRHs and MNOs.

where f is operating frequency of MNOs' RRHs and it is set to 900 MHz for red-colored RRH, 1800 MHz for green-colored RRHs and 2100 MHz for blue-colored RRHs in Fig. 2. It should be noted that this situation leads to different path loss values at the same distance which causes unfair RRH assignments in favor of RRHs with lower operating frequency. In order to avoid from this inconsistency, bias values under the consideration of operating frequencies need to be added into channel measurement reports of UEs associated with MNOs operating at higher frequencies. In order to have same path loss values for different operating frequencies at the same locations, bias values of 7.8479 dB for RRHs operating at 1800 MHz and 9.5932 dB for RRHs operating at 2100 MHz are used compared to RRHs operating at 900 MHz. We further assume that perfect channel state information (CSI) is available in receiver sides and used CSI instead of quantized CQIs values.

We compare the performance of *regret* with homogeneously assignment of RRHs as in Fig. 2. The used evaluation metric is RSS level which can be calculated by $\mathrm{HS-DSCH\ Power + Antenna\ Gain - Cable\ Loss - Path\ Loss}$ as a consequence of connection with those RRHs under the consideration of two scenarios. In first scenario, the average number of UEs associated with MNO-1, MNO-2 and MNO-3 is set to 100. On the other hand, the second scenario is more skewed and the average numbers of UEs associated with MNO-1, MNO-2 and MNO-3 are set to 10, 100 and 300, respectively. In both scenarios, UEs are randomly distributed.

Regret algorithm discussed in Section III maximizes the total values of CQIs (or RSSs in evaluations) in the network defined by (3) using the utility function (7). In the first scenario (see Fig. 3 where $\mathrm{F}(.)$ denotes cumulative density function (CDF) of given variable), the probability of RSS higher than -60 dBm is 0.12 with homogeneous RRH assignment whereas it can be observed that it is increased to 0.17 with the use of *regret*. When we turn to lower levels, the probability of RSS less than -80 dBm is 0.06 with proposed RRH assignment, however, this probability takes the value of 0.015 with homogeneous assignment. On the other hand, in the second scenario, the probability of RSS higher than -60 dBm is increased to 0.19. The benefits of our proposed algorithm is more clear in more skewed and imbalanced scenario. The results reveal that our proposed algorithm does not guarantee maximization of minimum RSS in the network, on the other hand, it provides higher probability values for higher signal levels. The expected value is also increased from -69.0737 dBm to -68.1894 dBm for scenario I and for the second scenario, this number becomes -67.1221 dBm with game theoretical approach.

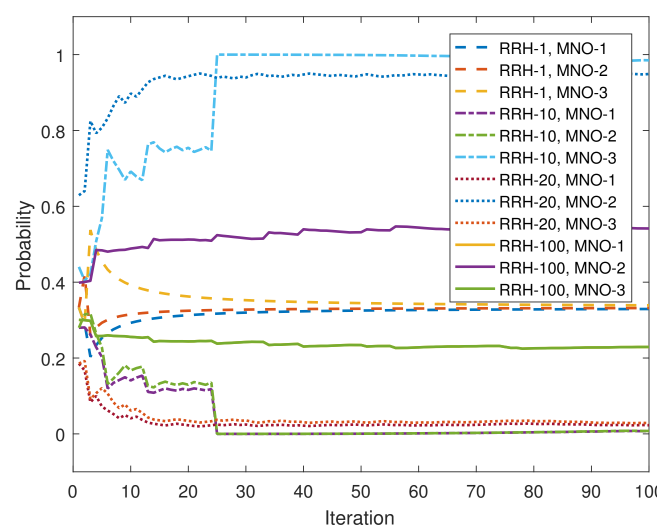

Fig. 5: Assignment probabilities of different RRHs to MNOs.

Fig. 4 depicts average values of regret vector defined in (8) with respect to different RRHs and MNOs using *regret* in the

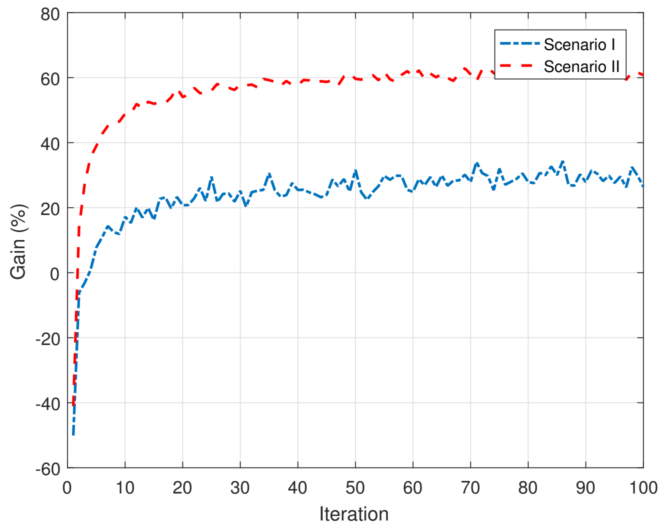

Fig. 6: Performance improvement of *regret* algorithm over homogeneous distribution.

network over 100 iterations. As a consequence, in Fig. 5, we take a look at the probability mass function (p.m.f), i.e. the assignment probabilities $\varrho_k^{\mathcal{I}_k}$ calculated in (10) of some of the RRHs to different MNOs for the *regret* algorithm. Fig. 5 represents the change in the p.m.f after 100 iterations for RRH-1, RRH-10, RRH-20, and RRH-100. The strategies are represented by the indices 1 to 3 in the legends, the iterations in x-axis and the probabilities of selecting these indices are given on the y -axis. At the initialization step, each RRH chooses initial MNO assignments or strategies with equal probability. Then, each RRH updates iteratively following *regret* algorithm, until the mixed strategy NE is achieved. For example, for RRH-10, the convergence occurs with probability of one into MNO-3, whereas for RRH-100, the convergence occurs with probability 0.55 into MNO-2 and with probability of ≈ 0.225 to MNO-1 and MNO-3. After 10 iterations, the probability of choosing MNO-3 for RRH-20 is higher than that for any other MNOs, although the other probabilities for MNO-1 and MNO-2 are not totally eliminated. After 30 iterations, all other probabilities are eliminated. A stationary point is reached when RRH-20 chooses MNO-3 in the 100-th iteration. Therefore, the existence of mixed strategy NE and the convergence toward mixed strategy NE in *regret* are illustrated by the curves in Fig. 5. Note that when probabilities of MNO selection converge in Fig. 5, the corresponding overall RSS values obtained by *regret* is shown in Fig. 3. Steady state is reached when all the users, i.e. RRHs select a MNO index with fixed probabilities.

We also investigate gain of *regret* learning algorithm over homogeneous distribution versus the number of iterations in Fig. 6 where it is calculated by $100\text{x}(\text{RSS}_{\text{Regret}}\ (\text{Watt}) - \text{RSS}_{\text{Homogeneous}}\ (\text{Watt}))/\text{RSS}_{\text{Homogeneous}}\ (\text{Watt})$. Fig. 6 shows that *regret* increases the performance with an amount of 30% for scenario I and 60% for scenario II with respect to homogeneous distribution value at the end of iterations.

V. Conclusion

In this paper, we proposed a *regret* based learning algorithm for RRH assignment in SDN-based multi-tenant network architecture using game-theoretical approaches. We consider two different scenarios with respect to distribution and total number of UEs associated with those MNOs. The performance of the proposed learning-based game theoretic mechanisms is evaluated in terms of obtained total RSS levels while considering traditional homogeneous RRH assignment as a benchmark. The results reveal the advantages of the proposed mechanisms over the traditional approach.

References

[1] K. Samdanis, X. Costa-Perez, and V. Sciancalepore, “From network sharing to multi-tenancy: The 5G network slice broker,” *IEEE Communications Magazine*, vol. 54, pp. 32–39, July 2016.

[2] 5G PPP Architecture Working Group, *View on 5G Architecture*, 7 2016. https://5g-ppp.eu/white-papers/.

[3] A. L. Peterson, Larry et. al.and Al-Shabibi, T. Anshutz, S. Baker, A. Bavier, S. Das, J. Hart, G. Palukar, and W. Snow, “Central office re-architected as a data center,” *IEEE Communications Magazine*, vol. 54, no. 10, pp. 96–101, 2016.

[4] O. Narmanlioglu and E. Zeydan, “Software-defined networking based network virtualization for mobile operators,” *Computers & Electrical Engineering*, vol. 57, pp. 134–146, January 2017.

[5] T. S. Buda, H. Assem, L. Xu, D. Raz, U. Margolin, E. Rosensweig, D. R. Lopez, M.-I. Corici, M. Smirnov, R. Mullins, *et al.*, “Can machine learning aid in delivering new use cases and scenarios in 5g?,” in *Network Operations and Management Symposium (NOMS), 2016 IEEE/IFIP*, pp. 1279–1284, IEEE, 2016.

[6] J. Pérez-Romero, J. Sánchez-González, O. Sallent, and R. Agustí, “On Learning and Exploiting Time Domain Traffic Patterns in Cellular Radio Access Networks,” in *Machine Learning and Data Mining in Pattern Recognition*, pp. 501–515, Springer, 2016.

[7] J. Moysen, L. Giupponi, and J. Mangues-Bafalluy, “A machine learning enabled network planning tool,” in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–7, Sept 2016.

[8] E. Zeydan, E. Bastug, M. Bennis, M. A. Kader, I. A. Karatepe, A. S. Er, and M. Debbah, “Big data caching for networking: moving from cloud to edge,” *IEEE Communications Magazine*, vol. 54, pp. 36–42, September 2016.

[9] S. Sozuer, E. Zeydan, and C. Etemoglu, “A new approach for clustering alarm sequences in mobile operators,” in *Network Operations and Management Symposium (NOMS), 2016 IEEE/IFIP*, pp. 1055–1060, IEEE, 2016.

[10] W. Verbeke, K. Dejaeger, D. Martens, J. Hur, and B. Baesens, “New insights into churn prediction in the telecommunication sector: A profit driven data mining approach,” *European Journal of Operational Research*, vol. 218, no. 1, pp. 211–229, 2012.

[11] P. Caballero, A. Banchs, G. de Veciana, and X. Costa-Perez, “Network Slicing Games: Enabling Customization in Multi-Tenant Mobile Networks,” *arXiv preprint arXiv:1612.08446*, 2016.

[12] S. Hart and A. Mas-Colell, “A simple adaptive procedure leading to correlated equilibrium,” *Econometrica*, vol. 68, no. 5, pp. 1127–1150, 2000.

[13] G. Scutari, D. P. Palomar, and S. Barbarossa, “Competitive design of multiuser mimo systems based on game theory: A unified view,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1089–1103, 2008.

[14] D. Fudenberg and J. Tirole, “Game theory MIT press,” *Cambridge, MA*, p. 86, 1991.

[15] H. P. Young, *Strategic learning and its limits*. OUP Oxford, 2004.

[16] H. Holma and A. Toskala, *WCDMA for UMTS: HSPA evolution and LTE*. John Wiley & Sons, 2010.

[17] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.