Cell Association with User Behavior Awareness in Heterogeneous Cellular Networks

Yao Sun, Gang Feng, *Senior Member, IEEE*, Shuang Qin, *Member, IEEE* and Sanshan Sun,

***Abstract*—In heterogeneous cellular networks (HetNets) with macro base station (MBS) and multiple small BSs (SBSs), cell association of user equipment (UE) affects UE transmission rate and network throughput. Conventional cell association rules are usually based on UE received signal-to-interference-and-noise-ratio (SINR) without being aware of other UE statistical characteristics, such as user movement and distribution. User behaviors can indeed be exploited for improving long term network performance. In this paper, we investigate UE cell association in HetNets by exploiting both individual and clustering user behaviors with aim to maximize long-term system throughput. We model the problem as a stochastic optimization problem, and prove that it is PSPACE-hard. For mathematical tractability, we solve the problem in two steps. In the first step, we investigate UE association for a specific SBS. We use restless multi-armed bandit model to derive association priority index for the SBS. In the second step, we develop an Index Enabled Association (IDEA) policy for making cell association decisions in general HetNets based on the indices derived in the first step. IDEA determines a set of admissible BSs for a UE based on SINR, and then associates the UE with the BS that has the smallest index in the set. We conduct simulation experiments to compare IDEA with other three cell association policies. Numerical results demonstrate the significant advantages of IDEA in typical scenarios.**

***Index Terms*—Cell Association, HetNets, Restless Multi-Armed Bandit, Throughput, User Behavior.**

I. INTRODUCTION

Data traffic demand in cellular networks has been growing at an exponential rate in recent years. To significantly increase network capacity in a cost-efficient way, a paradigm shift in network architecture from traditional single-tier homogeneous networks with high-power tower-mounted base stations towards multi-tier heterogeneous cellular networks (HetNets) is emerging [1]. HetNets are composed of traditional macro base stations (MBSs) overlaid with lower transmit power small base stations (SBSs) such as pico, femto and relay, which are usually deployed in hotspot areas to enhance network performance. In HetNets, a user equipment (UE) is allowed to be associated with either an MBS or an SBS due to overlapping coverage. Intuitively, cell association may to a great extend affect user transmission rate as well as overall system throughput performance, due to limited radio resources and interferences.

The cell association rule in traditional cellular networks, and up to LTE release-8, has been based on signal-to-interference-plus-noise-ratio (SINR) seen by the UE (called max-SINR), i.e., a UE associates itself with the BS that provides the strongest SINR [2]. The introduction of SBSs into traditional macro cells leads to a heterogeneous network architecture. When this max-SINR cell association policy applies, majority of UEs are associated with MBS because of the significant difference of transmit power between MBS and SBS. The UEs served by the heavily loaded MBSs may obtain low transmission rate due to limited radio resources. On the other hand, this cell association rule may also lead to waste of radio resources in SBSs and thus network throughput performance could be far from optimality. A known improvement of cell association rules for HetNets is Cell Range Expansion (CRE) based on SINR [3]. This approach tries to increase the coverage range of SBSs by adopting a bias to the received SINR from SBSs so that some UEs can be offloaded to SBSs from MBSs. However, CRE does not consider the load of small cells when offloading traffic, and thus some small cells may be overloaded. Furthermore, static bias cannot offset SBSs accurately for the scenario with dynamic UE distribution and movement. Considering network dynamics, UE mobility and multi-tier network architecture, cell association in HetNets becomes an essential yet challenging issue.

The cell association problem in HetNets is usually formulated as an integer programming with various objectives, such as load balance among BSs, maximization of system or individual UE throughput, fairness among UEs [4]–[9]. A number of mathematical tools, such as convex optimization, game theory, Markov chain, etc., can be used to solve the problem, and corresponding cell association policies can be designed. These policies could be optimal for a specific network snapshot where UEs locations, interferences, transmit power, etc. are fixed. However, long-term system performance of these polices may not be optimal as network dynamics, such as UE mobility and channel quality variation, are not taken into account. Indeed, user behaviors can be exploited to improve long-term system throughput performance. Moreover, in real mobile systems, some user behaviors could be predictable based on users location and schedule by some data analysis tools [10]–[13]. For example, the authors of [10] find a 93% potential predictability in user mobility across the whole user

This work was supported by the National Science Foundation of China under Grant number 61631005 and 61471089, and the Fundamental Research Funds for the Central Universities under Grant number ZYGX2015Z005.

Y. Sun, S. Qin and S. Sun are with the National Key Lab on Communications, University of Electronic Science and Technology of China, Chengdu 610051, China (e-mail: sunyao@std.uestc.edu.cn;blueqs@uestc.edu.cn;sanshansun@hotmail.com).

G. Feng is with the National Key Lab on Communications, University of Electronic Science and Technology of China, Chengdu 610051, China, and also with the Center for Cyber Security, University of Electronic Science and Technology of China, Chengdu 610051, China (e-mail: fenggang@uestc.edu.cn).

base. Besides the study on the prediction of user behaviors, some research work shows that user behaviors can be exploited to optimize cellular networks performance [14]–[16].

In HetNets, SBSs are usually deployed in hotspot areas, and user behaviors are quite different between hotspot and non-hotspot areas. Thus it is rather meaningful to exploit user behaviors in HetNets to make cell association decisions. In this paper, we investigate cell association of a two-tier HetNet with the objective to maximize the long-term system throughput by exploiting the characteristics of user behaviors. We consider both clustering and individual user behavior characteristics, and formulate the cell association problem as a stochastic optimization problem. We prove that the problem is PSPACE-hard, and solve it in two steps. First, we analyze the UE association for a specific SBS. We use restless multi-armed bandit (RMAB) model to derive the association priority index for the SBS, which corresponds to the long-term system throughput. Second, we develop an Index Enabled Association (IDEA) policy to make cell association decisions for UEs based on the indices derived in the first step. In IDEA, we determine a set of admissible BSs for a UE based on SINR, and then associate the UE with the BS that has the smallest index in the set. We conduct simulation experiments to compare IDEA with other three cell association policies. Numerical results demonstrate the advantages of IDEA in typical scenarios.

The remainder of the paper is organized as follows. We begin with an overview of related work in Section II. The system model and problem formulation are described in Section III. In Section IV, we study UE association rules for a specific SBS. According to the solutions, we propose a new cell association policy IDEA for HetNets in Section V. In Section VI we present simulation results and conclude the paper in Section VII.

II. Related Work

In recent years, many researchers have been addressing the cell association problem in HetNets. The authors of [17] present an extensive survey of the state-of-the-art in cell association algorithms. They systematically review the studies of user association in HetNets, massive MIMO networks, mmWave networks, and energy harvesting networks. Focusing on HetNets, most proposed cell association policies determine the access BS without considering user behaviors [4]–[9], while a little work takes it into account [18]–[20]. Indeed, only [20] explicitly exploits user behavior for the network selection in heterogeneous wireless networks. In the following we briefly describe these two categories of research work.

A. Conventional Cell Association Policies

The max-SINR policy used in in 3GPP release 8 associates a UE with the BS that provides the best SINR [7]. Similarly, most conventional policies usually make cell association decision for a specific network snapshot with aim to optimize the instantaneous network performance. One of the popular technical tools is game theory [4], [5], [8], [9]. The authors of [4] propose an auction-based algorithm to achieve load balance between MBS and femto BSs (FBSs). They first model the cell association problem as a graph matching problem, and use price to reflect the load of FBS. In the carefully designed auction mechanism they show that the optimal solution could be obtained to maximize the global utility. In [5], the authors formulate the cell association problem as a non-cooperative game, and then propose a distributed algorithm named RAT-game. They analyze the convergence and Pareto-efficiency of this algorithm. The authors of [8] and [9] also employ game theoretical approach to address network selection in heterogeneous wireless networks and femtocell networks association respectively. Besides, the authors of [21] propose an opportunistic user association for HetNets with two traffic types: human-to-human (H2H) and machine-to-machine (M2M). They formulate the user cell association as a bargaining problem, and exploit Nash Bargain Solution to obtain the association rule which can guarantee QoS of H2H traffic while providing fair resource allocation for M2M traffic. By using game approach we can only obtain a Nash equilibrium, but we do not know theoretically how close between the Nash equilibrium and the global optimality.

Optimization is another effective approach to address the cell association problem in HetNets [6], [7]. The authors of [6] formulate the cell association problem as an integer programming. By assuming that the users can be associated with multiple BSs simultaneously, the integer programming is transformed into a convex optimization problem. Then they use a logarithmic utility function as the optimization objective in order to achieve the fairness between different BSs and propose a distributed algorithm according to the dual problem. The authors of [7] study this problem with wireless resources and Quality of Service (QoS) constraints. They propose a distributed cell association algorithm with aim to minimize the global outage probability. Since the formulated problem is an integer or mixed integer programming, the optimal solution cannot be easily obtained. Moreover, this optimization approach usually solves the cell association problem for a static scenario where network settings and parameters are fixed, and the long-term network performance is not addressed.

Besides game theory and optimization, some other mathematical tools such as graph theory are also used for solving cell association problem. For example, in [22] and [23], the authors leverage graph theory to study the cell association with the aim of maximizing system throughput [22] and load balance [23] respectively.

B. User Behavior Aware Cell Association Policies

Recently, some researchers begin to discuss user behaviors in HetNets, and they try to exploit the characteristics of the user behaviors for energy efficient BS deployment [14], BS sleep mode design [15], multimedia broadcast [16], etc. to improve network performance. In addition, a few researchers begin to consider user behaviors in cell association [18]–[20]. The authors of [18] study handoff in HetNets by considering user behaviors such as UEs movement trajectory, position and speed. They use Markov Decision Process (MDP) to theoretically analyze the steady system state probability distribution, and based on that develop an optimal handoff policy. In [19], the authors propose a machine learning based cell association algorithm in HetNets in a predictable horizon. To the best of our knowledge, only the authors of [20] explicitly exploit the characteristics of individual user behaviors for network selection in heterogeneous wireless networks, consisting of LTE and Wi-Fi. They propose a two-layer game-theoretic framework to solve the problem by considering individual user behavior with aim to reduce unnecessary handoffs. They regard the social membership of individual user as a constraint in their game model. So each individual player selects network respecting to both his preference and social membership.

Moreover, some research work studies the cell association rules from a different perspective. The authors of [24] and [25] study the cell association for renewable energy powered HetNets with backhaul constraints [24] and QoS constraints [25] respectively. The authors of [26] study UE handoff schemes for LTE HetNets.

III. System Model and Problem Formulation

In this section, we first describe our network and user behavior models, and then formulate the cell association as a stochastic optimization problem.

A. Network Model

Consider a HetNet with M open access SBSs underlying an MBS. Let $\mathcal{M}$ be the set of BSs with the cardinality $|\mathcal{M}| = M+1$. We assume that a central controller is deployed in the MBS. The SBSs share the same frequency resources with the MBS [4]. An SBS can serve at most $\overline{N}$ UEs simultaneously due to resource constraint. We assume that each UE is allowed to be associated with only one BS at the same time [4]. Several hotspot areas are located inside the MBS coverage, as shown in Fig. 1.

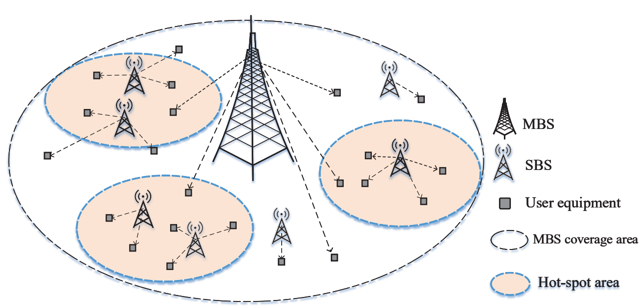


Fig. 1. Two-tier heterogeneous cellular network.

Let $\mathcal{N}$ be the set of UEs in the network. For a particular UE in $\mathcal{N}$, say UE n, we assume that the achievable transmission rate is determined by two factors: spectral efficiency and the fraction of wireless network resources allocated to the UE [4], [6]. The spectral efficiency for UE n is directly related to the received SINR of the serving BS j. In this paper we assume that the channel is flat and the transmit power of BS is allocated uniformly to each sub-channel. Hence the SINR of UE n associated with BS j at time t can be represented as

$$SINR_n^j(t) = \frac{P_{Tj} g_{jn}(t)}{\sum_{\substack{k\in\mathcal{M}\\ k\neq j}} P_{Tk} g_{kn}(t) + \sigma^2}, \ j \in \mathcal{M}, \quad (1)$$

where P_{Tj} denotes the transmit power of BS j; $g_{jn}(t)$ is the channel gain between UE n and BS j at time t; and σ^2 is the noise level. Thus the spectral efficiency for UE n at time t can be expressed as

$$r_n^j(t) = \log_2\left(1 + SINR_n^j(t)\right). \quad (2)$$

As each BS serves multiple UEs, the UEs associated with the same BS should share the bandwidth resources. For a given UE n, we use $c_n^j(t)$ to denote the bandwidth allocated by BS j at time t. Thus the achievable transmission rate for UE n associated with BS j at time t is $R_n^j(t) = c_n^j(t)\cdot r_n^j(t)$ [4], [6]. We assume that both the MBS and SBSs uniformly allocate the bandwidth resources to their served UEs. The achievable transmission rate can be expressed as

$$R_n^j(t) = \frac{B_j}{U_j}\cdot r_n^j(t), \quad (3)$$

where U_j is the number of UEs served by BS j, and B_j is the total bandwidth of BS j. As all the BSs use the same frequency, we use B to replace B_j.

B. User Behavior Model

In this paper, we consider both individual and clustering user behaviors in cell association. We focus on UE distribution characteristics and UE mobility pattern for clustering and individual user behaviors respectively.

Let us first discuss the clustering user behavior. Obviously the UE distribution density in hotspot areas is relatively higher than that in the other areas. Intuitively this uneven distributions can affect cell associations and thus system throughput. Similar to that in [15], [16], we use Gini coefficient to mathematically describe the degree of the distribution unbalance. Let $\rho(x)$ be a general UE distribution function shown in Fig.2, where x-axis and y-axis represent the cumulative fraction of BS and UE respectively. We use $h(x) = x$ to denote the complete homogeneous UE distribution curve. The Gini coefficient G based on $\rho(x)$ is defined as $G = \frac{A}{A+B}$, where A is the area between $h(x)$ and $\rho(x)$, and B is the area between $\rho(x)$ and coordinate axis. As $h(x)$ is deterministic, G is only related to $\rho(x)$, and it can be calculated by

$$G = \frac{A}{A+B} = 2A = 1 - 2B. \quad (4)$$

A larger $G (0 \leq G \leq 1)$ means that there are more UEs distributed in the hotspot area, while a smaller G means a more even UE distribution. Gini coefficient G is time-varying due to UE mobility, and we can periodically update it.

Next we discuss the individual user behavior which is reflected by UE mobility patterns. We use the mobility pattern which is very similar to a well-known model Straight-line Motion with Random Bouncing (sLRB) [27] where users move in random directions along a straight line with a constant speed, and once reaching the cell-edge, users bounce in random directions. In this mobility model, we assume that the speed is not a constant, and it changes at the decision times with a certain probability. The assumption on movement direction remain the same as that in sLRB model. Hence, both user movement speed and direction are implicitly taken into

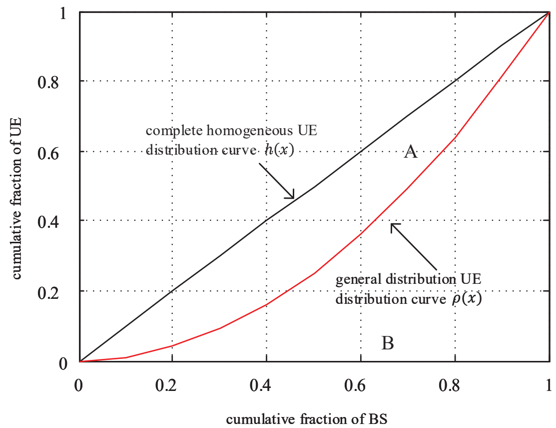


Fig. 2. User distribution curve in HetNets.

account in our mobility model. In detail, For a specific UE, say UE n, we assume that he moves to a specific direction with a random speed v_n. The movement speed changes at decision times with a certain probability, due to users schedule, users location, or even the transmission rate he is enjoying, etc.. UEs may have different speed transition probabilities according to their behavior characteristics. Similar to that in [28], we discrete UE speed into K levels. Let $\mathcal{K}_n = 1, 2, \cdots, K$ be the set of the speed levels of UE n and $\tau_n(t)$ be the speed level at time t. We assume that $\tau_n(t)$ can be changed at decision times, and then it will keep static until the next decision time. Thus it is a stochastic variable, and the transition probability, denoted by $\varphi^j_{w_n v_n}$, when associated with BS j can be expressed as

$$\varphi^j_{w_n v_n} = \Pr\{\tau_n(t+1) = v_n | \tau_n(t) = w_n, a_n(t) = j\}, \quad (5)$$

where $a_n(t)$ denotes the serving BS of UE n. The probability $\varphi^j_{w_n v_n}$ is related to the serving BS due to that the serving BS is related to the location and QoS of the UE which may affect the user mobility behavior. As it is not the key problem that we focus on, we use a simple statistical method to obtain the value of $\varphi^j_{w_n v_n}$. In more details, we can periodically collect and calculate the speed transition frequencies and use them to approximate the transition probabilities. Indeed the value of $\varphi^j_{w_n v_n}$ in current period is calculated from the data collected in the last period. $\varphi^j_{w_n v_n}$ can reflect characteristics of individual user behaviors. In general, we use two parameters to reflect the user behaviors in our system model, G is for clustering behaviors, and $\varphi^j_{w_n v_n}$ is for individual behaviors. Both G and $\varphi^j_{w_n v_n}$ are periodically updated.

After discussing the user behavior model, we analyze the relationship between the available transmission rate $R^j_n(t)$ and the user behaviors. First, $R^j_n(t)$ is directly affected by the user distribution and spectrum efficiency. Intuitively, high spectrum efficiency and spare user distribution could lead to a high transmission rate for a UE. Second, if the movement speed is relatively high for a UE, the channel quality and thus the available transmission rate may change quickly. On the other hand, if the user distribution varies greatly which means that the available bandwidth resource changes rapidly due to the resource sharing of the same BS, the transmission rate of the UE could also change greatly. Hence, the value of $R^j_n(t)$ is directly related to the user distribution and spectrum efficiency, and the variation trend is affected by UE movement and distribution.

C. Cell Association Problem

Based on user behaviors, we model the cell association problem as a stochastic optimization problem with aim to maximize long-term system throughput. We need to make cell associations for UEs at decision time t subject to resource constraints and user behaviors. We define binary variables $y^j_i(t) \in \{0,1\}$, $\forall (i,j) \in \mathcal{N} \times \mathcal{M}$ to indicate whether UE i is associated with BS j at decision time t. The problem can be formulated as

$$\mathbf{P1}: \ \max \ \sum_{t=1}^{T} \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{M}} y^j_i(t) R^j_i(t), \quad (6)$$

$$\text{s.t. } 0 \le \sum_{i\in\mathcal{N}} y^j_i(t) \le N_j, \forall j \in \mathcal{M}, \quad (6\text{-}1)$$

$$\sum_{j\in\mathcal{M}} y^j_i(t) = 1, \forall i \in \mathcal{N}, \quad (6\text{-}2)$$

$$y^j_i \in \{0,1\}, \forall (i,j) \in \mathcal{N} \times \mathcal{M}, \quad (6\text{-}3)$$

where N_j in (6-1) states the connection limitation of BS j. Considering the clustering user behavior parameter G, N_j can be expressed as

$$N_j = \begin{cases} \left\lfloor \left(\frac{G^2+1}{2}\right) \overline{N} \right\rfloor, & \textit{SBS } j \textit{ is in hotspot area} \\ \left\lceil \left(\frac{1-G^2}{2}\right) \overline{N} \right\rceil, & \textit{SBS } j \textit{ is in non-hotspot area} \\ |\mathcal{N}|, & \textit{the MBS} \end{cases} . \quad (7)$$

Hence, N_j is directly related to the user distribution. For example, when $G = 0$ which means a uniform UE distribution, the value of N_j is the same for all SBSs. When $G = 1$ which means all the UEs are distributed in the hot-spot area, N_j equals to the maximum connection capability $|\mathcal{N}|$ for the SBSs in hot-spot area, and 0 for the SBSs in non-hotspot area respectively. On the other hand, since the optimization objective in (6) is the long-term throughput in T, we cannot only consider a statistic network snapshot. In other words, we need to take user mobility into consideration to optimize long-term system throughput. Therefore, both the clustering and the individual user behavior are also taken into consideration in **P1**.

The objective of **P1** is to maximize the total system throughput in duration T. (6-1) ensures that the number of UEs associated with the same BS does not exceed the maximum BS connection limitation. (6-2) and (6-3) guarantee that each UE is associated with only one BS at a time.

Obviously it is difficult to solve stochastic optimization problem **P1** directly due to the following reasons. (1) $y^j_i(t)$ are binary variables. (2) Our optimization objective is a long-term metric: total system throughput in time T. The instantaneous optimal solution cannot guarantee the long-term performance. (3) Both user behavior parameters G and $\varphi^j_{w_n v_n}$ are time-varying. In the following, we show that even in the simplest scenario where the system consists of only one SBS and one MBS, the problem is already PSPACE-hard, which means that the time and memory consumption grow exponentially with problem scale. To make the problem mathematically tractable, we resort to a sub-optimal solution by solving the original problem in two steps. First, we derive the UE association priority index for a specific SBS. Second, we make cell association decisions in general HetNets based on the index obtained in the first step.

IV. UE Association for An SBS

As the first step in solving problem **P1**, we study UE association for a specific SBS (written as UAS problem for short) in the proposed system model. Based on stochastic optimization model RMAB, we model UAS problem as a *pseudo-RMAB* problem, where the constraint of actions is different from that in RMAB, and solve it by the primal-dual index heuristic algorithm. We first briefly describe RMAB problem.

A. Preliminaries: RMAB Problem

Multi-armed bandit (MAB) is one of decision theory especially in stochastic optimization problem. A MAB model consists of three elements: actions, states and rewards. The action chooses one of N arms at each decision time by the controller. Every arm has finite state space. We say that an arm is *active* when it is chosen at decision time t. The active arm i can generate an immediate reward $r_i(x)$ at time t in state $x(t)$, and then this arm will switch to another state y at the next decision time $t+1$ with probability P_{xy}. An arm is *passive* when it is not chosen at time t, and then no reward is gained and the state of this arm is frozen. Our aim is to maximize the expected reward or total discounted reward over a long-term horizon by finding the optimal policy that decides which arm to be chosen at each decision time.

There are several variants of the classical MAB model, one of them is RMAB proposed by Whittle [29]. In RMAB model, we allow m (m could be 1) active arms at each decision time, and the states of all arms can be changed (not only the active arms) at decision times. All the arms will generate a reward in two different ways: active and passive (the passive reward could be 0). The objective of RMAB is identical to that of the classical MAB model. It has already been proved that finding the optimal policy for RMAB is PSPACE-hard even in the special case of deterministic state transition situation and $m = 1$ [30]. In the following we formulate UAS problem as a *pseudo-RMAB* problem based on RMAB model.

B. UAS Problem Formulation

Consider a specific SBS, say SBS k, in the proposed system model, and let $\mathcal{N}$ be the set of UEs with cardinality N. Based on RMAB model, we formulate UAS problem respect to three aspects: actions, states and rewards.

First, we discuss the actions of our problem. We regard N UEs in the system as N arms in RMAB model. Let $a_n(t) \in \{0,1\}$ denote the action of UE n at time t, where $a_n(t) = 1$ means that UE n is associated with the SBS at time t otherwise the MBS. In other words, according to the RMAB model, $a_n(t) = 0$ means that arm n is passive at time t and active otherwise. The RMAB modeling requires that there are exactly m active arms at each decision time, which implies that in UAS problem, the SBS should always serve fixed number of UEs. This constraint is unrealistic, and thus we use the constraint that the number of UEs served by the SBS cannot exceed the connection limitation (defined in Section III.C):

$$0 \le \sum_{n=1}^{N} a_n(t) \le N_k,\ t = 1, 2\cdots \tag{8}$$

Due to this different constraint of actions, UAS problem cannot be modelled as an RMAB problem. Hence, we call it *pseudo-RMAB* problem.

Second, we discuss the state and the transition probability in UAS problem. For a specific UE n, the state of the UE at time t is determined by two factors: UE mobility speed and spectral efficiency of the SBS. The former is studied in Section III.B, and we here study the latter. Similarly, we discretize the spectral efficiency into L levels. Let $\mathcal{L}_n = 1, 2, \cdots, L$ be the set of the spectrum efficiency levels of UE n, and $l_n(t)$ be the level at time t. Thus $l_n(t)$ is a stochastic variable, and the transition probability $\psi^{a,w_n}_{g_n h_n}$ under a certain speed level w_n can be expressed as

$$\psi^{a,w_n}_{g_n h_n} = \Pr\left\{ l_n(t+1) = h_n \;\middle|\; \begin{array}{c} \tau_n(t) = w_n, l_n(t) = e_n, \\ a_n(t) = a \end{array} \right\}. \tag{9}$$

Note that speed level can affect the transition probability of spectral efficiency. We use the same method presented in Section III.B to obtain the value of $\psi^{a,w_n}_{e_n h_n}$.

With $\tau_n(t) \in \mathcal{K}_n$ and $l_n(t) \in \mathcal{L}_n$, we can define $\mathcal{S}_n = \mathcal{K}_n \times \mathcal{L}_n$ as the state space of UE n, where $\times$ is Cartesian product. We use vector $\boldsymbol{i_n} = (w_n, e_n) \in \mathcal{S}_n$ to denote the state of UE n which means that $\tau_n(t) = w_n$ and $l_n(t) = e_n$. We then derive the state transition probability of UE n.

***Proposition** 1: For UE n, the transition probability from state $\boldsymbol{i_n} = (w_n, e_n)$ to $\boldsymbol{j_n} = (v_n, h_n)$ under action $a_n(t) = a$ is $P^a_{\boldsymbol{i_n j_n}} = \psi^{a,w_n}_{e_n h_n} \cdot \varphi^a_{w_n v_n}$.*

Proof: According to the conditional probability formula we have

$$\begin{aligned} P^a_{\boldsymbol{i_n j_n}} &= \Pr\{S_n(t+1) = \boldsymbol{j_n} | S_n(t+1) = \boldsymbol{i_n}, a_n(t) = a\} \\ &= \Pr\left\{ \begin{array}{c} \tau_n(t+1) = v_n, \\ l_n(t+1) = h_n, \end{array} \middle| \begin{array}{c} \tau_n(t) = w_n, l_n(t) = e_n, \\ a_n(t) = a \end{array} \right\} \\ &= \Pr\left\{ l_n(t+1) = h_n \;\middle|\; \begin{array}{c} \tau_n(t) = w_n, l_n(t) = e_n, \\ a_n(t) = a, \tau_n(t+1) = v_n \end{array} \right\} \\ &\times \Pr\left\{ \tau_n(t+1) = v_n \;\middle|\; \begin{array}{c} \tau_n(t) = w_n, l_n(t) = e_n, \\ a_n(t) = a \end{array} \right\}. \end{aligned} \tag{10}$$

Since the transition probability of spectrum efficiency is independent with the speed of the next decision time, we have

$$\begin{aligned}&\Pr\left\{l_n(t+1)=h_n \,\middle|\, \begin{array}{c}\tau_n(t)=w_n, l_n(t)=e_n,\\ a_n(t)=a, \tau_n(t+1)=v_n\end{array}\right\}\\ =&\Pr\left\{l_n(t+1)=h_n\,|\tau_n(t)=w_n, l_n(t)=e_n, a_n(t)=a\right\}\\ =&\psi_{e_nh_n}^{a,w_n}.\end{aligned} \tag{11}$$

Moreover, the transition probability of speed is independent with the current spectrum efficiency, and we thus have

$$\begin{aligned}&\Pr\{\tau_n(t+1)=v_n|\tau_n(t)=w_n, a_n(t)=a, l_n(t)=e_n\}\\ =&\Pr\{\tau_n(t+1)=v_n|\tau_n(t)=w_n, a_n(t)=a\}\\ =&\varphi_{w_nv_n}^{a}.\end{aligned} \tag{12}$$

Combining (10) (11) and (12), we can express the state transition probability as

$$\begin{aligned}P_{\boldsymbol{i_n j_n}}^{a}&=\Pr\{S_n(t+1)=\boldsymbol{j_n}|S_n(t+1)=\boldsymbol{i_n}, a_n(t)=a\}\\ &=\psi_{e_nh_n}^{a,w_n}\cdot\varphi_{w_nv_n}^{a}.\end{aligned}$$

■

Third, we discuss the reward in UAS problem. Once a cell association decision is made, UEs will gain an immediate reward $R_n^{a_n(t)}(t)$ from the served BSs, which can reflect the immediate throughput in real communication system. Hence, we use the achievable transmission rate to define the reward. For a particular UE n, we can directly give the reward $R_n^{a_n(t)}(t)$ according to (3). Thus the system reward at time t is $\sum_{n=1}^{N} R_n^{a_n(t)}(t)$.

A policy Π can be written as a matrix, $\Pi = [\pi(1), \pi(2), \cdots, \pi(t)\cdots]$, where the N dimension vector $\pi(t)=[a_1(t), a_2(t), \cdots, a_N(t)]^T$ that satisfies (8), is the cell association policy at decision time t. The expected discounted reward under a specific policy Π is given by

$$E_\Pi\left[\sum_{t=1}^{\infty}\beta^t\left(\sum_{n=1}^{N}R_n^{a_n(t)}(t)\right)\right], \tag{13}$$

where β is a discount factor, and $0<\beta<1$. Since the total time T in problem **P1** is large enough compared with the decision period, we can transform the optimization objective from a sum of undiscounted reward over a finite horizon into a total discounted reward over an infinite time horizon (13). Moreover, as mentioned in [29], [31], the solutions of non-discounted problems could be obtained by letting the discount factor tend to one. Hence, we focus on the problem with discounted reward objective in the paper, similar to the way used in related work [29], [31].

Our objective is to find the optimal policy that maximizes the expected discounted reward provided that the initial state $S(0)$ is given. Note that the initial state $S(0)$ denotes the system state, and the system state space $\mathcal{S}$ is the product of the state spaces of all UEs, i.e. $\mathcal{S}=\mathcal{S}_1\times\mathcal{S}_2\times\cdots\times\mathcal{S}_N$, where $\times$ is Cartesian product. Thus, the problem can be expressed as

$$\mathbf{P2}:\ \max_{\Pi}\ E_\Pi\left[\sum_{t=1}^{\infty}\beta^t\left(\sum_{n=1}^{N}R_n^{a_n(t)}(t)\right)\middle|\, S(0)\right], \tag{14}$$

$$\text{s.t. } 0\le\sum_{n=1}^{N}a_n(t)\le N_k,\ t=1,2\cdots \tag{14-1}$$

$$a_n(t)\in\{0,1\},\ t=1,2\cdots \tag{14-2}$$

Similar to the illustration in Section III.C, we know that both individual and clustering user behaviors are considered in **P2**.

We now explain that how problem **P1** can be transformed into **P2** for studying the cell association between a specific SBS and the MBS. First, we introduce 0-1 variable $a_n(t)$ in **P2** to indicate whether the UE n is associated with the specific SBS. Hence, the variable $y_i^j(t)$ in **P1** is replaced by $a_n(t)$ in **P2**. Second, for the optimization objective, since there is only one SBS in **P2**, the notation $\sum_{j\in\mathcal{M}}$ in **P1** does not exist in **P2**. Moreover, we transform the optimization objective from a sum of undiscounted reward over a finite horizon into a total discounted reward over an infinite time horizon (13). Hence, the optimization objective becomes to (14). Third, the constraint (6-1) in **P1** is corresponding to (14-1) in **P2**, and (6-2), (6-3) in **P1** corresponds to (14-2) in **P2**.

C. Complexity Analysis of UAS Problem

In this subsection, we show that UAS problem **P2** is PSPACE-hard. PSPACE is a class of problems solvable in polynomial space [30]. In comparison, a decision problem is PSPACE-hard if any problem in PSPACE can be reduced to it in polynomial time [31]. PSPACE-hard problem is considered more intractable than NP-hard [30].

***Proposition** 2: UAS problem **P2** is PSPACE-hard.*

Proof: First, for a special case of UAS problem where the constraint of active arm is $\sum_{n=1}^{N}a_n(t)=0$, we just associate all the UEs with the MBS. Then we prove the PSPACE-hardness for a general UAS problem with the constraint $0<\sum_{n=1}^{N}a_n(t)\le N_k$.

Assume that UAS problem is not PSPACE-hand. In this case for all instances of UAS, we can find a polynomial-memory solution. Specifically, let $N_k=1$, for convenience, we denote this specific instance as UAS1. We denote the RMAB problem that the number of active arms m equals to 1 as RMAB1. The only difference between UAS1 and RMAB1 is the constraint of active arm: $0<\sum_{n=1}^{N}a_n(t)\le 1$ for UAS1, and $\sum_{n=1}^{N}a_n(t)=1$ for RMAB1.

First, all the instances in RMAB1 are included in UAS1. Second, as $a_n(t)\in\{0,1\}$, for the optimal solution of UAS1 is the same as RMAB1. Thus, RMAB1 can be reducible to UAS1. As we can solve UAS1 in polynomial memory, RMAB1 is not PSPACE-hard, which is contradictory to the proof in [30]. Thus, the assumption that UAS problem is not PSPACE-hand is not true. ■

Note that according to Proposition 2, we know our original problem **P1** is also PSPACE-hard.

D. Primal-dual Index Heuristic Algorithm

Borrowing the idea in solving RMAB problem in [32], we use the primal-dual index heuristic algorithm to solve the *pseudo-RMAB* problem **P2**. First, similar to Whittles idea [29], we use the total discounted number of active arms to relax (14-1) due to the PSPACE-hardness. It states that we restrict the total discounted number of arms do not exceed $N_k/(1-\beta)$ instead of the constraint of the exactly number at each decision time. Thus the constraint of active arms becomes

$$0 \leq \sum_{t=1}^{\infty} \beta^t \left(\sum_{n=1}^{N} a_n(t) \right) \leq \frac{N_k}{1-\beta}. \tag{15}$$

For solving the *pseudo-RMAB* problem we introduce new optimization variables $x^1_{\boldsymbol{j_n}}$ and $x^0_{\boldsymbol{j_n}}$, where $x^1_{\boldsymbol{j_n}}$ ($x^0_{\boldsymbol{j_n}}$) represents the expected total discounted time that UE n is in state j_n and active (passive). We have

$$x^1_{\boldsymbol{j_n}} = E\left[\sum_{t=1}^{\infty} I^1_{\boldsymbol{j_n}}(t) \beta^t \right], \tag{16}$$

where $I^1_{\boldsymbol{j_n}}(t)$ is an indicator defined as

$$I^1_{\boldsymbol{j_n}}(t) = \begin{cases} 1, & \textit{UE } n \textit{ is associated with SBS in state } \boldsymbol{j_n} \textit{ at time } t \\ 0, & \textit{otherwise} \end{cases}. \tag{17}$$

We use the total discounted time rather than the total time to define $x^1_{\boldsymbol{j_n}}$ and thus can ensure $x^1_{\boldsymbol{j_n}}$ to be finite.

Thus the constraint in (14-1) can be relaxed to

$$0 \leq \sum_{n=1}^{N} \sum_{\boldsymbol{j_n} \in \mathcal{S}_n} x^1_{\boldsymbol{j_n}} \leq \frac{N_k}{1-\beta}. \tag{18}$$

Besides this constraint of $x^1_{\boldsymbol{j_n}}$, for a particular UE n, there are also some constraints for the total discounted time of different states. We need to find the performance region of variables $x^0_{\boldsymbol{j_n}}$ and $x^1_{\boldsymbol{j_n}}$. For convenience, we introduce vector $\boldsymbol{x_n}$, whose elements are all the variables for UE n, and it can be written as

$$\boldsymbol{x_n} = \left(x^{a_n}_{\boldsymbol{j_n}} \right)_{\boldsymbol{j_n} \in \mathcal{S}_n, a_n \in \{0,1\}}. \tag{19}$$

We find that the performance region of $\boldsymbol{x_n}$ is Q_n according to Proposition 1 in [32],

$$Q_n = \left\{ \mathcal{R}_+^{||\mathcal{S}_n| \times \{0,1\}|} \middle| x^0_{\boldsymbol{j_n}} + x^1_{\boldsymbol{j_n}} = \alpha_{\boldsymbol{j_n}} + \beta \sum_{\boldsymbol{i_n} \in \mathcal{S}_n} \sum_{a_j} P^{a_j}_{\boldsymbol{i_n j_n}} x^{a_j}_{\boldsymbol{i_n}} \right\},$$

where $\alpha_{\boldsymbol{j_n}}$ is the probability that the initial state of UE n is $\boldsymbol{j_n}$. Note that $\alpha_{\boldsymbol{j_n}}$ is determined by $S(0)$, and it can only be 0 or 1. The performance region means that $\boldsymbol{x_n}$ is within the region of Q_n under any cell association policies. Hence, by introducing the performance region, we can define constraint (20-1) for $\boldsymbol{x_n}$. Therefore, problem (14) can be transformed into the following problem

$$\max \sum_{n=1}^{N} \sum_{\boldsymbol{j_n} \in \mathcal{S}_n} \sum_{a_n \in \{0,1\}} R^{a_n}_{\boldsymbol{j_n}} x^{a_n}_{\boldsymbol{j_n}}, \tag{20}$$

$$\text{s.t. } \boldsymbol{x_n} \in Q_n, \quad \forall n \in \mathcal{N}, \tag{20-1}$$

$$0 \leq \sum_{n=1}^{N} \sum_{\boldsymbol{j_n} \in \mathcal{S}_n} x^1_{\boldsymbol{j_n}} \leq \frac{N_k}{1-\beta}. \tag{20-2}$$

We next briefly describe the derivation process from (14) to (20). We first introduce new variables $x^{a_n}_{\boldsymbol{j_n}}$ defined in (16). From the definition of $x^{a_n}_{\boldsymbol{j_n}}$, the optimization objective (14) can be transformed into (20). Note that the given condition $S(0)$ in (14) is included as constraint (20-1). Then we find the performance region of $x^{a_n}_{\boldsymbol{j_n}}$ is Q_n when the initial state $S(0)$ is given. Hence, constraint (20-1) is obtained. Constraint (14-1) corresponds to (20-2) expressed by $x^{a_n}_{\boldsymbol{j_n}}$, and (14-2) does not exist for $x^{a_n}_{\boldsymbol{j_n}}$. With these derivations, problem (14) can be transformed into (20).

For each $\boldsymbol{x_n} \in Q_n$, there are exactly $|\mathcal{S}_n|$ constraints. For example, for UE 1 with state space $\mathcal{S}_1$, for each state $\boldsymbol{j_1} \in \mathcal{S}_1$, we have the constraint $x^0_{\boldsymbol{j_1}} + x^1_{\boldsymbol{j_1}} = \alpha_{\boldsymbol{j_1}} + \beta \sum_{\boldsymbol{i_n} \in \mathcal{S}_n} \sum_{a_j} P^{a_j}_{\boldsymbol{i_n j_1}} x^{a_j}_{\boldsymbol{i_n}}$. Since there are $|\mathcal{S}_1|$ states for UE 1, we can list all the $|\mathcal{S}_1|$ constraints in this way. Therefore, the total number of the constraints in problem (20) is $\sum_{n=1}^{N} |\mathcal{S}_n| + 1$.

By introducing a slack variable in constraint (20-2), we can obtain the equivalent problem in standard linear programming (LP) form for (20). For a standard LP, by using dual theory we can give the following dual program $\mathcal{D}$ of (20),

$$(\mathcal{D}) \ \min \ \sum_{j=1}^{N} \sum_{\boldsymbol{j_n} \in \mathcal{S}_n} \alpha_{\boldsymbol{j_n}} \lambda_{\boldsymbol{j_n}} + \frac{N_k}{1-\beta} \lambda, \tag{21}$$

$$\text{s.t. } \lambda_{\boldsymbol{j_n}} - \beta \sum_{\boldsymbol{i_n} \in \mathcal{S}_n} P^0_{\boldsymbol{j_n i_n}} \lambda_{\boldsymbol{i_n}} \geq R^0_{\boldsymbol{j_n}}, \ \forall \boldsymbol{j_n} \in \mathcal{S}_n, n \in \mathcal{N}, \tag{21-1}$$

$$\lambda_{\boldsymbol{j_n}} - \beta \sum_{\boldsymbol{i_n} \in \mathcal{S}_n} P^1_{\boldsymbol{j_n i_n}} \lambda_{\boldsymbol{i_n}} + \lambda \geq R^1_{\boldsymbol{j_n}}, \ \forall \boldsymbol{j_n} \in \mathcal{S}_n, n \in \mathcal{N}, \tag{21-2}$$

$$\lambda \geq 0. \tag{21-3}$$

Both problems (20) and (21) are LP which can be easily solved. Let $\{\overline{x}^{a_n}_{\boldsymbol{j_n}}\}$ and $\{\lambda^*_{\boldsymbol{j_n}}, \lambda^*\}$ be the optimal solutions to problem (20) and its corresponding dual problem $\mathcal{D}$ (21) respectively. We define $\overline{\gamma}^0_{\boldsymbol{j_n}}$ and $\overline{\gamma}^1_{\boldsymbol{j_n}}$ for UE n

$$\overline{\gamma}^0_{\boldsymbol{j_n}} = \lambda^*_{\boldsymbol{j_n}} - \beta \sum_{\boldsymbol{i_n} \in \mathcal{S}_n} P^0_{\boldsymbol{j_n i_n}} \lambda^*_{\boldsymbol{i_n}} - R^0_{\boldsymbol{j_n}}, \tag{22}$$

and

$$\overline{\gamma}^1_{\boldsymbol{j_n}} = \lambda^*_{\boldsymbol{j_n}} - \beta \sum_{\boldsymbol{i_n} \in \mathcal{S}_n} P^1_{\boldsymbol{j_n i_n}} \lambda^*_{\boldsymbol{i_n}} + \lambda^* - R^1_{\boldsymbol{j_n}}. \tag{23}$$

Then, for the SBS k, we define

$$\delta^k_{\boldsymbol{j_n}} = \overline{\gamma}^1_{\boldsymbol{j_n}} - \overline{\gamma}^0_{\boldsymbol{j_n}}. \tag{24}$$

as the index of UE n at state $\boldsymbol{j_n}$. Following proposition describes the meaning of the index in our problem.

***Proposition** 3: The index $\delta^k_{\boldsymbol{j_n}}$ states the decreasing rate of the total system throughput as a function of the elapsed time when UE n is associated with SBS k at state $\boldsymbol{j_n}$.*

Proof: According to the dual theory, we obtain that $\overline{\gamma}^1_{\boldsymbol{j_n}}$ and $\overline{\gamma}^0_{\boldsymbol{j_n}}$ are the optimal reduced costs with the following meaning: $\overline{\gamma}^1_{\boldsymbol{j_n}}$ ($\overline{\gamma}^0_{\boldsymbol{j_n}}$) is the decreasing rate of the optimization objective in problem (20) with the increasing of variable $x^1_{\boldsymbol{j_n}}$ ($x^0_{\boldsymbol{j_n}}$).

If UE n is associated with SBS k in state $\boldsymbol{j_n}$ at time t, on the one hand, $x^1_{\boldsymbol{j_n}}$ is increased by β^t according to

(16), and the optimization objective in (20) is decreased by $\beta^t \overline{\gamma}^1_{\boldsymbol{j_n}}$; on the other hand, $x^0_{\boldsymbol{j_n}}$ is decreased by β^t, and the optimization objective in (20) is increased by $\beta^t \overline{\gamma}^0_{\boldsymbol{j_n}}$. Hence, the optimization objective of problem (20) is totally decreased by $\beta^t \left(\overline{\gamma}^1_{\boldsymbol{j_n}} - \overline{\gamma}^0_{\boldsymbol{j_n}}\right)$ when UE n is associated with SBS k in state $\boldsymbol{j_n}$ at time t. As β and t are parameters which are not related to UE n, $\delta^k_{\boldsymbol{j_n}} = \overline{\gamma}^1_{\boldsymbol{j_n}} - \overline{\gamma}^0_{\boldsymbol{j_n}}$ is used to denote the decreasing rate. ■

The optimal solution to problem (21), $\{\lambda^*_{\boldsymbol{j_n}}, \lambda^*\}$, is affected by user behavior parameters $P^a_{\boldsymbol{i_n j_n}}$ and G. From (22), (23) and (24), we know that $\overline{\gamma}^0_{\boldsymbol{j_n}}$, $\overline{\gamma}^1_{\boldsymbol{j_n}}$ and thus the derived index $\delta^k_{\boldsymbol{j_n}}$ are directly related to $P^a_{\boldsymbol{i_n j_n}}$ and G. In other words, both the individual and clustering user behavior characteristics can affect the index computation, and thus the user cell association decisions.

In summary, through solving the two LP problems, we obtain the index defined in (24) of the SBS. According to the meaning of index, we know that for a specific UE, the smaller the index is, the more likely the UE is associated with the SBS. Note that the rationale behind the index computation is leveraging the relationship between user behavior and the associated BS to make cell association decisions. Based on the indices obtained, we will develop an Index Enabled Association (IDEA) policy for the general model of HetNets in the next section.

Note that the proposed algorithm can also work for obtaining the index for the non-discounted problem (mentioned in the paragraph after (13)) with the following optimization objective

$$\max_{\Pi} E_{\Pi} \left[\sum_{t=1}^{T} \sum_{n=1}^{N} R_n^{a_n(t)}(t) \middle| S(0) \right], \tag{25}$$

By letting $\beta = 1$, and using finite time instead of infinite time, the variable $x^1_{\boldsymbol{j_n}}$ defined in (16) and constraint (20-2) can be expressed as $x^1_{\boldsymbol{j_n}} = E\left[\sum_{t=1}^{T} I^1_{\boldsymbol{j_n}}(t)\right]$ and $0 \leq \sum_{n=1}^{N} \sum_{\boldsymbol{j_n} \in \mathcal{S}_n} x^1_{\boldsymbol{j_n}} \leq TN_k$ respectively. The other parts of the algorithm remain unchanged.

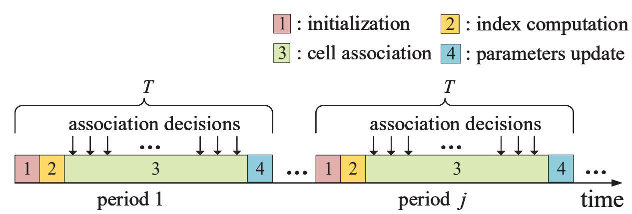
V. IDEA Cell Association Policy

In this section, we focus on the original problem **P1** to make cell association decisions for UEs in general HetNets. We first elaborate the procedure of IDEA policy, and then analyze the computational complexity and signaling overhead of IDEA policy.

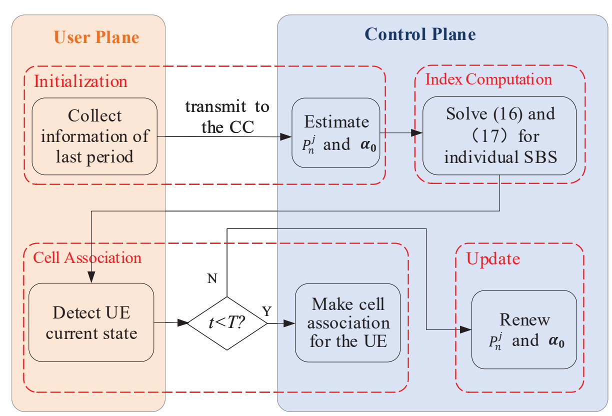
A. IDEA Policy

We now focus on the cell association for a general HetNet which contains multiple SBSs and one MBS. For each SBS, we solve a UAS problem to obtain an association index defined in (24). From Proposition 3, we know that the index represents the decreasing rate of the total system throughput as function of the elapsed time, which reflects the association priority between a specific SBS and the MBS. Note that the index is usually a real number instead of an integer. As the indices are derived for individual SBSs by compared with the same MBS, the association priority among SBSs are still valid. Thus, according to Proposition 3, we know that UEs are more likely to be associated with an SBS that has small index. Based on the indices, we develop the IDEA policy for general HetNets.

As we focus on long-term system throughput and user behavior parameters are time-varying, we need to update the parameters periodically in IDEA. We denote by T the length of parameter updating period. In each period, the timeframe of our proposed IDEA policy is mainly composed of four phases: initialization, index computation, cell association and user behavior parameters update, as shown in Fig.3 (a). Fig.3 (b) illustrates the four phases of IDEA policy in detail. IDEA policy involves both user plane and control plane operations. Initialization and cell association phases are performed in both planes, and the other two phases are performed only in control plane.



(a) Timeframe of IDEA policy



(b) IDEA policy in a specific period

Fig. 3. Illustration of IDEA policy

Referring to Fig.3 (b) and Algorithm 1, we describe the four phases of IDEA policy in detail. First, in the initialization phase, the central controller (CC) leverages the last period data to estimate user behavior parameters $G(t)$, $\mathbb{P}^j_n$, and system initial state $\boldsymbol{\alpha_0}$.

In index computation phase (lines 2-6 in Algorithm 1), the CC computes the indices for individual SBS and thus obtains a mapping function Φ_k between UE states and the indices, where $\Phi_k(\boldsymbol{j_n}) = \delta^k_{\boldsymbol{j_n}}$ means that the index of SBS k is $\delta^k_{\boldsymbol{j_n}}$ when UE n is associated with SBS k in state $\boldsymbol{j_n}$. From Algorithm 1 lines 2-6, we know that we use the same mapping functions in a specific period. Thus, the index computation is only executed once in a period.

In the cell association phase, once a specific UE n needs to make a handoff decision (line 7 in Algorithm 1), he should detect and notify the current state to the CC. The CC first selects the set of admissible SBSs, which is denoted by $\mathcal{A}_n$ and can be expressed by $\mathcal{A}_n = \{k|SINR_n^k \geq \gamma\}$, where γ is an SINR threshold parameter. The CC then maps the states to indices according to the mapping function Φ_k. Finally, the UE is associated with the MBS if $\mathcal{A}_n = \varnothing$, or the SBS $k^* = \arg\min_{k \in \mathcal{A}_n} \delta_{\boldsymbol{k_n}}^k$ otherwise. We repeat this phase for making all UE cell association decisions until the end of each period (line 2 of Algorithm 1).

In the updating phase, UEs send their own last period state transition frequency information as well as the location to the CC which in turn uses this information to update $\mathbb{P}_n^j$ and G as shown in Section III.B. Once the parameters are updated, the index mapping functions need to be re-computed (lines 4-5 of Algorithm 1).

For new arrival UEs, due to the lack of behavior information, we use the traditional maximum SINR rule to make cell association decisions in the first period, and then go into the updating phase to evaluate the user behavior parameters at the end of first period. After that, we use IDEA policy for these new arrival UEs in the following periods.

Algorithm 1 : IDEA policy

Input: Network topology; UE locations, movement direction and speed.

Output: UE association decisions.

Initialization:
1: Associate UE with BS provided the maximum SINR

Updating and Index Computation:
2: **if** $t = nT$ **then**
3: Input: the information of last period.
4: Calculate user behavior parameters G and $\mathbb{P}_n^j$.
5: Output: new mapping functions Φ_k for SBSs.
6: **end if**

UE Association:
7: **while** $SINR_n^k < \gamma$ **do**
8: **for** $k \in \mathcal{A}_n$ **do**
9: detect the state of UE n for SBS k
10: obtain index $\delta_{\boldsymbol{j_n}}^k$ from mapping function Φ_k.
11: **end for**
12: choose the target BS $k^* = \arg\min_{k \in \mathcal{A}_n} \delta_{\boldsymbol{k_n}}^k$.
13: **end while**
14: **go into** Index Computation and Update phase

B. Computational Complexity and Signaling Overhead

In IDEA policy, the major computational complexity lies in the index computation phase, in which we need to solve two LP problems with $\mathcal{O}(n^4)$ computational complexity in the worst case [33], where n is the number of variables. Thus, the computational complexity for a UAS problem is $\mathcal{O}\left((2|\mathcal{S}|)^4 + |\mathcal{S}|^4\right)$, i.e. $\mathcal{O}\left(|\mathcal{S}|^4\right)$, and thus $\mathcal{O}\left(M|\mathcal{S}|^4\right)$ for **P1** problem. Note that the indices are computed offline before cell association phase, and it is only computed once in a period. Thus, this computational complexity does not affect the real-time cell association decisions. In cell association phase, the indices of admissible BSs are sorted and thus the computational complexity of this phase is $\mathcal{O}\left(|\mathcal{A}_n|\right)$. Note that this computational complexity is similar to that of the traditional maximum SINR cell association policy, which also needs to sort SINR.

We next discuss the signaling overhead in IDEA policy. The UE that needs to make cell association decision notifies the states (both speed and the channel conditions from the UE to each SBS) to the CC, and the number of signaling exchanges needed is $M + 1$. Then the CC selects the set of admissible BSs $\mathcal{A}_n$ for the UE, and sends the corresponding indices to the UE with $|\mathcal{A}_n|$ signaling overhead. Finally, the UE is associated with the SBS with the smallest index. Thus, the total number of signaling exchanges needed is $M + 1 + |\mathcal{A}_n|$ for a UE to make a cell association decision. In comparison, for the traditional maximum SINR cell association policy, the number of signaling exchanges is $|\mathcal{A}_n|$. Thus, the cost of IDEA policy is the extra $M + 1$ signaling exchanges.

C. Impact of Individual and Clustering User Behaviors on Cell Association

In this subsection, we analyze the respective impact of user behavior parameters on system performance. First, we analyze the impact of the clustering user behavior parameter G when the individual user behavior parameter $\varphi_{w_n v_n}^j$ is fixed. From (7), we can see that G directly affects BS connection limitation parameter N_j. For a larger G, N_j is larger for hot-spot SBSs, and smaller for non-hotspot SBSs. Then, from Problem (20), we can see that the upper bound of feasible region is higher (lower) with larger (smaller) N_j. Since (20) is an LP, the optimal solution is obtained on the boundary of the feasible region. Therefore, a larger (smaller) value of $\sum_{n=1}^{N} \sum_{\boldsymbol{j_n} \in \mathcal{S}_n} {x^*}_{\boldsymbol{j_n}}^1$ can be obtained for those SBSs in hotspot (non-hotspot) area by solving Problem (20), where ${x^*}_{\boldsymbol{j_n}}^1$ is the optimal solution of (20). From the definition of $x_{\boldsymbol{j_n}}^1$ (16), a larger (smaller) $x_{\boldsymbol{j_n}}^1$ means that UE n is associated with SBS j for a longer (shorter) time. Hence, when $\varphi_{w_n v_n}^j$ is fixed, for a larger G, users are more likely to be associated with the SBSs which are in the hot-spot area.

Next, we discuss the impact of $\varphi_{w_n v_n}^j$. We can see that $\varphi_{w_n v_n}^j$ directly affects transition probability $P_{\boldsymbol{i_n j_n}}^a$, and thus the solution of (20). Therefore, the cell association rule is related to this parameter. However, it is hard to theoretically analyze how the parameter affect the system throughput. We will thus conduct some simulation experiments in Section VI to explore the relationship between UE movement speed and the proposed cell association rules.

VI. Numerical Results and Discussion

In this section, we compare IDEA with the other three typical cell association policies: conventional max-SINR, RAT-game [5] and learning-based SAMSRL [19] by using simulations. The max-SINR policy is indeed the conventional cell association policy which always associates UEs with the BS that can provide the maximum SINR. RAT-game is based on a non-cooperative game which uses the throughput as each UEs preference. Thus UE always selects the BS to increase their own individual throughput in this policy. SAMSRL policy is based on a reinforcement learning framework with the objective to maximize long-term system throughput. With the leaning mechanism, SAMSRL implicitly takes user behavior into account by updating reward, and thus optimizes the long-term performance.

A. Simulation Settings

We consider a two-tier HetNet which consists of an MBS, some SBSs and UEs. Within the coverage of MBS, there are three hotspot areas. The radius of the MBS, SBSs and the underlying hotspot area is 500, 50 and 150 meters, respectively. The MBS is located at the central of the network. The ratio of SBSs located in hotspot area and non-hotspot area is 0.6 and 0.4 respectively. An SBS is able to admit at most 6 UEs simultaneously. UEs are randomly distributed in the area with randomly generated user behavior parameter $\varphi^j_{w_n v_n}$. The transmit power of the MBS and SBSs is 46 and 20 dBm respectively. We assume that UE can achieve the Shannon capacity, i.e. the spectrum efficiency of UE n served by BS j is $r^j_n(t) = \log_2(1 + SINR^j_n)$ with the unit bits per second per Hertz. The bandwidth allocated to all the BSs is assumed to be 10MHz. We use $L(d) = 34 + 40\log(d)$ and $L(d) = 37 + 30\log(d)$ to model the path loss from a UE to the MBS and SBSs respectively. The noise power is set to -104dBm for all receivers based on the room temperature and bandwidth. The system parameters used in simulations are listed in Table II, which are similar to those in related work [6], [7]. Our numerical computations are implemented with MATLAB codes and carried out on a PC equipped with an Intel-i5 4 core 3.2GHz processor and 4G RAM.

TABLE I
SIMULATION PARAMETERS

Parameters	Value
Macrocell radius	500 m
Small cell radius	50 m
Hotspot area radius	150 m
Connection capability of SBS	6
Power of MBS	46 dBm
Power of SBS	20 dBm
SINR threshold	-10 dB
Bandwidth	10 MHz
Path loss model for MBS	$L(d) = 34+40\log(d)$
Path loss model for SBS	$L(d) = 37+30\log(d)$
Noise power	-104 dBm

B. Numerical results and discussions

In Experiment 1, we compare the total system throughput in 1000 seconds for the three cell association policies. We fix the number of SBSs and UEs to 30 and 100 respectively. The average UE movement speed is 5m/s. We conduct 100 independent experiments with random distributions of UEs and SBSs for computing the average system throughput. Fig.4 shows the total system throughput for the four cell association policies in 1000 seconds with different length of update period T. As max-SINR, RAT-game and SAMSRL do not take consideration of user behavior, T does not affect the performance of the three policies. From Fig.4 we can see that the system throughput of IDEA policy is approximately equal to that of max-SINR policy when $T = 10s$. This is because that the estimations of user behavior parameters are inaccurate due to the lack of history information in the short period time T. These inaccurate parameters degrade the system throughput of IDEA. We also find that the system throughput of IDEA policy increases with the length of period T, and when T exceeds 100s, we can see that the system throughput of IDEA is significantly higher than that of the other two policies. In more details, when $T = 200s$, the system throughput of max-SINR, SAMSRL, RAT-game and IDEA is 7.48×10^3, 7.95×10^3, 8.06×10^3 and 8.59×10^3, respectively. These numbers show that IDEA can improve system throughput by 15%, 8% and 7% when compared with max-SINR and RAT-game, respectively. Note that the system throughput is the sum of 1000 rate (uniformed in bit/Hertz) samples in 1000 seconds (so call long-term throughput as the optimization objective). hence, the achievable rate per user is about 0.08 bps/Hz, which is reasonable for cellular networks.

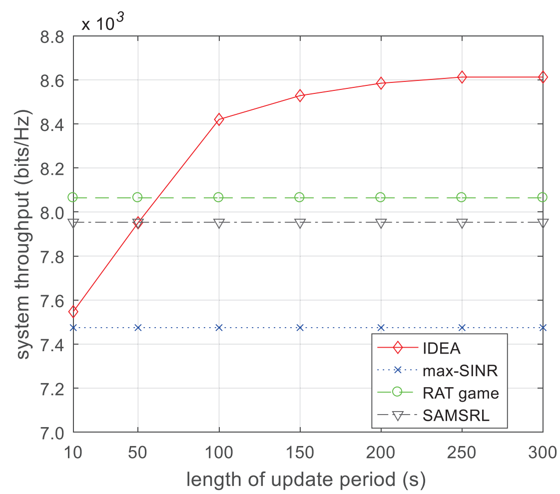


Fig. 4. Comparisons of system throughput vs. T.

In Experiment 2, we compare the number of handoffs in the four cell association policies with the same parameters in the first experiment. Fig.5 shows the number of handoffs for the four policies in 1000 seconds. We can see that the number of handoffs of SAMSRL is much smaller than that of the other three policies. This is because that SAMSRL inherently has constraints on redundant handoffs by using learning mechanism. We also find that the number of handoffs in IDEA policy increases rapidly when the length of period T is short, and it changes slightly when $T \geq 100s$. Specifically, when $T = 100s$, the total number of handoffs in 1000 seconds of SAMSRL, max-SINR, RAT game, and IDEA is 342, 577, 795 and 798, respectively. This result implies that significant system throughput gain can be accomplished with a relatively small compromise on the number of handoffs.

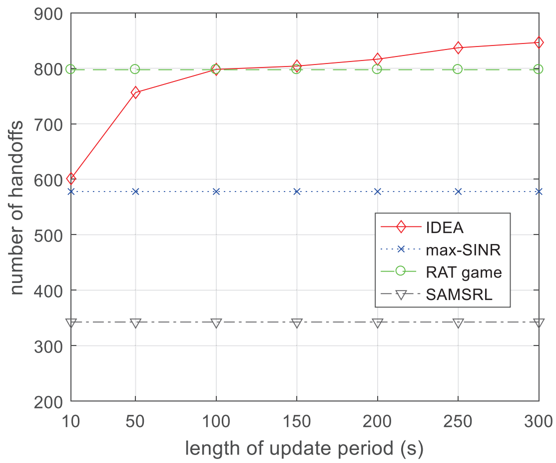


Fig. 5. Comparisons of the number of handoffs vs. T.

In Experiment 3, we compare the system throughput for different values of user distribution parameter G. We fix $T = 100s$ in this experiment and use other parameters which are used in Experiment 1. Fig.6 shows the system throughput in 1000 seconds for the four cell association policies as a function of G. From the results, we can see that the system throughput of all the four policies increases with G. From the theoretical analysis, we know that users are more likely to be associated with the SBSs which are in the hot-spot area for a larger G. Since the density of SBS in hot-spot area is higher than that in other area, UEs can obtain more bandwidth and thus achieve a higher system throughput. Moreover, we also find that the system throughput of IDEA is almost significantly higher than that of the other three policies due to the consideration of user behaviors. Only when G is larger than 0.5, IDEA and SAMSRL have similar throughput performance.

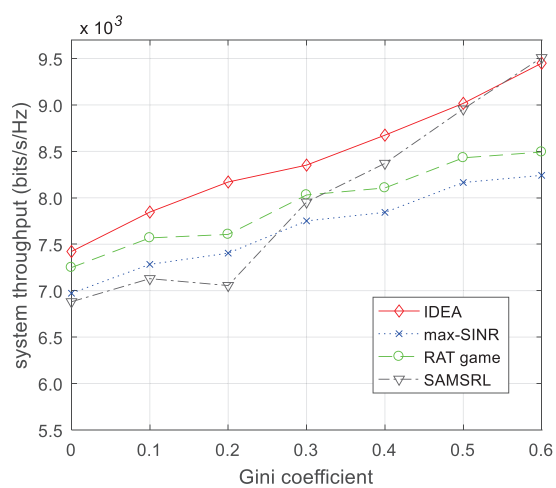


Fig. 6. Comparisons of system throughput vs. Gini coefficient.

In Experiment 4, we examine the effect of UE movement speed on system throughput. Again, we use the same parameters as those in Experiment 1, and fix $T = 100s$, $G = 0.3$. Fig.7 shows the system throughput in 1000 seconds for the four policies as a function of UE movement speed. From this figure we can see that IDEA achieves the highest system throughput under no mobility circumstance. This is because that besides movement speed, IDEA also takes user distribution characteristics into account. With the mean speed of UEs increasing from 2 to 6 m/s, IDEA always outperforms the other three policies in system throughput. When the mean speed exceeds 10 m/s, the system throughput of all the four policies is low due to the fast changing of channel quality.

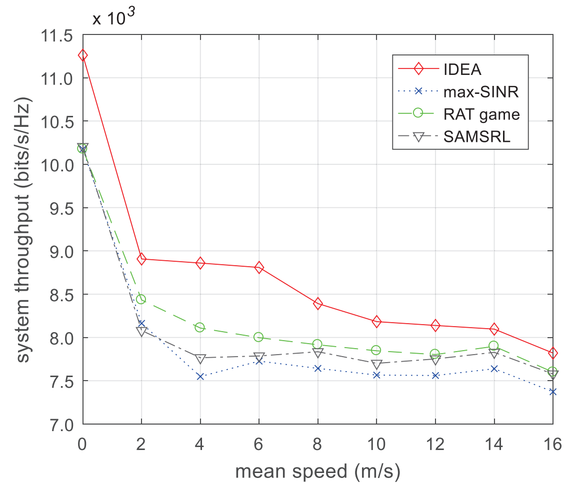


Fig. 7. Comparisons of system throughput vs. mean speed.

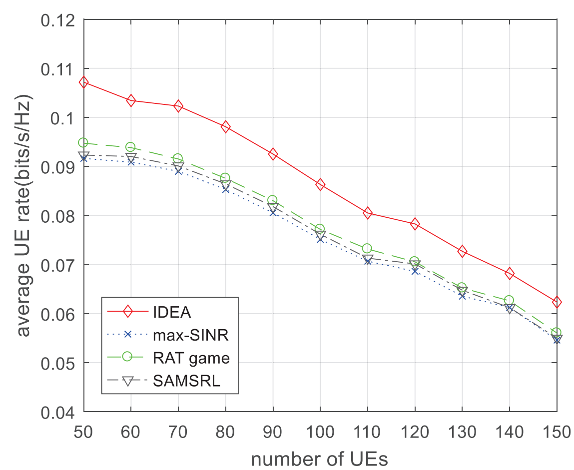


Fig. 8. Comparisons of average UE transmission rate vs. number of UEs.

In Experiment 5, we explore the relationship between UE transmission rate and the number of UEs while using fixed $T = 100s$ in IDEA. Fig.8 shows the average UE rate in 1000 seconds for the four policies as a function of the number of UEs. From this figure we can see that the average UE transmission rate of IDEA is always higher than that of the other three policies. We also find that the average UE transmission rate of the four policies decreases with the number of UEs, which is due to the bandwidth limitations in the system.

VII. Conclusions

In this paper, a user behavior aware cell association policy IDEA has been proposed for HetNets. In IDEA, both clustering and individual user behavior characteristics have been taken into account. IDEA makes cell association decisions according to not only instantaneous UE and network states but also UE mobility pattern and distribution characteristics. Thus, a high long-term system throughput can be achieved. Numerical results have demonstrated that IDEA policy can improve the system throughput by approximately 15% when compared with traditional cell association policies.

REFERENCES

[1] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A Survey on 3GPP Heterogeneous Networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10–21, 2011.

[2] J. G. Andrews, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Communications Magazine*, vol. 51, no. 3, pp. 136–144, mar 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6476878

[3] Qualcomm Europe, "Range expansion for efficient support of heterogeneous networks," *TSG-RAN WG1*, 2008.

[4] W. Wang, X. Wu, L. Xie, and S. Lu, "Femto-Matching : Efficient Traffic Offloading in Heterogeneous Cellular Networks," in *IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 325–333.

[5] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and M. Chiang, "RAT Selection Games in HetNets," in *IEEE Conference on Computer Communications (INFOCOM)*, 2013, pp. 998–1006.

[6] Q. Ye, B. Rong, Y. Chen, M. Al-shalash, C. Caramanis, and J. G. Andrews, "User Association for Load Balancing in Heterogeneous Cellular Networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2706–2716, 2013.

[7] H. Boostanimehr and V. K. Bhargava, "Unified and Distributed QoS-Driven Cell Association Algorithms in Heterogeneous Networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1650–1662, 2015.

[8] D. Niyato and E. Hossain, "Dynamics of Network Selection in Heterogeneous Wireless Networks : An Evolutionary Game Approach," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 2008–2017, 2009.

[9] J.-s. Lin and K.-t. Feng, "Femtocell Access Strategies in Heterogeneous Networks using a Game Theoretical Framework," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1208–1221, 2014.

[10] Chaoming Song, Z. Qu, N. Blumm, and Albert-László Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[11] X. Ge, J. Ye, Y. Yang, and Q. Li, "User Mobility Evaluation for 5G Small Cell Networks Based on Individual Mobility Model," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 528–541, 2016.

[12] Feichen Shen, "A pervasive framework for real-time activity patterns of mobile users," in *proceedings of 2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 2015, pp. 248–250. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7134038

[13] P. Vaka, Feichen Shen, M. Chandrashekar, and Yugyung Lee, "PEMAR: A pervasive middleware for activity recognition with smart phones," in *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 2015, pp. 409–414. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7134073

[14] X. Zhang, Y. Zhang, R. Yu, W. Wang, and M. Guizani, "Enhancing Spectral-Energy Efficiency for LTE-Advanced Heterogeneous Networks: A Users Social Pattern Perspective," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 10–17, 2014.

[15] Y. Huang, X. Zhang, J. Zhang, J. Tang, Z. Su, and W. Wang, "Energy-Efficient Design in Heterogeneous Cellular Networks Based on Large-Scale User Behavior Constraints," *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4746–4757, 2014.

[16] X. Zhang, R. Yu, Y. Zhang, Y. Gao, M. Im, G. C. Laurie, and W. Wang, "Energy-efficient multimedia transmissions through base station cooperation over heterogeneous cellular networks exploiting user behavior," *IEEE Wireless Communications*, vol. 21, no. August, pp. 54–61, 2014.

[17] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-k. Wong, R. Schober, and L. Hanzo, "User Association in 5G Networks : A Survey and an Outlook," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1018–1044, 2016.

[18] F. Guidolin, I. Pappalardo, A. Zanella, and M. Zorzi, "A Markov-based Framework for Handover Optimization in HetNets," in *Proceedings of Ad Hoc Networking Workshop (MED-HOC-NET), 2014 13th Annual Mediterranean. IEEE*, vol. 8, 2014, pp. 134–139.

[19] C. Dhahri and T. Ohtsuki, "Learning-based Cell Selection Method for Femtocell Networks," in *proceedings of 2012 IEEE Vehicular Technology Conference (VTC Spring)*, 2012, pp. 1–5.

[20] Y. Cao, D. Duan, X. Cheng, L. Yang, and J. Wei, "Dynamic Network Selection in HetNets : A Social-Behavioral (SoBe) Approach," in *Proceedings of IEEE Global Communications Conference(GLOBECOM)*, no. 4, 2014, pp. 4653–4658.

[21] D. Liu, Y. Chen, K. K. Chai, T. Zhang, and M. Elkashlan, "Opportunistic User Association for Multi-Service HetNets Using Nash Bargaining Solution," *IEEE Communications Letters*, vol. 18, no. 3, pp. 463–466, 2014.

[22] Y. Chen, J. Li, W. Chen, Z. Lin, and B. Vucetic, "Joint User Association and Resource Allocation in the Downlink of Heterogeneous Networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5701–5706, 2016.

[23] Y. Chen, J. Li, Z. Lin, G. Mao, and B. Vucetic, "User Association with Unequal User Priorities in Heterogeneous Cellular Networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7374–7388, 2016.

[24] Q. Han, B. Yang, G. Miao, C. Chen, and X. Wang, "Backhaul-Aware User Association and Resource Allocation for Energy-Constrained HetNets," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 580–593, 2017.

[25] Q. Han, B. Yang, C. Chen, and X. Guan, "Energy-aware and QoS-aware load balancing for HetNets powered by renewable energy," *Computer Networks*, vol. 94, pp. 250–262, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2015.10.025

[26] Y. Li, B. Cao, and C. Wang, "Handover Schemes in Heterogeneous LTE Networks: Challenges and Opportunities," *IEEE Wireless Communications*, vol. 23, no. 2, pp. 112–117, 2016.

[27] S. Vasudevan, R. N. Pupala, and K. Sivanesan, "Dynamic eICIC A Proactive Strategy for Improving Spectral Efficiencies of Heterogeneous LTE Cellular Networks by Leveraging User Mobility and Traffic Dynamics," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 4956–4969, 2013.

[28] Y. Zhang, L. Lei, H. Long, and K. Zheng, "Performance Analysis of User Association Policies in Small Cell Networks Using Stochastic Petri Nets," in *Proceedings of 2013 IEEE International Conference on Communications Workshops (ICC)*, 2013, pp. 1194–1198.

[29] P. Whittle, "Restless Bandits: Activity Allocation in a Changing World," *Applied Probability*, vol. 25, pp. 287–298, 1988.

[30] C. H.Papadimitriou and J. N.Tsitsiklis, "The Complexity of Optimal Queuing Network Control," *Mathematics of Operations Research*, vol. 24, no. 2, pp. 293 – 305, 1999.

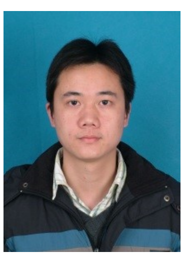
[31] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed Bandit Allocation Indices*, 2nd ed. John Wiley & Sons, 2011.

[32] D. Bertsimas and J. Niño-Mora, "Restless Bandits, Linear Programming Relaxations, and a Primal Dual Index Heuristic," *Operations Research*, vol. 48, no. 1, pp. 80–90, 2000.

[33] N. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica*, vol. 4, no. 4, pp. 373–395, 1984.

Yao Sun received the B.S. degree in Mathematical Sciences from the University of Electronic Science and Technology of China (UESTC). He is currently working towards his Ph.D. degree at National Key Laboratory of Science and Technology on Communications, UESTC. His research and study interests include intelligent access control, handoff and resource management in mobile networks based on machine learning and other data analytics.

Gang Feng (M'01, SM'06) received his BEng and MEng degrees in Electronic Engineering from the University of Electronic Science and Technology of China (UESTC), in 1986 and1989, respectively, and the Ph.D. degrees in Information Engineering from The Chinese University of Hong Kong in1998. He joined the School of Electric and Electronic Engineering, Nanyang Technological University in December 2000 as an assistant professor and was promoted as an associate professor in October 2005. At present he is a professor with the National Laboratory of Communications, University of Electronic Science and Technology of China. Dr. Feng has extensive research experience and has published widely in computer networking and wireless networking research. His research interests include resource management in wireless networks, next generation cellular networks, etc. Dr. Feng is a senior member of IEEE.



Shuang Qin received the B.S. degree in Electronic Information Science and Technology, and the Ph.D. degree in Communication and Information System from University of Electronic Science and Technology of China (UESTC), in 2006 and 2012, respectively. He is currently an associate professor with National Key Laboratory of Science and Technology on Communications in UESTC. His research interests include cooperative communication in wireless networks, data transmission in opportunistic networks and green communication in heterogeneous networks.

Sanshan Sun received the B.S. degree in Electrical Engineering and Automation from Sichuan Normal University in 2006 and the M.S. degree in Communication and Information Systems from Chongqing University of Posts and Telecommunications in 2009, respectively. He is currently a Ph.D. candidate in National Key Laboratory of Science and Technology on Communications in University of Electronic Science and Technology of China (UESTC). His research interests include resource management and game theory.