Wireless Network Association Game With Data-Driven Statistical Modeling

Yu-Han Yang, *Student Member, IEEE*, Yan Chen, *Senior Member, IEEE*, Chunxiao Jiang, *Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—The explosion in demand for wireless data services in recent years has triggered pervasive deployment of wireless networks. How to associate to one of the wireless networks in the best interest of a user is an essential problem to mobile computing. In this paper, we analyze a data set of wireless LAN traces collected from campus networks, from which we observe that the user arrival distribution is approximately Poisson distributed; the session time and the waiting time to switch network can be approximated by exponential distributions. Based on the data analysis, we formulate a wireless access network association game as a multidimensional Markov decision process with negative network externality, where the best response strategy is an approximate Nash equilibrium. A modified value iteration algorithm is proposed to search the best response strategy profile. Applying the proposed algorithm to the data-driven stochastic model, the best response strategy is shown to achieve a better individual expected utility while satisfying the individual rationality, and attain a near-optimal social welfare performance compared to other strategies such as the centralized method and the greedy algorithm.**

***Index Terms*—Wireless network association, game theory, statistical modeling, data set analysis, Markov decision process.**

I. Introduction

NOWADAYS, with the recent proliferation of wireless devices and the ubiquity of wireless networks, users can connect to WiFi wireless networks through hot-spots or access points (APs) in most public areas. As the macrocellular networks usually have a broader range of coverage, the WiFi networks are smaller in its reachable range but more densely deployed. Moreover, the development of femtocells [1] also arouses more choices for cellular service subscribers. Therefore, when a user attempts to access a wireless network, oftentimes he/she may encounter a decision to choose one of multiple wireless networks. In order to obtain a better performance, a user sometimes needs to decide whether to switch to another network during a session. From a user's viewpoint, the decision of network association can lead to different quality of service during the session. From the perspective of a service provider, better allocation of users can provide more efficient utilization of resources such as signal power, temporal and spatial bandwidth. Hence, the wireless access network association problem is becoming more and more important due to its frequent occurrence in our daily life and the influence to efficient resource utilization.

To tackle the wireless network association problem in a practical viewpoint, the model formulation has to take into account empirical study of user behavior, which is not possible without real-life data. The pattern and the statistical properties of user behavior can be extracted from massive amount of wireless LAN traces of APs available in various environments such as university campus, shopping malls, restaurants, coffee shops, airports, etc. In most current practical systems, the network association decision is often made based on the instantaneous signal-to-interference-plus-noise ratio (SINR) criterion, i.e., a user simply connects to the wireless network with the highest SINR. Such a strategy may be a good heuristic by considering the instantaneous utility, but it may not be optimal since SINR does not take into account the long term *negative network externality* [2], [3] caused by subsequent users' decisions. The negative network externality refers to the negative effect on a user caused by other users with the same strategy in a network. For example, the traffic congestion caused by the vehicles that choose the same route delays each vehicle's traveling time. The instantaneous information only reflects the current condition without considering the future utility, which can be significantly degraded if subsequent users make the same decision.

Recently, the wireless network association problem has attracted significant attention in the literature [4]–[18]. The tutorial in [6] provides a comprehensive survey on many existing methods in the literature, in which utility functions and different attributes such as bandwidth, delay, packet loss, etc., are summarized and compared. In [16], a detailed state-of-the-art of existing vertical handover decision mechanisms and decision schemes for heterogeneous wireless networks are categorized and summarized. One category of network association is based on centralized methods to optimize the system performance metrics such as sum rate, minimum rate, or proportional fairness [4], [5]. In [7], an analytic hierarchy process is applied to decide the relative weights of evaluative criteria set according to user preferences and service applications. In [8], Niyato et al. study a network-selection algorithm based on population evolution, which requires a centralized controller, and an algorithm based on reinforcement-learning, where a user can learn and adapt the decision on network selection to reach evolutionary equilibrium without any interaction with other users. In [5], the cell association and resource allocation are considered jointly, and a distributed algorithm via dual decomposition is

Manuscript received March 24, 2015; revised July 6, 2015; accepted August 24, 2015. Date of publication September 1, 2015; date of current version January 7, 2016. The associate editor coordinating the review of this paper and approving it for publication was Prof. Dusit Niyato.

Y.-H. Yang, Y. Chen, and K. J. R. Liu are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: yhyang@umd.edu; yan@umd.edu; kjrliu@umd.edu).

C. Jiang is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: jchx@tsinghua.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2475427

proposed to solve a logarithmic utility maximization problem. Localized cooperation is introduced in [17], where the global social welfare optimization is decomposed into sub-problems by exploiting the spatial distribution of networks with user demand awareness. In [18], an evolutionary game is formulated to model and investigate the adaptive service selection of users, and an evolutionary stable strategy is shown to be an effective solution. Another category of network association methods is characterized by game theory, which models strategic interactions among users using formalized incentive structures [19], [20]. In wireless communications and networking, game theory has been widely studied in many applications [20]–[24] including non-cooperative power control [22], cooperation stimulation [21], and spectrum allocation [23]. In [15], Aryafar et al. investigate the dynamics of network selection games in heterogeneous wireless networks and the convergence properties of these games. In [11], the network selection is modeled as a congestion game, where players make decisions simultaneously to optimize the interference and throughput. The network association problem in [10] is formulated as a non-cooperative game in which users selfishly minimize an association cost accounting for the path length and the path interference to reach the gateway.

While most of the existing works study the scenario that users make simultaneous decisions, in this paper, we consider the network association problem under the scenario where users make *sequential* decisions, and to obtain a better long term utility, users have to consider the negative network externality, i.e., the decisions of subsequent users, to determine his/her best response strategy. Sequential decisions considering the negative network externality effect are investigated in the Chinese Restaurant Game (CRG) [25]–[27], which studies the optimal decision and social learning with negative network externality but with a fixed number of users. In [27], [28], the dynamic CRG is proposed to allow users arriving and leaving stochastically. In our previous work [28], the wireless network selection problem is investigated without considering the strategy of switching to another network, i.e., a user has to stay in a network until departure once he/she associates with the network. In addition, the proposed model in [28] is not justified based on the real-life data set analysis.

In this paper, we further extend the dynamic CRG in [28] to incorporate the behavior of switching to another network. We also extract statistical properties of users' behaviors in wireless networks by analyzing a data set of wireless LAN traces collected from Dartmouth campus networks in a span of 4 months [29]. It is validated that the user arrivals in wireless access networks are approximately Poisson distributed. Previous work on WLAN trace analysis [30]–[36] focus on different aspects, such as uplink/downlink traffic modeling, user mobility patterns, and geographic distribution of users. In [33], [34], the authors validate the arrival processes of users as being time-varying Poisson processes based on first applying a nonlinear transformation of the arrival time, and then verifying the exponentiality of the transformed values. However, it is only shown that the statistical test is a necessary condition for a Poisson process, but the sufficiency may not be clear. In [35], Kullback-Leibler (KL) divergence is used to perform a Poisson distribution test for the arrival process of new sessions at each AP. In [36], the user occupancy distribution of an AP is shown to be a Poisson distribution. However, a Poisson user occupancy distribution does not imply a Poisson user arrival process. Based on the data set, the empirical probability distribution functions of the number of user arrivals are plotted in Figure 1, where different curves represent the number of user arrivals in different durations. It can be seen that the behavior of these probability distribution curves is very similar to the Poisson distribution with different mean values. Furthermore, the waiting time to departure, i.e., the duration of a session, and the waiting time to switch to another network appear to be exponential distributions.

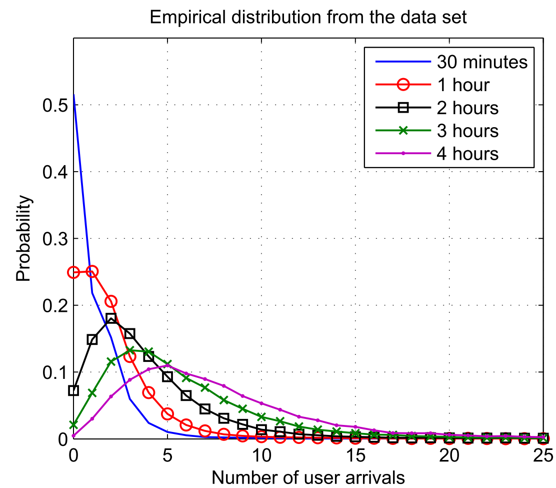


Fig. 1. The empirical probability distribution of the number of user arrivals in different durations measured from the data set.

With the statistical properties extracted from the wireless LAN traces, we are able to construct a stochastic model for the wireless access network system. Next, we show that the problem of finding the best response strategy profile of network association when arriving and the best response strategy of switching during a session is a multi-dimensional Markov decision process (M-MDP). A modified value iteration algorithm is proposed to obtain a solution of an ϵ-approximate Nash equilibrium. It is observed from numerical simulations that the strategy profile obtained by the proposed algorithm has a threshold structure, which allows a much smaller required space to store the strategy profile. Note that no theoretical proof is provided for the threshold structure. Simulation results demonstrate the efficiency and effectiveness of the proposed algorithm, i.e., while achieving the best response strategy for the individual, the proposed algorithm attains similar performance of social welfare compared to the maximum social welfare strategy.

A more sophisticated probabilistic model, such as the modified Poisson distribution used in [37], may be used to fit the data more accurately. However, a general process may not possess the Markovian property which plays a significant role in the system model such that the wireless access network association problem can be reduced into the M-MDP formulation. In this paper, since the empirical inter-arrival distribution can be approximated by an exponential distribution, the simplest Markovian arrival process, i.e., the Poisson process, is adopted to model the user arrival behavior.

We note that since the system model of the proposed wireless access network association game in this paper is quite general and not restricted to any particular technology, the proposed model and algorithms can be applied to many wireless technologies including macro-cellular, femto-cellular and WiFi WLANs given the assumptions such as the availability of the information of arrival rates and utility functions. For example, in a heterogeneous cellular system with macro and femto base-stations available, a cellular user can calculate the best response strategy profile in consideration of the strategies of other users, and then associates to the base-station accordingly to obtain a better long-term utility. WiFi WLANs are considered in this paper as an example of wireless access network selection to illustrate and validate the effectiveness of the proposed method.

The rest of the paper is organized as follows. In Section II, the system model of the wireless network system is introduced. Section III describes the formulation of the wireless access network association game, the expression of the expected utility, and the M-MDP. The analysis of the data set is contained in Section IV, in which we evaluate the probability distribution of user inter-arrival time, session time, and the waiting time to switch to another network. In Section V, data-driven simulation results are shown to demonstrate the performance of the proposed value iteration algorithm. Finally, the conclusion is drawn in Section VI.

II. System Model

In this section, we describe the system model of the wireless access network association game. With the statistical model of the user arrival being a Poisson process and the user departure following an exponential distribution, we can formulate the wireless access network system as follows. The system consists of K networks, and network k acts as a server of a finite capacity N_k, i.e., the network is able to simultaneously serve at most N_k users. For simplicity, it is assumed there is no buffer or waiting room when a network is fully occupied by users.

A user type is defined by the network reachability. With arrival rate $\bar{\lambda}_{\mathcal{K}}$, users of user type $\mathcal{K}$ can choose a network from $\mathcal{K}$, where $\mathcal{K}$ is a non-empty subset of all networks $\{1, \ldots, K\}$. For example, if $K = 2$, then $\mathcal{K}$ can be $\{1\}$, $\{2\}$, or $\{1, 2\}$. The user departure rate is denoted as $\bar{\mu}_0$ uniformly for all networks. We also define uniformly for all networks the network-switching rate $\bar{\mu}_1$, which means the rate that a user switches to another network from his current network.

We consider a discrete time Markov system, where the system state $\mathbf{s} = (s_1, \ldots, s_K)$ takes its value from the state space $\mathcal{S} = \{(s_1, \ldots, s_K)|s_k = 0, 1, \ldots, N_k, k = 1, \ldots, K\}$, where s_k represents the number of users in network k, for $k = 1, \ldots, K$. The duration of a time unit is T (seconds). The arrival probability of type $\mathcal{K}$ users, $\lambda_{\mathcal{K}}$, can be approximated as $1 - e^{-\bar{\lambda}_{\mathcal{K}} T} \approx \bar{\lambda}_{\mathcal{K}} T$. Similarly, the departure probability of a user is approximated as $\mu_0 = \bar{\mu}_0 T$, and the network-switching probability as $\mu_1 = \bar{\mu}_1 T$. To simplify the presentation, we consider $N_k = N, \forall k$ in the rest of the paper.

The arriving user's strategy profile $\boldsymbol{\sigma}^{\mathcal{K}} = \{\sigma_{\mathbf{s}}^{\mathcal{K}} | \forall \mathbf{s} \in \mathcal{S}\}$ is a mapping from the aggregate state space to the action space, i.e., $\boldsymbol{\sigma}^{\mathcal{K}} : \{0, 1, \ldots, N\}^K \mapsto \{1, \ldots, K\}$. The switching user's strategy profile $\boldsymbol{\gamma}^{\mathcal{K}} = \{\gamma_{\mathbf{s}}^{\mathcal{K}} | \forall \mathbf{s} \in \mathcal{S}, \forall \mathcal{K}\}$ is a mapping from $\mathcal{S}$ to the action space, i.e., $\boldsymbol{\gamma}^{\mathcal{K}} : \{0, 1, \ldots, N\}^K \mapsto \{1, \ldots, K\}$. The system transition probability of an arrival event is given by

$$P_{\text{sys}}(\mathbf{s} + \mathbf{e}_j | \mathbf{s}) = \sum_{\mathcal{K}} I_j(\sigma_{\mathbf{s}}^{\mathcal{K}}) \lambda_{\mathcal{K}} \tag{1}$$

where $\mathbf{e}_j$ denotes the vector with the j-th element as one and zeros otherwise, $\sigma_{\mathbf{s}}^{\mathcal{K}}$ denotes the strategy at state $\mathbf{s}$ and $\sigma_{\mathbf{s}}^{\mathcal{K}} = j$ means the strategy to enter network j. The indicator function $I_j(\sigma_{\mathbf{s}}^{\mathcal{K}})$ is defined as $I_j(\sigma_{\mathbf{s}}^{\mathcal{K}}) = 1$ if $\sigma_{\mathbf{s}}^{\mathcal{K}} = j$; otherwise, $I_j(\sigma_{\mathbf{s}}^{\mathcal{K}}) = 0$. At state $\mathbf{s}$, since there are s_j users in network k and each user has an independent departure probability, the probability that one user leaves network j is $s_j \mu_0$. Thus, the system transition probability of a departure event is given by

$$P_{\text{sys}}\left(\mathbf{s} - \mathbf{e}_j | \mathbf{s}\right) = s_j \mu_0, \ \ j = 1, \ldots, K. \tag{2}$$

The network-switching strategy for state $\mathbf{s}$ and user type $\mathcal{K}$ is denoted by $\gamma_{\mathbf{s}}^{\mathcal{K}}$, and $\gamma_{\mathbf{s}}^{\mathcal{K}} = j$ means the strategy to switch from network k to network j. The system transition probability of a network-switching event is then given by

$$P_{\text{sys}}\left(\mathbf{s} - \mathbf{e}_k + \mathbf{e}_j | \mathbf{s}\right) = I_j(\gamma_{\mathbf{s}}^{\mathcal{K}}) s_k \mu_1, \ \ j, k = 1, \ldots, K. \tag{3}$$

As described in (1)–(3), it is feasible to consider all possible types of users who can choose a subset of networks. However, such a detailed description would be very complicated and difficult to illustrate the key idea. To focus on analyzing the intrinsic effect between the system of network association and the strategy profile of equilibrium, in the rest of the paper, we consider the basic scenario where users belong to one of the two classes such that the presentation can be more clear and understandable. The users of type I arrive with arrival rate $\bar{\lambda}_0$ and these users are able to choose among one of the K networks. With arrival rate $\bar{\lambda}_k$, the users of type II can only choose network k, for $k = 1, \ldots, K$. Then, equation (1) can be rewritten as

$$P_{\text{sys}}\left(\mathbf{s} + \mathbf{e}_j | \mathbf{s}\right) = \lambda_j + I_j(\sigma_{\mathbf{s}}) \lambda_0, \ \ j = 1, \ldots, K, \tag{4}$$

where $\sigma_{\mathbf{s}}$ denotes the strategy profile of type II users. Similarly, equation (3) can be rewritten as

$$P_{\text{sys}}\left(\mathbf{s} - \mathbf{e}_k + \mathbf{e}_j | \mathbf{s}\right) = I_j(\gamma_{k,\mathbf{s}}) s_k \mu_1, \ \ j, k = 1, \ldots, K, \tag{5}$$

where $\gamma_{k,\mathbf{s}}$ denotes the network-switching strategy for state $\mathbf{s}$ and network k. Figure 2 shows the state transition diagram of a two-network system.

With (2), (4), and (5), the system transition probability of a staying event is given by

$$P_{\text{sys}}\left(\mathbf{s} | \mathbf{s}\right) = 1 - \sum_{j=0}^{K} \lambda_j - \sum_{j=1}^{K} s_j (\mu_0 + \mu_1). \tag{6}$$

Note that the duration T of a time slot should be chosen such that $\sum_{j=0}^{K} \lambda_j + \sum_{j=1}^{K} N_j (\mu_0 + \mu_1) \leq 1$, i.e.,

$$T \leq \left(\sum_{j=0}^{K} \bar{\lambda}_j + \sum_{j=1}^{K} N_j (\bar{\mu}_0 + \bar{\mu}_1) \right)^{-1}. \tag{7}$$

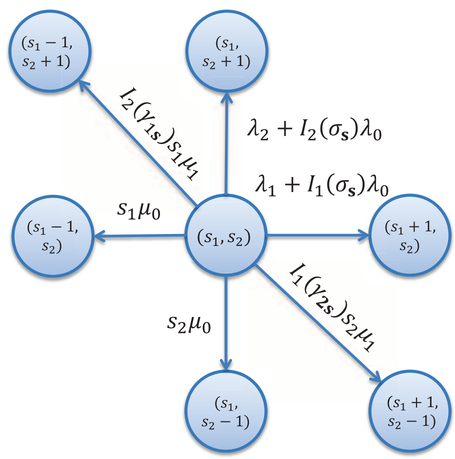


Fig. 2. State transition diagram of the wireless access network association system.

In one time slot, the utility obtained by a user in network k is defined by the function $R(s_k)$, which is a non-increasing function in s_k due to the negative network externality, i.e., the negative effect to the users in a network caused by the increasing number of users. For example, in a code division multiple access (CDMA) system where the available frequency spectrum is used at the same time by all users, $R(s_k)$ can be defined as the achievable data rate function, $\log\left(1+\frac{\mathsf{SNR}_k}{(s_k-1)\mathsf{INR}_k+1}\right)$, where SNR_k is the signal-to-noise power ratio, and INR_k is the interference-to-noise power ratio in network k. The increase of the number of users causes inter-user interference (IUI) to each user in the network. Such IUI results in a lower signal-to-interference-plus-noise power ratio (SINR) and thus a lower achievable data rate for each user in the network. In other scenarios where the available resource is allocated in an orthogonal way, e.g., time division multiple access (TDMA) for time resource allocation, frequency division multiple access (FDMA) for frequency resource allocation, or power control for total transmit power allocation. In these scenarios, the utility $R(s_k)$ can be defined by a simple fraction $\frac{C_k}{s_k}$, where C_k denotes the total amount of the entire available resource and $\frac{C_k}{s_k}$ is the amount of resource obtained by one user in the network.

III. Wireless Access Network Association Game

In this section, the wireless access network association game is formulated by first defining the utility function and deriving the expected utility function using the Bellman equation, based on which, the best response strategy is given. The network association problem is then shown to be a multi-dimensional Markov decision process, for which a modified value iteration algorithm is proposed.

A. Expected Utility

The expected utility of a user arriving and choosing network k to enter when the system state is $\mathbf{s}$ is denoted by $V_k(\mathbf{s})$, which can be expressed by definition as follows.

$$V_k(\mathbf{s}) = E\left[\sum_{t=0}^{\infty}(1-\mu_0)^t R_{k_t}(\mathbf{s}_t)\right], \tag{8}$$

where k_t denotes the network the user stays in at time slot t, with the initial condition $k_0 = k$. Since μ_0 is the probability that the user leaves the current network in one time slot, then $(1-\mu_0)$ is the probability that the user stays in the network in one time slot. Thus, the value $(1-\mu_0)$ can be regarded as the discounting factor for the future utility as time increases. The function $V_k(\mathbf{s})$ denotes the long-term utility of a user arriving and choosing network k to enter when the system state is $\mathbf{s}$. Thus, the function $V(\cdot)$ accounts for the decisions of subsequent users and thus the future number of users associated to the networks.

The expression in (8) can be simplified into a set of Bellman equations [38], i.e., for $k = 1, \ldots, K$,

$$V_k(\mathbf{s}) = R_k(\mathbf{s}) + (1-\mu_0)\sum_{k',\mathbf{s}'} P(k', \mathbf{s}'|k, \mathbf{s})V_{k'}(\mathbf{s}'), \tag{9}$$

where the transition probability $P(k', \mathbf{s}'|k, \mathbf{s})$ denotes the probability that in the current time slot, a user is in network k and the system state is at $\mathbf{s}$, and in the next time slot, the system state becomes $\mathbf{s}'$ and the user switches to network k' if $k' \neq k$, or the user keeps staying in the same network if $k' = k$. Considering different events as in the system transition probability, the conditional transition probability is given by

$$P\left(k', \mathbf{s}'|k, \mathbf{s}\right) = \begin{cases} \lambda_j + I_j(\sigma_{\mathbf{s}})\lambda_0, & \text{if } k' = k, \mathbf{s}' = \mathbf{s} + \mathbf{e}_j, \forall j, \\ s_j\mu_0, & \text{if } k' = k, \mathbf{s}' = \mathbf{s} - \mathbf{e}_j, \forall j \neq k, \\ (s_k - 1)\mu_0, & \text{if } k' = k, \mathbf{s}' = \mathbf{s} - \mathbf{e}_k, \\ I_j(\gamma_{k,\mathbf{s}})(s_k - 1)\mu_1, & \text{if } k' = k, \mathbf{s}' = \mathbf{s} - \mathbf{e}_k + \mathbf{e}_j, \forall j \neq k, \\ I_j(\gamma_{k,\mathbf{s}})\mu_1, & \text{if } k' = j, \mathbf{s}' = \mathbf{s} - \mathbf{e}_k + \mathbf{e}_j, \forall j \neq k, \\ 1 - \sum_{j=0}^{K}\lambda_j - \sum_{j=1}^{K} s_j(\mu_0 + \mu_1) + \mu_0, & \text{if } k' = k, \mathbf{s}' = \mathbf{s}, \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

Notice that there are slight differences in the departure probability and the switching probability between the system transition probability (2), (5) and the conditional transition probability (10).

B. Best Response Strategy

In the wireless access network association game, users adopt the best response strategy to maximize his own expected utility due to the selfish nature. A user makes a decision after he arrives and observes the system state $\mathbf{s}$. The strategy leads the user into certain network k and results in an expected utility $V_k(\mathbf{s} + \mathbf{e}_k)$. In subsequent time slots, the user may change from network k to another network based on $\gamma_{k,\mathbf{s}}$. When observing the state $\mathbf{s}$, the best response arriving strategy $\sigma_{\mathbf{s}}$ has to satisfy

$$\sigma_{\mathbf{s}} = \arg\max_j V_j(\mathbf{s} + \mathbf{e}_j), \forall \mathbf{s} \in \mathcal{S}. \tag{11}$$

Similarly, when observing the state $\mathbf{s}$, a switching user will choose the best response strategy $\gamma_{k,\mathbf{s}}$, which has to satisfy

$$\gamma_{k,\mathbf{s}} = \arg\max_j V_j(\mathbf{s} - \mathbf{e}_k + \mathbf{e}_j), \forall \mathbf{s} \in \mathcal{S}, \forall k. \quad (12)$$

It can be seen that with the arriving user's strategy profile satisfying (11) and the switching user's strategy profile (12), no user can obtain a higher expected utility by unilateral deviation to any other strategy [19]. Therefore, the strategy profile satisfying (9)–(12) is a Nash equilibrium of the stochastic game.

From (11) and (12), it can be observed that

$$\gamma_{k,\mathbf{s}} = \sigma_{\mathbf{s}-\mathbf{e}_k}, \forall \mathbf{s} \in \mathcal{S}, \forall k. \quad (13)$$

Thus, the best response switching strategy in network k at state $\mathbf{s}$ can be interpreted as the best response arriving strategy at state $\mathbf{s} - \mathbf{e}_k$, i.e., the state without the switching user in network k. In other words, the switching behavior can be equivalently considered as leaving the current network and arriving as an arriving user. From this perspective, the two best response strategy profiles are exactly the same, and the switching user's strategy $\gamma_{k,\mathbf{s}}$ in (10) can be replaced by $\sigma_{\mathbf{s}-\mathbf{e}_k}$.

C. Modified Value Iteration Algorithm

The problem of solving the strategy profile satisfying (9)–(12) is a Multi-dimensional Markov Decision Process (M-MDP) problem, in which multiple potential functions are associated with each system state. For a conventional MDP problem [38], there is only one single potential function, by directly optimizing which using the theory of dynamic programming (DP) [39], the optimal strategy can be found with a low complexity. In an M-MDP, the entanglement of the multiple potential functions makes the structure of the problem very complicated. Such a dependency can be expressed in a vector form:

$$\begin{bmatrix} V_1(\mathbf{s}) \\ V_2(\mathbf{s}) \\ \vdots \\ V_K(\mathbf{s}) \end{bmatrix} = \begin{bmatrix} R_1(s_1) \\ R_2(s_2) \\ \vdots \\ R_K(s_K) \end{bmatrix} + (1-\mu_0) \begin{bmatrix} \mathbf{p}_1^T(\mathbf{s}) \\ \mathbf{p}_2^T(\mathbf{s}) \\ \vdots \\ \mathbf{p}_K^T(\mathbf{s}) \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_K \end{bmatrix}, \quad (14)$$

where $\mathbf{p}_k(\mathbf{s})$ and $\mathbf{v}_k$ denote vectors comprising $P(k', \mathbf{s}'|k, \mathbf{s})$ and $V_k(\mathbf{s}')$ as elements, $\forall k$, $\forall \mathbf{s}$. The transpose operator is denoted by $(\cdot)^T$. For given k and $\mathbf{s}$, the vector $\mathbf{p}_k(\mathbf{s})$ consists of $K \times |\mathcal{S}|$ elements for different combinations of $(k', \mathbf{s}')$. (14) is formed by stacking (9) of different k's in a vector form for a given $\mathbf{s}$. DP cannot be directly applied in solving such a problem since the arriving strategy $\sigma_{\mathbf{s}}$ and the switching strategy $\gamma_{k,\mathbf{s}}$ are determined by comparing $V_k(\mathbf{s} + \mathbf{e}_k)$ for all k as in (11) and (12) instead of optimizing a single potential function. Note that different from the vector form given in [27], [28], the probability matrix in (14) is more general since it allows non-block-diagonal terms due to the switching behavior, while the probability matrix in [27], [28] only has block-diagonal terms.

As described in Section III-B, the best response strategy profile $\boldsymbol{\sigma}$ has to satisfy (11) given the expected utilities $\{V_k\}_{k=1}^K$. Given a strategy profile $\boldsymbol{\sigma}$, the expected utilities $\{V_k\}_{k=1}^K$ can be obtained using (9) or (14), where the conditional transition probability $P(k', \mathbf{s}'|k, \mathbf{s})$ is a function of the arriving strategy $\sigma_{\mathbf{s}}$ and the switching strategy $\sigma_{\mathbf{s}-\mathbf{e}_k}$. The expected utility of a user is influenced by other users' strategies through the transition probabilities as can be seen in the vector form (14). To find the best response strategy profile $\boldsymbol{\sigma}$ satisfying (9)–(12), we propose a modified value iteration algorithm to solve the problem by iteratively update the strategy profile and the expected utilities, i.e., at the n-th iteration, given the expected utilities, the strategy profile is updated as

$$\sigma_{\mathbf{s}}^{(n+1)} = \arg\max_k V_k^{(n)}(\mathbf{s} + \mathbf{e}_k), \forall \mathbf{s} \in \mathcal{S}. \quad (15)$$

The expected utility functions can be obtained by solving

$$V_k^{(n+1)}(\mathbf{s}) = R_k(s_k) + (1-\mu_0) \sum_{k',\mathbf{s}'} P^{(n+1)}\left(k', \mathbf{s}'|k, \mathbf{s}\right) \cdot V_{k'}^{(n+1)}(\mathbf{s}'), \forall \mathbf{s} \in \mathcal{S}, \forall k \in \{1, \ldots, K\}, \quad (16)$$

where the transition probability $P_k^{(n+1)}(\mathbf{s}'|\mathbf{s})$ is updated using the strategies obtained from (15), i.e.,

$$P^{(n+1)}\left(k', \mathbf{s}'|k, \mathbf{s}\right) = \begin{cases} \lambda_j + I_j(\sigma_{\mathbf{s}}^{(n+1)})\lambda_0, & \text{if } k' = k, \mathbf{s}' = \mathbf{s} + \mathbf{e}_j, \forall j, \\ s_j\mu_0, & \text{if } k' = k, \mathbf{s}' = \mathbf{s} - \mathbf{e}_j, \forall j \neq k, \\ (s_k - 1)\mu_0, & \text{if } k' = k, \mathbf{s}' = \mathbf{s} - \mathbf{e}_k, \\ I_j(\sigma_{\mathbf{s}-\mathbf{e}_k}^{(n+1)})(s_k - 1)\mu_1, & \text{if } k' = k, \mathbf{s}' = \mathbf{s} - \mathbf{e}_k + \mathbf{e}_j, \forall j \neq k, \\ I_j(\sigma_{\mathbf{s}-\mathbf{e}_k}^{(n+1)})\mu_1, & \text{if } k' = j, \mathbf{s}' = \mathbf{s} - \mathbf{e}_k + \mathbf{e}_j, \forall j \neq k, \\ 1 - \sum_{j=0}^K \lambda_j - \sum_{j=1}^K s_j(\mu_0 + \mu_1) + \mu_0, & \text{if } k' = k, \mathbf{s}' = \mathbf{s}, \\ 0, & \text{otherwise.} \end{cases} \quad (17)$$

In (16), the problem of the expected utilities involves a set of linear system, which consists of KN^2 unknown variables corresponding to $\{V_k^{(n+1)}(\mathbf{s}), \forall \mathbf{s}, \forall k\}$ and KN^2 equations, which can be solved by either matrix inversion or linear programming. Another approach is the value iteration algorithm [38], which first initializes $V_k^{(n+1)}(\mathbf{s})$ as an arbitrary value such as zero and iteratively updates itself using (16). Since the iteration function is a contraction mapping, it is guaranteed to converge to a unique fixed point. However, the convergence may be slow if the system space is large since it takes longer for the effect of a strategy to propagate through the whole system.

The proposed algorithm iteratively updates the strategy profile $\boldsymbol{\sigma}$ and the expected utilities $V_k(\mathbf{s})$ until converged. When the proposed algorithm converges, it is observed that there exists a threshold structure of the strategy profile. In [28], a theoretical proof of the threshold structure is given for the special case of $K = 2$ and $\mu_1 = 0$, i.e., in a two-network scenario with no switching strategy allowed. Although it is difficult to theoretically prove the threshold structure for the general cases, in Section V, by numerical simulations we have always observed a threshold structure of the strategy profile for all cases.

However, the strategy profile may not converge but oscillates due to the hard decision rule in (15). The non-convergence

TABLE I
MODIFIED VALUE ITERATION ALGORITHM

(i) Initialize: $V_k^{(0)}(\mathbf{s}) = 0$, $\forall k \in \{1, \ldots, K\}$, $\forall \mathbf{s} \in \mathcal{S}$. $\boldsymbol{\Sigma}_0 = \phi$.
(ii) **Loop** :
1. Update $\{\sigma_{\mathbf{s}}^{(n+1)}\}$ by (18).
If $\{\boldsymbol{\sigma}_{\mathbf{s}}^{(n+1)}\} = \{\boldsymbol{\sigma}_{\mathbf{s}}^{(n)}\}$, **then** stop loop.
else if $\{\boldsymbol{\sigma}_{\mathbf{s}}^{(n+1)}\} \in \boldsymbol{\Sigma}_n$, **then**
choose a $\{\boldsymbol{\sigma}_{\mathbf{s}}\} \in \bar{\boldsymbol{\Sigma}}_n$, and
let $\{\boldsymbol{\sigma}_{\mathbf{s}}^{(n+1)}\} = \{\boldsymbol{\sigma}_{\mathbf{s}}\}$.
end if
$\boldsymbol{\Sigma}_{n+1} = \boldsymbol{\Sigma}_n \cup \{\boldsymbol{\sigma}_{\mathbf{s}}^{(n+1)}\}$.
2. Update $\{P^{(n+1)}(k', \mathbf{s}'\|k, \mathbf{s})\}$ by (17).
3. Solve $\{V_k^{(n+1)}(\mathbf{s})\}$ in (16) by value iteration or linear programming.
Until $\bar{\boldsymbol{\Sigma}}_n = \phi$ or $\{\boldsymbol{\sigma}_{\mathbf{s}}^{(n+1)}\} = \{\boldsymbol{\sigma}_{\mathbf{s}}^{(n)}\}$.

occurs when the strategy of a state near a threshold of strategy change oscillates between different choices each time when the expected utility is updated. When such a situation happens, the expected utilities corresponding to different strategies are very close to each other. Hence, to tackle the problem, we relax the hard decision rule by allowing a small region of tolerance for switching among the strategies [40], which leads to the soft decision rule as follows.

$$\sigma_{\mathbf{s}}^{(n+1)} = \begin{cases} \sigma_{\mathbf{s}}^{(n)}, \\ \qquad \text{if } V_{\sigma_{\mathbf{s}}^{(n)}}^{(n)}(\mathbf{s} + \mathbf{e}_{\sigma_{\mathbf{s}}^{(n)}}) \geq \max_k V_k^{(n)}(\mathbf{s} + \mathbf{e}_k) - \epsilon, \\ \arg\max_k V_k^{(n)}(\mathbf{s} + \mathbf{e}_k), \\ \qquad \text{if } V_{\sigma_{\mathbf{s}}^{(n)}}^{(n)}(\mathbf{s} + \mathbf{e}_{\sigma_{\mathbf{s}}^{(n)}}) < \max_k V_k^{(n)}(\mathbf{s} + \mathbf{e}_k) - \epsilon, \end{cases} \tag{18}$$

where $\epsilon > 0$ is a small constant. Table I summarizes the proposed modified value iteration algorithm for the M-MDP, where $\boldsymbol{\Sigma}_n$ refers to the set of strategy profiles that have been examined at the n-th iteration, and $\bar{\boldsymbol{\Sigma}}_n$ denotes the complement set of $\boldsymbol{\Sigma}_n$, i.e., the set of strategy profiles that have not been examined.

The convergence of the modified value iteration algorithm can be easily shown by the monotonicity of $\{\boldsymbol{\Sigma}_n\}$, i.e.,

Lemma 1: $|\boldsymbol{\Sigma}_n|$ is strictly increasing in n, where $|\cdot|$ denotes the cardinality.

Proof: It can be seen from Table I that in each intermediate iteration, the line $\boldsymbol{\Sigma}_{n+1} = \boldsymbol{\Sigma}_n \cup \{\boldsymbol{\sigma}_{\mathbf{s}}^{(n+1)}\}$ must be executed, where the $\{\boldsymbol{\sigma}_{\mathbf{s}}^{(n+1)}\}$ is not in $\boldsymbol{\Sigma}_n$ since $\{\boldsymbol{\sigma}_{\mathbf{s}}^{(n+1)}\} \in \bar{\boldsymbol{\Sigma}}_n$. Thus, $\{\boldsymbol{\sigma}_{\mathbf{s}}^{(n+1)}\}$ is distinct from all the elements in $\boldsymbol{\Sigma}_n$. Therefore, the number of elements in the set $\boldsymbol{\Sigma}_n$ is strictly increasing in n. ■

Theorem 1: The proposed modified value iteration algorithm converges.

Proof: From Lemma 1, $\{|\boldsymbol{\Sigma}_n|\}$ is a strictly increasing sequence, and $|\boldsymbol{\Sigma}_n|$ has an upper bound since the number of strategy profiles is finite. Thus, the convergence of the sequence $\{|\boldsymbol{\Sigma}_n|\}$ follows. ■

The algorithm stops when an equilibrium is found or all the strategy profiles are searched. By definition, when the algorithm obtains a solution, the resulting strategy profile is an ϵ-approximate NE [19], in which the strategy at each state has an expected utility that is at most ϵ less than that of any other strategy. If the algorithm stops because all the strategies have been explored and none of them is an ϵ-approximate NE, then there exists no feasible solution. However, the existence of an ϵ-approximate NE solution can be guaranteed by first applying the mechanism design algorithm discussed in Section III-D to design a set of appropriate parameters, and then the iterative algorithm will find a feasible solution. Note that there may be multiple ϵ-approximate NEs especially for a larger ϵ when a larger region of tolerance is allowed for switching among the strategies.

The iterative algorithm is executed by each user to calculate the optimal strategy profile, i.e., the optimal strategy of each state. The calculation can be offline and stored as a table. When a user is in a system of networks, he/she can look up the table to make a decision by following the strategy profile. Since the strategy profile is computed by each user but not informed by a centralized server, the proposed algorithm is distributed. To perform the iterative algorithm, the users need to have the information such as the arrival and departure rates and the utility function.

As to the parameter measurement, the network system operator can perform the measurement continually, and update the arrival rate or switching rate averaged over a certain time period. Then the operator announces the parameters and each user can calculate the decision table when the rates are updated. We have observed in numerical simulations that if two sets of parameters are close to each other in their values, the resulting decision tables are also close to each other. This suggests that the previous decision table can be served as the initial decision table to start the proposed algorithm. Since the new decision table is close to the initial one, the convergence speed should be reasonably fast.

The number of system states as well as the computational complexity increase exponentially as the number of networks increases. For conventional MDP problems, the curse of dimensionality may be relieved by using re-enforcement learning techniques [41], [42] such as Q-learning [43] to approximate the optimal strategy. Similar approximation approach can be applied to alleviate the computational complexity in a higher-order M-MDP problem.

For more general cases (considering all possible types of users who can choose a particular subset of networks), the proposed algorithm can be extended by defining an appropriate strategy profile for all types of users and deriving the expected utility of each system state given a particular strategy profile. We can still apply the main idea of the proposed algorithm, which is to iteratively update the strategy profile and the expected utility.

D. Mechanism Design

In the previous subsection, the presented algorithm is performed by the users to search for the best response strategy profile given the system parameters, including the immediate utility function $R_k(s_k)$, the user arrival rate λ_k, the user departure rate μ_0, and the network-switching rate μ_1. On the other hand, for a network system operator, it is desirable to design a set of system parameters such that the resulting best response strategy profile is preferred to the overall network system. In the literature of game theory, such a scenario is called mechanism design, in which the system operator constructs an environment or a system setting by taking into account users' rationality and incentives to achieve the system's objective.

In this subsection, we introduce a mechanism design for the network association game. The network system operator performs the mechanism design to determine a set of system parameters, given which the users play the network association game and use the proposed modified value iteration algorithm to solve the game. The operator knows how the users will play the game and what the outcome will be. That is, when the system parameters are given, the resulting decision table is deterministic. Therefore, in the mechanism design, the operator can choose a set of system parameters in his own interest, which may be a certain objective such as network utilization or total throughput. Based on the objective, the mechanism design problem can be formulated to find out the adjustable system parameters, e.g., the network resource/capacity C_k, by means of resource allocation. In the following, we provide an example to demonstrate how to design a mechanism in the network system such that the resulting strategy profile is as desired.

Consider a network system with orthogonal resource allocation such as TDMA or FDMA, the utility can be modeled as a linear function $R_k(s_k) = \frac{C_k}{s_k}$, where C_k denotes the available resource in network k, and each of the s_k users in network k can obtain $\frac{C_k}{s_k}$ per unit time. Given the strategy profile $\boldsymbol{\sigma} = \{\sigma_{\mathbf{s}}, \forall \mathbf{s}\}$, the problem of designing C_k, i.e., managing appropriate resource to different networks, can be formulated as the following feasibility problem with variables $C_1, \ldots, C_K$ and $V_k(\mathbf{s}), \forall k, \forall \mathbf{s}$.

$$\mathcal{P}_{\text{MD}} : \text{Find } (C_1, \ldots, C_K) \tag{19}$$

$$\text{s.t. } V_k(\mathbf{s}) = R_k(s_k) + (1-\mu_0)\cdot \sum_{k',\mathbf{s}'} P(k', \mathbf{s}'|k, \mathbf{s}) V_{k'}(\mathbf{s}'), \forall k, \forall \mathbf{s}, \tag{20}$$

$$V_{\sigma_{\mathbf{s}}}(\mathbf{s} + \mathbf{e}_{\sigma_{\mathbf{s}}}) \geq \max_k V_k(\mathbf{s} + \mathbf{e}_k) - \epsilon, \forall \mathbf{s}. \tag{21}$$

Note that given the strategy profile, the conditional probability $P(k', \mathbf{s}'|k, \mathbf{s})$ is a constant in the above feasibility problem. Therefore, the constraints of the feasibility problem $\mathcal{P}_{\text{MD}}$ comprise $K \Pi_{k=1}^{K} N_k$ equalities and $\Pi_{k=1}^{K} N_k$ inequalities linear in the variables, and thus the problem is a linear programming problem, which can be solved in polynomial time using an interior-point algorithm.

For nonlinear utility functions such as $R_k(s_k) = \log\left(1 + \frac{\mathsf{SNR}_k}{(s_k-1)\mathsf{INR}_k+1}\right)$, it is possible to similarly formulate the feasibility problem with variables SNR_k and INR_k instead of C_k in (19). However, the resulting feasibility problem has non-convex constraints and hence it is difficult to solve. Optimization techniques such as convex approximation or global search may be applied but it is beyond the scope of this paper.

The formulation of $\mathcal{P}_{\text{MD}}$ is simply an example to illustrate the mechanism design and more sophisticated constraints or certain objective function can be added into the formulation to reflect the system requirement. For example, if the network system has a constraint of sum capacity, e.g., $\sum_k C_k \leq C_{\text{sum}}$, then such a constraint can be incorporated into the optimization problem $\mathcal{P}_{\text{MD}}$. The objective function of the optimization problem can be chosen based on a system goal, e.g., the social welfare (total throughput). The system parameters C_k can then be chosen to optimize the objective within the feasible region. The obtained solution can guarantee the existence of an ϵ-approximate NE, and users can apply the proposed algorithm to find out the best response strategy profile.

For ϵ big enough, every strategy profile could be an ϵ-approximate equilibrium. However, it is more desirable to find a small and feasible ϵ. One possible approach is to use bisection search. Since a larger ϵ makes the constraints in (21) less stringent, the feasibility solution of problem $\mathcal{P}_{\text{MD}}$ is monotonic with ϵ. Thus, the bisection search of ϵ along with the mechanism design algorithm can be used to find the smallest ϵ such that $\mathcal{P}_{\text{MD}}$ is feasible.

IV. Data Set Analysis

Previous work on WLAN trace analysis [30]–[33], [36] focus on different aspects, such as uplink/downlink traffic modeling, user mobility patterns, and geographic distribution of users. In this work, we are interested in the statistical modeling for the events related to the association between users and APs. Specifically, we aim to validate the probability distribution of the user arrival, the waiting time to departure, the waiting time to switch network. We adopt actual wireless network data drawn from CRAWDAD [29], a well known publicly available archive of wireless data resource for the research community, and analyze the probability distribution of user inter-arrivals, session time, and the switching frequency. In the following, we first introduce the basic information of the data set, our methodology, and the results of the analysis.

A. Data Set Description

The data set we use is the CRAWDAD Dartmouth campus WLAN trace [29], [32], [44], which includes syslog (system message log), SNMP (Simple Network Management Protocol polls), and tcpdump (TCP/IP packet analysis) during Fall term 2003 and Winter term 2004 in Dartmouth College. Both syslog and SNMP traces recorded the user association information with a timestamp, the user's MAC address, and the AP's name. However, we observe that sometimes a user's association record in the syslog traces is repeated for several times in a short period, and very often a user leaves without showing the record of a disassociation. As also noted in [32], most disassociation messages do not show a successful disassociate, but report an error that it attempts to disassociate with a wrong AP. Thus, it is rather difficult to uncover the true information of users' behavior by analyzing the syslog traces.

On the contrary, the SNMP traces, which collected the Simple Network Management Protocol (SNMP) polling every AP every 5 minutes, are more reliable for our purpose since each poll contains the instantaneous information of which user

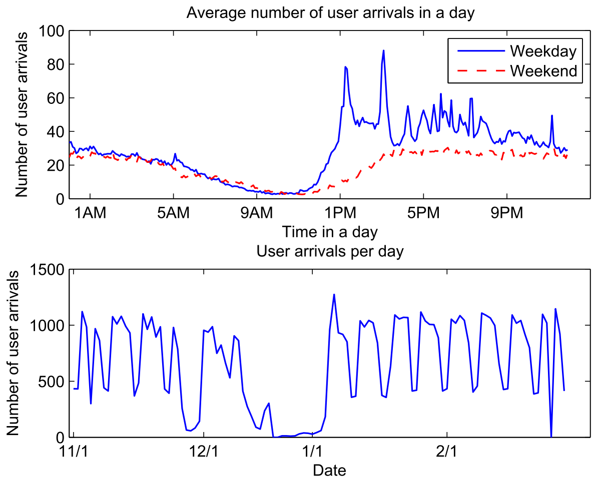


Fig. 3. The upper part shows the average number of user arrivals per hour in a weekday and in a weekend day; the lower part shows the number of user arrivals per day during the 4-month period.

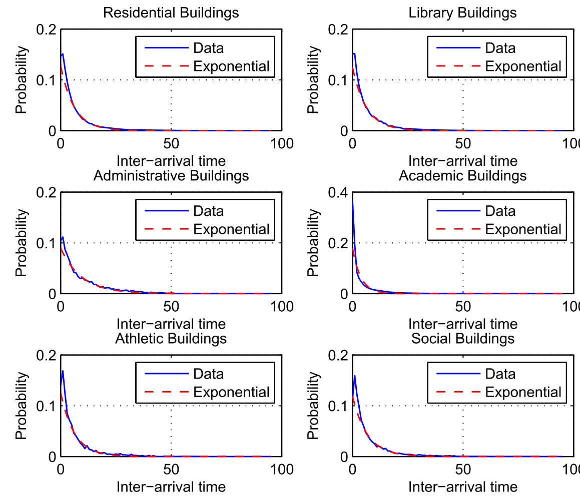


Fig. 4. The probability density function of the inter-arrival time versus the exponential distribution with the same mean value.

is currently connected to the AP. Although the 5 minutes period may be coarse at first sight, from our statistical analysis below, we find it sufficient for estimating the relevant parameters of the M-MDP system model. The traces were recorded by a central server using the Simple Network Management Protocol to poll each of the 560 APs in 6 different types of buildings (Academic, Administrative, Residential, Social, Library, and Athletic) on campus from November 1st, 2003 to February 28th, 2004. The SNMP query collected the AP-related information including the number of inbound and outbound bytes, packets and errors, and the users currently or recently associated with a given AP, and the user-related information including MAC and IP addresses, signal strength and quality, the number of inbound and outbound bytes, packets and errors.

B. Statistical Analysis

We plot the number of user arrivals during the 4-month period in the lower part of Figure 3, in which we can see that relatively fewer users arrive during weekends and holidays. The upper part of Figure 3 shows the average number of arrivals per hour in a weekday and a weekend day. As expected, the user arrivals occur more in the afternoon on weekdays than on weekends. Based on the above observation, in the following analysis, we only consider the abundant traces on weekdays to have richer and consistent data. Only traces between 9AM and 5PM are extracted so that the typical behavior during the daytime can be captured.

We define a user arrival by the event that a user is associated with an AP in a type of network and the user is not associated with any AP in the network in the past 2 time slots to take into account of the scenario that there may be a time slot when the user is not recorded by any AP but the user is switching from an AP to another AP. Similarly, a user departure is defined by the event that a user associated with an AP in a type of network becomes not associated with any AP in the network in the next 2 time slots. If a user is associated with multiple APs in one time slot, a switching event is defined to occur with a duration of 0. A switching event also occurs if the associated AP of a user has changed after k time slots to another AP before a departure event occurs. A session is then defined by the time between a user arrival and a departure with only switching events allowed in between.

Figure 4 shows the empirical probability density function (pdf) of the inter-arrival duration versus the theoretical exponential distribution with the same mean as the data set. It can be observed that the exponential distribution can provide a very good approximation to the empirical pdf for all 6 types of buildings. Compared to the theoretical exponential distribution, the empirical pdf tends to decrease faster in the middle range of the inter-arrival time, but when the inter-arrival time becomes larger, the tail of the empirical pdf stays longer. Such a tendency is especially prominent for Academic Buildings. We speculate that this may be due to the regular pattern of the activities on campus, where the durations of classes and break time are usually fixed. Hence, such a pattern may cause the user arrival event not as random than expected. Except this minor discrepancy, from Figure 4, the exponential distribution is still a satisfactory approximation.

In Figure 5, we plot the quantile-quantile plot [45] of the empirical probability mass function (pmf) of the number of user arrivals in 3 hours versus the theoretical Poisson distribution with the same mean value of the data set. The quantile-quantile (Q-Q) plot is a graphical method for comparing two probability distributions. If the two distribution are similar or linearly related, the points will approximately lie on a straight line. If the two distributions are exactly identical, the points on the Q-Q plot should lie on the line $x = y$. From the figure, we can see that the empirical pmf has a high similarity to a Poisson distribution. The distributions of Administrative Buildings and Academic Buildings are less similar to the Poisson distribution compared to the other types of buildings. We speculate that there might be more stationary users such as administrative staff in these two types of buildings while in the other types of buildings, most users might be students who usually stay in a particular type of buildings for a relatively short period of

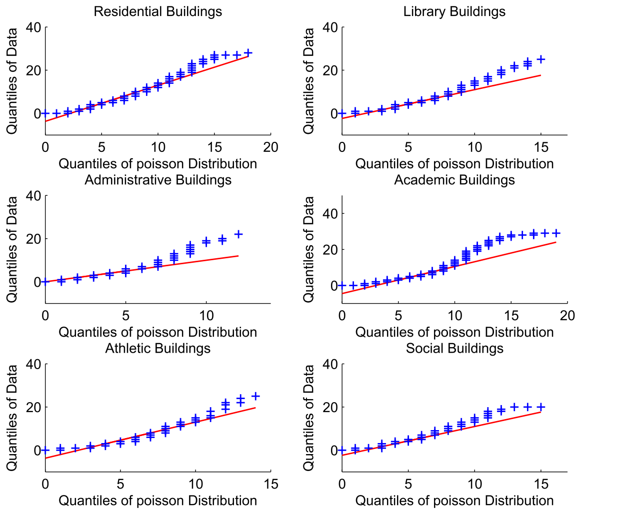


Fig. 5. The quantile-quantile plot of the probability mass function of the number of user arrivals in 3 hours versus the Poisson distribution with the same mean value.

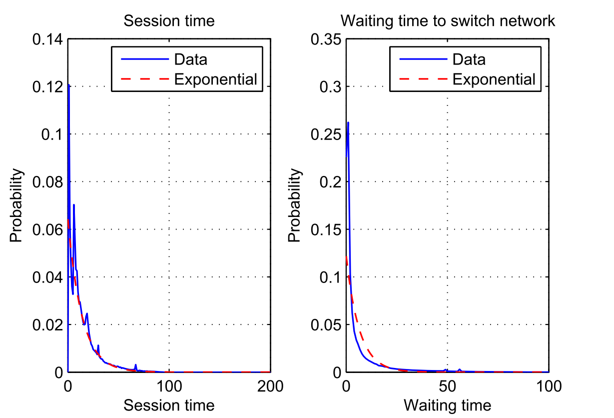


Fig. 6. The probability density functions of a session time and the waiting time to switch to another network.

time. Those stationary users might exhibit very different network association behavior, e.g., periodically leaving and joining the same network. Thus, the distributions of Administrative Buildings and Academic Buildings appear to be less similar to the Poisson distribution.

Figure 6 shows the pdfs of a session time and the waiting time to switch to another network. In each plot, we also compare the empirical curve with the exponential distribution with the same mean value. We can see that for the session time distribution, there are a few peaks which may indicate some fixed patterns of activities on campus. As discussed above, there may be a relatively high probability that the session time is equal to the duration of a class or break time between classes, e.g., 50 or 15 minutes. Except those peaks, the general trend of a session time still approximately follows an exponential distribution with the same mean value.

From the analysis of these statistical properties, we can model the realistic user arrivals as a Poisson distribution, and thus the inter-arrival time as an exponential distribution; the session time and the waiting time to switch network can be modeled as exponential distributions. Therefore, we have the

TABLE II
AVERAGED EMPIRICAL PARAMETERS FOR DIFFERENT TYPES OF CAMPUS NETWORKS

Types	$\bar{\lambda}^{-1}$	$\bar{\mu}_0^{-1}$	$\bar{\mu}_1^{-1}$
Residential	6.1460	12.7200	7.4541
Library	6.3522	12.2097	8.2054
Administrative	9.5474	14.0074	8.4835
Academic	3.7513	13.4151	6.3910
Athletic	6.2693	16.6079	6.3641
Social	6.7630	12.2915	6.4784
Overall	5.1096	13.9891	6.5283

Markov state model as described in Section II, in which a state in the wireless network system can be represented by the numbers of users in different networks without knowing the history of user arrivals due to the Markovian property. The departure probability μ_0 for a user can be approximated by the inverse of the mean session time, i.e., $\bar{\mu}_0 T$; the switching probability μ_1 can also be approximated by the inverse of the mean waiting time to switch network, i.e., $\bar{\mu}_1 T$.

The approximation error of the arrival and departure probabilities is affected by the practical duration of a time slot. The approximation is more accurate as the duration of a time slot is smaller. This can be seen by evaluating the approximation error of the probability, i.e,

$$\left|\left(1 - e^{-\lambda T}\right) - \lambda T\right| = \left|\sum_{n=2}^{\infty} \frac{(-\lambda T)^n}{n!}\right| \tag{22}$$

However, due to constraints such as the resolution of the data set, in practice, the duration of a time slot cannot be too small. In this paper, the data set collected the Simple Network Management Protocol (SNMP) polling APs in every 5 minutes, which is used as the duration of a time slot in the system model.

Table II summarizes the empirical average values of the parameters for the M-MDP model, including the mean inter-arrival time $\bar{\lambda}^{-1}$, mean session time $\bar{\mu}_0^{-1}$, and the mean switching time $\bar{\mu}_1^{-1}$, for different types of campus networks and the overall, that is, the average of all types of networks. Note that a unit time slot is 5 minutes. Thus, we may interpret the 'overall' row as: on average, every 26 minutes there is a user arrival event; each arrival stays for a session of 70 minutes in the network before departure; during a session, every 33 minutes the user switches to another AP.

V. DATA-DRIVEN SIMULATION

In this section, a data-driven numerical simulation is conducted for the wireless access association game described in Section III. Based on the data set analysis in Section IV, we adopt the system parameters such as users' arrival, departure, and switching rates from Table II. We note that practically λ_k, μ_0 and μ_1 can vary over time, and the optimal strategy should take into account all the instantaneous parameters. However, such time-varying system model is difficult to formulate and may result in computational intractability. Therefore, we consider the data-driven simulation with time-invariant parameters

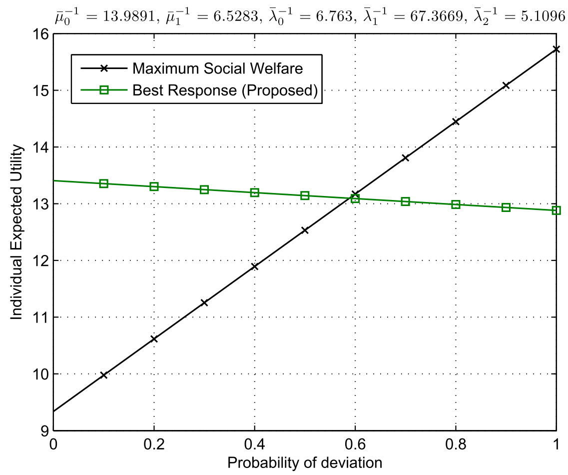


Fig. 7. The individual expected utility versus probability of deviation with the proposed best response strategy and the maximum social welfare strategy.

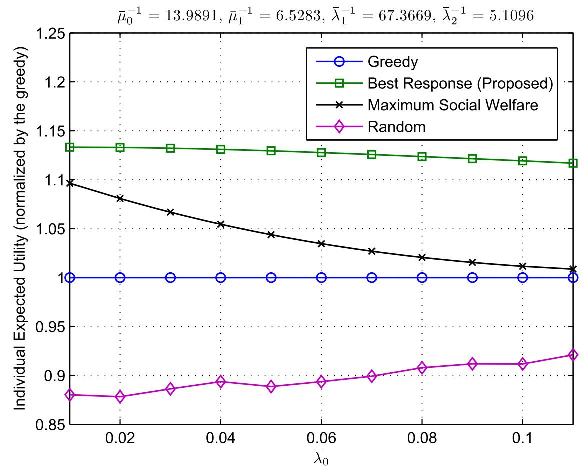


Fig. 8. Individual expected utility comparison in a 2-network system with different strategies including the greedy method, the proposed best response strategy, the centralized maximum social welfare strategy, and the random strategy.

obtained by averaging over daytime in weekdays to capture the typical behavior. The resulting solution in practical systems can be considered as an approximation of the optimal strategy considering time-varying parameters. Unless otherwise stated, in the following simulation, the parameters are chosen as $K = 2$, $\bar{\mu}_0^{-1} = 13.9891$, $\bar{\mu}_1^{-1} = 6.5283$, $\bar{\lambda}_0^{-1} = 6.763$, $\bar{\lambda}_1^{-1} = 67.3669$, $\bar{\lambda}_2^{-1} = 5.1096$, $T = 1$, $N = 4$, $\epsilon = 0.05$, where $\bar{\lambda}_1^{-1}$ is chosen such that $(\bar{\lambda}_1 + \bar{\lambda}_0)^{-1} = 6.1460$ to model that some users follow the strategy profile and the some users always enter one certain network. The utility function $R_k(s_k)$ is defined to be the achievable data rate $\log\left(1 + \frac{\mathsf{SNR}_k}{(s_k-1)\mathsf{INR}_k+1}\right)$, where $\mathsf{SNR}_k = 50$, $k = 1, 2$, and $\mathsf{INR}_k = 10$, $k = 1, 2$. In the following, we will compare the proposed best response strategy with other possible strategies including the random strategy, the myopic strategy, and the centralized strategy. The random strategy is to randomly (with uniform probability) select a network among all networks. The greedy strategy is to choose the network with the best immediate utility instead of the long term expected utility, i.e., $\sigma_{\mathbf{s}}^{\text{myopic}} = \arg\max_k R_k(s_k)$. The maximum social welfare strategy is the social welfare optimizer, i.e., the strategy profile that results in the globally maximum system throughput, the maximum amount of total achievable data rate from the entire network system, i.e., $\sum_{\mathbf{s}} \pi(\mathbf{s}) \sum_k s_k R_k(s_k)$, where $\pi(\mathbf{s})$ denotes the stationary probability at system state $\mathbf{s}$. In the simulation, the maximum social welfare strategy profile is found by exhaustive searching all possible strategy profiles. Since the complexity is in the order of $\mathcal{O}(K^{N^K})$, which is very high even for $K = 2$, we only simulate small N to demonstrate the comparison between different strategies.

In Figure 7, we verify the individual rationality by examining the relation between the deviation probability and the individual expected utility. It can be seen that if a user deviates from the proposed best response strategy profile, he/she can only obtain a worse individual expected utility; while for the maximum social welfare strategy profile, a user may be able to earn a better payoff by unilateral deviation to another strategy, since the objective of the maximum social welfare strategy is to optimize the social welfare without consideration of the individual rationality.

The individual expected utilities using different strategies are compared in Figures 8 by varying $\bar{\lambda}_0$. Using the greedy method as the baseline, the performance of each strategy is normalized with the corresponding value of the greedy method. Since each user optimizes his/her own expected utility, the proposed best response strategy as expected performs the best among all other strategies in terms of the individual expected utility. When $\bar{\lambda}_0$ is higher, i.e., more users who are able to choose among the networks, the maximum social welfare strategy provides worse individual expected utility due to the crowdedness of users and thus the conflict between maximizing the social welfare and the individual performance. Without taking into account any information, the random strategy is inferior to all others.

In the system model, two types of users are considered: type I users that can choose among all networks, and type II users that can only choose one network. The influence of the users of type II (that can choose only one network) is the crowdedness of that particular network. When there are more type II users, we can expect the proposed method outperforms relatively less than the myopic strategy in the individual expected utility compared to in the scenario where there are fewer type II users since the crowdedness of users in that network can hardly be changed by the type I users' decisions and this may make the proposed method to choose a network similar to the myopic strategy. As a result, the overall strategy profile of the proposed method and the myopic method may be very similar to each other.

In Figure 9, the social welfare performance (the system throughput, i.e., the sum of the expected utility of each user) of different strategies is compared by varying $\bar{\lambda}_0$. Since the maximum social welfare strategy is the global maximizer among all the strategy profiles, it attains the best performance with certainty. We can see that the proposed best response strategy is able to achieve a similar performance to the maximum social welfare strategy when $\bar{\lambda}_0$ is small, i.e., when the system is less crowded. When $\bar{\lambda}_0$ is higher, the performance becomes a bit worse but it is still better than the greedy strategy and the random strategy. It is interesting that although the proposed

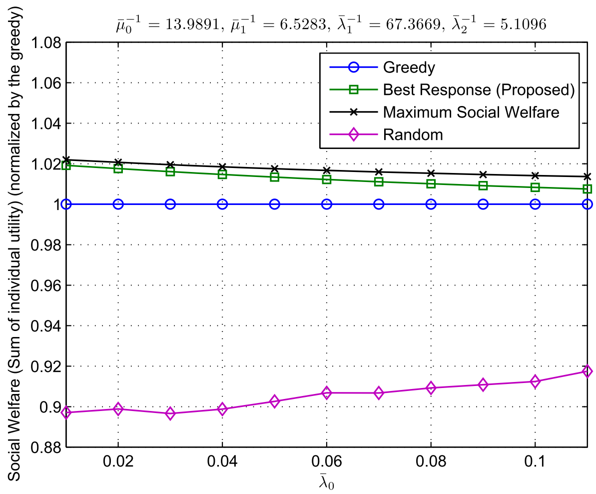


Fig. 9. Social welfare (sum expected utility) comparison in a 2-network system with different strategies including the greedy method, the proposed best response strategy, the centralized maximum social welfare strategy, and the random strategy.

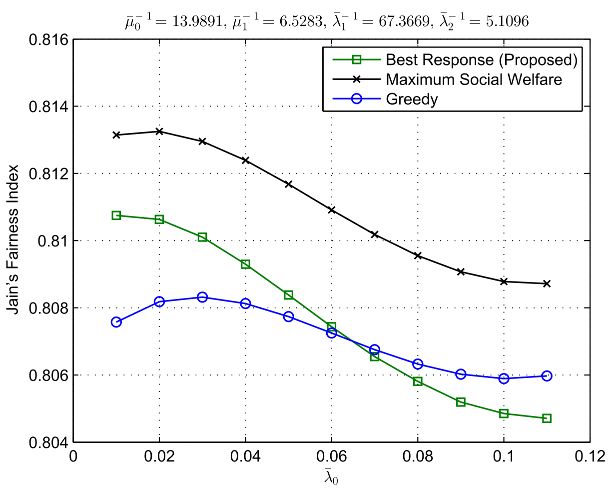


Fig. 10. Fairness comparison with different strategy including greedy method, the proposed best response strategy and the centralized maximum social welfare strategy.

best response aims to optimize each user's own expected utility by considering other users' strategies, it has a similar social welfare performance to the global optimum.

In Figure 10, the fairness of different approaches are compared in terms of Jain's fairness index [46], which is defined as

$$\mathcal{J}(\boldsymbol{\sigma}) = \frac{\left|\sum_{\mathbf{s}} \pi(\mathbf{s}) \sum_{k=1}^{K} s_k R_k(s_k)\right|^2}{\sum_{\mathbf{s}} \pi(\mathbf{s}) \left|\sum_{k=1}^{K} s_k R_k(s_k)\right|^2}, \tag{23}$$

where $\pi(\mathbf{s})$ denotes the stationary probability at system state $\mathbf{s}$. If the throughput distribution is more homogeneous, the resulting index is more close to 1. From the figure, the maximum social welfare strategy achieves the best fairness than the others. The proposed algorithm has better fairness compared to the greedy method when $\bar{\lambda}_0$ is low. When $\bar{\lambda}_0$ is higher, i.e., there are more type I users (users who can choose among the two networks), the proposed algorithm performs worse in fairness. The

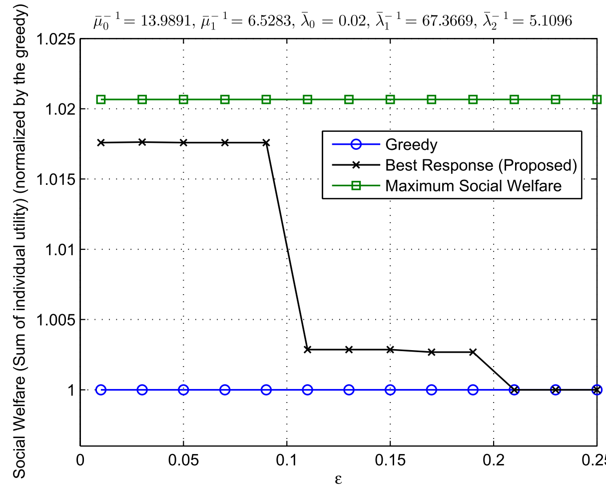


Fig. 11. The impact of ϵ on the social welfare performance of the proposed best response strategy.

relation between strategy and the distribution is quite complicated and is difficult to analyze. Intuitively, since the proposed algorithm is derived based on maximizing a user's long term utility due to the selfish nature and some states may have better inherent advantages than other states, the users at the advantageous states may be able to obtain better throughput, and consequently the distribution becomes less homogeneous. The greedy method is similar but just concerns only the immediate utility. On the other hand, the maximum social welfare strategy is defined to maximize the total achievable data rate from the entire network system and thus all the states are taken into consideration. Another observation is that the objective function of maximum social welfare strategy, $\sum_{\mathbf{s}} \pi(\mathbf{s}) \sum_k s_k R_k(s_k)$, is exactly the numerator of the Jain's fairness index in (23), so the maximization of the two objectives may be closely related.

Figure 11 demonstrates how the social welfare of the proposed best response strategy varies with respect to ϵ. We can observe that when ϵ increases, the social welfare performance of the proposed method becomes worse since a larger region of tolerance is allowed for switching between different strategies, and the resulting strategy profile is only guaranteed that the strategy at each state has an expected utility that is at most ϵ less than that of any other strategy. Figure 12 shows the impact of ϵ on the number of iterations for the strategy profile to converge using the proposed algorithm in Table I. It can be seen that in general a larger ϵ allows a smaller number of iterations to converge since the region of tolerance for changing the strategy profile is larger.

Figure 13 shows the feasible region of (C_1, C_2) in the mechanism design problem in (19), where the strategy profile $\{\sigma_{\mathbf{s}}\}$ is given as

$$\boldsymbol{\sigma} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix}, \tag{24}$$

where $[\boldsymbol{\sigma}]_{i,j} = \sigma_{(i,j)}$ denotes the strategy at state $\mathbf{s} = (i, j)$. Since the constraints are all linear in C_1 and C_2, the resulting

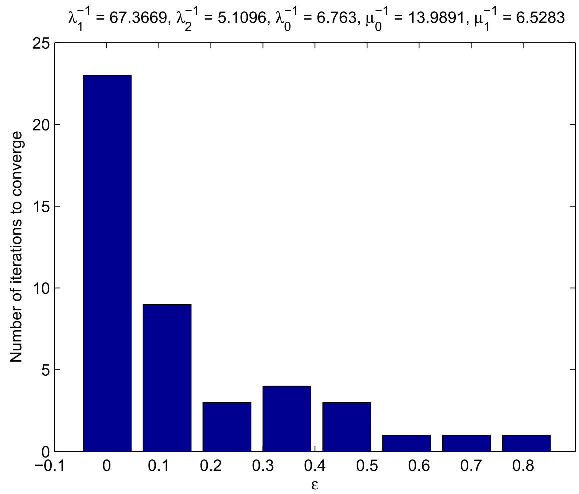


Fig. 12. The impact of ϵ on the number of iterations for the strategy profile to converge.

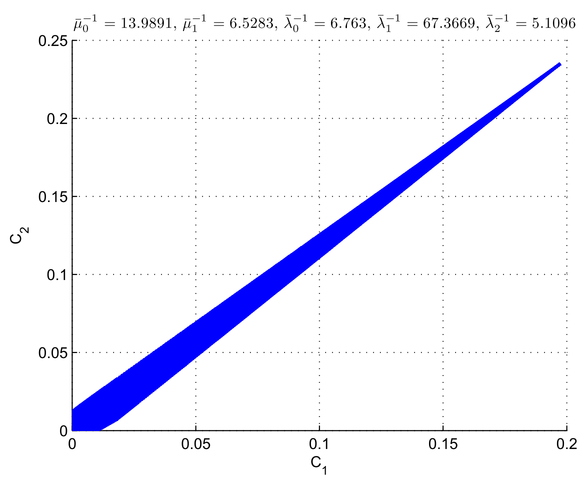


Fig. 13. The feasible region of (C_1, C_2) in the mechanism design problem $\mathcal{P}_{\text{MD}}$ in (19).

feasible region is a 2-dimensional convex region with piecewise linear boundaries. The system operator can then manage the available resource to design C_1 and C_2 such that the desired strategy profile is a best response for the users. Note that the feasible set may not always be non-empty. Thus, the mechanism design for the wireless network association can be used to check the existence of the best response strategy profile.

VI. Conclusions

In this paper, we first used the four months trace of 560 APs at Dartmouth College to validate the statistical characteristics of the user arrival process being Poisson, the session time, and the waiting time to switch network being exponential. Based on these observations, we constructed a Markov system model to investigate the relation between users' strategies and their expected utilities. It has been shown that finding the best response strategy, i.e., the approximate Nash equilibrium, requires solving a multi-dimensional Markov decison process. We proposed a modified value iteration algorithm to iteratively search for the solution.

Data-driven simulations were conducted to verify the individual rationality, i.e., unilateral deviation from the best response strategy only leads to a decrease of the individual expected utility. Compared with other strategies, the proposed best response strategy can achieve better individual expected utility while also has a similar performance in the social welfare (the sum of the individual expected utilities) to the maximum social welfare strategy.

References

[1] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: A survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008.

[2] W. H. Sandholm, "Negative externalities and evolutionary implementation," *Rev. Econ. Stud.*, vol. 72, no. 3, pp. 885–915, Jul. 2005.

[3] G. Fagiolo, "Endogenous neighborhood formation in a local coordination model with negative network externalities," *J. Econ. Dyn. Control*, vol. 29, no. 1–2, pp. 297–319, 2005.

[4] S. Corroy, L. Falconetti, and R. Mathar, "Cell association in small heterogeneous networks: Downlink sum rate and min rate maximization," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2012 pp. 888–892.

[5] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2706–2716, Jun. 2013.

[6] L. Wang and G.-S. Kuo, "Mathematical modeling for network selection in heterogeneous wireless networks—A tutorial," *IEEE Commun. Surv. Tuts.*, vol. 15, no. 1, pp. 271–292, First Quart. 2013.

[7] Q. Song and A. Jamalipour, "Network selection in an integrated wireless LAN and UMTS environment using mathematical modeling and computing techniques," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 42–48, Jun. 2005.

[8] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 2008–2017, May 2009.

[9] K. Mittal, E. M. Belding, and S. Suri, "A game-theoretic analysis of wireless access point selection by mobile users," *Comput. Commun.*, vol. 31, no. 10, pp. 2049–2062, 2008.

[10] A. Argento, M. Cesana, N. Gatti, and I. Malanchini, "A game theoretical study of access point association in wireless mesh networks," *Comput. Commun.*, vol. 35, no. 5, pp. 541–553, 2012.

[11] M. Cesana, N. Gatti, and I. Malanchini, "Game theoretic analysis of wireless access network selection: Models, inefficiency bounds, and algorithms," in *Proc. 3rd Int. Conf. Perform. Eval. Methodol. Tools*, 2008, pp. 6:1–6:10.

[12] A. Balachandran, P. Bahl, and G. Voelker, "Hot-spot congestion relief in public-area wireless networks," in *Proc. IEEE 4th Workshop Mobile Comput. Syst. Appl.*, 2002, pp. 70–80.

[13] L. Chen, "A distributed access point selection algorithm based on noregret learning for wireless access networks," in *Proc. IEEE Veh. Tech. Conf.*, May 2010, pp. 1–5.

[14] W. Shen and Q.-A. Zeng, "Cost-function-based network selection strategy in integrated wireless and mobile networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3778–3788, Nov. 2008.

[15] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and M. Chiang, "RAT selection games in HetNets," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 998–1006.

[16] A. Ahmed, L. M. Boulahia, and D. Gaïti, "Enabling vertical handover decisions in heterogeneous wireless networks: A state-of-the-art and a classification," *IEEE Commun. Surv. Tuts.*, vol. 16, no. 2, pp. 776–811, Second Quart. 2014.

[17] Z. Du, Q. Wu, P. Yang, Y. Xu, and Y.-D. Yao, "User-demand-aware wireless network selection: A localized cooperation approach," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4492–4507, Nov. 2014.

[18] K. Zhu, E. Hossain, and D. Niyato, "Pricing, spectrum sharing, and service selection in two-tier small cell networks: A hierarchical dynamic game approach," *IEEE Trans. Mobile Comput.*, vol. 13, no. 8, pp. 1843–1856, Aug. 2014.

[19] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA, USA: MIT Press, 1991.

[20] B. Wang, Y. Wu, and K. J. R. Liu, "Game theory for cognitive radio networks: An overview," *Comput. Netw.*, vol. 54, no. 14, pp. 2537–2561, Oct. 2010.

[21] W. Yu and K. J. R. Liu, "Game theoretic analysis of cooperation stimulation and security in autonomous mobile ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 5, pp. 507–521, May 2007.

[22] Z. Han, Z. Ji, and K. J. R. Liu, "Non-cooperative resource competition game by virtual referee in multi-cell OFDMA networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 6, pp. 1079–1090, Jul. 2007.

[23] Z. Ji and K. J. R. Liu, "Multi-stage pricing game for collusion-resistant dynamic spectrum allocation," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 182–191, Jan. 2008.

[24] K. J. R. Liu and B. Wang, *Cognitive Radio Networking and Security: A Game Theoretical View*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[25] C.-Y. Wang, Y. Chen, and K. J. R. Liu, "Chinese restaurant game," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 898–901, Dec. 2012.

[26] C.-Y. Wang, Y. Chen, and K. J. R. Liu, "Sequential chinese restaurant game," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 571–584, Feb. 2013.

[27] C. Jiang, Y. Chen, Y.-H. Yang, C.-Y. Wang, and K. J. R. Liu, "Dynamic Chinese restaurant game in cognitive radio networks," in *Proc. IEEE INFOCOM*, 2013, pp. 962–970.

[28] Y.-H. Yang, Y. Chen, C. Jiang, C.-Y. Wang, and K. J. R. Liu, "Wireless access network selection game with negative network externality," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5048–5060, Oct. 2013.

[29] D. Kotz, T. Henderson, I. Abyzov, and J. Yeo. (2009, Sep.). *CRAWDAD Data Set Dartmouth/Campus (v. 2009-09-09)* [Online]. Available: http://crawdad.cs.dartmouth.edu/dartmouth/campus

[30] A. Balachandran, G. M. Voelker, P. Bahl, and P. V. Rangan, "Characterizing user behavior and network performance in a public wireless LAN," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 30, no. 1, pp. 195–205, Jun. 2002.

[31] D. Schwab and R. Bunt, "Characterising the use of a campus wireless network," in *Proc. IEEE INFOCOM*, 2004, vol. 2, pp. 862–870.

[32] T. Henderson, D. Kotz, and I. Abyzov, "The changing usage of a mature campus-wide wireless network," in *Proc. ACM MobiCom*, 2004, pp. 187–201.

[33] M. Papadopouli, H. Shen, and M. Spanakis, "Modeling client arrivals at access points in wireless campus-wide networks," in *Proc. 14th IEEE Workshop Local Metrop. Area Netw.*, 2005, pp. 18–21.

[34] F. Hernández-Campos, M. Karaliopoulos, M. Papadopouli, and H. Shen, "Spatio-temporal modeling of traffic workload in a campus WLAN," in *Proc. 2nd Annu. Int. Workshop Wireless Internet*, 2006.

[35] W. Bao and B. Liang, "Insensitivity of user distribution in multicell networks under general mobility and session patterns," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6244–6254, Dec. 2013.

[36] Y.-C. Chen, J. Kurose, and D. Towsley, "A mixed queueing network model of mobility in a campus wireless network," in *Proc. IEEE INFOCOM*, 2012, pp. 2656–2660.

[37] H. Yu, D. Zheng, B. Y. Zhao, and W. Zheng, "Understanding user behavior in large-scale video-on-demand systems," in *Proc. ACM SIGOPS Oper. Syst. Rev.*, 2006, vol. 40, no. 4, pp. 333–344.

[38] M. L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming. Hoboken, NJ, USA: Wiley, 1994.

[39] D. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Belmont, MA, USA: Athena Scientific, 2007.

[40] J. Razavilar, K. J. R. Liu, and S. Marcus, "Jointly optimized bit-rate/delay control policy for wireless packet networks with fading channels," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 484–494, Mar. 2002.

[41] A. G. Barto, *Reinforcement learning: An Introduction*. Cambridge, MA, USA: MIT Press, 1998.

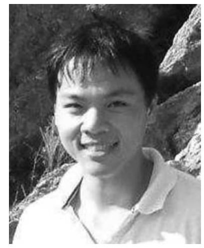
[42] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, pp. 237–285, 1996.

[43] C. J. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, no. 3–4, pp. 279–292, 1992.

[44] D. Kotz and K. Essien, "Analysis of a campus-wide wireless network," *Wireless Netw.*, vol. 11, no. 1–2, pp. 115–133, Jan. 2005.

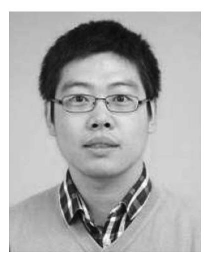
[45] M. B. Wilk and R. Gnanadesikan, "Probability plotting methods for the analysis for the analysis of data," *Biometrika*, vol. 55, no. 1, pp. 1–17, 1968.

[46] R. Jain, D.-M. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," Digital Equipment Corporation, Tech. Rep. DEC-TR-301, 1984.



Yu-Han Yang (S'06) received the B.S. degree in electrical engineering, two M.S. degrees in computer science and communication engineering, from National Taiwan University, Taipei, Taiwan, and the Ph.D. degree from the University of Maryland, College Park, MD, USA, in 2004, 2007, and 2013, respectively. Currently, he is with Google Inc., Mountain View, CA, USA. His research interests include wireless communications and signal processing. He was the recipient of Jimmy H. C. Lin Invention Award from the Department of ECE, University of Maryland, in 2013. He also received Study Abroad Scholarship from Taiwan government in 2009–2010.

Yan Chen (SM'14) received the Bachelor's degree from the University of Science and Technology of China, the M.Phil. degree from Hong Kong University of Science and Technology (HKUST), Hong Kong, and the Ph.D. degree from University of Maryland, College Park, MD, USA, in 2004, 2007, and 2011, respectively. His research interests include data science, network science, game theory, social learning and networking, and signal processing and wireless communications. He was the recipient of multiple honors and awards including Best Paper Award from IEEE GLOBECOM, in 2013, Future Faculty Fellowship and Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering in 2010 and 2011, respectively, Finalist of Dean's Doctoral Research Award from A. James Clark School of Engineering, the University of Maryland in 2011, and Chinese Government Award for outstanding students abroad in 2011.



Chunxiao Jiang (S'09–M'13) received the B.S. (Hons.) degree in information engineering from Beihang University, Beijing, China, and the Ph.D. (Hons.) degree from Tsinghua University (THU), Beijing, China, in 2008 and 2013, respectively. From 2011 to 2014, he visited the Signals and Information Group, Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA. He currently holds a Postdoctoral with the Department of Electrical Engineering, THU. His research interests include wireless communication and networking, and social networks. He was a recipient of the Best Paper Award from IEEE GLOBECOM in 2013, the Beijing Distinguished Graduated Student Award, the Chinese National Fellowship, and the Tsinghua Outstanding Distinguished Doctoral Dissertation in 2013.

K. J. Ray Liu (F'03) was a Distinguished Scholar–Teacher with the University of Maryland, College Park, MD, USA, in 2007, where he is Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of information and communications technology with recent focus on future wireless technologies, network science, and information forensics and security. He was a recipient of the 2016 IEEE Leon K. Kirchmayer Technical Field Award on graduate teaching and mentoring, the IEEE Signal Processing Society 2014 Society Award, and the IEEE Signal Processing Society 2009 Technical Achievement Award. Recognized by Thomson Reuters as a Highly Cited Researcher, he is a Fellow of AAAS. He is a Director-Elect of the IEEE Board of Director. He was President of the IEEE Signal Processing Society, where he has served as Vice President Publications and Board of Governor. He has also served as the Editor-in-Chief of the *IEEE Signal Processing Magazine*. He also received teaching and research recognitions from the University of Maryland including university-level Invention of the Year Award; and college-level Poole and Kent Senior Faculty Teaching Award, Outstanding Faculty Research Award, and Outstanding Faculty Service Award, all from A. James Clark School of Engineering.