# Dynamics of Network Selection in Heterogeneous Wireless Networks: An Evolutionary Game Approach

Dusit Niyato and Ekram Hossain, *Senior Member, IEEE*

***Abstract*—Next-generation wireless networks will integrate multiple wireless access technologies to provide seamless mobility to mobile users with high-speed wireless connectivity. This will give rise to a heterogeneous wireless access environment where network selection becomes crucial for load balancing to avoid network congestion and performance degradation. We study the dynamics of network selection in a heterogeneous wireless network using the theory of evolutionary games. The competition among groups of users in different service areas to share the limited amount of bandwidth in the available wireless access networks is formulated as a dynamic evolutionary game, and the evolutionary equilibrium is considered to be the solution to this game. We present two algorithms, namely, population evolution and reinforcement-learning algorithms for network selection. Although the network-selection algorithm based on population evolution can reach the evolutionary equilibrium faster, it requires a centralized controller to gather, process, and broadcast information about the users in the corresponding service area. In contrast, with reinforcement learning, a user can gradually learn (by interacting with the service provider) and adapt the decision on network selection to reach evolutionary equilibrium without any interaction with other users. Performance of the dynamic evolutionary game-based network-selection algorithms is empirically investigated. The accuracy of the numerical results obtained from the game model is evaluated by using simulations.**

***Index Terms*—Evolutionary equilibrium, evolutionary game theory, heterogeneous wireless access networks, Nash equilibrium, network selection, replicator dynamics.**

## I. Introduction

NEXT-GENERATION wireless networks will integrate different wireless access technologies such as the IEEE 802.16-based wireless metropolitan area networks (WMANs), cellular networks, and the IEEE 802.11-based wireless local area networks (WLANs) into a heterogeneous wireless network. Integration of these technologies will improve the performance of wireless connectivity and support seamless user mobility. While seamless mobility in a heterogeneous environment can be achieved through a vertical handoff in which a connection can be handed over among multiple wireless networks, network capacity can be improved through load balancing [1]. With multiple available wireless networks, wireless traffic load can be balanced to avoid congestion and performance degradation in any of the networks. Load balancing in a heterogeneous network can be achieved by using either a network-driven or a user-driven approach. With a network-driven approach, a centralized controller assigns network resources to the connections in a service area. However, in this approach, all available wireless networks must be tightly integrated, and large communication overhead could be incurred. Alternatively, with a user-driven load-balancing approach, network-selection algorithms are implemented at the user mobile. Such an approach may be preferred due to its low implementation complexity and low communication overhead.

We consider user-driven load balancing in a heterogeneous wireless network. A dynamic evolutionary game with multiple populations is used to analyze the dynamic behavior of the users for network selection. In particular, the game is formulated to logically model the competition among groups of users in the different service areas in which different numbers and different types of wireless technologies are available (i.e., WMAN, cellular network, and WLAN). Evolutionary equilibrium is considered as the solution to this competition. To obtain the solution, we present two algorithms, namely, population evolution and reinforcement learning algorithms for network selection. While the population evolution algorithm uses information about the users in the corresponding service area, the reinforcement learning algorithm utilizes only local knowledge obtained through learning to reach the evolutionary equilibrium. In addition, for comparison purposes, we formulate this competition as a noncooperative game for which the Nash equilibrium is considered as the solution. Extensive performance evaluation is performed, and the numerical results reveal the dynamics of network selection in a heterogeneous wireless network.

The major contributions of this paper can be summarized as follows.

- A game-theoretic approach is presented to solve the problem of network selection in the heterogeneous wireless access networks considering users with different wireless access service requirements. In particular, the theory of evolutionary games is used to investigate the dynamics of user behavior and solution in network selection.
- The solution to the network-selection problem obtained from the evolutionary game model is compared to the

Manuscript received June 16, 2007; revised November 4, 2007, February 18, 2008, and May 28, 2008. First published August 26, 2008; current version published April 22, 2009. This work was supported in part by TRLabs, Winnipeg, MB, Canada, thorough a scholarship and in part by the Natural Sciences and Engineering Research Council of Canada through a grant. The review of this paper was coordinated by Prof. B. L. Mark.

D. Niyato is with the School of Computer Engineering, Nanyang Technological University, Singapore 639798.

E. Hossain is with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada.

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/TVT.2008.2004588

Nash equilibrium solution obtained from a classical noncooperative game model.
- Both centralized and distributed algorithms are proposed to implement the proposed evolutionary game model for network selection.

The rest of this paper is organized as follows. Related works are reviewed in Section II. The system model for the heterogeneous wireless access network is described in Section III. Section IV presents an overview of evolutionary game theory. Section V presents the formulation of an evolutionary game model for the network-selection problem in a heterogeneous wireless access network. The population evolution and the reinforcement learning algorithm approaches for network selection are presented in Section VI. Section VII presents the performance evaluation results. Conclusions are stated in Section VIII.

## II. Related Work

### A. Heterogeneous Wireless Networks and Network Selection

The problem of integrating WLANs into the cellular wireless networks was investigated in the literature [2]–[4]. In [2], a hierarchical radio resource management framework was designed to support a seamless handoff between a WLAN and a cellular network. An admission control scheme for a vertical handoff in an integrated WLAN and code-division multiple access (CDMA) cellular network was proposed in [3], where an optimization problem was formulated to minimize call blocking probability while throughput and packet delay performances are maintained at the target level. In [4], a performance analysis model for an integrated cellular network and a WLAN was proposed. However, all these works ignored the issue of competition among users to access different types of wireless networks, and hence, the dynamics of network selection.

### B. Applications of Game Theory in Wired and Wireless Networking

Game theory was applied to solve the radio resource management problem in wireless networks. In [5], the admission and rate control problem for CDMA systems was formulated as a noncooperative game. The formulation considered the choice of a user to churn from one service provider to another. The decision on whether a new user can be admitted or not and the allocated transmission rate were determined from the Nash equilibrium. A similar approach was used in [6] to solve the admission control problem in WLANs. An evolutionary game was used to model the network routing problem in [7]. In this paper, users were modeled as a population who can choose the route of data flow to the destination. However, all these works considered only a single wireless access network.

### C. Pricing in Heterogeneous Wireless Networks

In a heterogeneous wireless access network, pricing can be used as a mechanism for resource allocation, admission control, and network selection. Three different approaches, namely, auction-based [8], optimization-based [9], and demand/supply-based [10] approaches, were applied for pricing in a heterogeneous network. In [8], mobile users bid for the radio resources from multiple radio access technologies by informing the service provider of their price and QoS requirements. Then, the service provider makes a decision on resource allocation in different wireless access networks to maximize the revenue. In contrast to the network-centric viewpoint, we take a user-centric viewpoint where users can independently select the access network, and the network-selection problem is modeled as an evolutionary game among groups of users.

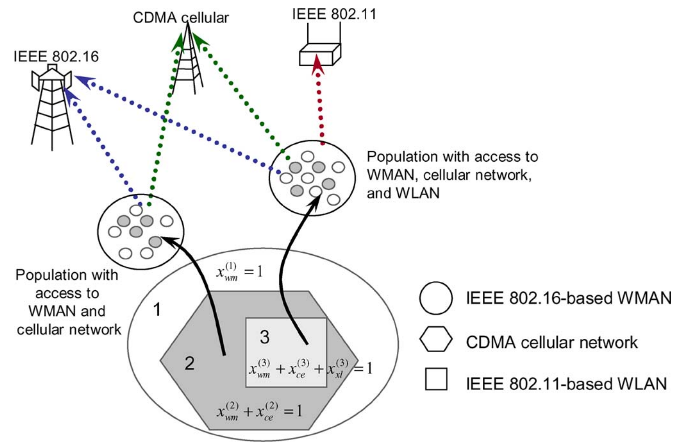


Fig. 1. Service areas under consideration in a heterogeneous wireless-access environment.

## III. System Model and Assumptions

### A. Wireless Access Networks

We consider a heterogeneous wireless access environment consisting of IEEE 802.16-based WMANs, CDMA cellular networks, and IEEE 802.11-based WLANs, as shown in Fig. 1. An IEEE 802.16-based WMAN is based on WirelessMAN-SC, which is a single-carrier interface [11]. We consider a wideband CDMA-based cellular wireless access system [12]. We consider an IEEE 802.11 WLAN radio interface with a distributed reservation-based medium-access-control (MAC) protocol, namely, *early backoff announcement (EBA)* [13], which is an enhanced version of distributed coordination function MAC and is able to reserve a certain amount of bandwidth for a particular user.

A mobile with multiple radio transceivers is able to connect to the different radio access networks. We consider a geographic region that is entirely covered by a WMAN base station (i.e., one IEEE 802.16 cell), partly covered by the cellular base station, and partly by the WLAN access point, as shown in Fig. 1. Users in the different service areas in this region have an access to different types and different numbers of wireless networks. In the system model, we do not consider the mobility of the users explicitly.

### B. Pricing Model and Service Class

We assume that the service providers use a linear pricing model. In particular, the price per connection/user is a linear function of the total number of connections/users in the corresponding coverage area (or cell). When a coverage area becomes congested, service providers in that area charge a higher price to gain more revenue. Furthermore, such a nondecreasing pricing function can avoid congestion in a wireless network [14]. A congestion pricing model with a similar spirit was used in [15]. If there are multiple classes of services, and if users in each service class have different preference on QoS performance, the framework for network selection developed in this paper can be used by users in each service class. In such a scenario, pricing could be different for different classes of users.

## IV. Evolutionary Game—An Overview

### A. Population, Evolution, and Solutions

In noncooperative game theory, a game can be described by a set of players, a set of strategies associated with each player, the payoff of each player given a chosen strategy, and the solution of all players. However, an evolutionary game extends the formulation of a noncooperative game by including the concept of population. This population is a group of individuals (i.e., players) in which the number of individuals can be finite or infinite. Further, in an evolutionary game model, there could be a single or multiple populations. The individuals from one population may choose strategies against individuals in another population. An evolutionary game defines a foundation to obtain the equilibrium solution for the game of the populations.

### B. Motivation of Using Evolutionary Game

Although evolutionary game theory was originally developed for biology [16], [17], its applications in other fields are growing due to the following reasons.

1) *Solution refinement*: In traditional game theory, the Nash equilibrium is the most popular solution. It ensures that a player cannot improve its payoff if none of the other players in the game deviates from the solution. However, when the solution to a noncooperative game has multiple Nash equilibria, a refined solution is required. Evolutionary equilibrium, which is based on the theory of evolutionary game theory, provides such a refined solution, and it ensures stability (i.e., group of players will not change their chosen strategies over time).
2) *Bounded rationality*: Unlike a classical single-play noncooperative game, in which all of the players make decisions that lead immediately to the desired solution, an evolutionary game involves players slowly changing their strategies to achieve the solution eventually.
3) *Dynamics in the game model*: An evolutionary game can explicitly capture the dynamics of interaction among the players in a population. In an evolutionary game, a player can observe the behavior of other players, learn from the observations, and make the best decision based on its knowledge. In addition, with replicator dynamics, the state of the game can be determined at a particular point in time, which is useful for investigating the trajectory (i.e., trend) of the strategies of the players while adapting their behavior to reach the solution.

### C. Replicator Dynamics and Evolutionary Equilibrium

In a dynamic evolutionary game, an individual from a population (i.e., a player in the game), who is able to reproduce (i.e., replicate) itself through the process of mutation and selection, is called a replicator. In this case, a replicator with a higher payoff can reproduce itself faster. When the reproduction process takes place over time, this can be modeled by using a set of ordinary differential equations called replicator dynamics. This replicator dynamics is important for an evolutionary game since it can provide information about the population (e.g., proportion of individuals who choose different strategies), given a particular point in time. This replicator dynamics is also useful to investigate the speed of convergence of strategy adaptation to reach a solution to the game.

In replicator dynamics, it is assumed that an individual chooses pure strategy $i$ from a finite set of strategies where the total number of available strategies in this set is $I$. Let $n_i$ denote the number of individuals choosing strategy $i$, and let the total population size be $N = \sum_{i=1}^{I} n_i$. The proportion of individuals choosing strategy $i$ is $x_i = n_i/N$, and it is referred to as the population share. The population state can be denoted by the vector $\mathbf{x} = [\, x_1 \quad \cdots \quad x_i \quad \cdots \quad x_I \,]$. The replicator dynamics can be defined as follows:

$$\dot{x}_i(t) = x_i(t)\left(\pi_i(t) - \overline{\pi}(t)\right) \tag{1}$$

where $\pi_i(t)$ is the payoff of the individuals choosing strategy $i$, and $\overline{\pi}(t)$ is the average payoff of the entire population. Based on the replicator dynamics, the evolutionary equilibrium is defined as the set of fixed points of the replicator dynamics that are stable. This evolutionary equilibrium is a desirable solution to the evolutionary game since when the population of players evolves over time (i.e., based on replicator dynamics), it will converge to the evolutionary equilibrium. Furthermore, at this evolutionary equilibrium, none of the individuals wants to change its strategy since its payoff is equal to the average payoff of the population.

## V. Evolutionary Game Formulation of the Network Selection Problem in Heterogeneous Wireless Networks

We model the network-selection problem by using a dynamic evolutionary game. In particular, given the capacity allocated to a certain class of users, users at the different (i.e., geographically separated) service areas compete to share the available bandwidth from different wireless networks. Here, an evolutionary game is used since it can capture the dynamics of network selection (i.e., strategy adaptation) based on the available information and bounded rationality of the users. That is, a user slowly changes the network (i.e., evolves) if its observed payoff is less than the average payoff of all users in the same group (i.e., users in the same service class in the same area). For this evolutionary game, the evolutionary equilibrium is considered as the solution, which ensures that all users in the same group receive identical payoff.

### A. Formulation of the Evolutionary Game

The evolutionary game for the network-selection problem in a heterogeneous wireless network can be described as follows.

- *Players*: For a particular service class, each user in each service area who can choose among multiple wireless access networks is a player of the game. For example, in Fig. 1, considering a particular service class, the players are the users in that service class in areas 2 and 3 who compete for bandwidth from WMAN, cellular network, and WLAN. Note that the users in area 1 are not involved in the game since the WMAN is the only wireless access network available to these users.
- *Population*: For a particular service class, the population in this evolutionary game refers to the set of users in a service area. We assume that the population corresponding to a service area is finite. In Fig. 1, users in area 2 form a population, and users in area 3 form another population.
- *Strategy*: The strategy of each user corresponds to the selection of a wireless access network. In Fig. 1, the set of strategies for the players in area 2 is $\{\mathrm{wm}, \mathrm{ce}\}$, while that for the players in area 3 is $\{\mathrm{wm}, \mathrm{ce}, \mathrm{wl}\}$.[1]
- *Payoff*: The payoff of a player is determined by his net utility.

To obtain the payoff, we use a concave utility function to quantify a user's satisfaction on achievable throughput. For a particular service class, the net utility of a user in area $a$ choosing network $i$ can be expressed as $\mathcal{U}(\mathcal{T}_i(n)) - \mathcal{P}_i(n)$, where $n$ is the total number of users in area $a$ choosing network $i$, $\mathcal{P}_i(n)$ is the pricing function, $\mathcal{T}_i(n)$ is the throughput of the user, and $\mathcal{U}$ denotes the utility function [18]. The throughput of each user can be computed from allocated capacity $C_i^{(a')}$ for the corresponding service class. For brevity, from now on, we will not refer to the service class, and when we refer to the users, they are assumed to be from a particular service class.

We assume that all users selecting network $i$ are allocated equal amounts of bandwidth from network $i$. Therefore, the net utility function can be defined as follows: $\pi_i^{(a')} = \mathcal{U}(C_i^{(a')}/(\sum_{a\in\mathbb{A}^{(a')}} n_i^{(a)})) - p_i \sum_{a\in\mathbb{A}^{(a')}} n_i^{(a)}$, where $n_i^{(a)}$ is the number of users in area $a$ choosing network $i$, $C_i^{(a')}$ is the network capacity in area $a'$ (i.e., total capacity associated with WMAN and/or cellular base station and/or WLAN access point), $p_i$ is the coefficient of linear pricing function used by network $i$ to charge a user, and $\mathbb{A}^{(a')}$ is the set of subareas in coverage area $a'$. Note that the coverage area of an access network may include multiple coverage areas of another access network (Fig. 1). For the WMAN coverage area in Fig. 1, this set can be defined as $\mathbb{A}^{(a')} = \{1, 2, 3\}$ since areas 1, 2, and 3 are in the coverage area of the WMAN base station.

[1]Throughout this paper, we use wm, ce, and wl to denote WMAN, cellular network, and WLAN, respectively. These abbreviations are applied to the variables associated with these wireless access networks.

Let $N^{(a)}$ denote the total number of users in area $a$ and $x_i^{(a)}$ the proportion of users choosing network $i$. The net utility can be expressed as follows:

$$\pi_i^{(a')}(\mathbf{x}) = \mathcal{U}\left(\frac{C_i^{(a')}}{\sum_{a\in\mathbb{A}^{(a')}} N^{(a)} x_i^{(a)}}\right) - p_i \sum_{a\in\mathbb{A}^{(a')}} N^{(a)} x_i^{(a)} \tag{2}$$

where $\mathbf{x}$ denotes the vector of the proportion of users choosing different networks in all areas. For the service areas in Fig. 1, the net utility of users in the coverage area of WMAN can be obtained from

$$\pi_{\mathrm{wm}}^{(1)}(\mathbf{x}) = \mathcal{U}\left(\frac{C_{\mathrm{wm}}^{(1)}}{n_{\mathrm{wm}}}\right) - p_{\mathrm{wm}} n_{\mathrm{wm}} \tag{3}$$

where $n_{\mathrm{wm}} = N^{(1)} + N^{(2)} x_{\mathrm{wm}}^{(2)} + N^{(3)} x_{\mathrm{wm}}^{(3)}$, and $\mathbf{x} = [\, x_{\mathrm{wm}}^{(2)} \; x_{\mathrm{wm}}^{(3)} \; x_{\mathrm{ce}}^{(3)} \,]$. Note that $\pi_{\mathrm{ce}}^{(2)}(\mathbf{x})$ and $\pi_{\mathrm{wl}}^{(3)}(\mathbf{x})$ can be obtained in a similar way.

### B. Replicator Dynamics and Evolutionary Equilibrium of Network Selection

We consider an evolutionary game of network selection in a heterogeneous wireless network where the group of users in area $a$ can choose among the available wireless access networks (i.e., select strategy $i$). The game is repeated, and in each period (i.e., in each generation), a user observes the net utility (i.e., payoff) of other users in the same area. Then, in the next period, the user adopts a strategy that gives a higher payoff. The speed of the user in observing and adapting the network selection is controlled by parameter $\sigma > 0$.

For a small period of time, the rate of strategy change is governed by the replicator dynamics, which is defined as follows:

$$\dot{x}_i^{(a)} = \sigma x_i^{(a)} \left(\pi_i^{(a)}(\mathbf{x}) - \overline{\pi}^{(a)}(\mathbf{x})\right) \tag{4}$$

where $\sigma$ is the gain for the rate of strategy adaptation. The average payoff of the users in area $a$ is computed from $\overline{\pi}^{(a)}(\mathbf{x}) = \sum_i x_i^{(a)} \pi_i^{(a)}(\mathbf{x})$. Based on this replicator dynamics of the users in area $a$, the number of users choosing network $i$ increases if their payoff is above the average payoff. It is impossible for a user to choose network $k$, which provides a lower payoff than the current payoff. This replicator dynamics satisfies the condition of $\sum_i \dot{x}_i^{(a)} = 0$. Therefore, if $\sum_i x_i^{(a)}(0) = 1$, then $\sum_i x_i^{(a)}(t) = 1 \;\forall t$, where $x_i^{(a)}(t)$ denotes the proportion of users in area $a$ choosing network $i$ at time $t$.

For the service areas shown in Fig. 1, the replicator dynamics can be expressed as follows:

$$\dot{x}_{\mathrm{wm}}^{(2)} = \sigma\left(x_{\mathrm{wm}}^{(2)} - \left(x_{\mathrm{wm}}^{(2)}\right)^2\right)\left(\pi_{\mathrm{wm}}^{(2)}(\mathbf{x}) - \pi_{\mathrm{ce}}^{(2)}(\mathbf{x})\right) \tag{5}$$

$$\begin{aligned}\dot{x}_{\mathrm{wm}}^{(3)} = \sigma x_{\mathrm{wm}}^{(3)} \Big(&\pi_{\mathrm{wm}}^{(3)}(\mathbf{x}) - x_{\mathrm{wm}}^{(3)}\pi_{\mathrm{wm}}^{(3)}(\mathbf{x}) - x_{\mathrm{ce}}^{(3)}\pi_{\mathrm{ce}}^{(3)}(\mathbf{x}) \\ &- \left(1 - x_{\mathrm{wm}}^{(3)} - x_{\mathrm{ce}}^{(3)}\right)\pi_{\mathrm{wl}}^{(3)}(\mathbf{x})\Big)\end{aligned} \tag{6}$$

$$\dot{x}_{\mathrm{ce}}^{(3)} = \sigma x_{\mathrm{ce}}^{(3)} \left( \pi_{\mathrm{ce}}^{(3)}(\mathbf{x}) - x_{\mathrm{wm}}^{(3)} \pi_{\mathrm{wm}}^{(3)}(\mathbf{x}) - x_{\mathrm{ce}}^{(3)} \pi_{\mathrm{ce}}^{(3)}(\mathbf{x}) - \left(1 - x_{\mathrm{wm}}^{(3)} - x_{\mathrm{ce}}^{(3)}\right) \pi_{\mathrm{wl}}^{(3)}(\mathbf{x}) \right). \quad (7)$$

### C. Evolutionary Equilibrium and Stability Analysis

We consider the evolutionary equilibrium as the solution to this network-selection game. An evolutionary equilibrium is a fixed point of the replicator dynamics. At this fixed point, which can be obtained numerically, payoffs of all users in area $a$ are identical. In other words, since the rate of strategy adaptation is zero (i.e., $\dot{x}_i^{(a)} = 0$), there is no user who deviates to gain a higher payoff.

To evaluate the stability at the fixed point $x_i^{(a)*}$, which is obtained by solving $\dot{x}_i^{(a)} = 0$, the eigenvalues of the Jacobian matrix corresponding to the replicator dynamics need to be evaluated. The fixed point is stable if all eigenvalues have a negative real part [19].

### D. Delay in Replicator Dynamics

In an actual heterogeneous wireless network, at the time when a user makes the decision on network selection, up-to-date information about a population (i.e., proportion of users choosing different strategies, $x_i^{(a)}$) may not be available. Therefore, a user must rely on historical information, which, again, may be delayed for a certain period. This delay can occur due to the communication latency among users or latency incurred at a centralized controller (e.g., a base station) to collect and feed back payoff information from/to every user. In this case, we assume that the network-selection decision that a user makes at time $t$ is based on the population information at time $t - \tau$ (i.e., delay for $\tau$ units of time). In this case, the replicator dynamics can be modified as follows:

$$\dot{x}_i^{(a)}(t) = \sigma x_i^{(a)}(t-\tau) \left( \pi_i^{(a)} \left(\mathbf{x}(t-\tau)\right) - \overline{\pi}^{(a)} \left(\mathbf{x}(t-\tau)\right) \right)$$

which is a delay differential equation. In order to obtain the solution to this delay differential equation, information available at time $t < 0$ needs to be defined. In this case, we assume that the user has information at time $t = 0$ that is used to compute the solution for $t < \tau$.

To obtain the solution to the differential equation corresponding to the replicator dynamics, we apply the Runge–Kutta *method* [23]. Note that the stability of this replicator dynamics with information delay can be analyzed using the Lyapunov method [21].

### E. Nash Equilibrium

An alternative of the evolutionary equilibrium is the Nash equilibrium. For this, a noncooperative game is formulated among the groups of users in different service areas. In this scenario, users in the same area collaborate with each other (i.e., form a group) to compete for the bandwidth with other groups of users in other areas. *This noncooperative game formulation is based on the group behavior of the users for network selection, while the evolutionary game formulation is based on the individual behavior of the users.*

A player of this noncooperative game corresponds to a group of users in area $a$. A strategy here corresponds to the proportion of users choosing network $i$, which is denoted by $x_i^{(a)}$. The payoff of a player is the total net utility from all users in the group (i.e., users in the same area). In particular, the payoff of a group of users in area $a$ is the total net utility of all users in the group choosing all different networks and can be expressed as follows:

$$\pi^{(a)}(\mathbf{x}^{(a)}, \mathbf{x}^{(-a)}) = \sum_i \mathcal{U} \left( \frac{C_i^{(a)} N^{(a)} x_i^{(a)}}{\sum_{a' \in \mathbb{A}^{(a)}} N^{(a')} x_i^{(a')}} \right) - N^{(a)} x_i^{(a)} p_i \sum_{a' \in \mathbb{A}^{(a)}} N^{(a')} x_i^{(a')} \quad (8)$$

where $\mathbf{x}^{(a)}$ denotes a vector of the proportion of users choosing different networks in area $a$, and $\mathbf{x}^{(-a)}$ is a vector of the proportion of users in all areas, except area $a$. We refer to $\mathbf{x}^{(a)}$ as the strategy of player $a$ and $\mathbf{x}^{(-a)}$ as the strategy of other players.

For the service areas shown in Fig. 1, the payoff of players in areas 2 and 3 can be expressed as follows:

$$\begin{aligned}
&\pi^{(2)} \left( \left[x_{\mathrm{wm}}^{(2)}\right], \left[x_{\mathrm{wm}}^{(3)}, x_{\mathrm{ce}}^{(3)}\right] \right) \\
&= \mathcal{U} \left( \frac{C_{\mathrm{wm}}^{(1)} N^{(2)} x_{\mathrm{wm}}^{(2)}}{n_{\mathrm{wm}}} \right) - N^{(2)} x_{\mathrm{wm}}^{(2)} p_{\mathrm{wm}}(n_{\mathrm{wm}}) \\
&\quad + \mathcal{U} \left( \frac{C_{\mathrm{ce}}^{(2)} N^{(2)} \left(1 - x_{\mathrm{wm}}^{(2)}\right)}{n_{\mathrm{ce}}} \right) - N^{(2)} \left(1 - x_{\mathrm{wm}}^{(2)}\right) p_{\mathrm{ce}}(n_{\mathrm{ce}}) \quad (9)
\end{aligned}$$

$$\begin{aligned}
&\pi^{(3)} \left( \left[x_{\mathrm{wm}}^{(3)}, x_{\mathrm{ce}}^{(3)}\right], \left[x_{\mathrm{wm}}^{(2)}\right] \right) \\
&= \mathcal{U} \left( \frac{C_{\mathrm{wm}}^{(1)} N^{(3)} x_{\mathrm{wm}}^{(3)}}{n_{\mathrm{wm}}} \right) - N^{(3)} x_{\mathrm{wm}}^{(3)} p_{\mathrm{wm}}(n_{\mathrm{wm}}) \\
&\quad + \mathcal{U} \left( \frac{C_{\mathrm{ce}}^{(2)} N^{(3)} x_{\mathrm{ce}}^{(3)}}{n_{\mathrm{ce}}} \right) - N^{(3)} x_{\mathrm{ce}}^{(3)} p_{\mathrm{ce}}(n_{\mathrm{ce}}) \\
&\quad + \mathcal{U} \left( C_{\mathrm{wl}}^{(3)} \right) - N^{(3)} \left(1 - x_{\mathrm{wm}}^{(3)} - x_{\mathrm{wl}}^{(3)}\right) p_{\mathrm{wl}}(n_{\mathrm{wl}})) \quad (10)
\end{aligned}$$

where $n_{\mathrm{wm}} = N^{(1)} + N^{(2)} x_{\mathrm{wm}}^{(2)} + N^{(3)} x_{\mathrm{wm}}^{(3)}$, $n_{\mathrm{ce}} = N^{(2)}(1 - x_{\mathrm{wm}}^{(2)}) + N^{(3)} x_{\mathrm{ce}}^{(3)}$, and $n_{\mathrm{wl}} = N^{(3)}(1 - x_{\mathrm{wm}}^{(3)} - x_{\mathrm{ce}}^{(3)})$.

For the competition on network selection, the *Nash equilibrium* gives a strategy profile (list of strategies, one for each player) with the property that no player can increase his payoff by choosing a different action, given other players' actions [22]. In this case, the Nash equilibrium is obtained by using the best response function of each of the players, which denotes the best strategy of a player given other players' strategies. In particular, the best response function of a group of users is obtained based on the maximization of total net utility given the strategies of other groups of users. This best response function can be defined as follows:

$$\mathcal{B}^{(a)}(\mathbf{x}^{(-a)}) = \arg\max_{\mathbf{x}^{(a)}} \pi^{(a)}(\mathbf{x}^{(a)}, \mathbf{x}^{(-a)}). \tag{11}$$

The vector $\mathbf{x}^{(a)*}$ denotes the Nash equilibrium of this game if and only if

$$\mathbf{x}^{(a)*} = \mathcal{B}^{(a)}(\mathbf{x}^{(-a)*}) \qquad \forall a \tag{12}$$

where $\mathbf{x}^{(-a)*}$ denotes the vector of best responses of all players, except player $a$. For the service areas shown in Fig. 1, the Nash equilibrium can be expressed as follows:

$$\begin{aligned}\mathbf{x}^{(2)*} &= \left[x_{\mathrm{wm}}^{(2)*}\right] = \mathcal{B}^{(2)}\left(\left[x_{\mathrm{wm}}^{(3)*}, x_{\mathrm{ce}}^{(3)*}\right]\right)\\ \mathbf{x}^{(3)*} &= \left[x_{\mathrm{wm}}^{(3)*}, x_{\mathrm{ce}}^{(3)*}\right] = \mathcal{B}^{(3)}\left(\left[x_{\mathrm{wm}}^{(2)*}\right]\right).\end{aligned} \tag{13}$$

## VI. Implementations of the Network Selection Algorithm

We present two approaches for dynamic evolutionary game-based network selection by each individual user in a heterogeneous wireless network. The first approach is based on population evolution in which payoff information of the users in a particular area is maintained by a centralized controller (e.g., base station). The second approach is based on reinforcement learning, in which each user tries different networks, observes the size of the allocated bandwidth and price from the chosen network, and changes the network selection if necessary.

### A. Population Evolution Approach

In this approach, there is a centralized controller to maintain payoff information of all users from the same area. The network-selection decision of each user is based on its current payoff and the average payoff of all users in the same area. This network-selection algorithm can be described as follows.

1) For all users, network $i$ is randomly chosen (i.e., $i \in \{\mathrm{wm}, \mathrm{ce}, \mathrm{wl}\}$).
2) **loop**
3) A user computes payoff $\pi_i^{(a)}$ from the size of allocated bandwidth and price by using (2). This payoff information is sent to the centralized controller.
4) The centralized controller computes average payoff $\overline{\pi}^{(a)} = (\sum_u \pi_i^{(a)})/(N^{(a)})$ for the users and broadcasts it back to the users. {$N^{(a)}$ is the total number of users in area $a$.}
5) **if** $\pi_i^{(a)} < \overline{\pi}^{(a)}$ **then**
6) **if** $rand() < (\overline{\pi}^{(a)} - \pi_i^{(a)})/(\overline{\pi}^{(a)})$ **then**
7) Choose network $j$, where $j \neq i$ and $\pi_j^{(a)} > \pi_i^{(a)}$.
8) **end if**
9) **end if**
10) **end loop** for all users in all groups

### B. Reinforcement Learning Approach

In an actual heterogeneous wireless access network, a centralized controller (as required by the above-described population evolution approach) may not be available. Therefore, each user has to learn and adapt its network-selection decision independently. In this case, a $Q$-learning approach [24], [25] which is a type of reinforcement learning (i.e., learning by interaction), is applied. With this ability to learn, complete payoff information of other users in the same or different areas is no longer required for network selection. In this algorithm, $Q$-value (i.e., $Q_i(t)$ is used to maintain the knowledge about each network, and the decision can be made based on this knowledge. The network-selection algorithm can be described as follows.

1) $Q_i(0) = 0$ {initialize $Q$-value associated with network $i$ for all users in all groups }
2) **loop**
3) **if** $rand() \leq \gamma$ **then**
4) Randomly choose network $i$ {Exploration step}.
5) **else**
6) Choose network $i^* = \arg\max_i Q_i(k)$ {Exploitation step}.
7) **end if**
8) User computes payoff $\pi_i^{(a)}$.
9) Update $Q_i(k+1) = (1-\lambda)Q_i(k) + \lambda(\pi_i^{(a)} + \beta \max_i Q_i(k))$.
10) **endloop** for all users in all groups

In this network-selection algorithm, a user performs the exploration step with probability $\gamma$, and $\lambda$ denotes the learning rate that is used to control the speed of adjustment of the $Q$-value. A new $Q$-value (i.e., $Q_i(k+1)$, which is the expected payoff for the future iterations), is obtained based on the previous value (i.e., $Q_i(k)$) along with the new observed payoff (i.e., $\pi_i^{(a)}$). Here, the new observed payoff is biased by the outcome of choosing the best action based on the available knowledge (i.e., $\max_i Q_i(k)$).

## VII. Performance Evaluation

### A. Parameter Setting

We consider a heterogeneous wireless network with the service areas shown in Fig. 1. For the IEEE 802.11 WLAN, the channel rate is 11 Mb/s, and the maximum saturation throughput achieved through EBA is assumed to be 7 Mb/s [13]. The total transmission rate in each CDMA cell is 2 Mb/s. For the IEEE 802.16-based wireless access, the transmission rate is 10 Mb/s in a single cell.

For performance evaluation, we first consider a system with a single class of users, and then, consider a system with two classes of users as well. The number of users in each area is $N^{(1)} = 10$, $N^{(2)} = 10$, and $N^{(3)} = 30$. We assume that $p_i = 0.01$. A linear utility function, namely, $\mathcal{U}(b) = u_i b$, is assumed where $b$ is the allocated bandwidth, and $u_i = 1$. For the replicator dynamics, we set $\sigma = 1$ [in (4)]. For the network-selection

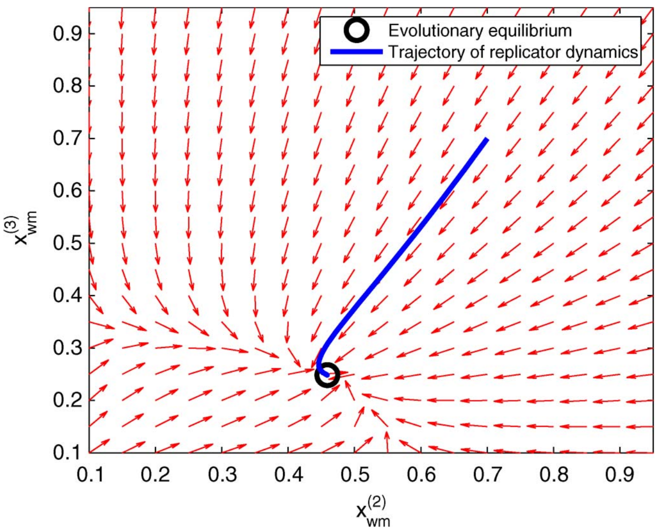


Fig. 2. Phase plane of replicator dynamics.

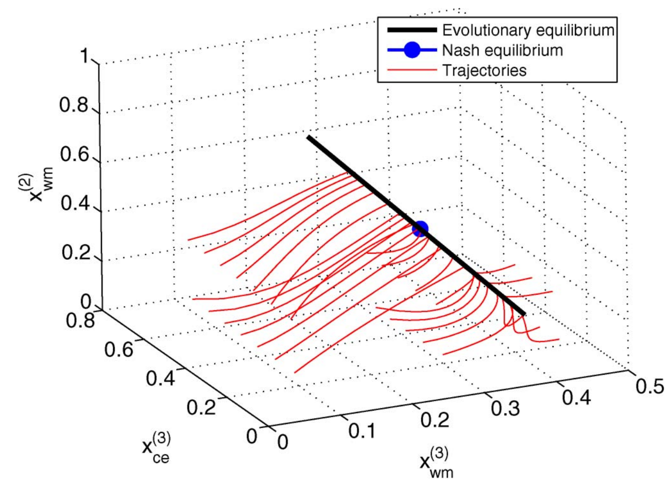


Fig. 3. Evolutionary equilibria and Nash equilibrium corresponding to the service areas shown in Fig. 1.

algorithm based on reinforcement learning, we assume that $\gamma = 0.1$, $\lambda = 0.1$, and $\beta = 0.2$. The initial proportion of users choosing each network (i.e., $x_i^{(a)}(0)$) is varied.

### *B. Numerical Results*

*1) Dynamic Behavior of User Population:* We first investigate the phase plane of the replicator dynamics in Fig. 2. In this case, we assume that none of the users in area 3 chooses the cellular network (i.e., $x_{\mathrm{ce}}^{(3)} = 0$). The proportions of users in areas 2 and 3 choosing WMAN (i.e., $x_{\mathrm{wm}}^{(2)}$ and $x_{\mathrm{wm}}^{(3)}$) are plotted. This phase plane shows the direction of the adaptation in network selection to the evolutionary equilibrium. For example, given an initial point $x_{\mathrm{wm}}^{(2)}(0) = x_{\mathrm{wm}}^{(3)}(0) = 0.7$, the trajectory (i.e., the thick solid line) follows the arrows to reach the equilibrium.

The basin of attraction consists of those proportion vectors for which the evolutionary equilibrium will be reached in the limit as $t \to \infty$. In this case, the basin of attraction, which is obtained from the phase plane, is the entire feasible region (i.e., $0 < x_{\mathrm{wm}}^{(2)}$ and $x_{\mathrm{wm}}^{(3)} < 1$).

*2) Evolutionary Equilibrium and Nash Equilibrium:* Fig. 3 shows the evolutionary equilibria corresponding to the service areas shown in Fig. 1. Here, the evolutionary equilibrium is given by a line in the 3-D space. In particular, there are a number of evolutionary equilibria that are stable (as observable from the many trajectories starting at different points). All of these equilibria provide identical average net utility to the users in all areas. This can be easily interpreted based on the following example. When most of the users in area 2 select WMAN (i.e., $x_{\mathrm{wm}}^{(2)} \to 1$), the individual net utility of users selecting WMAN in area 3 decreases (i.e., capacity of WMAN is shared by more users). Consequently, users in area 3 deviate to choose the cellular network or the WLAN, whichever is less congested. Due to this adaptation in network selection, the same average net utility can be maintained given different values of $x_{\mathrm{wm}}^{(2)}$ for multiple evolutionary equilibria. We observe that the slope of the line indicating the evolutionary equilibrium points is affected by the total number of users in each service area. However, the capacities of the access networks and the pricing coefficients affect only the location of the evolutionary equilibrium.

Fig. 3 also shows the Nash equilibrium, which is, in this case, a single point located on one of the evolutionary equilibria. This indicates that while there are an infinite number of solutions that the users in the same area are satisfied with (i.e., at the evolutionary equilibrium, all users in the same area have equal net utility), there is only a single solution such that none of the group of users in different areas can obtain a higher total net utility by adapting the proportions of users choosing the different available networks. Therefore, *the Nash equilibrium in this case can be considered as a refinement (i.e., a special case) of evolutionary equilibrium that can be achieved based a group behavior rather than an individual behavior as in an evolutionary game*. Note, however, that maintenance of this group behavior will involve some communication signaling among the group members (e.g., message from a user to join/leave the group).

*3) Evolutionary Equilibrium Under Different Utility Functions:* We investigate the evolutionary equilibrium obtained for a logarithmic utility function [18], namely, $\mathcal{U}(b) = \log(1 + b)$. We observe that the evolutionary equilibrium in this case is given by a straight line, and its location is affected by the utility function. Similar results are expected for other types of concave utility functions.

We consider a system with two classes of users. The utility functions for the users in classes 1 and 2 are linear and logarithmic, which correspond to constant-bit-rate (CBR) and best-effort traffic, respectively. The number of users for each class of users in each area is $N_1^{(1)} = N_2^{(1)} = 5$, $N_1^{(2)} = N_2^{(2)} = 5$, and $N_1^{(3)} = N_2^{(3)} = 15$. The pricing coefficients are $p_{i,1} = 0.01$ and $p_{i,2} = 0.0075$. We vary the capacity share $\phi_2$ for class 2 users, i.e., $C_2^{(a)} = \phi_2 C^{(a)}$ and $C_1^{(a)} = (1 - \phi_2) C^{(a)}$, where $C^{(a)}$ is the total capacity of the access networks in coverage area $a$. The variation in net utility is shown in Fig. 4. When the capacity share for class 2 users increases, the net utility of the users in this class increases.

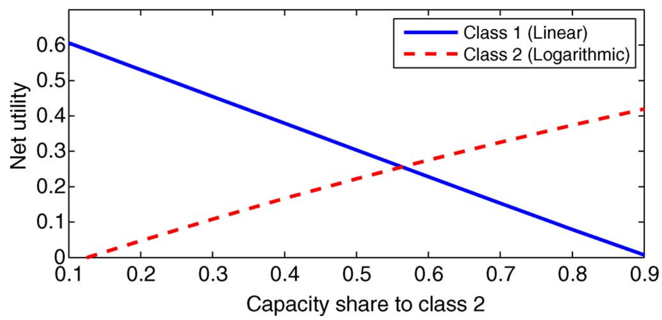


Fig. 4. Net utility of two classes of users under different capacity shares.

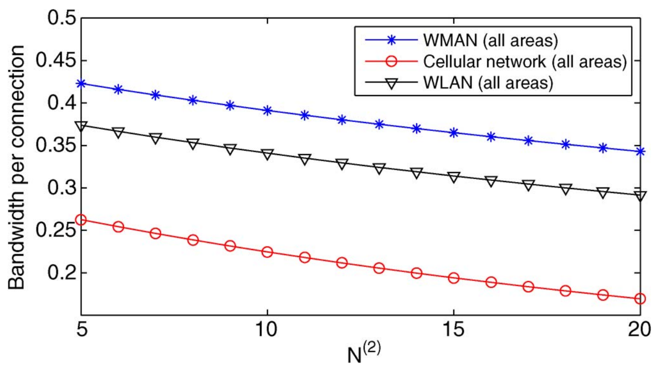


Fig. 5. Allocated bandwidth to a connection under different numbers of users.

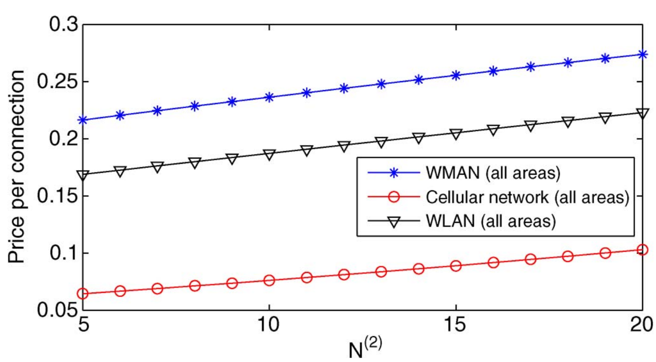


Fig. 6. Price per connection under different numbers of users.

*4) Adaptation of Evolutionary Equilibrium:* We vary the number of users in area 2, and the resulting size of bandwidth allocated to each user and price per connection at the evolutionary equilibrium are shown in Figs. 5 and 6, respectively. When the number of users in area 2 increases, traffic load in WMAN and cellular network increases. As a result, the allocated bandwidth per connection for the users choosing these networks in area 3 becomes smaller than that for a user choosing WLAN. The users in area 3 then deviate (i.e., churn) to WLAN. Therefore, as the number of users in area 2 increases, the amount of bandwidth allocated to users in area 2 as well as users in area 3 decreases. As the number of users increases, the price per connection increases. In addition, prices offered by all the networks increase due to the dynamic network selection by users in all areas. Note that the allocated bandwidth is the lowest for users choosing the cellular network, and the price offered by this network is also the lowest. Similar results are expected, even if the network topology varies.

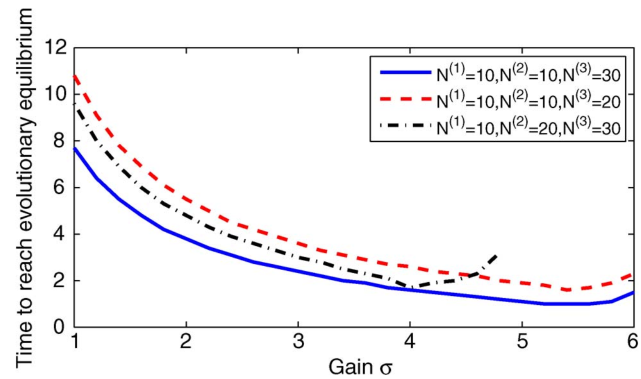


Fig. 7. Time to reach evolutionary equilibrium.

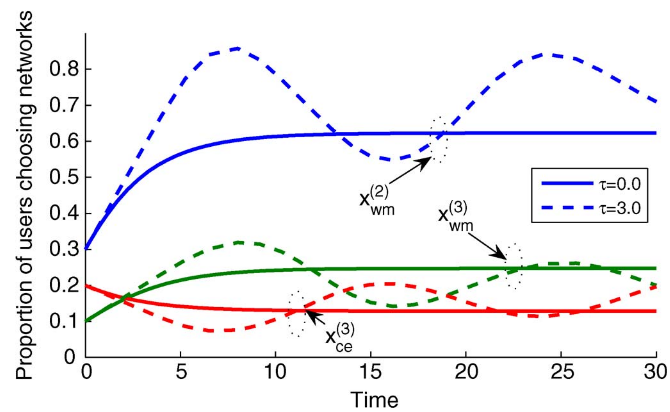


Fig. 8. Trajectories of strategy adaptation over time toward evolutionary equilibrium.

*5) Speed of Network Selection:* We investigate the impact of $\sigma$ [in (4)], which controls the speed of observation and network selection based on replicator dynamics. The time to reach the evolutionary equilibrium is shown in Fig. 7. When the gain $\sigma$ increases, the time to reach the equilibrium decreases since the proportion of users performing adaptation in each step becomes larger. However, as the gain $\sigma$ increases further, after a particular value, the time to reach the equilibrium increases. If the gain $\sigma$ is very large, the adaptation becomes unstable, and the evolutionary equilibrium may not be reached. We also observe that the time to reach the equilibrium is affected by the number of users in each area.

*6) Impact of Delay in Information Exchange:* Using replicator dynamics, we investigate the impact of $\tau$ (i.e., delay in exchanging information about population) on the dynamics of strategy adaptation. The trajectories of strategy adaptation over time toward the evolutionary equilibrium are shown in Fig. 8. When $\tau > 0$, we observe a fluctuating dynamics of strategy adaptation. The larger the delay is, the more the fluctuation there will be. In this case, a value of $\tau$ given by $\tau'$ can be found such that for $\tau > \tau'$, network selection never reaches the evolutionary equilibrium. This is due to the fact that when outdated/incorrect information is used by the users, the decisions tend to be inaccurate. We also observe that with $\tau > 0$, the system becomes less stable when the capacity of the network and/or the offered price (by the access networks) increases and/or when the total number of users decreases.

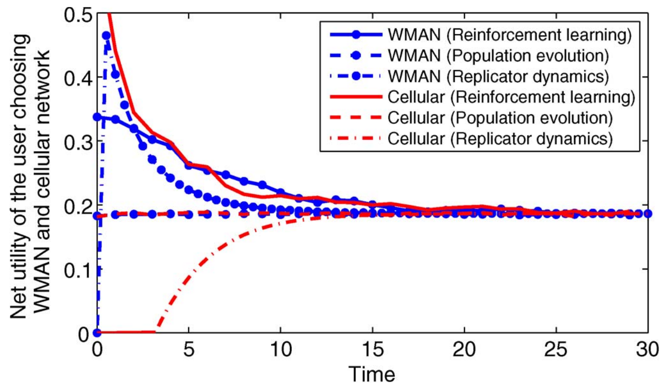


Fig. 9. Convergence of the network-selection algorithms to the evolutionary equilibrium.

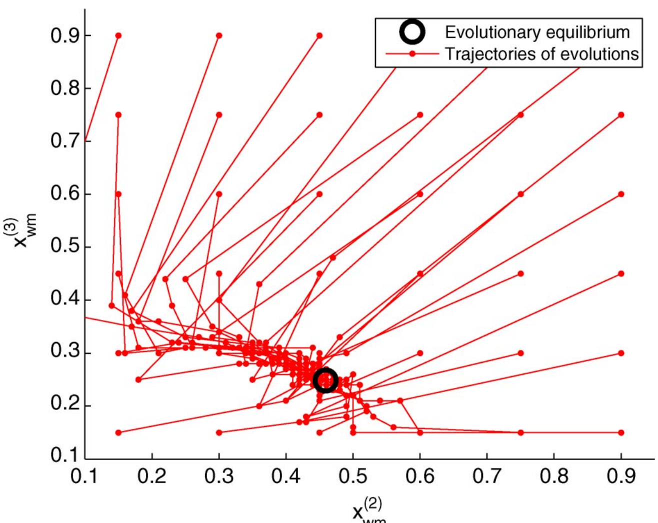


Fig. 10. Evolution of network selection based on population evolution algorithm.

*7) Convergence to Evolutionary Equilibrium:* The convergence properties of the network-selection algorithms based on population evolution and reinforcement learning are studied (in Fig. 9). The former algorithm converges to the equilibrium within less than ten iterations (i.e., with a net utility of 0.186 for a user). In contrast, the latter one requires more iterations to reach the equilibrium. Since the population evolution algorithm can utilize the average payoff information, it takes much less time to converge than does the reinforcement learning algorithm in which a user selects a network independently by using only its local payoff information obtained through exploration.

While the network-selection algorithm based on population evolution can reach the evolutionary equilibrium very fast, the reinforcement learning algorithm is attractive from a practical system implementation viewpoint. This is due to the fact that, a centralized controller to gather, process, and broadcast payoff information of the users in the service area may not be available in practice. In addition, the users may operate independently and may not want to share the payoff information with each other.

*8) Phase Plane of Population Evolution:* For the network-selection algorithm based on population evolution, we set $x_{\rm ce}^{(3)} = 0.0$ and vary $x_{\rm wm}^{(2)}$ and $x_{\rm wm}^{(3)}$. The evolution of strategy

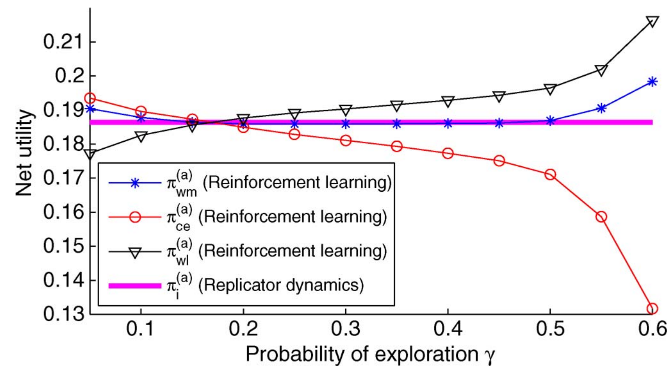


Fig. 11. Performance of reinforcement-learning-based network-selection algorithm.

adaptation (i.e., adaptation of the proportions of users choosing different networks) is shown in Fig. 10. With average payoff information, a user can adapt its strategy for network selection to reach the evolutionary equilibrium. In particular, if the current payoff of a user is lower than the average payoff of the users in the corresponding service class, the user may deviate from the current network to obtain a better payoff. We observe that the direction of adaptation in the proportions of users choosing different networks is similar to the phase plane of replicator dynamics shown in Fig. 2.

*9) Performance of the Reinforcement-Learning-Based Network Algorithm:* We evaluate the performance of reinforcement-learning-based network-selection algorithm under different values of $\gamma$ (i.e., probability of exploration). The payoff (i.e., net utility) of a user is shown in Fig. 11. In general, when $\gamma$ is small, a user spends only a small fraction of time in learning and gathering information about the available access networks. Since the knowledge is not complete, the payoffs of the users are not the same. Specifically, users choosing WLAN have the lowest payoff (which is lower than the average payoff) due to the largest number of users being in area 3 for the service area shown in Fig. 1. The payoffs of users choosing the cellular network and the WMAN are higher than the average payoff due to small number of users in areas 1 and 2.

When $\gamma$ is large, a user spends a large portion of time to learn by randomly choosing one of the available access networks, and, again, payoffs of the users are not identical. Specifically, the payoffs of users selecting WLAN is the highest (which is higher than the average payoff). Note that a value of $\gamma$ can be found such that the payoffs of all users choosing different networks are identical (e.g., with the assumed system parameters $\gamma \approx 0.16$ for the service areas shown in Fig. 1).

## VIII. Conclusion

Developed primarily to study the behavior of biological agents, evolutionary games can be used to analyze the competition among players with bounded rationality. We have investigated the dynamics of network selection in heterogeneous wireless networks using the theory of evolutionary games. The users in different service areas compete for bandwidth from different wireless networks (i.e., WMAN, cellular network, and

WLAN). A user selects the wireless access network based on its utility, which is a function of allocated bandwidth and price per connection. The dynamics of network selection has been mathematically modeled by the replicator dynamics that describes the adaptation in proportions of users choosing different access networks. The evolutionary equilibrium has been considered to be the stable solution for which all users receive identical net utility from accessing different networks.

We have alternatively formulated this network-selection problem as a noncooperative game among multiple groups of users using the same service class in different service areas, and the Nash equilibrium has been obtained as the solution to this game. We have proposed two algorithms, namely, population evolution and reinforcement learning algorithms for dynamic evolutionary game-based network selection. The network-selection algorithm based on population evolution utilizes information from all users in the same service area. On the other hand, in the reinforcement-learning-based algorithm, the users learn the performances and prices of different networks by interaction. Knowledge gained from learning is used to make the best decision for network selection.

## REFERENCES


[1] D. Niyato and E. Hossain, "Call admission control for QoS provisioning in 4G wireless networks: Issues and approaches," *IEEE Netw.*, vol. 19, no. 5, pp. 5–11, Sep./Oct. 2005.

[2] G. T. Karetsos, S. A. Kyriazakos, E. Groustiotis, F. D. Giandomenico, and I. Mura, "A hierarchical radio resource management framework for integrating WLANs in cellular networking environments," *IEEE Wireless Commun.*, vol. 12, no. 6, pp. 11–17, Dec. 2005.

[3] F. Yu and Y. Krishnamurthy, "Optimal joint session admission control in integrated WLAN and CDMA cellular networks with vertical handoff," *IEEE Trans. Mobile Comput.*, vol. 6, no. 1, pp. 126–139, Jan. 2007.

[4] W. Song, H. Jiang, and W. Zhuang, "Performance analysis of the WLAN-first scheme in cellular/WLAN interworking," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1932–1952, May 2007.

[5] H. Lin, M. Chatterjee, S. K. Das, and K. Basu, "ARC: An integrated admission and rate control framework for competitive wireless CDMA data networks using noncooperative games," *IEEE Trans. Mobile Comput.*, vol. 4, no. 3, pp. 243–258, May/Jun. 2005.

[6] Y.-L. Kuo, E. H.-K. Wu, and G.-H. Chen, "Noncooperative admission control for differentiated services in IEEE 802.11 WLANs," in *Proc. IEEE GLOBECOM*, Nov./Dec. 2004, vol. 5, pp. 2981–2986.

[7] S. Fischer and B. Vocking, "Evolutionary game theory with applications to adaptive routing," in *Proc. ECCS*, Nov. 2005, pp. 104–109.

[8] O. Sallent, J. Perez-Romero, R. Agusti, L. Giupponi, C. Kloeck, I. Martoyo, S. Klett, and J. Luo, "Resource auctioning mechanisms in heterogeneous wireless access networks," in *Proc. IEEE VTC Spring*, May 2006, vol. 1, pp. 52–56.

[9] W. Zhang, "Bearer service allocation and pricing in heterogeneous wireless networks," in *Proc. IEEE ICC*, May 2005, vol. 2, pp. 1367–1371.

[10] H. Chan, P. Fan, and Z. Cao, "A utility-based network selection scheme for multiple services in heterogeneous networks," in *Proc. IEEE Int. Conf. IWCMC*, Jun. 2005, vol. 2, pp. 1175–1180.

[11] *IEEE 802.16 Standard—Local and Metropolitan Area Networks—Part 16*, IEEE Std. 802.16a-2003, 2003.

[12] L. Xu, X. Shen, and J. W. Mark, "Fair resource allocation with guaranteed statistical QoS for multimedia traffic in wideband CDMA cellular network," *IEEE Trans. Mobile Comput.*, vol. 4, no. 2, pp. 166–177, Mar./Apr. 2005.

[13] J. Choi, J. Yoo, C. Kim, and S. Choi, "EBA: An enhancement of the IEEE 802.11 DCF via distributed reservation," *IEEE Trans. Mobile Comput.*, vol. 4, no. 4, pp. 378–390, Jul./Aug. 2005.

[14] J. Hou, J. Yang, and S. Papavassiliou, "Integration of pricing with call admission control to meet QoS requirements in cellular networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 13, no. 9, pp. 898–910, Sep. 2002.

[15] S. Yaipairoj and F. C. Harmantzis, "Auction-based congestion pricing for wireless data services," in *Proc. IEEE ICC*, Jun. 2006, vol. 3, pp. 1059–1064.

[16] R. A. Fisher, *The Genetic Theory of Natural Selection*. Oxford, U.K.: Clarendon, 1930.

[17] M. D. Sahlins, *Evolution and Culture*. Ann Arbor, MI: Univ. Michigan Press, 1970.

[18] Z. Jiang, Y. Ge, and Y. Li, "Max-utility wireless resource management for best-effort traffic," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 100–111, Jan. 2005.

[19] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory*. Berlin, Germany: Springer-Verlag, 2004.

[20] L. Tadj and A. Touzene, "A QBD approach to evolutionary game theory," *Appl. Math. Model.*, vol. 27, no. 11, pp. 913–927, Nov. 2003.

[21] F. Mazenc and S.-I. Nuculescu, "Lyapunov stability analysis for nonlinear delay systems," in *Proc. IEEE Conf. Decision Control*, Dec. 2000, vol. 3, pp. 2100–2105.

[22] M. J. Osborne, *An Introduction to Game Theory*. London, U.K.: Oxford Univ. Press, 2003.

[23] R. C. Robinson, *An Introduction to Dynamical Systems*. Englewood Cliffs, NJ: Prentice–Hall, 2004.

[24] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning)*. Cambridge, MA: MIT Press, 1998.

[25] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*. Cambridge, MA: MIT Press, 1998.



**Dusit Niyato** received the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Winnipeg, MB, Canada, in 2008.

He is currently an Assistant Professor with the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests are radio resource management in cognitive radio networks and broadband wireless access networks.



**Ekram Hossain** (S'98–M'01–SM'06) is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada.

His current research interests include the design, analysis, and optimization of wireless communication networks and cognitive radio systems. He is a co-author/co-editor for the books *Dynamic Spectrum Access and Management in Cognitive Radio Networks* (Cambridge University Press, 2009), *Heterogeneous Wireless Access Networks* (Springer, 2008), *Introduction to Network Simulator NS2* (Springer, 2008), *Cognitive Wireless Communication Networks* (Springer, 2007), and *Wireless Mesh Networks: Architectures and Protocols* (Springer, 2007).

Dr. Hossain serves as an Editor for the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *IEEE Wireless Communications*, and several other international journals. He is a registered Professional Engineer in the Province of Manitoba.