AP Association in a IEEE 802.11 WLAN

Kinda Khawam*, Johanne Cohen*, Paul Muhlethaler†, Samer Lahoud‡, Samir Tohme*
*University of Versailles - PRISM Laboratory, France
†INRIA, France
‡University of Rennes I - IRISA, France

***Abstract*—Nowadays, with the abundance of IEEE 802.11 access points (APs), a mobile user has the flexibility to choose one of several APs, each using a separate channel. Rather than relying on the simplistic standardized algorithm to select the AP, it would be preferable to use optimal algorithms that reduce the user data transfer time. In this paper, the AP selection process is apprehended as an ordinal potential game, which is a class of non-cooperative games known to possess at least one pure Nash Equilibrium (PNE). We put forward a fully decentralized algorithm based on replicator dynamics to attain those PNE. Further, to assess the loss in efficiency of the proposed selfish distributed algorithm, we compare its performances against a centralized optimal approach derived by solving a mixed integer linear program.**

I. Introduction

The IEEE 802.11x protocol is currently the standard for wireless LANs (WLANs) [1]. It has been widely deployed in airports, coffee shops and homes. Users scan the wireless channel in order to find the AP which shows the highest signal strength and associate to it. User throughput in each AP is determined by the number of other users associated to it as well as the physical rate being used. In particular, it has also been observed [2] that all the users connected to the same AP receive the same throughput. Resorting to non-cooperative game theory is quite suitable to model the way users compete selfishly for limited resources. Devising an optimal AP selection scheme depends on the existence of Nash equilibria for the present game. In this paper, we prove that the model at hand is an ordinal potential game, known to have at least one PNE. Furthermore, we prove that an algorithm based on replicator dynamics converges to the PNEs of our game. Finally, to quantify the efficiency loss of the distributed game approach, known as the price of anarchy, we compare its performance against a centralized approach where resource allocation is made in a way that satisfies all mobile users. It turns out that even though the distributed game results are sub-optimal, the acceptable discrepancy between the two sets of results and the inherent adaptability of the decentralized approach makes it really promising.

In [3] a study is made on fairness issues and how the load should be balanced using fractional association in a cooperative scenario. Usually, users have no particular incentive to cooperate with each other and would be interested in maximizing their individual payoffs. In [4], the case of non-cooperative users who decide on the optimal frame size and PHY rate to be used in order to maximize their individual throughputs is studied. The users are all assumed to be in a single cell and compete for throughput within that cell. Another work on non-cooperative association is [5], which provides a simulation study of the benefit of associating to the AP that would provide the best estimated link rate. Some results on cooperative association of users to different APs are provided in [6].

The rest of the paper is organized as follows. The system model and cost characterization in WiFi are given in Section II. The radio access selection scheme is presented as a non-cooperative ordinal potential game in Section III. The distributed learning algorithm based on replicator dynamics is presented in Section IV. The optimal centralized approach is given in V as well as an evaluation of the price of anarchy. The conclusion is given in Section VI.

II. The WLAN System Model

In the WiFi standard, the set of achievable rates is not continuous. Indeed, users transmit at different peak rates based on the signal strength received. The algorithm that selects the peak rate chooses a higher rate if the signal strength is good and progressively cuts down the rate as signal strength decays. This behavior results in a discrete set of achievable peak rates $\chi_{1,s} < \chi_{2,s} < ... < \chi_{M_s,s}$ where M_s is the maximum number of achievable rates for AP s (see Table I).

The instantaneous rate $\Re_{k,s}(t)$ that user k gets when connected to AP s depends on its location and varies with time t due to fading effects (mobility is not taken into account). Hence, we have the following:

$$\Re_{k,s}(t) = y_k(t) \cdot \chi_{k,s}$$

where y_k are i.i.d. random variables (of unit mean) that represent the impact of fast fading experienced by mobile user k. They follow an exponential distribution as we consider Rayleigh fading [7]. As in this work we assume stationarity, the time index is omitted in what follows.

In this paper, the uplink traffic is neglected which leads to a fair access scheme on the downlink channel. However, when a low rate user captures the channel, this user will use it for a long time, which penalizes high rate users and reduces the fair access strategy to a case of fair rate sharing (assuming a constant MAC frame size and neglecting the 802.11 waiting times (i.e., DIFS, SIFS, etc.) in comparison with transmission times). Consequently, the data rate of a WiFi user is given by:

$$R_k^s = \left[\sum_{i=1}^{n} \frac{\mathbb{1}_{\{i \text{ in AP } s\}}}{\Re_{i,s}}\right]^{-1} = \left[\sum_{i=1}^{n_s} \frac{1}{\chi_{i,s} \cdot y_i}\right]^{-1}$$

$\chi_{k,s}$	6	9	12	18	24	36	48	54

Table I
THE SET OF DISCRETE PEAK RATES (MBITS/S) IN 802.11G

where n is the total number of mobile users and n_s is the number of mobile users that simultaneously choose AP s. Using the Jensen inequality, the mean data rate in AP s has the following lower bound:

$$\mathbb{E}[R_k^s] \geq \left[\sum_{i=1}^{n_s} \frac{1}{\chi_{i,s}}\right]^{-1} \quad (1)$$

In the rest of the paper, the mean data rate in AP s will be approximated by its lower bound. Hence, using the lower bound on the mean data rate, the mean data transfer time in AP s for any user k is given by:

$$T_{k,s} = \frac{1}{\mathbb{E}[R_k^s]} = \sum_{i=1}^{n_s} \frac{1}{\chi_{i,s}} \quad (2)$$

Note that the index k can be omitted as all users have the same data transfer time.

III. NON-COOPERATIVE GAME FOR AP SELECTION

Non-Cooperative game theory models the interactions between players competing for a common resource. Hence, it is well adapted to radio resource management modeling. For each state of the system, defined by the number of mobile users n, we define a multi-player game $\mathcal{G}$ between those n mobile users present in an area with m APs. In this model, there is a sequence of one-stage games, each corresponding to a given state of the system, defined by the number of mobile users. Whenever a new mobile is admitted in the system, the game is played again with an additional player. Mobile users are assumed to make their decisions without knowing the decisions of other users.

We present here the general framework: $M = \{1,..,m\}$ is the set of available APs, $N = \{1,..,n\}$ is the set of players. Each user is a player that has to pick one AP among the m available APs. The strategy of player k is denoted by x_k such that $x_k = s$ if player k chooses AP s. Hence, $X = (x_k)_{k\in N} \in \mathcal{S} = \{1, 2, \cdots, m\}^N$ is a pure strategy profile. $\mathcal{S}$ is the space of all profiles.

We will define the cost function of a user i that selected strategy $x_k = s$ for any $s \in \mathcal{S}$ as follows:

$$c_k(x_k = s, X_{-k}) = \sum_{i:x_i=s} \frac{1}{\chi_{i,s}} \quad (3)$$

Where X_{-k} denotes the vector of strategies played by all other users except user k. Note that the cost function is the mean data transfer time given in Equation (2).

A. The Nash Equilibrium

In a non-cooperative game, an efficient solution is obtained when all players adhere to a Nash Equilibrium (NE). An NE is a profile of strategies in which no player will profit from deviating from its initial strategy unilaterally. Hence, it is a strategy profile where each player's strategy is an optimal response to the other players' strategies.

In general, finite games are not guaranteed to have a PNE. Nevertheless, they possess a mixed NE where each player has to continually change its AP selection according to a distribution probability over the strategy set. Mixed equilibria have practical issues. They lead to a situation where each mobile user has to be simultaneously connected to more than one AP and to split its traffic over those APs. Whenever a PNE exists, an equilibrium can be reached where every mobile user is consigned to only one AP.

B. Ordinal Potential Game

Ordinal Potential games (OPG) form a special class of normal form games where the unilateral change of one user's strategy x_i to x_i' results in a change of its cost function that is paralleled by a change of a so-called potential function $\phi : \mathcal{S}^n \to \mathbb{R}$ as follows:

$$c_i(x_i, X_{-i}) < c_i(x_i', X_{-i}) \Leftrightarrow \phi(x_i, X_{-i}) < \phi(x_i', X_{-i})$$

An OPG allows for at least one PNE which avoids costly traffic splitting between available APs.

Proposition 3.1: The game $\mathcal{G}$ is an ordinal potential game.

Proof: Let $\mathcal{P}(\mathcal{S})$ be the power set of $\mathcal{S}$. We denote by ϵ the following:

$$\epsilon = min\left\{\left|\sum_{i\in X} \frac{1}{\chi_{i,s}} - \sum_{i\in X'} \frac{1}{\chi_{i,s}}\right| : (s,s') \in \mathcal{S}^2, (X,X') \in \mathcal{P}^2(N)\right\}$$

We will define for any $s \in \mathcal{S}$ and any $i \in N$,

$$\alpha_{i,s} = max(1, \frac{1}{\epsilon})\frac{\chi_{max}}{\chi_{i,s}}$$

where χ_{max} is the greatest available peak rate.

Now, we will define our potential function which maps a profile $X = (x_1, x_2, \ldots, x_n)$ to a real number:

$$\phi(X) = \sum_{s\in\mathcal{S}} \left[4^{\sum_{i:x_i=s}\alpha_{i,s}}\right] \quad (4)$$

Further, we will define a function $f(x)$ such that:

$$f(x) = 4^{\alpha+x} + 4^{x-\gamma} - 4^x - 4^{x-\gamma+\beta}$$

Where $\alpha \geq 1$, $\beta \geq 1$, and γ such that $\alpha \geq \beta - \gamma + 1$. We prove below that for any real $x > 0$, we have $f(x) > 0$.

$$\begin{aligned} f(x) &= 4^{x-\gamma} + 4^{\alpha+x}(1 - 4^{-\alpha} - 4^{\beta-\alpha-\gamma}) \\ &\geq 4^{x-\gamma} + 4^{\alpha+x}(1 - 1/4 - 4^{\beta-\alpha-\gamma}) \\ &\geq 4^{x-\gamma} + 4^{\alpha+x}(1 - 1/4 - 1/4) > 0 \end{aligned}$$

The first inequality is obtained through the relation $\alpha \geq 1$. The second inequality is deduced from the relation $\alpha \geq \beta - \gamma + 1$.

Let X and X' be two profiles which only differ in the strategy of one player z. Finally, we will prove that $c_z(X') - c_z(X) > 0$ if and only if $\phi(X') - \phi(X) > 0$.

We assume that $c_z(X') - c_z(X) > 0$. By definition of parameter ϵ, we have $c_z(X') - c_z(X) \geq \epsilon$. Moreover, by definition of parameter $\alpha_{i,s}$, we get :

$$\alpha_{z,x'_z} + \sum_{\substack{i:x_i=x'_z,\\ i\neq z}} \alpha_{i,x'_z} - \alpha_{z,x_z} - \sum_{\substack{i:x_i=x_z,\\ i\neq z}} \alpha_{i,x_z} \geq 1 \quad (5)$$

$$\phi(X') - \phi(X) = 4^{\alpha_{z,x'_z} + \sum_{i:x_i=x'_z,\, i\neq z} \alpha_{i,x'_z}} + 4^{\sum_{i:x_i=x_z,\, i\neq z} \alpha_{i,s}} - 4^{\alpha_{z,x_z} + \sum_{i:x_i=x_z,\, i\neq z} \alpha_{i,x_z}} - 4^{\sum_{i:x_i=x'_z,\, i\neq z} \alpha_{i,x'_z}}$$

It remains to check that $\phi(X') - \phi(X) = f(x)$ where

- $x = \sum_{i:x_i=x'_z,\, i\neq z} \alpha_{i,x'_z}$
- $\alpha = \alpha_{z,x'_z}$ (by definition of α_{z,x'_z}, we have $\alpha > 1$)
- $\beta = \alpha_{z,x_z}$,
- $\gamma = x - \sum_{i:x_i=x_z,\, i\neq z} \alpha_{i,s} > 1$

Note that from Equation (5), we have $x + \alpha \geq x - \gamma + \beta + 1$.

We can apply the same result for the case where $c_z(X') - c_z(X) < 0$. ■

IV. Distributed Learning of PNE

Implementing a practical distributed AP selection policy to reach PNE is not straightforward and must be carried out carefully. In this paper, we resort to replicator dynamics ([9]) to learn Nash equilibria.

1) Replicator Dynamics: A *mixed strategy* $q_k = (q_{k,1}, q_{k,2}, \ldots, q_{k,m})$ corresponds to a probability distribution over pure strategies. In other words, pure strategy s is chosen with probability $q_{k,s} \in [0,1]$, with $\sum_{s=1}^{m} q_{k,s} = 1$. Let K_k be the simplex of mixed strategies for user k. Any pure strategy s can be considered as a mixed strategy e_s, where vector e_s denotes the unit probability vector with s^{th} being a component unity, hence a corner of K_k.

Let $K = \prod_{i=1}^{n} K_k$ be the space of all mixed strategies. A *strategy profile* $Q = (q_1, ..., q_n) \in K$ specifies the (mixed or pure) strategies of all players. Following classical convention, we write $Q = (q_k, Q_{-k})$, where Q_{-k} denotes the vector of strategies played by all other players.

The general distributed algorithm is the following:

Input: $q(0) = (q_1(0), ..., q_n(0))$ any vector of probabilities.

At each round t, every user k:

- selects AP s with probability $q_{k,s}(t)$. This leads to an outcome $r_k(s)$ for user k.
- Updates $q_k(t)$ as follows:

$$q_k(t+1) = q_k(t) + b \cdot F_k^b(r_k(t), x_k(t), q_k(t)) \quad (6)$$

In the general version of the depicted algorithm, the function $F_k^b(r_k(t), x_k(t), q_k(t))$ can be very broad (some conditions are nevertheless defined in [9]) and b is a real number smaller than 1. The cost $r_k(t)$ (obtained at step t by user k as a consequence of the selected strategy) can be completely random. In practice, the mobile users will rely on the beacon frame (sent at least every 100ms) to evaluate their performances (through their received power level) and hence their cost function.

Algorithms of this form are fully distributed as decisions made by users are completely decentralized: at each time step, user k only needs to know its own cost $r_k(t)$ and mixed strategy q_k. In our model, r_k is equal to $1 - c_{k,s}$. The function F_k^b is given by:

$$F_k^b(r_k(t), x_k(t), q_k(t)) = \gamma(r_k(t)) \cdot (\mathbb{1}_{\{x_k=s\}} - q_k(t)) \quad (7)$$

where $\gamma : R \to [0,1]$ is some affine decreasing function. The proof of convergence of this replicator is given in the Appendix as well as the proof of the following proposition:

Proposition 4.1: For any initial condition where $\forall k \in N$ and $\forall s \in M$, $q_{k,s} \neq 0$ or $q_{k,s} \neq 1$, the considered learning algorithm converges to a PNE.

Thus, according to Proposition 4.1, a single AP is selected when the distributed algorithm converges.

2) Simulation Results: We run different simulations with various numbers of users and APs. In the first set of simulations, we take 4 APs and 20 users. In Figure 1, we depict the strategy dynamics q_k of two randomly selected users (users 2 and 4) as a function of the number of iterations. We can see that the users' strategies converge to either 0 or 1 opting for one single AP instead of load balancing their traffic between several APs. We recorded this behavior through the extensive simulations we performed (further, the result of convergence to PNE is proven in the appendix).

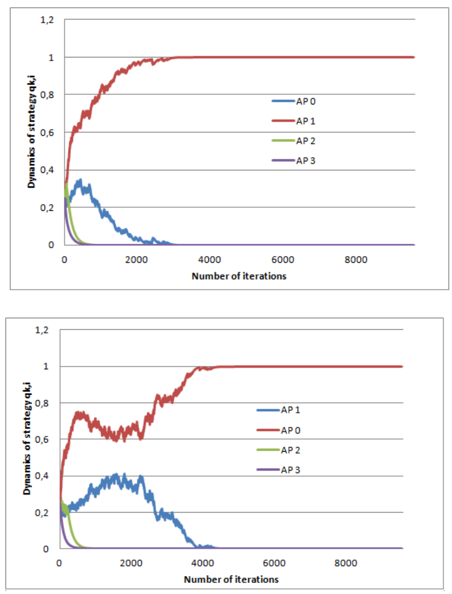
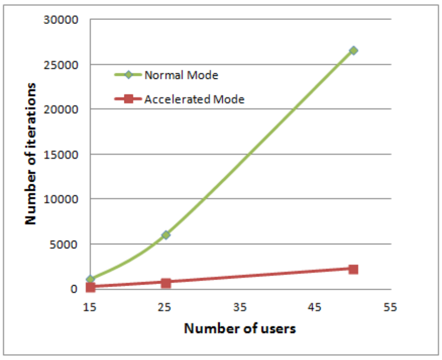
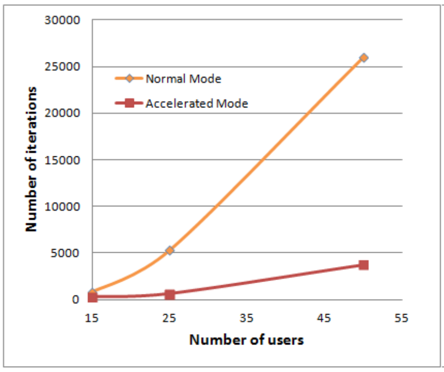


Figure 1. Replication Dynamics: strategy updates for 2 random users

We also notice the slow convergence of the algorithm (approximately 9500 iterations for the present case) which hinders the benefits of a distributed approach. However, we see through the extensive simulations we ran that the convergence is relatively fast at the beginning of the algorithm but it slows down drastically half way through. At that point, the AP that will be ultimately selected by each user is clearly designated (we can see this behavior in Figure 1 around 4000 iterations) and it is useless to pursue the AP selection process. Hence, we propose an accelerated mode of the algorithm such that whenever for a given user the probability of selecting any AP surpasses 0.8, convergence is assumed to be reached and the user chooses this AP. In Figure 2, we run 25 simulations to compare the number of iterations of the accelerated mode vs. the normal mode for respectively 4 and 5 APs as a function of the number of active users. We can see the tremendous improvement in the accelerated mode for comparable performances.



(a) 4 APs



(b) 5 APs

Figure 2. Normal Mode vs. Accelerated Mode

V. The Price of Anarchy

In this section, we quantify the loss in efficiency suffered when a distributed scheme is adopted rather than a centralized optimization.

A. Optimal Centralized Approach

Unlike the distributed approach where precedence is given to the individual user's interests, resource allocation may be performed in a way that favors the overall system performance. We do so by introducing a *centralized approach*, where the system assigns the traffic of each mobile user to the APs in order to minimize the total network cost. We first introduce a binary variable $\theta_{k,s}$ that indicates whether user k has selected AP s or not. The cost incurred by user k for choosing AP s proposed in (2) is modified as follows:

$$T_{k,s} = \frac{1}{\chi_{k,s}} + \sum_{k' \in N, k' \neq k} \frac{\theta_{k',s}}{\chi_{k',s}}.$$

where

$$\theta_{k,s} = \begin{cases} 1 & \text{if user } k \text{ is associated to AP } s, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the individual cost of user k is given by:

$$T_k(\theta_k) = \sum_{s=1}^{m} \left(\frac{\theta_{k,s}}{\chi_{k,s}} + \sum_{k' \in N, k' \neq k} \frac{\theta_{k',s} \cdot \theta_{k,s}}{\chi_{k',s}} \right) \quad (8)$$

Where θ_k is the vector whose components are $\theta_{k,s}$.

Our goal consists in finding the optimal user association (designated by $\theta_{k,s}$ for each user $k \in N$ and AP $s \in M$) that minimizes the total network cost. The latter, denoted by C_{tot}, is defined as the sum of mobile users' individual costs and is given by $\sum_{k=1}^{n} T_k(\theta_k)$.

Thus, the global approach can be formulated as an optimization problem $(\mathcal{P})$ introduced below:

$$(\mathcal{P}): \text{ Min } C_{tot}(\theta_{k,s}, k \in N, s \in M) = \sum_{s \in M, k \in N} \left(\frac{\theta_{k,s}}{\chi_{k,s}} + \sum_{k' \in N, k' \neq k} \frac{\theta_{k',s} \cdot \theta_{k,s}}{\chi_{k',s}} \right) \quad (9)$$

Subject to:

$$\sum_{s \in M} \theta_{k,s} = 1 \quad \forall k \in N \quad (10)$$

$$\theta_{k,s} \in \{0, 1\} \quad \forall (k, s) : k \in N, s \in M \quad (11)$$

Constraints (10) ensure that a given user is connected to only one AP. Constraints (11) are the integrality constraints for the decision variables $\theta_{k,s}$.

1) Optimal Solution: Problem $(\mathcal{P})$ is a binary non-linear optimization that consists in minimizing the objective function (9) subject to the constraints (10) and (11). In this section, we propose two methods to compute the optimal solution of $(\mathcal{P})$, (i) using an exhaustive search algorithm and (ii) converting $(\mathcal{P})$ into a Mixed Integer Linear Programming (MILP).

a) Exhaustive search: The exhaustive search algorithm explores all the possible solutions to compute the minimum of the objective function. We note that the time complexity of this exhaustive search algorithm depends on the number of users and the number of APs. More precisely, the objective function should be evaluated for each value of $\theta_{k,s}$. As $\theta_{k,s}$ = 0 or 1 for each $s = 1, ..., M$ and k= $1, ..., N$, the time complexity for computing the minimum value of the objective function is in $\mathcal{O}(2^{M \cdot N})$. Thus, the exhaustive search is computational intensive, and rapidly becomes intractable for large scenarios.

b) Mixed Integer Linear Programming: In this section, we explain how to convert our non linear optimization problem $(\mathcal{P})$ into a MILP. A MILP problem consists of a linear objective function, a set of linear equality and inequality constraints and a set of variables with integer restrictions. Generally, MILP problems are solved using a branch-and-bound approach based on linear-programming. The idea of this approach is to solve Linear Program (LP) relaxations of the MILP and to look for an integer solution by branching and bounding on the decision variables provided by the LP relaxations. Thus, in a branch-and bound approach the number of integer variable determines the size of the search tree and influences the running time of the algorithm. Hence, we should try to keep the number of variables with restrictions (binary or integer) low.

To linearize our objective function, we replace the non linear terms by new variables and additional inequality constraints, which ensures that new variables behave according to the non linear terms they are replacing. Particularly, in the objective function (9) we replace each quadratic term $\theta_{k,s} \cdot \theta_{k',s}$ by a new variable $z_{k,k',s}$ and add the following three inequalities to the set of constraints:

$$z_{k,k',s} - \theta_{k,s} \leq 0 \quad \forall (k,k',s) : s \in M, k \neq k' \in N \qquad (12)$$

$$z_{k,k',s} - \theta_{k',s} \leq 0 \quad \forall (k,k',s) : s \in M, k \neq k' \in N \qquad (13)$$

$$\theta_{k,s} + \theta_{k',s} - z_{k,k',s} \leq 1 \quad \forall (k,k',s) : s \in M, k \neq k' \in N \qquad (14)$$

Inequalities (12) and (13) ensure that $z_{k,k',s}$ is equal to zero when either $\theta_{k,s}$ or $\theta_{k',s}$ are equal to zero, while the inequalities (14) force $z_{k,k',s}$ to be equal to one if $\theta_{k,s}$ and $\theta_{k',s}$ are equal to one. In addition, we give the bound constraints for the variables $z_{k,k',s}$ which are introduced during the linearization process:

$$0 \leq z_{k,k',s} \leq 1 \quad \forall (k,k',s) : s \in M, k \neq k' \in N \qquad (15)$$

Finally, our MILP problem $(\mathcal{P}')$ is given by:

$$Min\ C_{tot}(\theta_{k,s}, z_{k,k',s}) = \sum_{k \in N, s \in M} \left(\frac{\theta_{k,s}}{\chi_{k,s}} + \sum_{k' \in N, k' \neq k} \frac{z_{k,k',s}}{\chi_{k',s}} \right) \qquad (16)$$

Subject to the constraints:(10)-(15).

B. Simulation Results

In Figure 3, we illustrate the mean time necessary to send a data unit for all users as a function of the number of active users present in the system for both the centralized optimal approach and our algorithm based on replicator dynamics using a 95 percent confidence interval. We find an expected improvement in the optimal approach in comparison with the decentralized approach, especially for large numbers of users. However, the acceptable discrepancy between the two sets of results and the low degree of system complexity of the decentralized approach makes it an attractive solution.

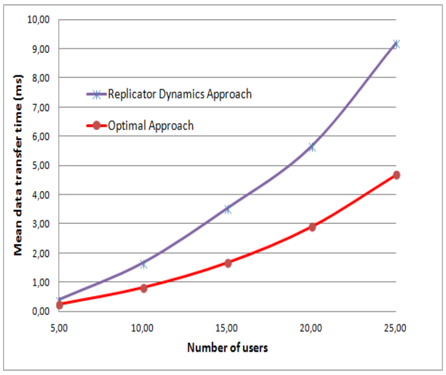


Figure 3. Mean data transfer time for 4 APs

VI. Conclusion

Nowadays, with the abundance of diverse air interfaces in the same operating area, advanced radio resource management is vital to take advantage of the available system resources. This paper focuses on such a network selection problem in the context of IEEE 802.11 WLANs where several access points provide connection service to users. We formulate this problem as a non-cooperative game where each user tries to minimize its cost function, defined as the data transfer time. We then conduct an analysis of the formulated game and propose an access point selection algorithm based on replicator dynamics. The proposed algorithm, which can be implemented based on local observation, is especially suitable in decentralized adaptive learning environments such as wireless access networks. Finally, the simulation results highlight the effectiveness of the proposed algorithm to achieve high system efficiency compared with an optimal centralized scheme.

Appendix

Theorem A.1: Let G be an instance of game $\mathcal{G}$. The learning algorithm, for any initial condition in $K - K^*$, always weakly converges to a Nash Equilibrium.

Proof: Let $Q = (q_1, q_2, \ldots, q_n)$ be a mixed profile of the game. According to Sastry et al. [12], the Linear Reward-Inaction algorithm converges weakly towards a replication dynamic:

$$\frac{dq_{i,s}}{dt}(Q) = q_{i,s} \left(\mathbb{E}[\ c_i\ |Q\ ] - \mathbb{E}[\ c_i\ |q_{i,s} = 1, Q_{-i}\ ] \right) \qquad (17)$$

This equation, called the (multi-population) replicator dynamics, is well-known to have its limit points related to Nash equilibria (through the so-called Folk's theorem of evolutionary game theory [11]). More precisely, we have the following theorem:

Theorem A.2: The following are true for the solutions of Equation (17): (i) All Nash equilibria are stationary points. (ii) All strict Nash equilibria are asymptotically stable. (iii) All stable stationary points are Nash equilibria.

From [9], the limit for $b \to 0$ of the dynamics of stochastic algorithms is some Ordinary Differential Equations (ODE) whose stable limit points, when $t \to \infty$ (if they exist), can only be NEs. Hence, if there is convergence for the ordinary differential equation, then one expects the replicator dynamic algorithm to reach an equilibrium. Moreover, in [8], Coucheney *et al.* proved that such NEs are pure.

Let us see if the continuous dynamic converges with stability arguments. We denote by $\pi(X|Q)$ the probability that all players choose AP x_i according to the mixed profile Q:

$$\pi(X|Q) = \prod_{i=1}^{n} q_{i,x_i}. \tag{18}$$

where $X = (x_1, x_2, \ldots, x_n)$.

We define the following function $\Phi : \mathcal{K} \to \mathbb{R}$

$$\Phi(Q) = \sum_{X \in \mathcal{S}^n} \pi(X|Q)\phi(X) \tag{19}$$

Let us study the evolution of function $\Phi(Q)$ over time. We focus on $\frac{d\Phi}{dt}(Q)$. By definition, we have :

$$\frac{d\Phi}{dt}(Q) = \sum_{i=1}^{n}\sum_{s\in S}\frac{\partial \Phi}{\partial q_{i,s}}\frac{dq_{i,s}}{dt}(Q)$$

First, we compute $\frac{\partial \Phi}{\partial q_{j,s}}(Q)$:

$$\frac{\partial \Phi}{\partial q_{j,s}}(Q) \quad = \quad \textstyle\sum_{X_{-j}\,\in\,\mathcal{S}^{n-1}} \pi(X_{-j}|Q)\phi(x_j = s, X_{-j})$$

Hence, we have the following:

$$\begin{aligned}
&\frac{d\Phi}{dt}(Q) = \sum_{i\in N}\sum_{s\in S}\frac{\partial \Phi}{\partial q_{i,s}}\frac{dq_{i,s}}{dt}(Q)\\
&= \textstyle\sum_i \sum_{s\in S} q_{i,s}\frac{\partial \Phi}{\partial q_{i,s}}\left(\mathbb{E}[\ c_i\ |Q\ ] - \mathbb{E}[\ c_i\ |q_{i,s} = 1, Q_{-i}\ ]\right)\\
&= \sum_i \sum_{(k,s)\in\mathcal{S}^2} q_{i,k}q_{i,s}\frac{\partial \Phi}{\partial q_{i,s}}\cdot\\
&(\mathbb{E}[\ c_i\ |q_{i,k} = 1, Q_{-i}\ ] - \mathbb{E}[\ c_i\ |q_{i,s} = 1, Q_{-i}\ ])\\
&= \sum_i \sum_{k<s} q_{i,k}q_{i,s}\left(\frac{\partial \Phi}{\partial q_{i,s}}(Q) - \frac{\partial \Phi}{\partial q_{i,k}}(Q)\right)\cdot\\
&(\mathbb{E}[\ c_i\ |q_{i,k} = 1, Q_{-i}\ ] - \mathbb{E}[\ c_i\ |q_{i,s} = 1, Q_{-i}\ ])
\end{aligned}$$

Now we will prove that $(\frac{\partial \Phi}{\partial q_{i,k}}(Q) - \frac{\partial \Phi}{\partial q_{i,s}}(Q))$ has the same sign as $\mathbb{E}[\ c_i\ |q_{i,k} = 1, Q_{-i}\ ] - \mathbb{E}[\ c_i\ |q_{i,s} = 1, Q_{-i}\ ]$.

$$\begin{aligned}
&\tfrac{\partial \Phi}{\partial q_{i,k}}(Q) - \tfrac{\partial \Phi}{\partial q_{i,s}}(Q) =\\
&\textstyle\sum_{X_{-j}\in\mathcal{S}^{n-1}} \pi(X_{-j}|Q)\left(\phi(x_j = k, X_{-j}) - \phi(x_j = s, X_{-j})\right)
\end{aligned}$$

W.l.o.g., we assume that $\mathbb{E}[\ c_i\ |q_{i,k} = 1, Q_{-i}\ ] - \mathbb{E}[\ c_i\ |q_{i,s} = 1, Q_{-i}\ ] > 0$. Let $\epsilon' = max(1, \frac{1}{\epsilon})\chi_{max}$. We will show that $4^{\epsilon'\mathbb{E}[\ C_k\ |q_{i,k}=1,Q_{-i}\ ]} + 4^{\epsilon'\mathbb{E}[\ C_s\ |q_{i,k}=1,Q_{-i}\ ]} - 4^{\epsilon'\mathbb{E}[\ C_s\ |q_{i,s}=1,Q_{-i}\ ]} - 4^{\epsilon'\mathbb{E}[\ C_k\ |q_{i,s}=1,Q_{-i}\ ]}$ is positive.

Recall that in the proof of Proposition 3.1, for any positive, fixed constants $\alpha \geq 1$, $\beta \geq 1$, and γ such that $\alpha \geq \beta - \gamma + 1$, the function $f(x) = 4^{\alpha+x} + 4^{x-\gamma} - 4^{x} - 4^{x-\gamma+\beta}$ in positive for any x. Furthermore, $f(x)$ is convex as $f''(x) = ln(4)\cdot f'(x) = (ln(4))^2 \cdot f(x) > 0$ for any x.

We set $x = \epsilon'\mathbb{E}[\ C_k\ |q_{i,s} = 1, Q_{-i}\ ]$, $\gamma = x - \epsilon'\mathbb{E}[\ C_s\ |q_{i,s} = 1, Q_{-i}\ ]$ and $\alpha = \alpha_{i,k}$. By definition of parameters ϵ and ϵ', we can check that $\alpha \geq 1$, $\beta \geq 1$, and γ such that $\alpha \geq \beta - \gamma + 1$.

Accordingly, we have $4^{\epsilon'\mathbb{E}[\ C_k\ |q_{i,k}=1,Q_{-i}\ ]} + 4^{\epsilon'\mathbb{E}[\ C_s\ |q_{i,k}=1,Q_{-i}\ ]} - 4^{\epsilon'\mathbb{E}[\ C_s\ |q_{i,s}=1,Q_{-i}\ ]} - 4^{\epsilon'\mathbb{E}[\ C_k\ |q_{i,s}=1,Q_{-i}\ ]} > 0$

Since f is convex, using Jensen's inequality, we have:

$$\begin{aligned}
&\sum_{X_{-i}\in\mathcal{S}^{n-1}} \pi(X_{-i}|Q)(4^{\alpha_{i,k}+\sum_{j:x_j=k}\alpha_{j,k}} + 4^{\sum_{j:x_j=s}\alpha_{j,s}}\\
&- 4^{\alpha_{i,s}+\sum_{j:x_j=s}\alpha_{j,s}} - 4^{\sum_{j:x_j=k}\alpha_{j,k}}) > 0
\end{aligned}$$

So we can conclude that $\frac{\partial \Phi}{\partial q_{i,k}}(Q) - \frac{\partial \Phi}{\partial q_{i,s}}(Q) > 0$, and therefore $\frac{d\Phi}{dt}(Q) \leq 0$.

Thus Φ is nondecreasing along the trajectories of the replication dynamics. Thus, due to the nature of the learning algorithm, all solutions of the ODE (17) remain in the strategy space if initial conditions $\in [0, 1]$. From (A), we know that $\frac{d\Phi(Q^*)}{dt} = 0$ implies that $\forall i \in N \forall s, k \in \mathcal{S}$:

$$q^*_{i,s} = 0, or\ (\mathbb{E}[\ c_i\ |q_{i,k} = 1, Q_{-i}\ ] = \mathbb{E}[\ c_i\ |q_{i,s} = 1, Q_{-i}\ ])$$

Such a Q^* is consequently a stationary point of the dynamics.

Since from Theorem A.2, all stationary points that are not NEs are unstable, Proposition 4.1 holds.

Thus all solutions have to converge to some stationary point corresponding to NE. We can deduce that the learning algorithm, for any initial condition in $K - K^*$, always converges to a NE of instance G. This concludes the proof of this theorem.

■

REFERENCES

[1] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, ANSI/IEEE Std 802.11, 1999 Edition.

[2] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, Performance Anomaly of 802.11b, in Proceedings of IEEE INFOCOM 2003, San Fransisco, CA, March 2003.

[3] Y. Bejerano, S.-J. Han, and L. E. Li, Fairness and load balancing in wireless LANs using association control, in Proceedings of ACM Mobicom, Philadelphia, PA, September 2004.

[4] E. Altman, A. Kumar, D. Kumar, and R. Venkatesh, Cooperative and non-coooperative control in IEEE 802.11 WLANs, in Proceedings of the International Teletraffic Congress, Beijing, August 2005.

[5] T. Korakis, O. Ercetin, S. Krishnamurthy, L. Tassiulas, and S. Tripathi, Link Quality based Association Mechanism in IEEE 802.11h compliant Wireless LANs, in Workshop on Resource Allocation in Wireless Networks (RAWNET), April 2005.

[6] A. Kumar and V. Kumar, Optimal Association of Stations and APs in an IEEE 802.11 WLAN, in Proceedings of of the National Conference on Communications (NCC), IIT Kharagpur, January 2005.

[7] Sklar B., Rayleigh Fading Channels in Mobile Digital Communication Systems, IEEE Communications Magazine, Vol. 35 Iss.7, 1997.

[8] Coucheney P. , Touati,C., and Gaujal B. Fair and efficient user-network association algorithm for multitechnology wireless networks. In Proc. of the 28th conference on Computer Communications miniconference (INFOCOM), 2009.

[9] Bournez O, Cohen J., Stochastic Learning of Equilibria in Games: The Ordinary Differential Equation Method, submitted. see paper in http://webloria.loria.fr/ jcohen/i.php/Main/HDR

[10] Weibull, J.W., 1995. Evolutionary Game Theory. MIT Press, Cambridge, MA.

[11] Hofbauer, J. and Sigmund, K. (2003). Evolutionary game dynamics. Bulletin of the American Mathematical Society, 4:479–519.

[12] P. S. Sastry and V. V. Phansalkar and M. A. L. Thathachar, Decentralized Learning of Nash Equilibria in Multi-Person Stochastic Games With Incomplete Information, IEEE Transactions on Systems, Man, and Cybernetics, 1994, vol. 24, 769-777.