Dynamic Resource Allocation of Delay Sensitive Users Using Interactive Learning over Multi-carrier Networks

Hsien-Po Shiang, Wenchi Tu, and Mihaela van der Schaar

Abstract—In this paper, we study distributed solutions for joint power-spectrum resource allocation among delay sensitive users over multi-carrier networks. Our work differs from prior research in two ways. First, unlike prior works that only consider the impact of effective throughput of a user, our work also considers the source traffic characteristics of applications and applies queuing analysis to quantify the packet transmission delay, which is especially important to a delay sensitive user. Secondly, most prior works focus on the equilibrium of the decentralized solutions that usually require global information in a static network. Instead, we focus on interactive adaptation methods for users to dynamically maximize their expected utility based on the local observed information. We propose an interactive learning framework for users to adapt their power/channel selection to the wireless networks with interference coupling among users. Our simulation results show that the proposed interactive learning approach with information exchange among users improves the total utility significantly.

***Index Terms*— dynamic resource allocation; power control; multi-carrier network; interactive learning; information exchanges.**

I. INTRODUCTION

Joint power and spectrum resource allocation research has been investigated extensively in multi-carrier wireless networks [1][2]. It is well-known that in a multi-user system, the resource allocation problem is complicated since the mutual wireless interference among users results in a nonconvex optimization problem. In the interference channels, the water-filling approaches [1][2] can only provide power allocation that is asymptotically optimal when the number of users is large. Alternatively, a centralized optimization approach provided in [3] proposes a dual method to solve the nonconvex optimization problem for maximizing the overall throughput. However, centralized solutions are not desirable in practice, since the centralized solutions tend to be complex and not scalable as the network size grows. Moreover, the centralized solutions require the propagation of global information back and forth to a common coordinator, thereby incurring delay that may be unacceptable for delay sensitive applications. Hence, it is essential to study distributed solutions for such power-spectrum resource allocation problems.

Prior distributed power allocation research [4][5] focuses on constructing a power control game where each user possesses its own utility function. Various energy-efficient utility functions are developed. For instance, users maximize a ratio of throughput over the transmitted power (measured in bits/joule) instead of merely the throughput in [4][6]. However, for delay sensitive applications, these works only consider the impact of the effective throughput of a user, while ignoring the important impact of the source traffic characteristics of such applications. In this paper, we apply queuing analysis to quantify the packet transmission delay, which is impacted by both the effective throughput as well as the source traffic characteristics of delay sensitive users.

In addition, due to the informationally-decentralized nature of the wireless network, we approach this energy-efficient power control problem from a different angle. First of all, most of the game-theoretic works in wireless communication assume that users are able to obtain required global information to play in the game. However, in practice, users usually make decisions based on local observed information. Investigating the information availability and its impact on the resulting utility become an important issue in a wireless network. Moreover, unlike prior distributed power allocation research that focuses on the convergence to the Nash equilibrium [4][5], we emphasize the strategic learning approaches for users to adapt their frequency channel and power level selection to the dynamics of wireless networks. There are several reasons why we focus on the dynamic adaptation from the user side. First, even though the equilibrium exists, without sophisticated mechanism design [7], the equilibrium may only lead to a suboptimal result due to the myopic observation of the users. Moreover, in a dynamic wireless network, the wireless network environment and the source traffic may change before users transmission strategies converge to equilibrium. Hence, we study the strategic learning methods of a user to adapt its power/frequency channel selection to the dynamic wireless network with mutual interference coupling among users.

In this paper, focusing on the dynamic adaptation from the user side, we propose an interactive learning framework for distributed power control of delay sensitive users in multi-carrier system. We consider not only the impact of the effective throughput over the wireless network, but also the source traffic characteristics including the source rates and the delay deadlines of the applications. In the proposed framework, wireless users are able to learn the behaviors of their major interference sources and strategically adapt their power/frequency channel selections. Based on the observed information, different types of learning approaches can be

Manuscript received Sep. 21 2007. This work was supported by ONR project.

adopted. Based on the impact on the expected utility function, different learning efficiencies can be quantified in different network scenarios.

This paper is organized as follows. In Section II, we present our network settings and formulate the distributed power-spectrum resource allocation problem with interference coupling in the wireless network. We next introduce our interactive learning framework for multiple delay sensitive users in Section III. The adopted interactive learning approach is driven by the observed information of a user. In Section IV, we characterize the observed information. Based on the observed information, we introduce two types of interactive learning approaches. Section V gives the simulation results and Section VI concludes the paper.

II. Network Settings and Problem Formulation

A. PHY/MAC layer network settings

We assume that there are V network users in the wireless network (v =1,…, V), which are organized in an ad-hoc manner. Each network user v has a source node n_v^s and a destination node n_v^d that can establish a direct communication connection in order to transmit a delay-sensitive data stream, i.e. $v = \{n_v^s, n_v^d\}$. Assume that the users in the wireless network may utilize multiple frequency channels for transmission. Let $\mathcal{F}$ represent a set of all possible frequency channels. Without losing generality, we assume that only frequency channels in the set $\mathcal{F}_v \in \mathcal{F}$ are available to the user v . Each user tends to maximize its own utility function by selecting appropriate frequency channels and transmitted power levels in the selected channels. Assume that a network user v transmit its application through only one [1] of the available frequency channels $f_v \in \mathcal{F}_v$ with a power level P_v below the maximum power limit $P_v^{\max}$, i.e. $0 \le P_v \le P_v^{\max}$. In this paper, we assume that the transmitter can only select discrete power levels in a set $\mathcal{P}_v$. Hence, we define the action of a user v as $A_v = [f_v, P_v] \in \mathcal{A}_v = \mathcal{F}_v \times \mathcal{P}_v$. Let $\mathbf{F} = [f_v, v = 1,\ldots,V]$ and $\mathbf{P} = [P_v, v = 1,\ldots,V]$ represent the selected frequency vector and power vector across all the users.

The Signal-to-Interference-Noise Ratio (SINR) experienced by a user v in frequency channel f_v depends on the user's action A_v and the action of all the other users, denoted as A_{-v}. Hence,

$$\gamma_v(A_v, A_{-v}) = \gamma_v(\mathbf{F},\mathbf{P}) = \frac{G_{vv}(f_v)P_v}{N_0 W_{f_v} + \sum_{v' \neq v, f_{v'} = f_v} G_{v'v}(f_v)P_{v'}}, \quad (1)$$

where $G_{v'v}(f_v)$ represents the channel gain from the transmitter (the source node $n_{v'}^s$) of the user v' to the receiver (the destination node n_v^s) of the user v , which is related to the distance of the two nodes and channel characteristics. To calculate the SINR value, the channel gain matrices $\mathbf{G}(f) = [G_{v'v}(f)]_{V\times V}$ for all frequency channels $f \in \mathcal{F}_v$ needs to be determined.

Let T_v and p_v represent the physical transmission rate and packet error rate of user v using the frequency channel f_v. Denote $B_v = T_v(1 - p_v)$ as the corresponding effective throughput. T_v and p_v are estimated by the MAC/PHY layer link adaptation [8], which can be modeled as sigmoid functions of the SINR $\gamma_v(A_v, A_{-v})$:

$$p_v(f_v, \gamma_v(A_v, A_{-v})) = \frac{1}{1 + \exp(\zeta(\gamma_v(A_v, A_{-v}) - \delta))}, \quad (2)$$

$$B_v(f_v, \gamma_v(A_v, A_{-v})) = \frac{T_v(f_v)}{1 + \exp(-\zeta(\gamma_v(A_v, A_{-v}) - \delta))}, \quad (3)$$

where ζ and δ are empirical constants corresponding to the modulation and coding schemes for a given packet length L_v for user v .

B. Queuing model for delay sensitive applications

Assume that a delay sensitive application with delay deadline d_v is sent by the user v through the network using the average input rate R_v (bits/sec). Assume that the user v maintains a queue with an infinite buffer size in the application layer. We model the packet arrival process using a Poisson process. The packet arrival rate is hence defined as $\lambda_v = R_v / L_v$ (packet/sec). The packet protection scheme at the MAC is assumed to be similar to the Automatic Repeat Request protocol in IEEE 802.11 networks [9]. Then, the transmission time of a packet can be modeled as a geometric distribution [10]. For simplicity, we approximate the queuing model as M/M/1 queue with the service rate $\mu_v(A_v, A_{-v}) = B_v(f_v, \gamma_v(A_v, A_{-v})) / L_v$ (packet/sec). We denote the end-to-end delay of transmitting the delay sensitive application through the network as $D_v(A_v, A_{-v})$. The average end-to-end delay can be obtained by

$$E[D_v(A_v, A_{-v})] = \frac{1}{\mu_v(A_v, A_{-v}) - \lambda_v}, \text{ for } \mu_v(A_v, A_{-v}) > \lambda_v. \quad (4)$$

Using the M/M/1 queuing model, the probability that the packet of the user v can be received before the delay deadline d_v is

$$\mathrm{Prob}\{D_v(A_v, A_{-v}) \le d_v\} = \begin{cases} 1 - \exp(-\dfrac{d_v}{E[D_v(A_v, A_{-v})]}), & \text{for } \mu_v(A_v, A_{-v}) > \lambda_v \\ 0 & \text{, otherwise} \end{cases}, \quad (5)$$

C. Utility function

The behaviors of the users highly depend on their utility functions. In this paper, as in [4][6], we assume the users tend to maximize their energy-efficient utility functions (measured in bits/joule). The difference is that we consider the end-to-end packet loss due to the expiration of the delay deadline for delay sensitive applications, which is not only impacted by the effective throughput but also the source traffic characteristics. The utility function of a user v is defined as

[1] In [4], it is shown that the energy-efficient utility function (measured in bits/joule) is maximized when the user only transmits over its "best" frequency channel. The "best" frequency channel is the channel that requires the least power P_v to achieve a certain target SINR γ_v^{tar}.

$$u_v(A_v, A_{-v}) = \frac{\lambda_v \times \text{Prob}\{D_v(A_v, A_{-v}) \le d_v\}}{P_v}. \quad (6)$$

This utility function reflects the expected number of bits that are successfully received (rather than transmitted) per joule of energy consumed. Figure 1 illustrates the utility function of a user v using different power $0 \le P_v \le P_v^{\max}$ in a selected frequency channel f_v with fixed interference.

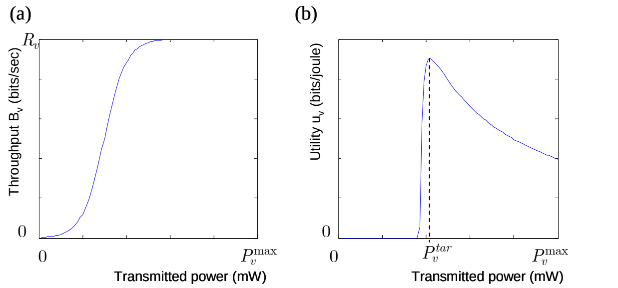

Fig. 1. (a) Throughput B_v vs. P_v in a selected frequency channel f_v with fixed interference. (b) Utility u_v vs. P_v in a selected frequency channel f_v with fixed interference.

Figure 1 shows that the utility function remains 0 unless the transmitted power is high enough to support a sufficient throughput. A user v needs to maintain a sufficient throughput $B_v(f_v, \gamma_v(A_v, A_{-v})) > R_v$ so that the service rate of the queue $\mu_v(A_v, A_{-v}) > \lambda_v$ to keep the probability $\text{Prob}\{D_v(A_v, A_{-v}) \le d_v\} > 0$. Hence, the operational region each of a delay sensitive user needs to satisfy the condition $B_v(f_v, \gamma_v(A_v, A_{-v})) > R_v$. Note that there exists an optimal power level $P_v^{tar}(f_v)$ to maximize the utility function.

D. Problem formulation

Conventionally, the optimization problem of a user v can be formulated as:

$$A_v^{opt} = [f_v^{opt}, P_v^{opt}] = \arg\max_{A_v \in \mathcal{A}_v} u_v(A_v, A_{-v}). \quad (7)$$

As mentioned before, due to the interference coupling in the PHY/MAC layer, the utility u_v depends on the action A_v taken by the user v and also on the actions of other users A_{-v}. In order to select an optimal action A_v^{opt}, user v needs to evaluate its utility $u_v(A_v, A_{-v})$ given the action A_{-v}. Note that the action A_v^{opt} in equation (7) is optimal only when the action A_{-v} is fixed. However, in practice, after the action A_v^{opt} is selected by user v, the actions of other users A_{-v} will also change due to the mutual interference coupling in the wireless network. Hence, interactive adaptation over time is necessary for user v to keep maximizing its utility. Moreover, the optimization solution is only valid when other users do not change their actions in a static network condition. In a dynamic wireless environment with interference coupling, equilibrium may not exist and the network can change faster than the decision making of a user before it converges. Dynamic adaptation to the changing network is more important in practice than just identifying the equilibrium. Hence, we focus on how users make their own decisions based on the information they observed. Note that the observed information is usually incomplete (either due to insufficient measurement of SINR, buffer length, or incomplete information exchange from localized users due to the informationally decentralized nature of the wireless networks). In summary, the action A_v^t at time t is determined based on the observed information o_v^{t-1}, which implicitly model the actions A_{-v}. The problem is now reformulated as:

$$A_v^t = [f_v^t, P_v^t] = \arg\max_{A_v \in \mathcal{A}_v} E[u_v(A_v, o_v^{t-1})]. \quad (8)$$

The question is what information should be observed and how to model the actions A_{-v} of other users from the observed information o_v^{t-1} to evaluate $E[u_v]$. For example, the evaluation of the utility $E[u_v]$ can be performed by observing the SINR $\gamma_v(A_v, A_{-v})$ in PHY/MAC layer or by direct information exchange from other users. In Section IV, we will discuss the observed information o_v^{t-1} in more details. Figure 2 illustrates the mutual interference coupling in the dynamic wireless network. In the next section, we provide an interactive learning framework for a delay sensitive user v to model A_{-v} and make such adaptive decisions.

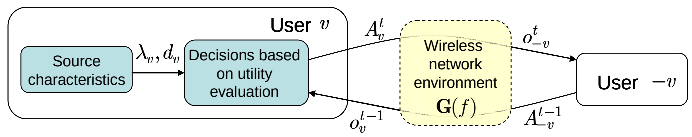

Fig 2. The mutual interference coupling in the dynamic wireless network at time t.

III. Interactive Learning Framework for Dynamic Resource Allocation for Delay Sensitive Users

The goal of the user v in our interactive learning framework is to adapt the action $A_v^t = [f_v^t, P_v^t]$ given the observed information o_v^{t-1}. We first present two propositions to simplify the learning problem for delay sensitive users with the utility defined in Section II.*C*.

<u>*Proposition 1:*</u> Assume the users that select a certain frequency channel f form a set $\mathbf{N}_f$. The target SINR value $\gamma_v^{tar}(f), v \in \mathbf{N}_f$ that jointly maximizes $u_v(f), v \in \mathbf{N}_f$ is the unique positive solution of $\gamma \dfrac{\partial B_v(f,\gamma)}{\partial \gamma} = \dfrac{L_v}{d_v}(F_v(f,\gamma) - 1)$, where

$$F_v(f,\gamma) \equiv \exp(B_v(f,\gamma)\frac{d_v}{L_v} - \lambda_v d_v). \quad (9)$$

<u>*Proof:*</u> Given the channel model $B_v(f,\gamma)$ for the frequency channel f in equation (3), user v, $v \in \mathbf{N}_f$ can apply queuing analysis with the application characteristics R_v, L_v and d_v. From equation (4) and (5), we have $\text{Prob}\{D_v \le d_v\} = 1 - \dfrac{1}{F_v}$. The optimality condition of $\dfrac{\partial u_v}{\partial P_v} = 0$ becomes $-P_v \dfrac{\partial}{\partial P_v}(\dfrac{1}{F_v}) = (1 - \dfrac{1}{F_v})$. The left hand side can be derived as $\gamma_v \dfrac{\partial B_v}{\partial \gamma_v}\dfrac{d_v}{L_v}\dfrac{1}{F_v}$, since $P_v \dfrac{\partial \gamma_v}{\partial P_v} = \gamma_v$. By

multiplying F_v to both sides, we have the optimality condition in Proposition 1 and the corresponding γ_v^{tar} that maximizes the utility function u_v. □

Proposition 1 suggests that user v using the frequency channel f should adapt the target power level $P_v^{tar}(f)$ to the observed interference to support the target SINR value $\gamma_v^{tar}(f)$. If the target SINR $\gamma_v^{tar}(f)$ requires a power higher than $P_v^{\max}$ (the interference in the frequency channel is too high), $P_v^{tar}(f) = P_v^{\max}$. Next, we further determine the frequency selection of the user v given the observed interference in channel $f \in \mathcal{F}_v$.

Proposition 2: Let $F_v^{tar}(f) = F_v(f, \gamma_v^{tar}(f))$. Given the observed interference and the corresponding target $P_v^{tar}(f)$, the optimal action A_v^* of a user v is

$$f_v^* = \arg\min_{f \in \mathcal{F}_v}\{P_v^{tar}(f) \times \frac{F_v^{tar}(f)}{F_v^{tar}(f) - 1}\}$$
$$\text{and } P_v^* = P_v^{tar}(f_v^*). \quad (10)$$

Proof: From Proposition 1, maximizing $u_v = \dfrac{\lambda_v}{P_v}(1 - \dfrac{1}{F_v})$ leads to the solutions in equation (10). □

Modeling the action A_{-v}, a user v can estimate the observed interference, and thereby selects the frequency channel f_v^* and power level P_v^* iteratively to match the SINR value $\gamma_v(A_v, A_{-v})$ to the target SINR $\gamma_v^{tar}(f_v^*)$. In this paper, we rely on the strategic learning approaches [11] to model the action A_{-v}, which is driven from the observed information o_v^{t-1}. Assume $S_v^t(A)$ represents the probability that a user v takes A as its action at time t. The strategic learning enables users to learn their transmission strategies $\mathbf{S}_v^t = [S_v^t(A), \text{ for } A \in \mathcal{A}_v] \in \mathcal{S}_v$, where $\mathcal{S}_v$ is a set of probability distributions over all feasible actions $A \in \mathcal{A}_v$.

Figure 3 provides the block diagram of a two-user case as an example of our interactive learning framework. The strategic learning approach is driven by the observed information o_v^{t-1}, which will be discussed in detail in the next section.

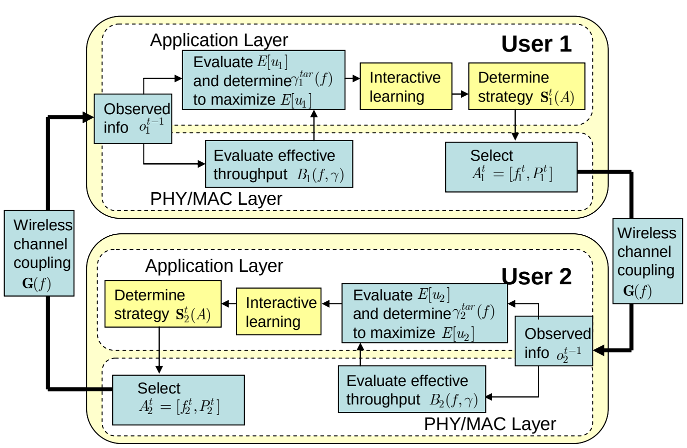

Fig. 3. System block diagram for the interactive learning for the dynamic power level/frequency channel adaptation with two users.

IV. Observed Information for Interactive Learning

A user can apply different strategic learning methods to evaluate the expected utility $E[u_v(\mathbf{S}_v^t, o_v^{t-1})]$ in a dynamic network environment. Different information types drive the learning approach of the users that result in different expected utility. In our problem settings, we define the "entire" history of the information in the network at time slot t as

$$h^t = \{\gamma_v^s, \mathbf{G}_v^s, A_v^s, E[D_v^s], \text{for } v = 1,\ldots,V, s = 0,\ldots,t\}.$$

From the history, the observed information of a user v at time t is a subset of the entire history $o_v^t \subseteq h^t$. Note that the user v at a certain time slot s can observe the information in two types: private information $o_v^{t,priv} = \{\gamma_v^s, E[D_v^s], s = 0,\ldots,t\}$ or public information $o_{-v}^{t,pub} = \{\mathbf{G}_u^s, A_u^s, \text{for } u \in \Omega_v, s = 0,\ldots,t\}$, where Ω_v represents a set of neighbor users of user v. Note that a user can also observe a subset of $o_v^{t,priv}$ or $o_{-v}^{t,pub}$ (not necessarily observe information at each decision time slot). Depending on the type of information, two classes of strategic learning approaches can be adopted:

1) Payoff-based learning without information exchange from other users

Without information exchange from other users, a user learns its transmission strategy from the experienced payoff (experienced utility values). The observed private information can be $o_v^{t,priv} = \{\gamma_v^s, s = 0,..,t\}$ or $o_v^{t,priv} = \{E[D_v^s], s = 0,\ldots,t\}$ to evaluate the expected utility. The payoff history reinforces the best response transmission strategy [11]:

$$\mathbf{S}_v^t = \arg\max_{\mathbf{S}_v \in \mathcal{S}_v} E[u_v(\mathbf{S}_v, o_v^{t-1,priv})]. \quad (11)$$

$$A_v^t = Rand(\mathbf{S}_v^t), \quad (12)$$

where $Rand(\mathbf{S}_v^t)$ represents a random selection based on the probability distribution $\mathbf{S}_v \in \mathcal{S}_v$. This class of learning is referred as the reinforcement learning [11], which is a type of payoff-based learning. Note that the actions of other users are embedded in the measurement of the experienced SINR values γ_v^s in PHY/MAC layer or expected delay $E[D_v^s]$ in application layer.

2) Model-based learning with information exchange from other users

To model the other users' behavior when public information $o_{-v}^{t,pub}$ is exchanged, a user directly counts the frequency with which other users select certain actions [11]. Denote the strategy set of other users as $\mathbf{S}_{-v}^t = \{\mathbf{S}_u^t, u \in \Omega_v\}$. Based on $\mathbf{S}_{-v}^{t-1}$ at time t, a user selects the best response strategy as

$$A_v^t = \arg\max_{A_v \in \mathcal{A}_v} E[u_v(A_v, \mathbf{S}_{-v}^{t-1})]. \quad (13)$$

This class of learning is referred as fictitious play [11], which is a model-based learning.

Due to the space limit, we only briefly describe the insight of the two learning approaches and omit the implementation details of using the interactive framework in Figure 3.

V. Simulation Results

We simulate an ad hoc wireless network shown in Figure 4 with 5 users (with distinct source and destination pairs) and 3 frequency channels. The frequency channels are accessible for all the users, i.e. $\mathcal{F}_v = \mathcal{F}$, for all v. Each user can choose its power level P_v from a set $\mathcal{P} = \{20, 40, 60, 80, 100\}$ (mW). Hence, there are a total of 15 actions A_v (3 channel selections with 5 power levels). For the physical layer parameters, the channel gain between different network nodes can be modeled using $G_{vv'} = K_0 \times ({dis_0}/{dis_{vv'}})^\alpha$, where $dis_{vv'}$ represents the distance from the transmitter of the user v to the receiver of the user v', and $K_0 = 5 \times 10^{-4}$, $dis_0 = 10$, $\alpha = 2$ are constants. For the application layer parameters, we set the average packet length $L_v = 1000$ bytes, $R_v = 500$ Kbps, and $d_v = 200$ msec for all the users.

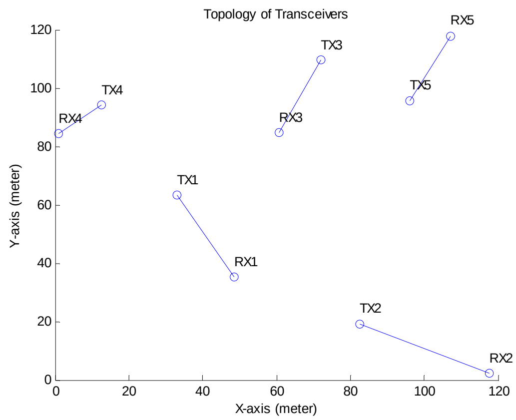

Fig. 4. Topology settings for the simulation

Figure 5 shows the simulation results of the total utility (summation over 5 users) versus the physical bandwidth T_v of frequency channels. Since the required rates of the applications are fixed to $R_v = 500$ Kbps, the total utility will saturate as the bandwidth increases. We simulate three different cases with different information types – 1) no learning, 2) reinforcement learning (the payoff-based learning), and 3) fictitious play (the model-based learning). Each simulation result is averaged over 500 time slots in the dynamic network settings with mutual interference in Equation (1). The results show that the reinforcement learning with private information $o_v^{t,priv}$ outperforms the no learning case (actions are iteratively selected in a myopic manner as in [4]). Moreover, with public information exchange $o_{-v}^{t,pub}$, users are able to exploit the resource more efficiently, since the users can one step further predict the behaviors of other interference sources in the network. The total utility can be significantly improved compared to the other two cases.

VI. Conclusion

In this paper, we provide an interactive learning framework for users to adapt their frequency channel and power level selections in multi-carrier wireless networks. The results show that depending on the observed information, interactive learning significantly improves the performance of delay sensitive users. Note that in this paper, we assume the available information is observed at each decision time slot. However, in practice, the information can be observed with other observation frequencies. The adaptive information observation and its impact on the interactive learning performance form an interesting area to look at.

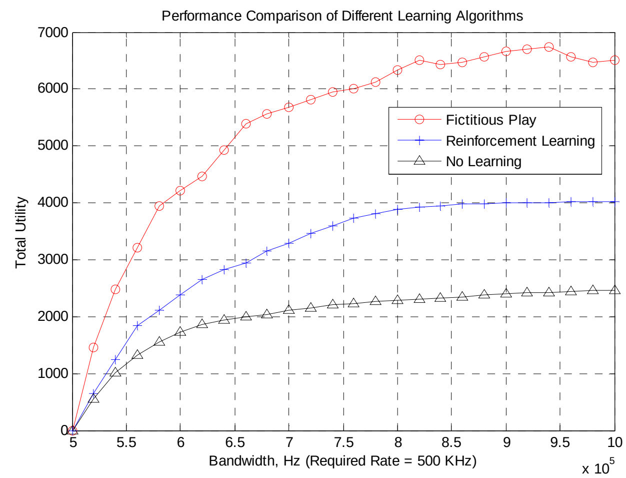

Fig. 5 Performance comparison of different learning algorithms

References

[1] S. Vishwanath, S. A. Jafar, and A. Goldsmith, “Adaptive resource allocation in composite fading environments,” *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pp. 1312-1316, San Antonio, TX, Nov 2001.

[2] Z. Shen, J. G. Andrews, and B. L. Evans, “Optimal power allocation in muli-user OFDM systems,” *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pp. 337-341, San Francisco, CA, Dec 2003.

[3] W. Yu, R. Lui, “Dual Methods for Nonconvex Spectrum Optimization of Multi-carrier Systems,” *IEEE Transactions on Communications*, vol. 54, no. 7, July 2006.

[4] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz, “A game-theoretic approach to energy-efficient power control in multi-carrier CDMA systems,” *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 24, pp. 1115-1129, June 2006.

[5] W. Yu, G. Ginis, and J. M. Cioffi, “Distributed multi-user power control for digital subscriber lines,” *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 20, pp. 1105-1115, June 2002.

[6] D. J. Goodman and N. B. Mandayam, “Power control for wireless data,” *IEEE Personal Communications*, vol. 7, pp. 48-54, Apr 2000.

[7] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, “Pricing and power control in a multi-cell wireless data network,” *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 19, pp. 1883-1892, Oct 2001.

[8] D. Krishnaswamy, “Network-assisted Link Adaptation with Power Control and Channel Reassignment in Wireless Networks,” *3G Wireless Conference*, pp. 165-170, 2002.

[9] T. S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, 2002.

[10] A. G. Konheim, “A Queuing Analysis of Two ARQ Protocols,” *IEEE Transactions on Communications*, vol. com-28, no. 7, July 1980.

[11] H. P. Young, *Strategic Learning and its Limits*, Oxford University Press, NY 2004.