Int. J. Electron. Commun. (AEÜ) 67 (2013) 817–826

Contents lists available at SciVerse ScienceDirect

International Journal of Electronics and Communications (AEÜ)

journal homepage: www.elsevier.com/locate/aeue

A stochastic game framework for joint frequency and power allocation in dynamic decentralized cognitive radio networks

Xiu Liu [a,b], Guoru Ding [a,*], Yang Yang [a], Qihui Wu [a], Jinlong Wang [a]

[a] College of Communications Engineering, PLA University of Science and Technology, Nanjing, Jiangsu 210007, China
[b] Comprehensive Training Base of Guangzhou Military Area, Guilin, Guangxi 541002, China

ARTICLE INFO

Article history:
Received 6 December 2012
Accepted 10 April 2013

Keywords:
Cognitive radio networks
Frequency allocation
Power control
Stochastic game
Multi-agent learning

ABSTRACT

Cognitive radio networks (CRNs) have been recognized as a promising solution to improve the radio spectrum utilization. This article investigates a novel issue of joint frequency and power allocation in *decentralized* CRNs with *dynamic* or *time-varying* spectrum resources. We firstly model the interactions between decentralized cognitive radio links as a stochastic game and then proposed a strategy learning algorithm which effectively integrates multi-agent frequency strategy learning and power pricing. The convergence of the proposed algorithm to Nash equilibrium is proofed theoretically. Simulation results demonstrate that the throughput performance of the proposed algorithm is very close to that of the centralized optimal learning algorithm, while the proposed algorithm could be implemented distributively and reduce information exchanges significantly.

1. Introduction

1.1. Background and open issues

Cognitive radio networks (CRNs) have been recognized as a promising solution to improve the radio spectrum utilization [1]. Decentralized CRNs, in which no centralized entity exists, recently have been of great interest due to high feasibility, scalability, infrastructure-independence, etc. [2–5].

To enable decentralized CRNs, one of the main challenges is how to effectively utilize the spectrum resource and control the transmission power of distributively deployed CRUs. The difficulty mainly results from the following two facts: on one hand, in a decentralized CRN, the resource allocation of each CRU is coupled with each other, i.e., the selection of frequency band and transmit power of one CRU affects the selection of other CRUs. On the other hand, in a CRN without any centralized coordinator, it is intractable for each CRU to obtain enough information about the environment state and the strategy selection of other CRUs.

1.2. Related work and motivation

To tackle the above open issues, game theory is a powerful theoretical tool that is good at studying interactions among self-decision individuals [6]. Therefore, the research of game theory-based wireless resource allocation has attracted worldwide research interests (see, e.g. [7–9]). It is noted that most existing studies focus on the non-CRN or static wireless environment and thus the resource allocations in those cases are usually modeled as static games, while in dynamic CRNs, the spectrum resources (i.e., the of PU spectrum occupation) may change randomly over time, space and frequency [10,11].

Stochastic game (SG) [12,13] is a promising tool to study the interactions among self-decision individuals in a *dynamic* environment. Recently, there are some studies applying SG to wireless networks modeling and analysis (WNMA) [14–17]. Specially, in [14], a SG-based wireless resource allocation framework is firstly proposed, where the best-response learning algorithm is designed for the users to predict the impact of current actions on future performance. In [15], the transmission rate adaptation problem of each CRU is formulated as a general-sum Markovian dynamic game with a delay constraint. In [16], a SG framework is proposed for anti-jamming defense. Very recently, a robust distributed power control algorithm is proposed in [17] based on repeated stochastic game with learning automaton. Although SG-based WNMA has been researched extensively recently, to the best knowledge of the authors, the issue of tailoring an effective SG model for*joint frequency and power allocation in dynamic decentralized CRNs* is still underdeveloped.

From a theoretical perspective, the existing SG models can be classified into three groups: fully cooperative, fully competitive and mixed SGs. Fully cooperative SGs mean that all agents have the same return such as Team-Q [18]. Fully competitive SGs require that the sum return of all agents is zero such as Minmax-Q [19].

* Corresponding author.
E-mail address: dingguoru@gmail.com (G. Ding).

1434-8411/$ – see front matter © 2013 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.aeue.2013.04.002

Mixed SGs do not impose any requirements on returns such as Nash-Q [20], CE-Q [21], and Asymmetric Q [22]. We find there are three major limitations restricting a direct application of existing SG models to the problem of joint frequency and power allocation in dynamic decentralized CRNs: (1) *Overstrict requirements*: for example, the unique optimal joint actions is required in Team-Q and the existence of saddle point is required in Nash-Q or CE-Q; (2) *Too much information exchanges*: all existing SGs have an explicit or implicit assumption that each agent know the perfect information of other agents, which needs global information exchanges in practice; and (3) *Unacceptable implementation complexity*: strategies in existing SGs are discrete. One possible way is adopting discrete power levels instead of continuous transmission power, but the size of state-action space will be very huge, which brings in unacceptable implementation complexity (see Section 4.2 for details).

In a nutshell, a good algorithm for the problem of interest in this article should contain the following four merits: (1) *Joint action learning*: in multi-agent environment, due to actions of all agents affecting each other, joint action learning can avoid conflict effectively and realize the high efficiency allocation; (2) *Small state-action space*: as mentioned before, convergence speed is determined by the size of state-action space; (3) *Limited information exchanges*: information exchanges are great burden to distributed CR network; and 4) *Moderate constraints on Q function*: too many constraints on Q function will restrict the practical application.

1.3. Contributions

Motivated by the observations above, in this article we investigate the issue of joint frequency allocation and power control in dynamic decentralized CRNs using stochastic game. The main contributions are summarized as follows:

- *A joint frequency and power allocation stochastic game (JFPA-SG) theoretical framework is proposed* for dynamic decentralized CRNs. The proposed JFPA-SG framework provides an elegant mathematical model to characterize the evolutionary dynamics of the environment state of the primary network and to study the coupled strategy learning for the CRN with multiple interactive CRUs. In addition, this JFPA-SG framework is also able to depict the heterogeneous characteristic of the spectrum resource availability for different CRUs by considering that CRUs located at different locations may experience different occupancy behaviors of the PUs (see Fig. 1(a)).
- *A multi-agent frequency learning with power pricing (MAFLPP) algorithm is developed* to learn the discrete frequency strategy and continuous power strategy jointly and reduce the information exchanges among agents at the same time. With multi-agent frequency learning, distributed agents learn the optimal frequency strategies in dynamic environment. With power pricing, agents sharing the same frequency are able to control their transmission power to improve their overall throughput without any information exchanges.
- *A practical implementation scheme of application the proposed MAFLPP algorithm into large scale CRNs is also designed*, which resolves the common computational complex problem of multi-agent reinforcement learning (MARL) algorithms when they are implemented in large scale networks.
- *In-depth numerical simulations are provided* to demonstrate the effectiveness of the proposed MAFLPP algorithm. It is observed that the proposed MAFLPP learning algorithm, which is a fully decentralized algorithm with very limited information exchanges, significantly outperforms the random allocation algorithm and single agent reinforcement learning algorithm, and obtains comparable throughput performance with the optimal centralized learning algorithm.

1.4. Organization and notations

The rest of this article is organized as follows. Section 2 presents the system model and problem statement. JFPA-SG is proposed in Section 3. In Section 4, we develop the MAFLPP algorithm. The convergence of the MAFLPP algorithm is proofed in Section 5. In Section 6, we present the simulation results, followed by conclusions in Section 7.

To facilitate the readers, some key abbreviations in this article are summarized as follows: cognitive radio networks (CRNs), cognitive radio users (CRUs), cognitive radio links (CRLs), primary users (PUs), stochastic game (SG), joint frequency and power allocation stochastic game (JFPA-SG), reinforcement learning (RL), signal agent reinforcement learning (SARL) and multi-agent frequency learning with power pricing (MAFLPP).

2. System model and problem statement

We consider a decentralized dynamic CRN consisting of M primary users (PUs) and N distributed cognitive radio links (CRLs). We assume that each PU has a licensed channel and each CRL corresponds to a pair of CRU transceivers. An example scenario with $M=2$, $N=3$ is depicted in Fig. 1. As shown in Fig. 1(a), the circles around PUs stand for their interference regions, hence CRL1 is in the interference range of PU1, CRL3 is in the interference range of PU2 and CRL2 is in the interference ranges of both PU1 and PU2. Fig. 1(b) shows the evolutionary dynamics of the occupying states of PUs. Let $s_{i,k}=0$, $i\in\{1, 2\}$, $k\in\{1, 2, \ldots, t\}$ denote that channel i is idle in the kth time slot and $s_{i,k}=1$ denote channel i is occupied in that time slot. We further assume that the occupying state of each PU on its licensed channel follows a discrete time Markov process and take the joint occupying state $(s_{1,k}, s_{2,k})$, $k\in\{1, 2, \ldots, t\}$ as the environment state of the CRN. We consider that the occupying states of PUs are independent with each other and the state transition probability of PU i or channel i is $P(s_{i,k}, s_{i,k+1})$, thus the joint state transition probability of PUs can be given as

$$P\left[(s_{1,k}, s_{2,k}), (s_{1,k+1}, s_{2,k+1})\right] = P(s_{1,k}, s_{1,k+1})P(s_{2,k}, s_{2,k+1}), \qquad (1)$$

where k indicates the time slot.

For a given environment state of the network as illustrated in Fig. 1, the optimal frequency allocation is straightforward. For instance, if the state of a time slot is $(s_{1,1}, s_{2,1})=(0, 1)$, which means channel 1 is idle and channel 2 is occupied by PU2, the best frequency option of CRL2 and CRL3 is channel 1 regardless of CRL1's actions. Accordingly, CRL1 should select channel 2. As CRL2 and CRL3 select the same channel, effective power control is needed to mitigate mutual interference. Due to the state dynamic of the channels, the best frequency allocations for CRLs in different time slots should be different. Fig. 1(c) provides the best frequency allocations of the three CRLs in the CRN, which corresponds to the evolutionary dynamics of the occupying states of PUs shown in Fig. 1(b).

It is noted that apparently we can easily given the optimal frequency allocation for the CRN as illustrated in Fig. 1 at each time slot, however, there is an implicit assumption that it needs a centralized entity such as base station (BS) who knows the perfect information of the whole network and the exact future state of primary channels before it making decisions.

Consequently, to design the optimal frequency allocation, not mention to joint frequency and power allocation, is a very difficult problem for decentralized dynamic CRNs. To tackle this challenge, while considering the dynamic and heterogeneous characteristic