Jamming-Resilient Wideband Cognitive Radios with Multi-Agent Reinforcement Learning

Mohamed A. Aref, Communications and Information Sciences Laboratory (CISL), Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, USA
Sudharman K. Jayaweera, Communications and Information Sciences Laboratory (CISL), Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, USA

ABSTRACT

This article presents a design of a wideband autonomous cognitive radio (WACR) for anti-jamming and interference-avoidance. The proposed system model allows multiple WACRs to simultaneously operate over the same spectrum range producing a multi-agent environment. The objective of each radio is to predict and evade a dynamic jammer signal as well as avoiding transmissions of other WACRs. The proposed cognitive framework is made of two operations: sensing and transmission. Each operation is helped by its own learning algorithm based on Q-learning, but both will be experiencing the same RF environment. The simulation results indicate that the proposed cognitive anti-jamming technique has low computational complexity and significantly outperforms non-cognitive sub-band selection policy while being sufficiently robust against the impact of sensing errors.

KEYWORDS

Anti-Jamming, Multi-Agent Reinforcement Learning, Q-Learning, Stochastic Game, Wideband Autonomous Cognitive Radios

1. INTRODUCTION

The cognitive radio (CR) concept emerged as an evolution of software-defined radio (SDR) whose original purpose was to address the issue of interoperability and the ability to incorporate new signal processing algorithms (Jayaweera, 2014). In this context, the CR can be viewed as an SDR platform equipped with cognition abilities as shown in Figure 1. Just as human perception, the antennas in Figure 1 works as the sensory organs of the cognitive radios. On the other hand, the cognitive engine works as the human’s brain that may be able to acquire knowledge and make decisions (Jayaweera, 2014; Kinsner, 2009; Wang, 2014).

Operating over wideband spectrum find increasing relevance in aerospace, military and commercial communications applications. However, conventional transceivers may not be capable of operating with mobility over such a wideband spectrum. Wideband autonomous cognitive radios (WACRs) promise multi-mode, multi-band radios with the ability to capitalize on such wide spectrum opportunities. Furthermore, they have the ability of autonomous decision-making (Cervantes et al., 2013) and self-learning. They can optimally self-reconfigure to adapt to the user needs and surrounding RF environment in real-time (Jayaweera, 2014; Aref, Jayaweera, & Machuzak, 2017). Indeed, the key to such autonomous operation is the radio’s ability to sense and comprehend its operating environment.
One of the most challenging security threats in which WACRs can be a great asset is jamming attacks. Jamming is malicious signal transmissions generated by an outside source that aims to disrupt the reliable communications. In practice, however, there may be multiple WACRs simultaneously operating over the same spectrum range leading to a multi-agent environment in which each WACR will need to avoid both malicious jammer as well as the transmissions of other radios. This scenario may be modeled as a stochastic game, an extension of Markov Decision Processes (MDPs), in which interactions among different agents is considered (Aref, Jayaweera, & Machuzak, 2017). In this context, a WACR may use multiagent reinforcement learning (MARL) to solve the stochastic game by learning an optimal, or near-optimal, policy to keep its communication link unjammed (Aref, Jayaweera, & Machuzak, 2017; Schwartz, 2014).

MARL has previously been proposed in the literature for anti-jamming communications in cognitive radio (CR) networks. For instance, Lo and Akyildiz (2012) proposed a stochastic general-sum game for modeling the jammed control channels. The objective was to obtain an optimal control channel allocation strategy for CRs to avoid jamming attacks using Win-or-Learn-Fast (WoLF) principle (Bowling & Veloso, 2002). The approach in (Lo & Akyildiz, 2012) considered the effect of sensing errors, however it was limited only to control channels. Wang et al. (2011) used minimax Q-learning to find anti-jamming policies for secondary users (SUs) in multi-channel CR systems. The CR and the jammer in (Wang et al., 2011) were treated as two equally knowledgeable learning agents. One of the drawbacks of the proposed algorithms in (Wang et al., 2011) is that it assumed perfect sensing.

Gwon et al. (2013) formulated a competing stochastic game by dividing the network into two sub-networks: the ally network and the enemy network. The objective of each of the two sub-networks is to achieve the maximum spectrum utilization while jamming the opponent transmission as much as possible. Several reinforcement learning techniques were proposed: Minimax-Q, Nash-Q and Friend-or-Foe Q-learning. This work was extended in (Gwon et al., 2015) for the case of time-varying channel rewards. A new algorithm based on online convex programming was introduced in (Gwon et al., 2015) to obtain an optimal strategy that achieves the best steady-state channel rewards.

Most recently, a cognitive anti-jamming stochastic game model was proposed in (Aref & Jayaweera, 2017) to enable a WACR evading a jammer signal that sweeps across the spectrum of interest to the radio as well as transmissions of other radios. The advantage of the proposed model in (Aref & Jayaweera, 2017) was that it enables the WACR to switch its operating sub-band before getting jammed, compared to previously proposed anti-jamming techniques that switch the operating sub-band only after getting jammed. Although the performance of the proposed learning policy was

---

**Figure 1. Cognitive radio as an evolution of software defined radio (SDR)**
Design and Implementation of a Cognitive User-Support System for Skin Progress Analysis Using a Smart Phone
[www.igi-global.com/article/design-and-implementation-of-a-cognitive-user-support-system-for-skin-progress-analysis-using-a-smart-phone/88991?camid=4v1a](http://www.igi-global.com/article/design-and-implementation-of-a-cognitive-user-support-system-for-skin-progress-analysis-using-a-smart-phone/88991?camid=4v1a)

Cognitive Garment Panel Design Based on BSG Representation and Matching
Shuang Liang, Rong-Hua Li and George Baciu (2012). *International Journal of Software Science and Computational Intelligence* (pp. 84-99).
[www.igi-global.com/article/cognitive-garment-panel-design-based/67999?camid=4v1a](http://www.igi-global.com/article/cognitive-garment-panel-design-based/67999?camid=4v1a)