2022 IEEE Wireless Communications and Networking Conference (WCNC)

# Deep-Q Reinforcement Learning for Fairness in Multiple-Access Cognitive Radio Networks

Zain Ali, Zouheir Rezki*, Hamid Sadjadpour
Electrical and Computer Engineering Department, Baskin School of Engineering,
University of California, Santa Cruz, USA
Emails: zali4@ucsc.edu, zrezki@ucsc.edu, hamid@soe.ucsc.edu

2022 IEEE Wireless Communications and Networking Conference (WCNC) | 978-1-6654-4266-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/WCNC51071.2022.9771661

***Abstract*—This work presents a deep-Q reinforcement learning (DQ-RL) framework to achieve fairness in multi-access cognitive radio (CR) systems. The proposed framework provides fast solution and is robust to channel dynamics. Further, to remove the computational overhead and the burden to feedback thousands of weights from the secondary receiver (SR), we propose a solution where the process of learning is carried out at the secondary transmitters (STs). The simulations show that by using the proposed technique, a good level of fairness is achievable with an outage probability of the primary system less than 0.04. We also provide the comparison of the proposed technique with a brute-forcing optimization method, and show the fairness gain of the proposed framework compared to the rate maximization model.**

## I. Introduction

Cognitive radio (CR) enhances the spectral efficiency of the system by allowing secondary users (SUs) to share the spectral resources of the primary users (PUs) [1]. In simple CR networks, a single secondary transmitter (ST) shares the channel of a primary transmitter (PT). A system that can accommodate multiple STs on the same channel could offer better service. However, contrary to single SU case, when multiple STs share the same channel, the optimization problem becomes non-convex. For these problems, the conventional optimization frameworks can not be used since convergence is not guaranteed.

Machine-learning-based frameworks have been shown to provide fast solutions as compared to conventional techniques. However, the standard supervised machine learning techniques require a large set of data for training the model [2]. Further, if the parameters of the system change and re-training is required, we need to reproduce the training set for the new parameters [3]. Recently, different reinforcement learning (RL) frameworks have been shown to provide fast solutions and do not need any other techniques to provide the training data [4]. In RL, the training agents learn by taking actions and receiving a reward from the environment that quantifies the merit of the action.

Several works in the literature have proposed RL solutions for optimization problems. The authors in [5] designed a Q-learning framework for cooperative sensing in overlay CR networks. Then, to maximize the energy efficiency of the underlay CR system, a Q-learning framework was designed in [6]. The proposed technique provides good results, however the scheme only learns to optimize power allocation for the current channels and may need to retrain every time the channel gains change. The work in [7] proposed a robust DRL-based spectrum access framework that does not require retraining when the channel gains change. However, the solution in [7] requires centralized training for a distributed system, hence thousands of weights of the deep neural networks (DNNs) are transferred to all the users after the completion of the training phase.

The problem of achieving fairness in different communication systems has been considered broadly in the literature. The authors in [8] optimized resource allocations to achieve fairness in orthogonal frequency division multiple access systems. The problem of unfairness becomes more serious in interference limited systems where multiple STs are assigned the same channel. Thus, many STs may not be allowed to transmit to prevent outage and/or to maximize the system rate. Considering the problem of fairness in CR systems, the authors in [9] proposed a centralized resource allocation framework to enhance the fairness in rates of the STs. In the considered system, the STs are immune to interference from other STs. Thus, the problem of power allocation becomes convex. Further, the authors proposed a Lagrangian-dual-based solution which is not guaranteed to converge in the case when multiple STs are assigned the same channel because the problem becomes non-convex. The authors in [10] achieved fairness at the wireless powered STs subject to the interference threshold of the primary system. Considering multiple STs on the same channel, the authors in [11] optimized power allocation to achieve fairness in the rates of all the STs in the system. To solve the non-convex optimization problem, the authors employed a sequential-quadratic-programming-based iterative technique. The techniques in [8]- [11] are iterative in nature and require a lot of time to provide the solution. The issue with these approaches is that once the channel gains of the system change, we need to re-run the optimization.

In this work, which is a shorter version of [12], we propose a deep-Q reinforcement learning (DQ-RL) framework to achieve fairness in the rates of all the STs sharing a channel in an underlay CR setting. The proposed technique does not require the generation of a large data-set for training and provides fast solution. Further, once the training is complete, the framework provides excellent results for new values of channel gains,

* This material is based upon work supported by the National Science Foundation (CAREER) under Grant No. 2114779.