Large Project Proposal: Spectrum Utilization Through Reinforcement Learning

By Shane Flandermeyer and Geoffrey Dolinger

Why the topic is interesting:

Shane: Cognitive radar is a rapidly growing area of research in the radio frequency (RF) community. I have spent much of my time as an undergraduate researcher implementing radar signal processing algorithms on software-defined radio (SDR) systems, which have been a key enabler in the evolution of cognitive radar due to their flexibility and ability to adapt in real-time. Many adaptive waveform selection and resource management algorithms I have studied (e.g., sense-and-avoid) adapt to the system's surroundings but lack a learning component, resulting in degraded performance when they fail to recognize patterns in the data. Reinforcement learning techniques like Deep Q-Networks and actor-critic networks are promising alternatives to traditional adaptive algorithms that I would like to explore in this project.

Geoff: During my career with the Airforce, I spent the last 4 years as a section chief for an innovation group that focused on Command and Control(C2) applications. This domain has experienced major challenges in growing complexity of information gathering, data processing, abstraction of data into usable forms, and effective decision recommendations and/or actions. I completed my Masters in 2012 with a focus in neural networks as applied to control systems and have followed the field during my career. Through my studies I believe that the C2 domain would benefit immensely from cognitive radar and explainability. I have a personal interest in all aspects of the cognitive radar problem but specifically the exploration of Reinforcement Learning to train deep intelligent radar agents to improve performance and predictions given contested and noisy environments.

Questions/problems this project addresses:

Wireless devices utilize the electromagnetic spectrum to function, and the rise of commercial telecommunications technology such as the 4G/5G and the internet of things has caused the spectrum to become overly crowded. In most cases, these commercial devices are primary users of the spectrum, and secondary users such as radar can only access the spectrum if they do not interfere with other users. It has become increasingly necessary to develop cognitive radar systems which adapt their transmitted waveform to "conform" with their surroundings. At the same time, the development of multi-function antenna arrays has made it possible to perform multiple radar tasks on a single device. However, the resources needed to complete these tasks (e.g., computing power, antenna beams, time) are limited, so efficient resource management is required to fully utilize the available hardware.

How we will address the problem:

We plan to use reinforcement learning techniques such as Markov Decision Processes (MDP), Deep Q-Networks (DQN) and/or actor-critic networks to train a cognitive agent to address the spectrum sharing and resource management problems. We will first train this agent in a simulation environment using reinforcement learning libraries such as Keras and Matlab's reinforcement learning toolbox, then we will implement the best-performing algorithm with real RF hardware in the loop. We do not have access to a multi-function antenna array, so our resource management will only be examined in simulation. More details about our simulation/hardware implementations and performance metrics are given in the next section.

Project Activities:

- First, we will develop a simplified spectrum simulator in which we assume the cognitive radar is the secondary user. This simulator will discretize the available bandwidth into bins, where each bin is given a binary class label based on whether an emitter is already present in the band. This will allow us to find the best learning techniques without having to also develop an algorithm to identify available frequency bands. The congested space will be occupied by either a random distribution or simulated activity by spectrum users.
- In this stage, we will use two primary metrics to evaluate performance of each technique: the number of collisions with other users and the number of missed opportunities for transmission. A collision occurs when the radar occupies the same frequency at the same time as another user, and a missed opportunity occurs when there is an unused frequency band that the radar chooses not to occupy.
- Next, we will develop a more realistic spectrum model where the waveforms of other emitters are taken into consideration. In this model, it will be necessary to implement a spectrum sensing algorithm to determine if the bands are occupied. By simulating the actual waveforms, we will also be able to evaluate the reinforcement methods in terms of their radar operation performance. We can also use these radar performance metrics to train and evaluate resource management algorithms that determine the "optimal" time allocations for task scheduling.
- Having determined which of the learning methods offers the best joint radar/spectrum sharing performance, we will implement the algorithm on an NI USRP X310 software-defined radio system. We anticipate that this will be the most complex stage of the project, so we will initially use recordings of congested spectrum environments to ease the constraints needed for a fully adaptive real-time implementation. To our knowledge, there are no existing (open source) real-time implementations of a cognitive radar in the literature.

Topic Paragraphs:

<u>Note (Shane)</u>: The paragraph below is slightly modified from my initial submission. Here, I am focusing on reinforcement learning rather than deep learning.

Shane: Traditional adaptive radio frequency (RF) systems perform signal processing at the receiver to improve performance and increase efficiency, taking advantage of their surroundings without altering them. Cognitive RF systems, on the other hand, use information from the environment to optimize both receiver processing and the parameters of the transmitted signal/waveform. Thus, cognitive systems operate based on the perception-action cycle of cognition in which the system collects information about its surroundings (perception), then tailors its transmitted waveform to fit the needs of its mission (action). Since the transmission has an impact on the system's surroundings, the perception-action process is repeated, and a feedback loop is formed between the system and its environment. This framework raises several open research questions on how the system should utilize the information it collects, and which parts of the radio processing architecture should be optimized using cognition. Reinforcement learning has been introduced as a candidate solution to these problems, and deep reinforcement learning algorithms have been successfully used to improve the waveform design process, dynamically manage RF system resources, and share the electromagnetic spectrum between competing devices. I plan to explore existing reinforcement learning approaches that have been used to solve problems involving cognitive radio.

Geoff: The Department of Defense and specifically the US Air Force has had an extreme growth in challenges related to the domain of Command and Control (C2). C2 experiences major challenges in growing complexity of information gathering, data processing, abstraction of data into usable forms, and effective decision recommendations and/or actions. I intend to investigate, research, and implement methods for radar cognition to address some of these challenges. Radar Cognition is separated into 3 components: a) radar systems perceiving their environment, b) radar systems capable of taking actions to improve their perception or influence the environment, and c) radar systems action decisions based on their perceptions. Machine Learning (ML) serves as a promising set of tools to address multiple aspects of the radar cognition problem. I intend to explore existing machine learning methods applied to radar cognition as well as ML techniques that show promise for radar cognition problems. The primary topic I plan to utilize spatiotemporal ML analysis of radar data (e.g., CNN/RNN methods) in connection with radar actions (e.g., radar system controls) to train learning agents to improve detection or tracking of targets. Primarily, I will explore deep learning methods to train cognitive agents such as reinforcement learning and artificial evolution.

References:

S. Durst and S. Brüggenwirth, "Quality of service based radar resource management using deep reinforcement learning," 2021 IEEE Radar Conference (RadarConf21), 2021, pp. 1-6, doi: 10.1109/RadarConf2147009.2021.9455234.

P. Pulkkinen, T. Aittomäki, A. Ström and V. Koivunen, "Time Budget Management in Multifunction Radars Using Reinforcement Learning," 2021 IEEE Radar Conference (RadarConf21), 2021, pp. 1-6, doi: 10.1109/RadarConf2147009.2021.9455344.

C. E. Thornton, R. M. Buehrer, A. F. Martone and K. D. Sherbondy, "Experimental Analysis of Reinforcement Learning Techniques for Spectrum Sharing Radar," 2020 IEEE International Radar Conference (RADAR), 2020, pp. 67-72, doi: 10.1109/RADAR42522.2020.9114698.

C. E. Thornton, M. A. Kozy, R. M. Buehrer, A. F. Martone and K. D. Sherbondy, "Deep Reinforcement Learning Control for Radar Detection and Tracking in Congested Spectral Environments," in IEEE Transactions on Cognitive Communications and Networking, vol. 6, no. 4, pp. 1335-1349, Dec. 2020, doi: 10.1109/TCCN.2020.3019605.

M. Kozy, J. Yu, R. M. Buehrer, A. Martone and K. Sherbondy, "Applying Deep-Q Networks to Target Tracking to Improve Cognitive Radar," *2019 IEEE Radar Conference (RadarConf)*, 2019, pp. 1-6, doi: 10.1109/RADAR.2019.8835780.

S. Ak and S. Brüggenwirth, "Avoiding Jammers: A Reinforcement Learning Approach," *2020 IEEE International Radar Conference (RADAR)*, 2020, pp. 321-326, doi: 10.1109/RADAR42522.2020.9114797.

M. Altmann, P. Ott, N. C. Stache, D. Kozlov and C. Waldschmidt, "A Cognitive FMCW Radar to Minimize a Sequence of Range-Doppler Measurements," *2020 17th European Radar Conference (EuRAD)*, 2021, pp. 226-229, doi: 10.1109/EuRAD48048.2021.00065.