Multi-Agent Adversarial Attacks for Multi-Channel Communications

Extended Abstract

Juncheng Dong, Suya Wu, Mohammadreza Soltani, Vahid Tarokh
Duke University
Durham, U.S.A
{juncheng.dong,suya.wu,mohammadreza.soltani,vahid.tarokh}@duke.edu

ABSTRACT

Recently Reinforcement Learning (RL) has been applied as an anti-adversarial remedy in wireless communication networks. However studying the RL-based approaches from the adversary's perspective has received little attention. Additionally, RL-based approaches in an anti-adversary or adversarial paradigm mostly consider single-channel communication (either channel selection or single channel power control), while multi-channel communication is more common in practice. In this paper, we propose a multi-agent adversary system (MAAS) for modeling and analyzing adversaries in a wireless communication scenario by careful design of the reward function under realistic communication scenarios. In particular, by modeling the adversaries as learning agents, we show that the proposed MAAS is able to successfully choose the transmitted channel(s) and their respective allocated power(s) without any prior knowledge of the sender strategy. Compared to the single-agent adversary (SAA), multi-agents in MAAS can achieve significant reduction in signal-to-noise ratio (SINR) under the same power constraints and partial observability, while providing improved stability and a more efficient learning process. Moreover, through empirical studies we show that the results in simulation are close to the ones in communication in reality, a conclusion that is pivotal to the validity of performance of agents evaluated in simulations.

KEYWORDS

RL; Communication; Multiagent System; Adversary Attacks

ACM Reference Format:
Juncheng Dong, Suya Wu, Mohammadreza Soltani, Vahid Tarokh. 2022. Multi-Agent Adversarial Attacks for Multi-Channel Communications: Extended Abstract. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 3 pages.

1 INTRODUCTION

Recently reinforcement learning (RL) has been successfully applied for designing algorithms for defending against adversary attacks in a hostile wireless communication environment [1, 3, 6, 7]. However, using the RL-based methods from the adversary's perspective has received little attention. This is especially important since a better understanding of adversaries' behavior can lead to better design of defense mechanisms as well as active defense. Furthermore, the current endeavor for utilizing the RL-based methods either in anti-adversary or adversary paradigm mostly consider communication within a single channel (either channel selection or single channel power control), while multi-channel communication is more common in real scenarios. In this work, we propose a multi-agent adversary system (MAAS) based on RL for modeling and analyzing adversaries (e.g., jammers) in a wireless communication scenario by a careful design of the reward function for realistic multi-channel communication scenarios. Through extensive simulations, we show that the proposed MAAS learns to choose the transmit channel and the most efficient power allocation for attack without any prior knowledge of the sender strategy. In particular, our results demonstrate that using MAAS and implicit collaboration between the adversaries can provide better performance and success rates compared to the single-agent adversary (SAA) case. In addition, MAAS is fault-tolerant and robust to the failure of some agents since other agents may take over the failed agent's duty and continue its operations.

2 MULTI-AGENT ADVERSARY SYSTEM (MAAS)

2.1 Communication Model

We first propose our communication model. In our scenario, we assume there is a pair of sender/receiver and there are N available channels for sending the signal (see figure **??**). We assume that the communication between the sender and the receiver and also the adversaries happens only at discrete time steps $t = 0, 1, 2, \ldots$. At each time step t, the sender is allowed to choose multiple channels to send its signal to the receiver. The amount of power allocated to each channel is assumed to be fixed during the time that channel is used for communication. We also assume that there are M adversaries (indexed by the index $j = 0, 1, ..., M-1$), each selects one channel out of N available channels and a power level from a set $P = \{P_0, P_1, \ldots, P_K\}$, where $0 \le P_0 < P_1 < \cdots < P_K$). This means that the adversary j's action at time t is a 2-dimensional vector $a_j^{(t)} = [C_j^{(t)}, P_j^{(t)}]^T$, where $C_j^{(t)}$ is the channel selected by adversary j, and $P_j^{(t)} \in P$ is its selected power level at time t. In our scenario, the goal is that adversaries can select their actions to maximize decrease of the quality of communication (QoC) between the sender and the receiver. The (QoC) at each time step t is defined by the signal-to-interference-plus-noise ratio (SINR):

$$\text{SINR}^{(t)} = \frac{P_S^{(t)} * h_S}{\eta + \sum_{j=0}^{M-1} P_j^{(t)} * h_j * I(C_j^{(t)} = C_S^{(t)})}, \quad (1)$$

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

where η denotes the communication noise, and h_S and h_j are power gains for the sender and adversary j, respectively. In order to evaluate the performance of an adversary in our communication model, we define *Success of Attack* (SA) as a binary-valued function as follows:

$$SA = \mathbb{1}(SINR^{(t)} < \tau SNR^{(t)}), \tag{2}$$

where τ denotes a pre-defined threshold, and SNR (signal-to-noise-ratio) is the maximal achievable SINR for the receiver, calculated as $SNR = \frac{P_s * h_s}{\eta}$ (i.e., the SINR without any interference). Also, $\mathbb{1}(condition)$ denotes the indicator function defined as 1 if the condition is true, and 0 otherwise. Moreover, We define the *Success Rate of Attack* (SRA) as the ratio of the number of SA over a time interval $T > 0$:

$$\text{SRA} = \frac{\sum_{t=0}^{T-1} \mathbb{1}(SINR^{(t)} < \tau SNR^{(t)})}{T}. \tag{3}$$

The choice of τ is problem-specific. However, as we will see in the experimental results, the change of τ will not have a dramatic effect in changing the SRA value of the proposed MASS.

2.2 Design of Reward Function

Our reward function includes two parts: the portion of channels blocked by adversaries, and the power cost incurred by adversaries for attacking the communication between the sender and the receiver.

- **Channels blocked by adversaries.** This can be computed by the decrease in the Shannon channel capacity calculated as $B * \left(\log_2(1 + \text{SNR}^{(t)}) - \log_2(1 + \text{SINR}^{(t)})\right)$, where B denotes the bandwidth of the channel.
- **Power cost.** The reward function should include a term, indicating the cost of power to penalize adversaries if they use the allocated power inefficiently. We consider a constant power cost, $Cost_{power}$ for all the adversaries and assume that the cost is known to adversaries throughout the communication.

Hence, the total reward function for the whole multi-agent system is given by:

$$R^{(t)} = B * \left(\log_2(1 + \text{SNR}^{(t)}) - \log_2(1 + \text{SINR}^{(t)})\right) - C_{\text{power}} * \sum_{j=0}^{M-1} P_j^{(t)}, \tag{4}$$

where P_j is the power used by adversary j.

2.3 RL for Multi-Agent Adversary System

We choose the Double Deep Q-learning with prioritized experience replay as RL agent for each adversary in MAAS for faster adaption to the environment [4, 5]. Similar to existing methods [7], we use the SINR at time $t-1$ as the state for the environment at time t and each adversary in MARL have its own reward for distributed training at time t defined as follows:

$$R_j^{(t)} = B * \left(\log_2(1 + \text{SNR}^{(t)}) - \log_2(1 + \text{SINR}^{(t)})\right) - Cost_{power} * P_j^{(t)}. \tag{5}$$

For training of the MAAS, we follow the distributed learning paradigm in which there is no central entity to coordinate the information exchange between adversaries. In addition, all adversaries interact with the environment simultaneously and can observe the SINR values sequentially. They also select their actions independently from each other. Each adversary j is equipped by an experience memory MEM_j for storing the adversary's experience for faster learning, and a pair of actor network Q_j^{actor} and target network Q_j^{target}) which are initialized randomly at the beginning of communication. At the very beginning ($t = 0$), adversaries make random actions. At any other time $t > 0$, each adversary selects its action $a_j^{(t)}$ using its actor network with the current state $s^{(t)}$ (the SINR value from the last time step, $t-1$) as input. The sequence of SINR values is also used for computing individual rewards $r_j^{(t)}$ which is used to train actor networks and target networks $Q_j^{(actor)}$ and $Q_j^{(target)}$.

3 SIMULATIONS

In this section, we present our empirical studies to demonstrate the performance of our proposed MAAS in both single-channel and multiple-channel wireless communication with four types of transmitters and four benchmark adversaries (Please see [2] for details of the simulations.). Figure 1 illustrate the overall performance of adversaries in different scenarios. MAAS also has a consistently increased success rate of attacks compared to all other benchmark adversaries regardless of threshold τ.

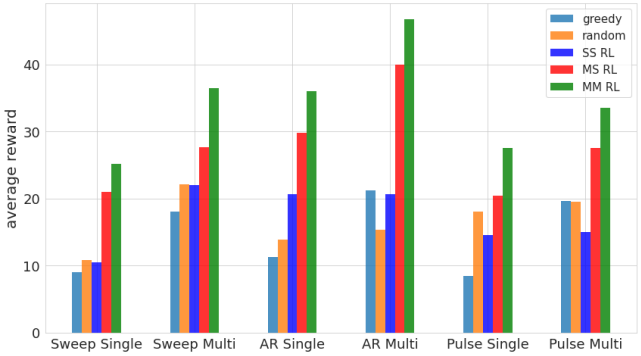

Figure 1: MAAS consistently gains over the other adversaries in all cases even though we enforce its total power to be the same as the MSRL.

4 CONCLUSION

We have proposed a MARL-based multi-agent adversary system (MAAS) along with a system model for multi-channel communication. The MAAS has shown an outstanding performance compared to various baselines, even with power constraints and without partial observability. The proposed MAAS has its value in active defense as well as understanding the behaviors of multi-agent RL adversaries in designing the defense mechanisms.

5 ACKNOWLEDGEMENT

This work was supported in part by Air Force Research Lab Award #FA 8750-20-2-0504.

REFERENCES

[1] Ye Chen, Yanda Li, Dongjin Xu, and Liang Xiao. 2018. DQN-based Power Control for IoT Transmission against Jamming. *IEEE Vehicular Technology Conference* 2018-June (jul 2018), 1–5. https://doi.org/10.1109/VTCSpring.2018.8417695

[2] Juncheng Dong, Suya Wu, Mohammadreza Sultani, and Vahid Tarokh. 2021. Multi-Agent Adversarial Attacks for Multi-Channel Communications. (2021).

[3] Xin Liu, Yuhua Xu, Luliang Jia, Qihui Wu, and Alagan Anpalagan. 2018. Anti-jamming Communications Using Spectrum Waterfall: A Deep Reinforcement Learning Approach. *IEEE Communications Letters* 22, 5 (may 2018), 998–1001. https://doi.org/10.1109/LCOMM.2018.2815018 arXiv:1710.04830

[4] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. 2016. Prioritized Experience Replay. arXiv:1511.05952 [cs.LG]

[5] Hado van Hasselt, Arthur Guez, and David Silver. 2015. Deep Reinforcement Learning with Double Q-learning. arXiv:1509.06461 [cs.LG]

[6] Jianliang Xu, Huaxun Lou, Weifeng Zhang, and Gaoli Sang. 2020. An Intelligent Anti-jamming Scheme for Cognitive Radio based on Deep Reinforcement Learning. *IEEE Access* 8 (2020), 202563–202572. https://doi.org/10.1109/ACCESS.2020.3036027

[7] Pei Gen Ye, Yuan Gen Wang, Jin Li, Liang Xiao, and Guopu Zhu. 2020. Fast Reinforcement Learning for Anti-jamming Communications. In *2020 IEEE Global Communications Conference, GLOBECOM 2020 - Proceedings*. https://doi.org/10.1109/GLOBECOM42002.2020.9322486 arXiv:2002.05364v1