Unmasking Vulnerabilities: Adversarial Attacks against DRL-based Resource Allocation in O-RAN

Yared Abera Ergu*, Van-Linh Nguyen*, Ren-Hung Hwang†, Ying-Dar Lin†, Chuan-Yu Cho‡, Hui-Kuo Yang‡
*Dept. of Computer Science and Information Engineering, National Chung Cheng University (CCU), Chiayi, Taiwan
†College of Artificial Intelligence, National Yang Ming Chiao Tung University (NYCU), Taiwan
‡ Information and Communications Research Lab, Industrial Technology Research Institute (ITRI), Taiwan
{yared111p, nvlinh}@cs.ccu.edu.tw, {rhhwang, ydlin}@nycu.edu.tw, {ares, hgyang}@itri.org.tw

ICC 2024 - IEEE International Conference on Communications | 978-1-7281-9054-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICC51166.2024.10623131

***Abstract*—The rapid advancement of wireless networks towards Artificial Intelligence (AI)-driven solutions attracts many vendors to build resilient and intelligent capabilities for Open Radio Access Networks (O-RAN). However, besides the benefits of achieving flexibility and intelligence, openness in native AI-driven O-RAN functions is also the target of severe AI-related security threats, e.g., adversarial attacks. This work addresses the security matter for the AI-powered solutions in the physical layer of O-RAN, specifically within the context of deep reinforcement learning (DRL)-based resource allocation. We introduce a new adversarial attack variant that manipulates the environment parameters and misleads the agent's observation during the inference phase. The attack can cause incorrect allocation decisions and significant degradation in the transmission data rate. Our evaluation results show that the attack degrades user data and packet delivery rates by up to 40% and 77.74%, respectively, particularly in ultra-low-latency services. We also found that the major weakness of DRL-driven radio resource allocation is the environment observation stage, where a group of compromised users or jammers can spoof noises and signal power to mislead environment interaction. In our context, the proposed policy infiltration attack is the most efficient approach to cause sustained network inefficiencies or reduced throughput for benign users.**

***Index Terms*—O-RAN, Adversarial Attacks, Resource Allocation, Policy Infiltration Attacks, Deep Reinforcement Learning**

I. Introduction

Open-RAN (O-RAN) has become one of the most popular platforms in radio access networks (RAN). The O-RAN Alliance includes major global operators and suppliers, such as AT&T and NTT DoCoMo. In addition, O-RAN has attracted the attention of industry and research teams due to its potential to improve flexibility and interoperability in next-generation networks [1], [2]. O-RAN can reduce the cost of network deployment and operation by leveraging commodity off-the-shelf hardware, virtualization, and multi-vendor environments. This new architecture will open opportunities for smaller suppliers and new startups to enter the market and provide diverse solutions. Recently, numerous teams have started incorporating artificial intelligence capabilities into critical O-RAN functions, aiming to actualize the vision of networks that are autonomous and self-optimizing [3], [4]. According to [5], AI will first be implemented in near-real-time radio intelligence controllers (xApps) and RAN service management (rApps).

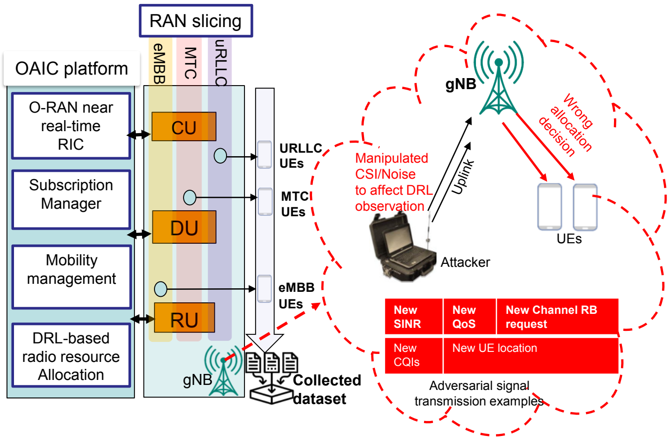

Fig. 1. O-RAN Architecture & DRL Resource Allocation with Adversarial Attacks where an attacker broadcasts false signal requests and noises to poison DRL's state space and affect the subsequent actions.

However, adversarial attacks against AI models and related security risks are of the utmost concern for the reliability of AI systems [3]. The O-RAN specification also describes this new threat [6], namely attacks against AI/ML models used for inference and control in xApps and rApps. In these attack types, the adversary can gain unrestricted control of one or more O-RAN nodes by manipulating the uplink(UL) signals and best-state observations stored in the SMO/non-RT RIC to generate real-time feeds synthetic data. These attacks may cause AI/ML solutions to output incorrect predictions or make wrong control decisions, resulting in performance degradation and, even worse, network connection interruptions [3].

This study addresses new adversarial attacks on AI-based radio resource allocation in O-RAN. There are several reasons for our choice to address this critical matter. First, the UL signal patterns from user equipment, which are the source of radio resource allocation training, can be easily manipulated by fake user equipment or jamming signals. It should be noted that counterfeit user equipment can easily be implemented using software-defined radio devices. Secondly, resource allocation is critical in ensuring network performance and optimal resource utilization in O-RAN. Adversarial attacks can manipulate allocation decisions, leading to resource allocation errors, as illustrated in Fig. 1. This may interfere with network stability, reduce overall system performance, and potentially lead to congestion, resource waste, and even network outages. Finally, incorrect resource allocation caused by adversarial attacks may cause cascading effects at different levels of the O-RAN architecture. For example, resource allocation errors in a certain part of the network may affect the end-user experience, leading to dissatisfaction with network quality or the knock-on effect of causing the entire area to be unable to connect correctly. To the best of our knowledge, this is the first attempt to conduct adversarial attacks against DRL-based resource allocation at the physical layer of the O-RAN platform.

A. State-of-the-Art Studies

Numerous studies stress the need to investigate diverse adversarial attack strategies on AI-driven solutions in wireless networks, including stealthy attacks, evasion, poisoning, signal modulation deception, and jamming attacks [7]. Table I summarizes related studies compared to our work. There are two primary attack approaches: white-box attacks and black-box attacks. In the first attack, the attacker has full knowledge (the architecture, parameters, and training data) and access to the target model. For example, the authors in [8] present a first attempt to craft white-box adversarial attacks against CNN-based radio signal classification tasks. In the second attack type, the attacker has limited or no access to the target model's architecture, parameters, or internal workings. For example, the authors in [9] proposed a black-box adversarial attack on DNN-based communications that involved leveraging Universal Adversarial Perturbations (UAPs) and the Perturbation Generation Model (PGM). The crafted attacks were subtle yet potent, designed to query the signal decoding and use the successful decoding responses to generate adversarial examples. However, achieving the delicate balance between attack potency and stealth required careful constraint adjustment.

TABLE I
COMPARISON OF THE STATE-OF-THE-ART STUDIES

Paper	CSI Input	Attack Method	Attack Target	Disrupt Network
[9]	✓	PGM via UAP	Signal decoding	×
[8]	✓	UAP	Signal classification	×
[10]	✓	MI-FGSM/PGDM	Power allocation	✓
[11]	×	PYH-Jamming	Channel access & power regulation	×
[12]	×	FGSM/PGD	Interference classification	×
Ours	✓	**FGSM/PIA**	**Radio resource allocation**	✓

As summarized in Table I, close to our work, the authors in [11] proposed a black-box approach to exploit signal interference and misleading reinforcement learning agents in dynamic channel access and power regulation. However, they assume that the attacker must have high capability in exploration and imbalanced training, hindering sustained attacks. In another work, the researchers in [10] orchestrated an adversarial evasion attack on DNN-powered power allocation within the realm of massive Multiple-Input Multiple-Output (maMIMO) to coerce the system into generating infeasible power allocation solutions. A major disadvantage is the infeasibility of capturing the true dynamism of the real-world state space, which is in a perpetual state of flux. In a recent study [12], adversarial poison attacks revealed vulnerabilities in data-driven xApps for interference classification. These attacks manipulated data in a shared RIC, leading to potential performance degradation.

B. Contributions

Unlike prior work, our study focuses on penetrating the policy decision of the deep reinforcement learning algorithm and adjusting user equipment UL signal broadcasting. The goal is to create wrong environment observation that contributes to the agent's chaotic allocation decision at the base station. This research marks our initial endeavor to address adversarial attacks on the physical layer of the enterprise O-RAN platforms. We implement an adversarial attack to evaluate the AI-powered control functions' reliability and demonstrate the model's capacity to withstand subtle policy infiltration attacks. Our key contributions are summarized as follows.

- We present a novel adversarial attack termed Policy Infiltration Attack (PIA) targeting specific decision policy in Proximal Policy Optimization-based resource allocation approaches within O-RAN frameworks.
- We design and implement a policy infiltrator that can craft subtle yet effective perturbations to the state space of the DRL agent, causing it to generate sub-optimal actions.
- Uncovering vulnerabilities in state-of-the-art Proximal Policy Optimization-based resource allocation, we conducted thorough experiments on PIA's performance and robustness across diverse UE traffic demands. Our evaluation included comparing the model's behavior under different scenarios. Also, this exploration of vulnerabilities opens a new research frontier for robust and resilient deployment of AI models in O-RAN at the physical layer.

The remainder of this paper is organized as follows. Section II presents the system model and problem formulation. Section III details our proposed models. The evaluation results of the system are shown in Section IV. Finally, the conclusion is summarized in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

This work considers a wireless communication environment with M MIMO-based Base Stations (BS) serving N mobile single-antenna Users Equipment (UE) with time-varying path loss and channel conditions. The sets of BSs and UEs are denoted as $\mathcal{M} = \{1, 2, \ldots, m, \ldots, M\}$, $\mathcal{N} = \{1, 2, \ldots, n, \ldots, N\}$. The base stations interface via E2 through the O-RAN interface manager, as illustrated in Fig. 1.

The base stations transmit a bundle of Resource Block Groups (RBG), the smallest resource unit allocated to a user, consisting of 12 consecutive subcarriers. Each BS has a set of $\mathcal{G}$ RBGs, denoted as $\mathcal{G}$, $\mathcal{G} = \{1, 2, \ldots, g, \ldots, G\}$. The available bandwidth is divided into ζ orthogonal slices to avoid interference and implement fair and efficient scheduling policies. We define $h^{m,g,n}$ to denote a binary notion of whether BS m has allocated resource block g to user n. Moreover, let ξ^n represent the channel coefficient for user n on slice ζ, and the transmission power assigned to the g-th RBG of BS m at time t is denoted as $P^{m,g}$.

The Signal-to-Interference-plus-Noise Ratio (SINR) between BS m and user n on RBG g at time t, defined as $\Psi^{m,g,n}$, is formulated by $\Psi^{m,g,n} =$

$$\frac{h^{m,g,n} \cdot \xi^{m,n} \cdot P^{m,g}}{\sum_{m' \in \mathcal{M}, m' \neq m} \sum_{n' \in \mathcal{N}, n' \neq n} h^{m',g,n'} \cdot \xi^{m',n'} \cdot P^{m',g} + \sigma^2}, \tag{1}$$

where $\sum_{m' \in \mathcal{M}, m' \neq m} \sum_{n' \in \mathcal{N}, n' \neq n} h^{m',g,n'} \cdot \xi^{m',n'} \cdot P^{m',g}$ denotes signal interference from other BSs/UEs (except BS m, UE n), σ^2 represents the noise variance.

The transmission capacity of BS m on the resource block g is given by

$$C^{m,g} = \beta_g \cdot \log_2(1 + \sum_{n \in \mathcal{N}} \Psi^{m,g,n}), \tag{2}$$

Here, β_g represents the assigned bandwidth for resource block g. We presume that user traffic adheres to Poisson arrivals with a designated mean arrival rate. Accumulated traffic is stored in a transmission buffer of limited capacity. If the queued data length exceeds the buffer's capacity, any surplus data will be discarded. The transmission rate of BS m on resource block g is articulated by

$$R^{m,g} = \begin{cases} C^{m,g}, & \text{if } C^{m,g}T < \sum_{n \in \mathcal{N}} h^{m,g,n} L^n \\ \frac{L^n}{T}, & \text{if } C^{m,g}T \geq \sum_{n \in \mathcal{N}} h^{m,g,n} L^n \end{cases} \tag{3}$$

where, L^n denotes the remaining transmission block size, T is the time duration slot, and ξ^n signifies the importance factor for user n on g. We also assume optimized power allocated $P^{g,n}$ with the consideration of the necessary constraints for each user n in slice ζ as described in [10]. Ultimately, the goal of the system is to optimize the overall transmission rate across all M BSs, articulated as:

$$\begin{aligned} & \max_{P^{g,n}, h^{m,g,n}} \sum_{m \in \mathcal{M}} \sum_{g \in \mathcal{G}} R^{m,g}, \\ & \text{s.t. (1) - (3)}, \\ & P_{\min} \leq P^{m,g} \leq P_{\max}, \quad \forall m, g, \\ & h^{m,g,n} \in \{0, 1\}, \quad \forall m, g, n, \\ & \sum_{n \in \mathcal{N}} h^{m,g,n} = 1, \quad \forall m, g. \end{aligned} \tag{4}$$

A. Resource Allocation Model

The resource allocation problem is to find an optimal policy for each agent that maximizes its expected cumulative reward over time while satisfying the QoS constraints and ensuring coordination among agents. It corresponds to selecting a signal correlation matrix for each user and transmits strategy $\{S_1, \ldots, S_{nr}\}$ in compliance with the power constraints. The loss function is designed to facilitate the generation of adversarial examples, empowering the adversary to choose a potent attack mechanism for effectively deceiving the model with high probability and subtlety. The loss function of a learning model with is given by

$$\mathcal{L}_{adv}(\theta, \phi, x, \delta) = \mathcal{L}_{origin}(\theta, \phi, x) - \mathcal{L}_{origin}(\theta, \phi, x + \delta) \tag{5}$$

Here, θ denotes a policy network, and ϕ means a value network. The objective is to find a perturbation δ to add to the input observations x and maximize the loss $\mathcal{L}_{adv}(\theta, \phi, x, \delta)$ to encourage perturbations that lead to an increase in the original loss.

B. Problem Formulation

We formulate an adversarial resource allocation problem within a PPO agent architecture, where an agent interacts with an environment defined by wireless channel conditions, user demands, Key Performance Indicator(KPI) requirements, and available resources. The objective function of the agent is the traffic-based allocation of radio resources to users. The agent determines values $h^{m,g,n}$ and β_g, key parameters influencing both UL and DL transmissions. $h^{m,g,n}$ controls UL resource allocation, impacting the distribution of resources from UE to the BS m. Adjustments to $h^{m,g,n}$ can impact resource quantity, transmission power, and other UL-related factors, consequently affecting the user's UL communication capacity [13]. Conversely, β_g governs DL resource allocation, focusing on the distribution of resources from the BS m to the UEs.

State Space (S): The state space encompasses several key elements to observe over time t, including the transmission rate $R_t^{m,g}$, transmission power $P_t^{m,g}$, relevant Channel Quality Information (CQI) amplitude CSI data as $\mathcal{H}_t^{m,g}$, the length of queuing data in the buffer denoted by $L_t^{m,g}$, and an indicator $h^{m,g,n}$ embedded within the state. We formally define the state of the resource allocation problem by

$$S_t^m = \{h^{m,g,n}, \mathcal{H}_t^{m,g}, R_t^{m,g}, P_t^{m,g}, L_t^{m,g} \mid g \in \mathcal{G}, n \in \mathcal{N}\} \tag{6}$$

where n is the number of users and m is the BS serving at the moment. The attacker's state space alterations introduce inaccuracies, significantly impacting the DRL agent's perception.

Action space (A): The action space, denoted as A_t^m, comprises the set of possible actions that the agent can take at time t for BS m. Each action $a_t^{m,g}$ is a vector representing adjustments to resource allocation parameters in S for user n that is expressed by

$$a_t^{m,g} = \{k_0, \ldots, k_{N-1} \mid k_n \in h_t^{m,g,n}\}, A_t^m = \{a_t^{m,g} \mid g \in \mathcal{G}\} \tag{7}$$

Reward ($\Re$): The reward is a composite measure of the overall throughput allocated to each User Equipment's (UE) slice by BS m. The reward at time step t, denoted as $\Re_t$, is calculated by

$$\Re_t = \sum_{n \in \mathcal{N}} \omega_n \cdot R_t^{m,g}, \tag{8}$$

where ω_n represents the weight assigned to each user n, considering factors such as current downlink buffer size, transmission rate, and the ratio of granted to requested Physical Resource Blocks (PRBs). In the adversarial attack, the reward is computed from the perturbed state space, resulting in an altered course of action due to the induced perturbation.

III. Proposed Policy Infiltration Adversarial Attack against Resource Allocation Model

This section details the adversarial attacks considered to attack the baseline model and the policy infiltrator algorithm designed to mislead the reward policy of the learning agent.

A. Attack model

In O-RAN [3], the 7.2x interface isn't encrypted on the control plane due to timing complexities. This opens up for impersonation attacks and potential data compromise. Additionally, the S-Plane is susceptible to performance degradation through malicious interference with synchronization infrastructure. In this work, an inference-based adversarial attack agent is designed against the pre-trained PPO-based resource allocation model, which employs an actor-critique network in a continuous action space of the wireless network environment.

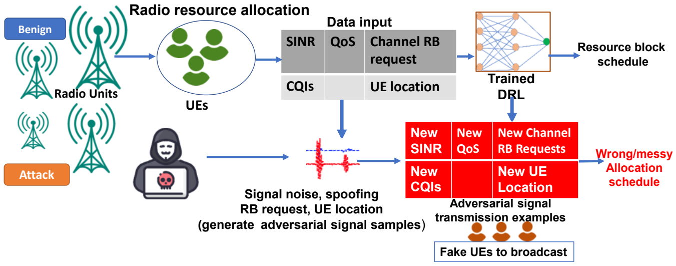

Fig. 2. An adversarial attack on the DRL-based resource allocation model in O-RAN: xApp uses DRL to allocate radio resources to users based on the environment, which includes the wireless channels, handover demands, the QoS requirements, and resource availability. An adversary attacks to perturb the channel state estimation and disrupt the resource allocation decisions.

Perturbations are constrained as $\|\eta\|_\infty \leq \epsilon$, where η is a given parameter. This constraint ensures that adversarial perturbations remain within a minimal error range, evading detection by the base station (BS) and staying within the standard tolerance range. In Figure 2, we visualize adversarial attacks on the DRL-based resource allocation model within the O-RAN framework. An xApp employs DRL to allocate radio resources to users based on environmental factors. However, adversaries aim to perturb the channel state estimation by introducing fake UEs to broadcast adversarial signals, deliberately leading the agent to wrong/messy allocation decisions.

B. Adversarial Attack Algorithms

The adversarial attack aims to impair system performance by distorting resource allocation choices through state space manipulation. The agent uses slice inputs as observations, modifies the policy settings, and tricks the agent with wrong signal information into making subtle incorrect decisions. Below, we outline the two algorithms employed for this purpose.

*1) **Fast Gradient Sign Method (FGSM):*** FGSM is a one-step adversarial attack technique that maximizes the loss function $\mathcal{L}$ [12] in a more efficient way. In resource allocation, FGSM is crafted to perturb RGB extracted from CSI $\mathcal{H}_x$ for UE n to generate an adversarial channel matrix resulting in higher loss $\mathcal{L}$ that misleads the model to provoke incorrect decisions. This adversarial channel matrix $\mathcal{H}_x^{adv}$ is represented by δ:

$$\delta = \epsilon \cdot \text{sign}(\nabla_x \mathcal{L}(\theta, x, y)), \tag{9}$$

where $\nabla_x \mathcal{L}(\theta, x, y)$ is the gradient of the loss function $\mathcal{L}(\theta, x, y)$ with respect to the input data x and ground truth label y .

*2) **Policy Infiltration Attack(PIA):*** leverages a proficient adversarial attack agent crafted to infiltrate the policy dictating UE's throughput requirements and available resource blocks (RBs). Our strategy employs imperceptible random noise ϵ to perturb the state space, effectively exploiting vulnerabilities in the model and impeding its capability to achieve the intended objectives. Unlike the widely-used FGSM, PIA aims to penetrate the behavior of a pre-trained model. Secondly, the PIA dynamically adjusts input iteratively for *episode* iterations, incorporating gradient information in each step, and breaks the loop only if the perturbation exceeds a predefined threshold. Third, this method is appropriate when a more complex and persistent attack is required, with the ability to overcome defenses against single-step attacks.

Algorithm 1: Policy Infiltrator Attack(PIA)

```
   Data: $\mathcal{H}_x \rightarrow f$, [$\mathcal{H}_x$ is the CSI input data fed to the model $f$]
   Initialization: Set $\delta_x = X, \epsilon$ (random $\delta_x$, threshold),
   Max_Iter $episode = 1000$, $\epsilon = 0.01$, threshold=1.1
   Result: Adversarial Input $\delta$
1  Calculate Baseline $f$ Prediction:
2  Baseline_prediction $= y$
3  Iterative Perturbation Generation:
4  for $i = 1$ to $episode$ do
5      $\nabla_x J(\delta_x, \text{baseline\_prediction}) =$
         $\nabla_x (\text{baseline\_prediction} \cdot \mathcal{L}(\theta, \cdot)) \rightarrow$ Calc-Gradient
6      $\delta_x \leftarrow \delta_x + \epsilon \cdot \text{sign}(\nabla_x J(\delta_x, \text{baseline\_prediction})) \rightarrow$
         Update Adversarial Input
7      $\delta_x \leftarrow \text{clip}(\delta_x, (\mathbf{X} - \epsilon, \mathbf{X} + \epsilon)) \rightarrow$ Clip Perturbation
8      infiltrated_prediction =f($\delta_x$)
9      Check Stopping Criteria:
10     if infiltrated_prediction $\neq$ baseline_prediction or
         $\|\delta_x - \mathbf{X}\|_\infty > \epsilon$ then
11         Break out of the loop
12     end
13 end
14 Perturbed $\Re \leftarrow A|S$
```

Algorithm 1 shows that the pseudo-code for the PIA algorithm targets to penetrate the behavior of a pre-trained model, denoted as f. We establish the baseline model (Lines 1-3) and leverage it to train the adversarial attack model. This adversarial training continues through *episode* iterations (Lines 4-10), intending to mislead the baseline model consistently. The perturbed data is used to gauge the agent's decision subtly. The training process persists until the adversarial model can effectively induce the baseline model to produce abnormal decisions, as shown in (Lines 10-12). This method strategically disrupts the model's behavior while staying hidden in the shadow, ensuring the baseline model continues to operate with the induced perturbation.

Deceptive Encouragement of Exploration

In our study, adversarial attacks are framed as a deceptive optimization problem, i.e., misleading the target of transmission rate in Equation 4, where the attacker strategically adjusts the power transmission, UE location, and resource allocation selection assigned to a compromised UE and influence ω_n in Equation (8). Let $Q(\omega)$ represent the performance metric to be optimized. The attacker collaborates with Users' Equipment (UEs) to provide fabricated feedback, modifying the true performance metric to a deceptive one, $Q_{\text{deceptive}}(\omega)$. The DRL agent, aiming to maximize its reward, follows policies that maximize $Q_{\text{deceptive}}(\omega)$, leading to suboptimal resource allocation strategies (π^*). This deviation from the network's true optimization objective causes biased resource allocation decisions, impacting users' QoS, which can be defined by

$$\pi^* = \arg\max_{\pi} Q_{\text{deceptive}}(\omega) \tag{10}$$

IV. Performance analysis

We use a pre-trained model for attack testing, which is evaluated on a complex network simulation scenario in a crowded urban region employing cutting-edge technology (5G) and real-world (OpenCelliD) Colosseum O-RAN Dataset. The model is trained on 7 GB of training data comprising various performance metrics (throughput, bit error rate), system state information (transmission queue size, SINR, and QoS), and resource allocation strategies (slicing and scheduling policies). This training data is collected through a total of 89 hours of experiments conducted on the world's largest wireless network emulator, Colosseum [2].

TABLE II
Experiment Setup

Parameter	Value
Network Type	5G
Number of BSs	4
Number of UEs	40
UL Frequency	1.02 GHz
DL Frequency	0.98 GHz
Channel Bandwidth	3 MHz
BS Locations	0.11 km
Slicing	Multi-slice(eMBB, URLLC,mMTC)
Scheduling Policies	(PF), (WF), (RR)
UE Allocation	Static allocation to slices
Mobility	Time-varying UEs
Time Granularity	500 ms

For optimization, in this work, scheduling paradigms use Proportionally Fair (PF), Waterfalling (WF), and Round-Robin (RR) based on UE QoS requirements. These paradigms simulate versatile scheduling and adaptive resource allocation in target network slices. This approach also ensures a nuanced and flexible optimization aligned with the unique characteristics of individual network slices in practice. And DRL Agent implementation integrated as xApp running in near real-time RIC as shown in Table II.

A. Evaluation Results

In this study, we conducted a performance analysis on the pre-trained resource allocation model while considering the presence of adversarial agents. We meticulously extracted numerical features from the reward log to facilitate a thorough comparison of essential evaluation metrics. The impact of these attacks can be observed from Table III, the data rate measure under no attack and after the attack. The table illustrates data rates in Megabits per second (Mbps) within a multi-slice scenario, where the User Equipment (UEs) are statically assigned to specific network slices based on QoS: (i) Enhanced Mobile Broadband (eMBB), representing users requesting video traffic; (ii) Machine-Type Communications (MTC) for sensing applications, and (iii) Ultra-Reliable Low Latency Communications (URLLC) for latency-constrained applications such as V2X. Consequently, the agents are rewarded based on specific KPI requirements. Initially, the eMBB and MTC agents are trained to maximize UE throughput, while URLLC agents focus on minimizing latency by allocating resources as quickly as possible.

TABLE III
Data Rate Measure under No Attack and After Attack

Slice	No attack	UPA/FGSMA	PIA	Degradation (%)	
	Mbps	Mbps	Mbps	UPA	PIA
eMBB	3.39	3.01	2.36	12	29
mMTC	0.14	0.12	0.12	26	22
uRLLC	0.05	0.04	0.03	20	40

We found that the severity of the adverse effects on UE caused by the count of unsuccessful attempts to deliver Resource Blocks (RBs) to the UEs is proportional to the allocation's failure to meet the latency requirements specified by the UEs. This relationship arises due to the extended delays induced by adversarial attacks in fulfilling resource allocation requests. Consequently, packets are compelled to be dropped due to timeouts. This effect is particularly critical in low-latency services, such as uRLLC, where extended communication delays may potentially lead to serious consequences, such as accidents involving high-speed vehicles on highways. Furthermore, it is noteworthy that the PIA attack exhibits a higher efficiency in diminishing UE data rates compared to the FGSM attack, as seen in Table III. This efficiency disparity can be attributed to PIA's direct influence on the DRL policy (well-targeted attack), while FGSM necessitates time to influence the process through DRL observations (brute-force attack).

TABLE IV
Model Performance Metrics under Different Scenarios

Scenario	Packet-DR%	Successful RBA	Average-CR
No Attack	82.58	52511	55.22
FGSM	82.01	40834	42.57
PIA	22.26	9473	9.96

In our performance assessment across various model scenarios, we analyzed critical metrics such as Packet Delivery Rate (PDR), which represents the ratio of successful Resource

Block (RB) allocations to the total RB allocation attempts by the agents. Average Cumulative Reward(Avg-CR) represents UE data rate satisfaction over time, emphasizing compliance with preset minimum bit-rate requirements among all model-generated rewards. We also examined Successful RB Allocations for each slice, a measure of the agent's request reply relay, calculated as the count of successful RB allocations per slice. Moreover, we gauged the overall performance impact on the resource allocation model through user data rate. Evaluation results in Table IV hint that the 'No Attack' scenario demonstrates the DRL-driven resource allocation model's commendable performance, featuring a high packet delivery rate, and successful RBG allocation. However, in the 'After Attack' scenario, the PIA attack contributes to a 77.74% (i.e., = 100-22.26) packet delivery rate decline, indicating significant quality of service degradation. Successful PRB allocations also declined, leading to a 40% drop in the data transmission rate. Notably, the PIA attack proves more efficient than the FGSM attack, evidenced by a greater drop in the reward, as shown in Figures 3, 4, 5 (indicate the outcomes under three traffic slices). In an extensive evaluation, we compared our system performance with that of [12], which closely aligns with our work. While the attack performance on the user data rate of the two methods is competitive, as shown in Table V, our attack targets the interaction phase of the DRL, instead of CNN-based signal classification. Further, the prior work employed a larger perturbation magnitude ϵ (0.04/0.05 vs 0.01), underscoring the high cost of performing perturbation than ours. Two attacks reaffirm the importance of efficient defense tactics for potentially harmful adversarial attacks at the physical layer.

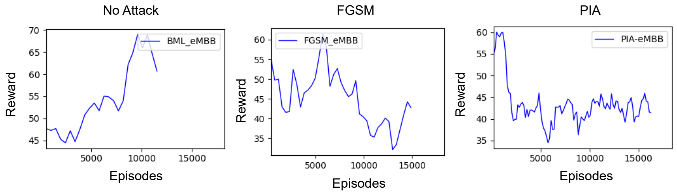

Fig. 3. Reward accumulation before and after the attack on eMBB slice

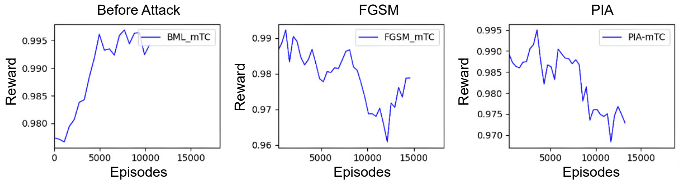

Fig. 4. Reward accumulation before and after the attack on mMTC slice

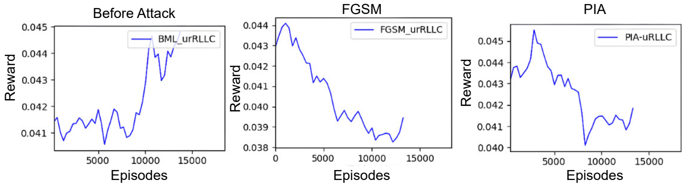

Fig. 5. Reward accumulation before and after the attack on uRLLC slice

TABLE V
UE DATA RATE COMPARISON BEFORE AND AFTER THE ATTACK

Scenario	[Ours] ϵ=0.01			Study [12] ϵ=0.04, ϵ=0.05		
	No Attack	FGSM	PIA	No Attack	FGSM	PGD
10 UEs	0.8	0.7	0.6	1.6	0.7	0.8
20 UEs	1.7	1.6	1.2	1.7	0.7	0.9
30 UEs	2.5	2.2	1.8	2.0	0.8	1.0
40 UEs	3.4	3.0	2.4	2.4	1.0	1.2

Data rate in this evaluation is measured by Mbps.

V. Conclusion

In this paper, we've conducted adversarial attacks on AI-driven resource allocation, a recognized threat in AI-native O-RAN systems. Our study represents a white-box method to perform adversarial attacks on Deep Reinforcement Learning (DRL)-based resource allocation at the physical layer of the O-RAN platform. Through these attacks, we've revealed vulnerabilities in the DRL-based resource allocation model, demonstrating its susceptibility to perturbations in the state space by compromised user equipment (UE). These insights underscore the need to address these issues before deploying AI-aided functions. A promising direction for further research is combining conventional attacks (e.g., DoS signaling) to disrupt the AI-driven radio resource allocation functions. Also, effective countermeasures against these adversarial threats can be exciting research topics for future work.

References

[1] L. Bonati, S. D'Oro, M. Polese, S. Basagni, and T. Melodia, "Intelligence and learning in o-ran for data-driven nextg cellular networks," *IEEE Communications Magazine*, vol. 59, no. 10, pp. 21–27, 2021.

[2] L. Bonati, M. Polese, S. DOro, S. Basagni, and T. Melodia, "Openran gym: Ai/ml development, data collection, and testing for o-ran on pawr platforms," *Computer Networks*, vol. 220, p. 109502, 2023.

[3] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding o-ran: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1376 – 1411, 2023.

[4] M. Dryjaski, . Kuacz, and A. Kliks, "Toward modular and flexible open ran implementations in 6g networks: Traffic steering use case and o-ran xapps," *Sensors*, vol. 21, no. 24, 2021.

[5] B. Tang, V. K. Shah, V. Marojevic, and J. H. Reed, "Ai testing framework for next-g o-ran networks: Requirements, design, and research opportunities," *IEEE Wireless Communications*, vol. 30, no. 1, pp. 70–77, 2023.

[6] O.-R. W. G. 11, "O-ran security threat modeling and remediation analysis 4.0," *WG11.O-RAN-Threat-Model-v04.00*, 2023.

[7] E. Habler, R. Bitton, D. Avraham, D. Mimran, E. Klevansky, O. Brodt, H. Lehmann, Y. Elovici, and A. Shabtai, "Adversarial machine learning threat analysis and remediation in open radio access network," 2023.

[8] M. Sadeghi and E. G. Larsson, "Adversarial attacks on deep-learning based radio signal classification," pp. 213–216, 2019.

[9] A. Bahramali, M. Nasr, A. Houmansadr, D. Goeckel, and D. Towsley, "Robust Adversarial Attacks Against DNN-Based Wireless Communication Systems," p. 126140, 2021.

[10] B. R. Manoj, M. Sadeghi, and E. G. Larsson, "Downlink power allocation in massive mimo via deep learning: Adversarial attacks and training," 2022.

[11] F. Wang, M. C. Gursoy, and S. Velipasalar, "Adversarial reinforcement learning in dynamic channel access and power control," 2021.

[12] N. N. Sapavath, B. Kim, K. Chowdhury, and V. K. Shah, "Experimental study of adversarial attacks on ml-based xapps in o-ran," 2023.

[13] H. Zhang, H. Zhou, and M. Erol-Kantarci, "Team learning-based resource allocation for open radio access network (o-ran)," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 4938–4943.