Efficient Adversarial Attacks Against DRL-Based Resource Allocation in Intelligent O-RAN for V2X

Yared Abera Ergu, *Graduate Student Member, IEEE*, Van-Linh Nguyen, *Member, IEEE*, Ren-Hung Hwang, *Senior Member, IEEE*, Ying-Dar Lin, *Fellow, IEEE*, Chuan-Yu Cho, Hui-Kuo Yang, Hyundong Shin, *Fellow, IEEE*, and Trung Q. Duong, *Fellow, IEEE*

***Abstract*—Artificial intelligence (AI) is projected to be a critical part of open radio access networks (O-RAN) to enable intelligence for connectivity management in smart vehicle-to-everything (V2X) networks and vehicle road cooperation systems. However, the openness and dependence of AI models on massive volumes of data render them subject to serious security vulnerabilities, such as adversarial attacks. This study investigates security issues in O-RAN's near real-time RAN intelligent controller (RIC), with an emphasis on deep reinforcement learning (DRL)-based resource allocation. We introduce a novel attack manipulating environmental observations to mislead AI agents, resulting in erroneous allocations and decreased physical resource block (PRB) transmission rates for various vehicular communications. We also discover flaws where compromised users or signal jammers can fake signal power to trick the AI agent's state observation. This can lead to a policy infiltration attack that makes the network performance drop significantly. Evaluation results show up to a 40% decline in user data rates, a 77.74% reduction in packet delivery rates, and significant disruptions in ultra-reliable and low-latency communications (uRLLC) services such as remote driving and connected automated vehicles. The policy infiltration attack causes a 20% increase in packet losses and up to 150% delay overall. The attack efficiency emphasizes the need for adversarial training in protecting AI-driven applications, which should be addressed in future O-RAN security specifications and AI-powered vehicular networks.**

Received 17 March 2024; revised 6 August 2024; accepted 18 September 2024. Date of publication 27 September 2024; date of current version 16 January 2025. This work was supported in part by the National Science and Technology Council (NSTC) of Taiwan under Grant 111-2222-E-194-007-MY2 and Grant 112-2221-E-194-017-MY3, and in part by the Advanced Institute of Manufacturing with High-tech Innovations (AIM-HI) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MoE) in Taiwan. The review of this article was coordinated by Dr. Phone Lin. *(Corresponding author: Van-Linh Nguyen.)*

Yared Abera Ergu is with the Department of Computer science and Information Engineering, National Chung Cheng University (CCU), Minhsiung 621301, Taiwan (e-mail: yared111p@cs.ccu.edu.tw).

Van-Linh Nguyen is with the Department of Computer science and Information Engineering, National Chung Cheng University (CCU), Minhsiung 621301, Taiwan, and also with the Advanced Institute of Manufacturing with High-tech Innovations, National Chung Cheng University (CCU), Minhsiung 621301, Taiwan (e-mail: nvlinh@cs.ccu.edu.tw).

Ren-Hung Hwang and Ying-Dar Lin are with National Yang Ming Chiao Tung University (NYCU), Tainan 300093, Taiwan (e-mail: rhhwang@nycu.edu.tw; ydlin@nycu.edu.tw).

Chuan-Yu Cho and Hui-Kuo Yang are with Industrial Technology Research Institute, Hsinchu 310401, Taiwan (e-mail: ares@itri.org.tw; hgyang@itri.org.tw).

Hyundong Shin is with Kyung Hee University, Yongin-si 17104, South Korea (e-mail: hshin@khu.ac.kr).

Trung Q. Duong is with Memorial University, St. John's, NL A1C 5S7, Canada, and also with Queen's University Belfast, BT7 1NN Belfast, U.K. (e-mail: tduong@mun.ca).

Digital Object Identifier 10.1109/TVT.2024.3466511

***Index Terms*—Adversarial attacks, deep reinforcement learning, O-RAN, policy infiltration attacks, resource allocation.**

I. Introduction

OPEN radio access network (O-RAN) has emerged as a frontrunner in RAN platforms, revolutionizing the telecommunications landscape with its open and interoperable architecture. The O-RAN Alliance brings together key worldwide operators and suppliers, including AT&T, China Mobile, Deutsche Telekom, NTT DoCoMo, and Ericsson, as well as other software communities. Furthermore, O-RAN has attracted the attention of industry and academic due to its potential for improvement and the surging demand for next-generation intelligent networks and ultra-reliable low-latency applications such as vehicle road cooperation [1], [2], [3]. Recently, several teams have explored integrating advanced artificial intelligence/machine learning (AI/ML) capabilities into core O-RAN functions [4], [5], e.g., in near real-time intelligent controller (RT-RIC) as xApps and RAN service management as rApps. This move is driven by the vision of creating networks that are truly autonomous and self-optimizing [6], [7], [8]. Unlike conventional resource allocation, AI-based solutions can address the challenges of heterogeneous devices and diverse traffic demands [9], [10], [11]. For example, the mmLBRA algorithm in [9] is a reinforcement learning-based approach that leverages O-RAN's RICs for intelligent decision-making. The method can enhance load balancing and data rate in dynamic environments such as vehicular networks and vehicle road cooperation systems.

However, security threats against AI-driven solutions are severe but have not yet been well-explored regarding the O-RAN platform for vehicle-to-everything (V2X) [12]. Adversarial attacks against AI models and related security risks are first mentioned in several preliminary studies [13]. The recent O-RAN specification also describes this new threat [14], namely attacks against AI/ML models used for inference and control 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 service management and orchestration (SMO)/non-RT RIC to generate real-time feeds of synthetic data. These attacks may lead to AI/ML systems producing inaccurate bandwidth demand/data rate choices, which can significantly degrade communication quality or even lead to disconnection in systems such as remote driving and vehicle road cooperation, as illustrated in Fig. 1.

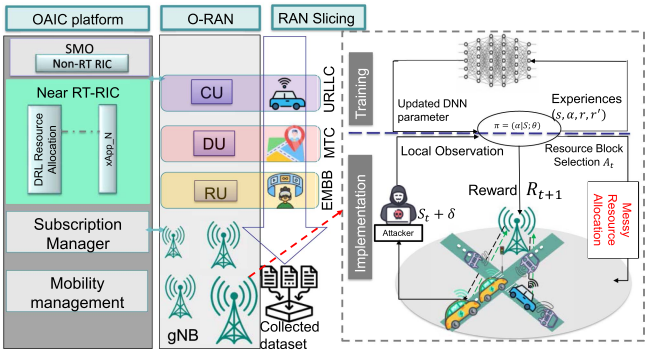


Fig. 1. O-RAN architecture for V2X networks integrating DRL-based resource allocation and illustrating potential adversarial attack points. The attacker sends fake signals and noises to disrupt DRL's decision-making process, leading to incorrect resource allocation.

A sophisticated unauthorized device can manipulate the model by generating deceptive wireless signals resembling those of authorized devices [15]. Third-party-developed xApps in O-RAN also raise security concerns due to their AI models' queries to system information (e.g., channel state, access request distribution) [16]. However, none of the studies specify the matter of adversarial attacks on O-RAN platform for V2X.

A. Motivation

With evolving cybersecurity challenges in O-RAN networks, this research explores potential vulnerabilities to attack the DRL-based resource allocation models in O-RAN, specifically focusing on lower-layer split (LLS) 7-2x and Near-RT RIC for various vehicular communications. We aim to demonstrate the feasibility of launching attacks against core control applications in O-RAN, even with the protection of LLS. The proposed attacks also highlight the lack of a clear functional split between near-RT RIC and radio units (RU), which can pose security disparities. The goal is to improve the security of O-RAN in the next standards and AI-driven vehicular networks by identifying the particular weaknesses of the existing O-RAN platform, using examples from V2X applications. Given that the community is moving towards intelligent cellular networks, this work can provide a new view of security risks, particularly in critical applications like vehicular networks for intelligent transportation systems.

B. Contributions

This work presents novel approaches to penetrate policy decisions within deep reinforcement learning algorithms used in O-RAN for V2X communications. Our research focuses on exploiting vulnerabilities in open interfaces to modify uplink signals, corrupting the agent's state observations during the inference step. This leads to chaotic resource allocation decisions with potentially cascading effects across the O-RAN architecture and low-latency V2X applications. To the best of our knowledge, this is the first comprehensive study of adversarial attacks against DRL-based resource allocation at the physical layer of enterprise O-RAN platforms. This approach presents unique challenges due to O-RAN's open interfaces and public source repository, which makes it easier for adversaries to introduce malicious inputs or penetrate the system hyper-parameters. Additionally, the V2X scenario introduces complexities such as high mobility, safety-critical operations, and ultra-low latency requirements, making resource allocation particularly sensitive to the rapid changes in signal quality. In short, the study's key contributions are:

- The first attempt to propose a novel gradient-based adversarial attack technique targeting proximal-policy optimization (PPO)-based resource allocation xApps in O-RAN, namely Policy Infiltration Attack (PIA). PIA achieves a 40% success rate in disrupting optimal resource allocation while maintaining a low perturbation magnitude in the normalized state space compared to the previous studies, demonstrating its efficiency. PIA also demonstrates a 77.74% packet loss for the uRLLC slice, a significant decline in quality of service (QoS), compared with legendary methods like projected gradient descent.
- The first attempt to provide a comprehensive study of the attack on subtle policy and impactful perturbations to DRL-powered resource allocation techniques in O-RAN. The proposed simulation environment accurately models O-RAN architecture and V2X communication patterns, incorporating realistic channel models and network topologies (e.g., urban vehicular scenarios). The infiltrator achieves a temporal consistency of 2.5 milliseconds delay in dynamic V2X environments, significantly outperforming the existing attack methods.
- The first attempt to provide extensive experiments of adversarial attacks on a fine-tuned pre-trained DRL model for O-RAN V2X. The evaluation metrics include: (1) resource allocation efficiency: PIA reduces allocation (RBA) efficiency by approximately 82% which shows an increase in allocation fluctuations under PIA compared to benign conditions (2) QoS impact: critical V2X application experience a 150% increase in latency and a marked decrease in connectivity stability. These results demonstrate the severe implications of the proposed attacks on V2X, particularly for safety-critical applications like collision avoidance and platooning.
- Open-source contribution: to foster further research and reproducibility, we have committed to releasing the source code[1] to the community. The codebase includes the implementation of attack algorithms, simulation environments, and evaluation tools. The framework in the codebase provides a baseline benchmark for future studies in securing AI-driven O-RAN systems, especially for vehicular communications.

The paper is organized as follows. Section II critically analyzes relevant approaches in related works. Section III explains our model consideration and problem formulation. Section IV details our proposed PIA algorithm and other attack techniques. Section V presents experimental results and analysis, comparing them to existing work. Finally, Section VI concludes the research and suggests future research directions.

[1][Online]. Available: https://github.com/Jaredabera/adversarialdrlORAN.git

TABLE I
SUMMARY OF SEVERAL STATE-OF-THE-ART RELATED STUDIES WITH THE PROPOSED METHOD AND FEATURE COMPARISON

Study	Disrupt Network	Network Target	CSI Input	Attack Methods	AI/ML Approach	Attack Target App	Ingress Pathway	Impact Severity	Limitations
[16]	×	O-CLOUD	×	FGSM/PGD	DNN	Interference Classifier xApp	Backhaul-Malicious Third-party xApp	Low	Constrained to white-box, O-RAN sec-cert oversight
[20]		CLOUD	✓	UAP	VT-CNN2	Signal Classifier	Tx-Sync Input-Specific Attack		Whitebox-centric attack, O-RAN non-compliant
[21]		CLOUD	×	Min.power adv-attack	VT-CNN2	AMC Classifier	Insider, Model-driven		Simulated constraints crave proof for attack validation
[22]	✓	CLOUD	×	Random IRS Phases Shifting	DRL-CT/FDRL	Tx, Rx pairs & IRS Controller	Compromised IRS Controller	Medium	Neglects security in O-RAN for IRS and gNB link
[23]		MC-MIMO	✓	FGSM/MI-FGSM/PGDM	DNN	Power Allocation	Distorted UE Locations		Solely white-box, disregarded O-RAN sec-cert
[24]		MEC servers	×	BMP Attack	-	Bearer Context Configuration	Compromised hosts		Suboptimal security measures, over looking O-RAN sec-cert standards
[25]		UE	×	Jamming	GANs	Spectrum Sharing and NW Slicing	Jamming with Surrogate Inference		No experimental validation; overlooks NexGen NWs
[19]		CLOUD	×	Poisoning via backdoor	DRL	Slicing and Scheduling Policies	Compromise E2 interface	High	Impact and trade-offs not exhaustively studied
[26]		CLOUD	✓	PGM via UAP	GAN	Signal Decoder	Perturbation Infusion via Jamming		Non ORAN sec-cert; precise constraints needed for potent stealth attacks
[27]		CLOUD	×	PYH-Jamming	DQN	Channel Access & Power Regulation	Adversarial Jamming		Costly exploration; imbalanced training hampers sustained attacks
Ours		O-RAN/V2X	✓	**FGSM/PIA/PGD**	**MADRL-PPO**	**Radio Resource Allocation**	**Co-opted and Compromised UEs**	**High**	PIA, solely tested in this domain; O-RAN sec-cert compliant

O-RAN: Open Radio Access Networks; sec-cert: O-RAN Security Certification Consideration; AMC: Automatic Modulation Classification; Tx-Sync: Transmitter Synchronous Input-Specific Attack; Min.power adv-attack: Minimum Power-Driven Adversarial Attack; BMP: Bearer Migration Poisoning; PGM: Perturbation Generation Model; PGD: Projected Gradient Descent; FGSM: Fast Gradient Sign Method; UAP: Universal Adversarial Perturbations.

II. RELATED WORKS

Early investigations have identified several adversarial attack methods aimed at AI-powered solutions in wireless communications. The attack tactics include covert attacks, evasion strategies, poisoning methods, signal modulation trickery, and jamming attacks [15], [17], [18], [19].

The comparison of relevant papers in Table I provides insights into attack techniques, targets, disruptive potential, and constraints of different attack approaches compared to our study. We also categorize them based on the target of the network attack (cloud or radio unit) and the likely consequences (disruption). The attack landscape is divided into two main paradigms: white-box attacks and black-box attacks. White-box attacks provide the adversary with full access to the target model's architecture, parameters, and training data, allowing for precise creation of adversarial perturbations. Black-box attacks occur when the attacker has restricted or no access to the internal mechanisms of the target model. For performance degradation categories, the authors in [20] provide a black-box adversarial attack method on DNN-based communication systems using universal adversarial perturbations (UAPs) and the perturbation generation model (PGM). These attacks are subtle and powerful but require precise modification to maintain a delicate balance between strength and stealth. The study [21] emphasizes the susceptibility of deep learning models to white-box adversarial attacks in automatic modulation classification (AMC) systems. Proposed low-power adversarial attack technique to create imperceptible perturbations that significantly reduce the performance of deep learning-based AMC systems.

Besides performance degradation, disrupting the network is also a potential target. The authors in [24] introduced bearer migration poisoning (BMP), an attack targeting near-real-time RIC components in the O-RAN framework. A bearer is a logical connection between a user and the network. BMP disrupts the user plane traffic path, leading to issues like routing black holes by exploiting vulnerabilities in bearer migration procedures. This manipulation of RIC perception causes network anomalies and significantly increases signaling costs, leading to catastrophic drops. However, the study overlooks the growing demand for AI-native xApps security certification (sec-cert) in both the real-time RIC units under development to initiate such attacks. In a separate study, Manoj et al. [23] conducted an adversarial evasion attack targeting DNN-powered power allocation within the domain of massive multiple-Input multiple-output (maMIMO), resulting in the generation of impractical power allocation solutions. A significant drawback lies in the inability to accurately capture the dynamic nature of real-world state spaces, which are constantly evolving. In prior research on signal classification jamming [25], a black-box attack approach was adopted, utilizing received signal strength indicator (RSSI) values to train a surrogate model when direct access to the enhanced spectrum classifier's (ESC) deep learning classifier was unavailable. This pragmatic strategy aims to emulate the behavior of the target model. The study builds upon previous investigations by examining adversarial attacks in two crucial scenarios: spectrum sharing with incumbent users and network slicing.

In close alignment with this study, the authors in [27] introduce an approach to exploit signal interference to mislead reinforcement learning agents involved in dynamic channel access and power regulation. However, their methodology assumes a high level of attacker capability in exploration, and the training imbalance poses challenges in sustaining attacks. Similarly, the study in [22] outlines an attack strategy targeting compromised intelligent reflecting surfaces (IRS), involving manipulation of phase shifts in the constituent elements. The adversary can initiate this attack by transmitting malicious commands to the IRS controller through a wireless link. The attack presupposes the adversary's access to channel state information (CSI) for direct and reflecting links, which can be acquired through eavesdropping on pilot signals.

This work considers a white-box attack in the context of intelligent O-RAN for V2X architecture. While the attacks may require a strong assumption of the model and parameter access, there are three merits to justify our selection. First, the proposed attacks indicate the common insider threats to target specific vehicular services or customers. Second, the techniques highlight vulnerabilities of DRL-based V2X apps in open-source frameworks like O-RAN. Compared to common adversarial attacks (e.g., FGSM) and the preliminary work [28], this research further enriches the understanding of real-world scenarios with massive user requests, considering outage and

TABLE II
TABLE: NOTATIONS

Symbol	Section	Description
r		rewards
γ	III	discount factor
p		transition probabilities
$\mathcal{M}$		Number of base stations (BSs)
$\mathcal{N}$		Number of user equipment (UEs)
$\mathcal{G}$		Available Resource block groups (RBGs)
ζ	III	Orthogonal slice
g		Assigned resource block to user n
$h_{m,g,n}$		Allocation indicator g of m to UE n
$\xi_{m,n}$		Channel Coefficient for UE n on slice ζ
$P_{m,g}$		Trans.Power to the g-th RBG of BS m
$\Psi_{m,g,n}$		SINR between m and UE n on RBG g
$C_{m,g}$		Transmission capacity of m on RBG g
$R_{m,g}$		Transmission rate of m on RBG g
α		Learning rate for the DRL model
β		Assigned bandwidth for g
$\tau_{m,n}$		Latency for each user-RU pair
$PL_{m,n}$		packet loss rate $< 0.01\%$
$\mathcal{L}_{adv}$		Loss function of the adversary
$\mathcal{R}$		Reward function for the DRL model
θ		Parameters of the target network
H_x	III-D	Channel matrix–CSI
ϕ		Model prediction
$\mathcal{A}$		Set of actions on $\mathcal{S}$
∇_δ		Gradient of δ with respect to $\mathcal{L}$
$\mathcal{S}$		Set of states for the environment
$\mathcal{L}(\theta,\cdot)$		Loss function of $f(\cdot)$
$f(\cdot;\theta)$		Model for resource allocation task
η		Adversarial perturbation constraint
ϵ		Perturbation magnitude
δ	IV-B	Adversarial input
$\pi(\cdot;\theta)$		Policy function of the DRL model
$\nabla_{X,M}$		Gradient of X with respect to M
ω^n		Weight assigned to UE n based on QoS

latency constraints crucial for V2X applications. Finally, this comprehensive approach establishes a foundation for conducting robust defense mechanisms. If a defense model can prevent white-box attacks, the risks of black-box attacks without access to hyper-parameters can be mitigated.

III. System Model and Problem Formulation

The notations used in this work are listed in Table II. Our research is focused on markov decision processes (MDPs) for deep reinforcement learning (DRL) in radio resource allocation (xApp), using Bellman's equation as a key principle [29]. The MDP tuple $(S, A, \gamma, p, r, h_{m,g,n})$ serves as a comprehensive representation, encapsulating key elements: state space (S), action space (A), discount factor (γ), transition probabilities (p), rewards (r), and the initial allocation state distribution $h_{m,g,n}$. The following subsections detail our system model and problem formulation.

A. System Model

In this study, we assume a wireless communication scenario with a downlink orthogonal frequency-division multiple-access system where $\mathcal{M}$ base stations (BS) cater to $\mathcal{N}$ single-antenna onboard units, also known as On-Board Equipment (OBE), in connected and automated vehicles. The network is also subdivided further to be dynamically responsive to distinct slices based on different quality of service (QoS) requirements, where each slice can serve up to N vehicles. The set of BSs and vehicles are denoted as $\mathcal{M} = \{1, 2, \ldots, m, \ldots, M\}$, and $\mathcal{N} = \{1, 2, \ldots, n, \ldots, N\}$, respectively. The base stations connect OBE nodes/users via E2 using the O-RAN interface manager [5], as shown in Fig 1.

Suppose that uplink user transmissions consist of uplink user data, random-access requests, user control channels, and sounding reference signals. The $\mathcal{M}$ transmit a bundle of Resource Block Groups (RBG), the smallest resource unit allocated to a vehicle (or user phone sitting in the car), consisting of 12 consecutive subcarriers. Each BS has a set of G RBGs, denoted as $\mathcal{G}$, $\mathcal{G} = \{1, 2, \ldots, g, \ldots, G\}$. The available bandwidth is divided into ζ slices to avoid interference and implement slice-specific, fair, and efficient scheduling policies. We also define $h_{m,g,n}$ to denote a binary notion of whether the BS m has allocated RBG g to user n, i.e., 1; otherwise 0. Let $\xi_{m,n}$ represent the channel coefficient for user n on slice ζ of the BS m, and the transmission power assigned to the g-th RBG of BS m at time t is denoted by $P_{m,g}$.

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 $\Psi_{m,g,n} =$

$$\frac{h_{m,g,n} \cdot \xi_{m,n} \cdot P_{m,g}}{\sum_{m' \in \mathcal{M}, m' \neq m} \sum_{n' \in \mathcal{N}, n' \neq n} h_{m',g,n'} \cdot \xi_{m',n'} \cdot P_{m',g} + \sigma^2}, \quad (1)$$

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 represents the noise variance.

The transmission capacity of BS m on the RBG g to user n is given by Shannon capacity formula

$$C_{m,n} = \beta_g \cdot \log_2 \left(1 + \sum_{n \in \mathcal{N}} \Psi_{m,g,n}\right), \quad (2)$$

where β_g represents the assigned bandwidth for RBG 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 RBG g, denoted as $R_{m,g}$, is defined as follows

$$R_{m,g} = \begin{cases} C_{m,g}, & \text{if } C_{m,g}T < \sum_{n \in \mathcal{N}} h_{m,g,n} L_n, \\ \frac{L_n}{T}, & \text{if } C_{m,g}T \geq \sum_{n \in \mathcal{N}} h_{m,g,n} L_n \end{cases} \quad (3)$$

where L_n denotes the remaining transmission block size, T is the time duration slot. We also assume optimized power allocated $P_{g,n}$ with consideration of the necessary constraints for each vehicles n in the slice ζ as described in [23]. In addition, we also have incorporated these QoS metrics into the optimization problem. Ultimately, the goal of the baseline system in this case is to optimize the overall transmission rate across all BSs and is expressed as

$$\begin{aligned} &\max_{P_{m,g}, h_{m,g,n}} \sum_{m \in \mathcal{M}} \sum_{g \in \mathcal{G}} R_{m,g}, \quad \text{s.t. (1)}-\text{(3)}, \\ &P_{\min} \leq P_{m,g} \leq P_{\max}, \; h_{m,g,n} \in \{0, 1\}, \\ &\sum_{n \in \mathcal{N}} h_{m,g,n} = 1, \; C_{m,n} \geq C_{\min}, \tau_{m,n} \leq \tau_{\max}, \\ &PL_{m,n} < 0.01\%, \quad \forall m \in \mathcal{M}, g \in \mathcal{G}, n \in \mathcal{N}, \end{aligned} \quad (4)$$

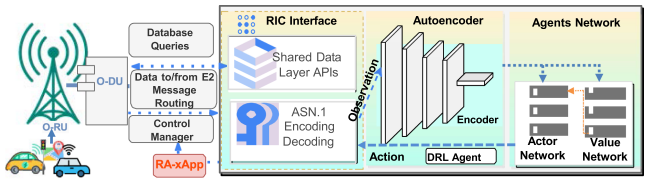


Fig. 2. Structure of the resource allocation xApp in O-RAN, showing RIC interface, Policy Engine with DRL, database, and E2 interface integration.

where the transmission rate of BS m on RBG g, denoted as $R_{m,g}$, depends on the channel capacity and the remaining block size. $P_{\min}$ and $P_{\max}$ set the minimum and maximum power limits for each g. The constraint $\sum_{n\in N} h_{m,g,n} = 1$ ensures that each g can be allocated to only one user at a time. And, where $C_{\min}$ is the minimum achievable data rate from (2), $\tau_{m,n}$ is the latency for each user-RU pair, and $PL_{m,n}$ is the packet loss rate. This explicit inclusion of QoS metrics ensures that our optimization not only maximizes overall system performance but also maintains the stringent requirements of URLLC and V2X applications.

B. Resource Allocation xApp for Maximizing User Access Network Performance for Various Vehicular Communications

Resource allocation involves dividing radio resource blocks (PRB), such as bandwidth, among network users based on their traffic characteristics and QoS needs. The slices can be used for enhanced mobile broadband (eMBB) services, massive machine type communication (mMTC), or ultra-reliable low latency communication (uRLLC) requirements. Each slice focuses on certain efficiency objectives and their related applications, such as in-vehicle entertainment, parking management/toll collection, and remote driving/collision avoidance. The network operator may use the resource allocation xApp to prioritize various sorts of vehicle services. Resource allocation in O-RAN is done via the E2 interface, allowing the RIC to make immediate changes to the distribution of radio resources across the network. The xApp transmits E2 instructions to the O-DU, which is equipped with a medium access control (MAC) scheduler, to regulate the policy for slice management, as seen in Fig. 2. Generally, the E2 Interface Manager acts as the communication nexus between the xApp and the O-DU via the E2 interface. It provides connectivity and essential APIs for seamless message exchange between components. The resource allocation issue aims to identify the most effective technique for maximizing network resource use over time, while ensuring adherence to QoS restrictions and fairness among users ($C_{m,n} \geq C_{\min}, \tau_{m,n} \leq \tau_{\max}$). It involves choosing a signal correlation matrix for each user and sending strategy $\{S_1, \ldots, S_{nr}\}$ while adhering to power limitations. This allocation feature guarantees the most efficient use of network resources according to current circumstances and requirements. The following subsection outlines a DRL-based resource allocation model.

C. Adaptation of DRL Framework

This study adapts the DRL framework from [1], which acts as both a scheduling policy selector and resource allocator. The MDP-based model chooses between proportionally fair (PF) and and round-robin (RR) policies, implicitly determining resource allocation. Based on the fundamental structure, we examine its vulnerabilities in vehicular networks, specifically how attacks influence resource allocation choices in O-RAN settings. This research also seeks to build a particular V2X scenario, which increases the difficulty of managing network dynamics during training, to assess the learning model's flexibility to variable UE resource demands [30]. The adaptation process allowed us to leverage the robust foundation of the previous work on pure AI-based resource allocation models in O-RAN while tailoring the DRL agents to study the security vulnerabilities in specific V2X applications.

D. Problem Formulation

In the context of O-RAN resource allocation, DRL is the most common model to enable dynamic and intelligent allocation of resources [17], [18], [19]. Based on the power of experienced learning and intelligent trial-exploration process in DRL, the agent can enhance the accuracy of resource allocation as well as utilize energy consumption through accurate decisions [27], [31], [32]. This study aims to address the challenges of adversarial attacks against DRL-based resource allocation in the physical layer of enterprise O-RAN platforms, specifically in Internet of Vehicles (IoV) contexts. The problem is characterized by the interplay of real-time processing demands, dynamic network topologies, and ultra-low latency requirements in vehicular networks. The physical layer's signal processing and encoding/decoding processes require near real-time operations, demanding attackers to swiftly identify vulnerable parameters in resource allocation algorithms, given rapidly changing channel environments. A high density of vehicles in V2X further complicates the attack landscape. The target is to exploit these complexities by targeting CQI falsification, manipulating uplink signal patterns/powers, and injecting adversarial noise during V2X communication initialization. Through modeling these interconnected factors to develop subtle attack techniques, the study demonstrates how to exploit the vulnerabilities of O-RAN's open interfaces, ultimately unmasking critical weaknesses in potential O-RAN-based vehicular networks.

To model DRL learning in dynamic vehicular networks, Markovian principles provide an appropriate framework [10], [33]. The agent in this study utilizes a PPO algorithm to adjust to the changing resource requests. Various factors such as wireless conditions, user needs, QoS standards, and resource availability affect the adaptive allocation decisions in this scenario. The main objective of the PPO model is to optimize the allocation of radio resources to vehicle OBEs depending on traffic request volumes. We assume that two pivotal decision variables, namely $h_{m,g,n}$ and β_g, play critical roles in both uplink (UL) and downlink (DL) transmissions. The binary variable $h_{m,g,n}$ controls the allocation of uplink resources, deciding how resource block groups are distributed from BS m to the onboard unit (OBU) of vehicle n. Modifications to $h_{m,g,n}$ affect resource availability, transmission power, and other uplink-related variables, influencing the user's uplink communication capability. Parameter β_g determines the fair and efficient distribution of resources from BS m to OBUs in the context of the downlink.

State Space (S): The state space encompasses several key elements observed over time t, including the traffic load $R^t_{m,g}$, transmission power $P^t_{m,g}$, relevant CSI data as channel conditions $\mathcal{H}^t_{m,g}$, the length of queuing data in the buffer denoted by $L^t_{m,g}$, and an indicator $h_{m,g,n}$. We formally define the state of the resource allocation problem by

$$S^t_m = \left\{h_{m,g,n}, \mathcal{H}^t_{m,g}, R^t_{m,g}, P^t_{m,g}, L^t_{m,g} \mid g \in \mathcal{G}\right\}, \quad (5)$$

where n is the number of vehicles (or users) and m is the BS serving at the moment. The attacker's state space alterations are adjusted to infuse the wrong observation, modifying the DRL agent's perception.

Action space (A): The DRL agent selects a comprehensive action, labeled as A^t_m, based on the current state, retrieving buffer data [bytes], bit rate allocation [Mbps], and total granted resource blocks for each action. The vector $a^t_{m,g}$ represents modifications to resource grants per request for each user n in slice S. The expression for $a^t_{m,g}$ is

$$a^t_{m,g} = \left\{n_0, \ldots, n_{K-1} \mid n_k \in h^t_{m,g,n}\right\}, A^t_m = \left\{a^t_{m,g} \mid g \in \mathcal{G}\right\} \quad (6)$$

Reward($\mathcal{R}$): The reward is a composite measure of the overall throughput allocated to all N vehicles sliced by BS m. The reward at time step t, denoted as $\mathcal{R}^t$, is calculated by

$$\mathcal{R}^t = \sum_{n \in \mathcal{N}} \omega_n \cdot R^t_{m,g}, \quad (7)$$

where ω_n represents the weight assigned to each user n. The network's performance (e.g., throughput, latency, user experience) serves as the reward, guiding the DRL agent's learning process.

This continuous loop allows the DRL agent to learn optimal spectrum allocation strategies in real-time, adapting to constantly changing network conditions. 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. Specifically, the main goal is to create adversarial inputs $x + \delta$ that disrupt the input space x. This involves strategically leveraging the model's responses, including tracking changes in its loss and clipping them to the crafted perturbation to dampen the effect, to infer potential model parameters without explicit knowledge of the underlying model constructs. In this case, we aim to augment the loss of the original model by adding additional perturbations. The adversarial loss function $\mathcal{L}(\theta, x, \delta, \phi)$ is designed to increase the difference between the baseline model's estimate and the ground truth. The initial loss is denoted by $\mathcal{L}(\theta, x, \phi)$, while the loss after perturbation is $\mathcal{L}(\theta, x + \delta, \phi)$. The adversarial loss function is defined as

$$\mathcal{L}_{\text{adv}}(\theta, x, \delta, \phi) = \mathcal{L}(\theta, x + \delta, \phi) - \mathcal{L}(\theta, x, \phi), \quad (8)$$

where the objective is to maximize the difference between the loss incurred by the baseline model (BML) when presented with the original input x and the loss when exposed to the perturbed input $x + \delta$. The presented formulation underscores the nature of a black-box attack, where the adversarial input $x + \delta$ is methodically modified across iterations to analyze the model's response in detail. The iterative refining process focuses on understanding the behavior of the model rather than revealing its underlying parameters explicitly. Through this process, the attack progressively tailors the adversarial input to exploit vulnerabilities in the model's decision-making, increasing the effectiveness of crafted adversarial inputs. The goal is to systematically test how the model reacts to changes and improve the adversarial input without knowing the specific model parameters. This deliberate investigation enables the creation of adversarial inputs that efficiently influence the model's decision bounds. The following section presents a semi-black box attack model.

IV. Proposed Policy Infiltration Adversarial Attacks Against DRL-Based Resource Allocation

O-RAN architecture significantly expands the attack surface of traditional RAN systems by introducing new interfaces (A1, E2, O1, O2, Open FH) and functions (SMO, Non-RT RIC, Near-RT RIC). This expansion, coupled with the integration of third-party applications like xApps in the Near-RT RIC, creates numerous entry points for potential attacks. The open nature of O-RAN, while promoting innovation, also facilitates whitebox attacks, as adversaries gain unprecedented access to system internals. Recent security analyses have identified several high-risk components, including Non-RT RIC, Near-RT RIC, and their associated apps (rApps and xApps), highlighting the vulnerability of this open architecture to sophisticated threats [14], [34], [35], [36]. The use of containerization, virtualization, and open-source software further compounds these risks, making O-RAN implementations potentially more susceptible to both known and zero-day vulnerabilities.

This section details the feasibility of adversarial attacks considered to attack the baseline model and the policy infiltrator algorithms designed to mislead the reward policy of the learning agent in DRL-based resource allocation methods.

A. Attack Model Feasibility, Goals, and Attack Workflow

We investigate weaknesses in the existing O-RAN framework's openness, specifically concentrating on a scenario where an authorized xApp sends control actions to the radio units or next-generation node B (gNB). The xApp uses RAN data and retrieves stored data from the service data layer (SDL) in the near-RT RIC to make intelligent decisions. The near-RT RIC structure's openness makes it vulnerable to security risks, especially via the unencrypted 7.2x interface on the control plane [6], [15]. The O-RAN Alliance has raised vague security concerns regarding the E2 interface, necessitating further certification [37]. The vulnerabilities provide opportunities for impersonation attacks and data compromise. The S-Plane might experience performance deterioration due to interference with synchronization infrastructure. Commercial-off-the-shelf (COTS) software-defined radios (SDR) and open-source cellular software may be used to create hostile user equipment (UE) or OBEs. The attacker can use OBE with a legitimate subscriber network identification, such as a globally unique temporary user equipment identification (GUTI). The attacker can also manipulate the cellular protocol stack to conduct availability attacks by sending counterfeit radio resource control (RRC) requests to deplete resources or penetrate targeted base stations [17]. In addition to open interface vulnerabilities, the separation of the gNB/RU results in more signaling messages between its disaggregated parts compared to a single, undivided gNB/RU. The UE registration process necessitates additional contact between O-DU and O-CU. Attackers can exploit this by impersonating a different UE and reducing the number of registration requests, or

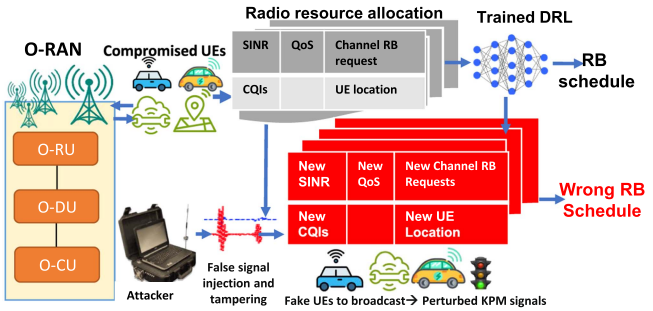


Fig. 3. Adversarial attack scenario on DRL-based resource allocation xApp in O-RAN.

by creating a botnet of compromised UEs (such as smartphones with malware) to generate numerous fake signaling notifications or service requests. Attackers can also influence UE registration and handover operations by creating false RRC communications from a hacked O-DU to the O-CU. As a result, there has been a significant increase in requests to the serving units, adding to the challenge of handling a large number of requests in vehicular networks, particularly in urban areas [18], [25], [38].

Our research focuses on adversarial attack strategies targeting the pre-trained PPO-based resource allocation model through O-RAN's open interfaces. These interfaces allow the integration of specialized AI/hardware accelerators at cell sites, enhancing computational power for generating complex perturbations. In addition, we exploit compromised and counterfeit UE to subtly influence resource allocation decisions, introducing a novel security dimension to wireless networks. A schematic picture depicting our proposed policy infiltration attack against the O-RAN resource allocation model can be seen in Fig. 3. We expose weaknesses in a legitimate xApp inside the RIC to adversarial attacks, where outside entities manipulate the original input to cause erroneous allocation choices. The final target is to deplete network performance by inducing congestion and raising delay. The perturbations are restricted by the inequality $\mathcal{L}(\theta, x, \delta, \phi) \leq \epsilon$, with the aim of influencing resource allocation with malicious intent, such as favoring certain users. Here, ϵ represents a predetermined threshold value. This constraint guarantees that adversarial perturbations are limited to a small error range, avoiding detection by the agent and keeping within the conventional tolerance range. We illustrate the workflow of adversarial attacks on the DRL-based resource allocation model in the O-RAN framework in Fig. 3. An xApp utilizes dynamic resource allocation to assign radio resource blocks to users according to environmental conditions and request density. Adversaries attempt to disrupt channel state estimates by inserting false UEs to transmit adversarial signals, intentionally causing the agent to make incorrect RBG allocation decision.

B. Details of the Adversarial Attack Algorithms

The adversarial attack aims to degrade system performance by distorting resource allocation choices via alterations induced in the state space, with potentially severe consequences in V2X communications. The attack causes minor but significant wrong judgments by using slice inputs as observations, altering learned policy state, and purposely providing the agent with false signal

Algorithm 1: FGSM Attack's Pseudo Code.

1 **Input:** Resource allocation model f, Clean state $\mathcal{H}_t^m$, Resource allocation targets $R_t^{m,g}$, Perturbation budget δ=0.01, Number of users $N = 40$
2 **Output:** Perturbed state $\mathcal{H}_{\text{adv}}$
3 **Procedure:**
4 Initialize $\mathcal{H}_{\text{adv}} = [\,]$
5 **for** *each user* n *in* $\mathcal{N}$ **do**
6 Calculate $\nabla_{\mathcal{H}_m^t} J(f(\mathcal{H}_m^t), R_{m,g}^t)$
7 Compute $\text{sign}(\nabla_{\mathcal{H}_m^t} J(f(\mathcal{H}_m^t), R_{m,g}^t))$
8 Scale $\delta \cdot \text{sign}(\nabla_{\mathcal{H}_m^t} J(f(\mathcal{H}_m^t), R_{m,g}^t))$
9 Update $\mathcal{H}_{\text{adv}}^{m,n} = \text{clip}(\mathcal{H}_{m,n}^t + \delta \cdot \text{sign}(\nabla_{\mathcal{H}_m^t} J(f(\mathcal{H}_m^t), R_{m,g}^t)), R_{m,g}^t)$
10 Append $\mathcal{H}_{\text{adv}}^{m,n}$ to $\mathcal{H}_{\text{adv}}$
11 **end**
12 **return** $\mathcal{H}_{\text{adv}}$

updates. These attacks are particularly concerning in the V2X context due to the safety-critical nature of vehicular communications [39].

In V2X scenarios, adversarial attacks can have far-reaching implications: (1) *safety-critical communications:* attacks targeting resource allocation for collision avoidance or emergency vehicle priority systems could lead to increased accident risks, (2) *platooning disruption:* adversarial perturbations in the state space could destabilize vehicle platoons by affecting inter-vehicle communication resources, (3) *traffic management interference:* false information injection could lead to suboptimal resource allocation for V2I (Vehicle-to-Infrastructure) communications, impacting traffic flow and efficiency, (4) *compromised autonomous driving:* attacks on V2X resource allocation could degrade the performance of cooperative perception and decision-making.

We implement and analyze three adversarial attack algorithms in the context of V2X communications:

1) Fast Gradient Sign Method (FGSM): FGSM is a one-step adversarial attack technique that maximizes the loss function $\mathcal{L}_{\text{adv}}(\theta, x, \delta, \phi)$ [40] in a more efficient way. In resource allocation, FGSM is crafted to perturb RBG extracted from CSI $\mathcal{H}_x$ for vehicle n to generate an adversarial channel matrix, resulting in a higher loss $\mathcal{L}$ that misleads the model to provoke incorrect decisions, as used in [16]. This adversarially perturbed state $\mathcal{H}_{adv}$ as an example is represented by

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

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

Algorithm 1 presents an example of attack pseudo-code. The FGSM technique is developed to generate an evasion pattern against a resource allocation model represented as f. The main goal is to generate adversarial perturbations that can be applied to input states. This technique helps to see how these disturbances affect the model's decision-making on resource allocation. Lines 1–3 set up the resource allocation model (f) as the starting point and initialize the clean state ($\mathcal{H}_m^t$). The perturbation budget (ϵ) is established to regulate the magnitude of modifications. Lines 4–12 cycle across each user in the network, creating variations to the original state using the model's loss gradient. The perturbations are intentionally crafted to slightly impact judgments about resource allocation. Clipping guarantees that the disturbed states

Algorithm 2: 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 δ_x, η), Max_Iter $iterations = 15000$, $\epsilon = 0.01$, η=1.1
Result: Adversarial Input δ
1 **Calculate Baseline** f **Prediction:**
2 **Baseline_prediction** $= f(\phi)$
3 **Iterative Perturbation Generation:**
4 **for** $i = 1$ ***to*** $iterations$ **do**
5 $\nabla J(\delta_x, f(\phi)) = \nabla J\left(f(\phi) \cdot \mathcal{L}(\theta, \cdot)\right) \rightarrow$ **Calc-Gradient**
6 $\delta_x \leftarrow \delta_x + \epsilon \cdot \text{sign}(\nabla_x J(\delta_x, f(\phi))) \rightarrow$ **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* $\neq (f(\phi)$ ***or*** $\|\delta_x - \mathbf{X}\|_\infty > \epsilon$ **then**
11 **Break** out of the loop
12 **end**
13 **end**
14 Perturbed $\mathcal{R} \leftarrow \mathcal{A}|\mathcal{S}$

stay within acceptable limits. The effect of these disrupted conditions on the resource allocation model is evaluated by examining their impact on the agent's decision-making process. The process continues through many iterations until the adversarial model successfully causes the baseline model to make anomalous judgments (Lines 10–12). This induced perturbation disrupts the baseline model's behavior while remaining concealed. The outcome is a set of altered states ($\mathcal{H}_{\text{adv}}$), revealing possible weaknesses or sensitivities in the resource allocation model.

2) Policy Infiltration Attack (PIA): This method leverages a proficient adversarial attack agent crafted to infiltrate the policy dictating vehicle throughput requirements and available resource blocks (RBs). Our technique uses undetectable random noise ϵ to disrupt the state space, effectively taking advantage of weaknesses in the model and hindering its capacity to reach the desired goals. PIA, unlike FGSM, is designed to analyze the behavior of a pre-trained model. The PIA modifies input repeatedly for a specified number of $iterations$, using gradient information in each step, and stops the process if the perturbation is beyond a set threshold. The gradient of the loss function for the PIA is computed as shown in 14 Line 5. Furthermore, termination occurs only when the perturbation exceeds a predefined threshold. This guarantees that the attacks stay within a specified range of perturbation intensity, achieving a careful equilibrium between attack effectiveness and subtlety. This strategy is suitable for executing intricate and sustained attacks that can bypass defenses designed to counter single-step attacks.

Algorithm 2 displays the pseudo-code for the attack algorithm designed to mimic the behavior of a pre-trained model represented as f. The method creates the initial model (Lines 1–3) and uses it to train the adversarial attack model. The adversarial training proceeds through $iteration$ iterations (Lines 4–10) with the goal of consistently misleading the baseline model. The integration of perturbed data is strategically designed to evaluate the agent's decision-making surreptitiously. The training process persists until the adversarial model can effectively induce the baseline model to produce abnormal decisions (Lines 10–12). The PIA uses a white-box approach to attack a pre-trained model f by iteratively modifying the adversarial input δ_x. The adjustment actively seeks to maximize the discrepancy between the original and perturbed losses by adhering to the guidance of the loss function $\mathcal{L}(\theta, \cdot)$ (Line 5), as shown in (8). The iterative refinement of δ_x probes the model's behavior without direct access to internal parameters, enhancing the effectiveness of the adversarial samples. The stopping criteria involve checking differences between infiltrated and baseline predictions and perturbation thresholds. Encouraging significant leaps or learning explosions in policy updates could lead to risky outcomes, akin to dropping off a specific network slice and resulting in suboptimal policies and incorrect radio resource allocation actions.

We evaluated the sensitivity of the PIA with varying ϵ magnitude and even demonstrated its fooling capability in a slice-specific environment. By varying ϵ, we explore how sensitive the model's predictions are to small changes in the input. Let δ_x be the adversarial perturbation applied to the original input $\mathcal{H}_x$. The perturbed input is given by $\mathcal{H}'_x = \mathcal{H}_x + \delta_x$. We constrain δ_x such that $\|\delta_x\|_\infty \leq \epsilon$, where $\|\cdot\|_\infty$ denotes the maximum absolute value norm. We examined how the success rate of creating adversarial cases fluctuates with variations in ϵ. The PIA algorithm's well-targeted policy penetration and flexibility make it a strong and efficient weapon for sophisticated adversarial attacks, surpassing older methods in situations requiring a greater degree of attack complexity and persistence.

In short, PIA is a strategy of deceptive incentives for exploration. We consider adversarial attacks as a deceptive optimization problem aimed at deceiving the intended transmission rate, as shown in (4). The attacker manipulates the power transmission, UE location, and resource allocation choices for a compromised UE in a deliberate manner, which impacts ω_n in (7). Let $Q(\omega)$ be the performance measure that has to be optimized. The attacker collaborates with vehicle OBEs and users to provide fabricated feedback, modifying the true performance metric to a deceptive one, $Q_{\text{deceptive}}(\omega)$. The DRL agent seeks to maximize its reward by adhering to rules that maximize $Q_{\text{deceptive}}(\omega)$, resulting in poor resource allocation methods (π^*). This deviation from the network's true optimization objective causes biased resource allocation decisions, impacting users' QoS, which can be defined by

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

In the V2X context, the biased resource allocation decision could manifest as manipulating the perceived urgency of different communication types, leading to suboptimal prioritization of safety-critical messages or degradation of QoS for V2X applications such as collision avoidance.

3) Projected Gradient Descent (PGD): The PGD attack improves the existing FGSM method by integrating a powerful multi-step gradient-based approach [41]. By using a tiny step size α, the altered input is mapped onto the $\mathcal{L}$ that encompasses the initial clean input chosen inside the ϵ-neighborhood of $\mathcal{H}_x$. The iterative procedure is guided by the gradient of the loss function to maximize the model's loss inside a ϵ-neighborhood. The iterative process for a N-step PGD attack is specified for the n-th iteration as follows:

$$\mathcal{H}_{x_0} = \mathcal{H}_x, \tag{11}$$

$$\mathcal{H}_n = \text{clip}\left[\mathcal{H}, \epsilon\right]\left\{\mathcal{H}_{n-1}, +\alpha \cdot \text{sign}\left(\nabla_{\mathcal{H}_{n-1}} J(\theta, \mathcal{H}_{n-1}, \phi)\right)\right\}, \tag{12}$$

$$\mathcal{H}_{\text{adv}} = \mathcal{H}_N, \tag{13}$$

where clip$[\mathcal{H}_x, \epsilon](\cdot)$ is element-wise clipping to $[\mathcal{H}_x - \epsilon, \mathcal{H}_x + \epsilon]$ so that the outcome stays in $\mathcal{L}$ and ϵ-neighborhood of $\mathcal{H}_x$. θ is the model's parameters, $\mathcal{H}_x$ is the input to the model as state space, and ϕ is the target label the adversary aims to fool.

Unlike PIA, PGD employs a direct and forceful optimization technique that centers on a series of repetitive phases for perturbation. PIA disrupts a model's behavior subtly by encouraging deception, whereas PGD aligns with the domain dynamics, introducing persistent, multi-step systemic adjustments. The fundamental difference is that PGD is characterized by its stealthiness and focus on a continuous adjustment method across several domains, while PIA adapts to the pre-trained resource allocation model being considered, directing the loss path towards deception of the model.

The PGD algorithm methodically focuses on the resource allocation state ($\mathcal{H}_x^m$) for every user in the network. The pseudocode of Algorithm 3 demonstrates an iterative process where parameters are adjusted by calculating the model's loss gradient (Line 6) and applying perturbations strategically (Lines 7–10) within a specified range (Line 11). The process iterates across a set number of iterations (Lines 4–12), enabling the model to adjust to introduced perturbations. Resource allocation feasibility is assessed at each stage (Lines 13–15), documenting any instances of infeasibility. The result is an adversarial state $\mathcal{H}_{\text{adv}}$ that illuminates potential weaknesses in the model under adversarial conditions (Line 16). The PGD tactic is distinct from the PIA technique. PIA alters the model's behavior to generate unconventional judgments (Lines 4–12), aiming to stay concealed while operating despite the disruption generated. PGD can be seen as a special case of PIA attack but under a small scope (by using clipping/sampling techniques). PGD is suitable when we have limited time and resources to penetrate and imitate all the policies of the DRL agent.

These algorithms demonstrate the vulnerability of DRL-based resource allocation in O-RAN for V2X applications, particularly in safety-critical and time-sensitive scenarios. Our analysis of these attacks underscores the need for robust defense mechanisms to ensure the reliability and safety of vehicular communications in the face of potential adversarial interventions.

Algorithm 3: PGD Attack.

Data: Clean state $\mathcal{H}_x^m$, Target resource allocation model f, Perturbation budget $\epsilon = 0.01$, Step size $\alpha = 0.1$, Number of iterations $iterations = 15000$
Result: Adversarial State $\mathcal{H}_{\text{adv}}$
1 **Initialization:** $\mathcal{H}_{\text{adv\_list}} = [\,]$;
2 **for** $i = 1$ *to* $iterations$ **do**
3 **for** *each user* n *in the network* **do**
4 **for** *each resource allocation parameter* θ **do**
5 Calculate the gradient of the loss w.r.t. $\mathcal{H}_x$ and metric θ:
6 $\nabla J(f(\mathcal{H}_x), \mathcal{H}_x, \theta) = \frac{\partial J(f(\mathcal{H}_x), \mathcal{H}_x, \theta)}{\partial \mathcal{H}_{x,n,p}}$;
7 Compute the perturbation as a sign of the gradient:
8 $\delta = \text{sign}(\nabla J(f(\mathcal{H}_x), \mathcal{H}_x, \theta))$;
9 Add scaled perturbation to the resource allocation parameter:
10 $\mathcal{H}_{\text{adv},n,p} = \mathcal{H}_{x,n,p} + \alpha \cdot \delta$;
11 Clip $\mathcal{H}_{\text{adv},n,p}$ to stay within the valid range:
12 $\mathcal{H}_{\text{adv},n,p} = \text{clip}(\mathcal{H}_{\text{adv},n,p}, \mathcal{H}_{x,n,p} - \epsilon, \mathcal{H}_{x,n,p} + \epsilon)$;
13 **end**
14 **end**
15 Calculate the perturbed state $\mathcal{H}_{\text{adv}}$;
16 **if** *resource allocation is infeasible* **then**
17 $\mathcal{H}_{\text{adv\_list}}$.append$(i)$;
18 **end**
19 **end**
20 **return** $\mathcal{H}_{\text{adv}}$

TABLE III
COMPARISON OF FEATURES SUPPORTED IN THREE ATTACK TYPES

Feature name	FGSM	PIA	PGD
Attack type	White-box attack	White-box attack	White-box attack
Requirement	No compromised UE	Compromised UE	Compromised UE
Affect	State input	Policy	Policy
Iteration	Infinity loop	Max iterations	Fixed budget
Return Value	Gradient descent	Policy imitation	Typical policy
Running Time	Fast	Slow	Medium
Attack efficiency (Successful rate)	Low	High	Medium but high damage
Required Memory	Low	High	High

C. Discussion on Exploiting Vulnerabilities in Intelligent O-RAN for Multiple Vehicular Services

The physical layer's openness in intelligent O-RANs presents a notable and urgent risk, demonstrated by three specific attack vectors: manipulation of CSI, interference with DRL model policies, and expedited penetration of the DRL system to enhance attack efficacy. The first attack, known as FGSM, takes advantage of the vulnerability of CSI to manipulation by introducing meticulously designed noise, presenting a real danger without undermining the integrity of the gNB or the AI model. This strategy, which does not rely on strong preconceptions, offers a very practical opportunity for exploitation. The second attack, known as PIA, attempts to influence the decision-making process of the DRL model by altering power transmission amplitudes. Meanwhile, the third attack, PGD, intends to enhance penetration speed to increase the damage caused by the PIA-based attack.

Table III provides a comprehensive overview of the key distinctions among these attack methods, highlighting their respective features, requirements, and efficiencies. All three attacks are categorized as white-box or semi-white block attacks, affording the attacker partially or completely access to the system, including model parameters. The white-box adversary can be seen as the strongest adversarial attack, i.e., the adversary can directly craft adversarial samples on the target model. Although the attack requires an assumption that we can compromise vehicles (e.g., stolen vehicles), this attack type reveals an important note to build a powerful resilient DRL-resource allocation model (beyond this study) that can resist our adversarial attacks. If an AI model can survive our strong attacks, it is trivial to resist less effective attack types like black-box attacks, where the attacker has limited or no access to the target model's architecture and parameters.

In short, FGSM focuses on state input (noise impact) and environmental observation (state changes over time). This attack has low efficiency but requires less memory to launch. In contrast, PIA and PGD prioritize policy imitation, requiring more memory resources for storing and analyzing policy changes. However, these last attacks offer heightened efficiency by either

TABLE IV
EXPERIMENT SETUP

Parameter	Value	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
Scheduling Policies	(PF), (WF), (RR)	Mobility	Time-varying UEs
Slice	**Traffic Class**	**Min QoS** (C_min)	**Latency** τ_{max}
0	eMBB	1 Mbps per UE	100 ms
1	MTC	30 pkt/s of 125 bytes per UE	500 ms
2	URLLC	10 pkt/s of 125 bytes per UE	1 ms

TABLE V
HARDWARE SPECIFICATIONS USED IN THE EXPERIMENT

Resource	Specification	Details
CPU	Intel CoreTM i7-10700F @2.9GHz	Multi-core processor
RAM	16GB	DDR4 2666 MHz
GPU	NVIDIA RTX4090 24G SUPER	GDDR6
CUDA Version	12.2 with 1408 CUDA cores	GPU computing toolkit

increasing damage or reducing allocation performance. Nevertheless, achieving such efficiency comes at the cost of extended training time, necessitating prolonged iterations and increased running times.

V. PERFORMANCE ANALYSIS

This work uses a pre-trained resource allocation model [1] to execute attack testing, assessing its performance in a highly populated metropolitan zone. The study also uses the real-world dataset from the Colosseum O-RAN dataset (OpenCelliD) [1]. The model, which was trained on an extensive seven GB dataset, includes a variety of performance metrics (such as throughput and bit error rate), system state information (such as transmission queue size, SINR, and QoS), and resource allocation methods (including slicing and scheduling algorithms). The training data was systematically collected during 89 hours of trials conducted on Colosseum, a renowned wireless network simulator [2].

This work utilizes established scheduling paradigms in the O-RAN platform [2], including proportionally fair (PF), waterfalling (WF), and round-robin (RR), to ensure a fair distribution of resources across various slices. These scheduling techniques allow flexible scheduling and adaptive resource allocation in the target network slices. This also guarantees that the optimization is nuanced and adaptable, taking into account the specific properties of different network slices in reality. Table IV shows the default experiment setup and parameters used in the evaluation. The model is trained and tested using the hardware parameters provided in Table V.

A. Performance Evaluation Scenarios

In this study, we conducted a performance analysis on the pre-trained resource allocation model while considering the presence of adversarial agents. The study focused on identifying the strengths and weaknesses of each attack strategy and the associated trade-offs. Numerical characteristics were retrieved from the reward log for a comprehensive comparison of key assessment criteria including packet delivery rate, successful resource block (RB), average cumulative reward, and user data rate. The effects of these attacks can be seen in Table VI, which shows the data rate before and after the attack. The table displays data rates in megabits per second (Mbps) in a multi-slice scenario, where UEs are statically allocated to network slices:

TABLE VI
DATA RATE MEASURE UNDER NO ATTACK AND AFTER ATTACK

Slice	No attack	FGSM	PIA	PGD	Degradation (%)		
	Mbps	Mbps	Mbps	Mbps	FGSM	PIA	PGD
eMBB	3.39	3.01	2.36	2.55	12	29	18
mMTC	0.14	0.12	0.12	0.12	26	22	22
uRLLC	0.05	0.04	0.03	0.03	20	40	40

(i) eMBB, representing users requesting video traffic or infotainment; (ii) mMTC for sensing applications or parking management; and (iii) URLLC for latency-constrained applications such as remote driving and collision avoidance. Therefore, the agents get rewards based on satisfying specific key performance indicator (KPI) requirements. Initially, the eMBB and mMTC agents are trained to maximize UE throughput, while URLLC agents focus on minimizing latency by allocating resources as quickly as possible.

B. Measurement Metrics for Evaluations

For measurement metrics, we use packet delivery rate (PDR), which is the ratio of successful RB allocations to total RB allocation attempts by the agents. Besides, we use average cumulative reward (Avg-CR) to measure UE data rate satisfaction over time, emphasizing compliance with predefined minimum bit-rate criteria across all model-generated incentives. We also use a successful RBG allocation rate for each slice, which is a measure of the agent's request reply relay determined by counting the number of successful RB allocations per slice. We assessed the performance effect of the resource allocation mechanism by analyzing user data rates.

C. Attack Performance of Three Models on Data Rate

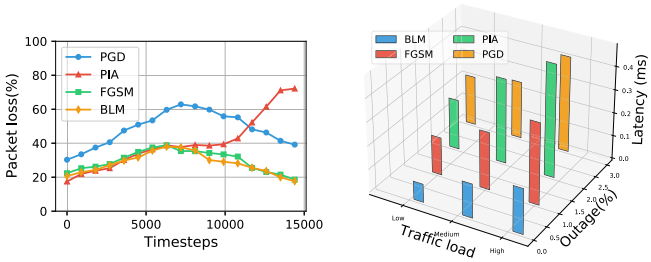
The intensity of the negative impact on connection quality produced by the number of failed attempts to deliver RBs to UEs is proportional to the allocation's failure to fulfill the three-slice latency criteria. This relationship arises due to the extended delays induced by adversarial attacks in fulfilling resource allocation requests. As a result, packets are forced to be discarded because of timeouts. In uRLLC service, delays may have major consequences, such as disrupting connections and causing high-speed vehicles to collide on highways. Table VI shows that the PIA attack outperforms the FGSM and PGD attacks in terms of reducing UE data rates. This efficiency discrepancy is because PIA has an instantaneous impact on DRL policy (precise attack), but FGSM takes time to influence the process via DRL observations (brute-force attack). In the performance comparison of the three attacks, we found that PIA has a more negative effect on PDR than PGD does. Compared to the FGSM attack, PGD reduces PDR by 37.83% and decreases the data rate significantly, reducing resource allocation efficiency in the O-RAN xApp.

D. Attack Performance of Three Attacks on Packet Drop/Delay

This subsection analyzes the effects of the attacks on the QoS and quality of experience (QoE) of different network slices, such as eMBB, mMTC, and uRLLC. It also shows how the attacks exploit network congestion and interference to cause delays and errors in resource allocation, especially for low-latency

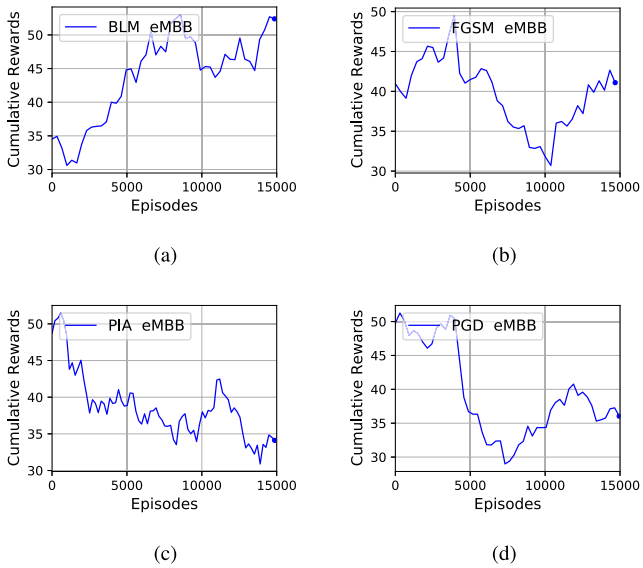
TABLE VII
MODEL PERFORMANCE METRICS UNDER DIFFERENT SCENARIOS

Scenario	PDR%	Successful RBA	Avg-CR	Attack Severity
No Attack	82.58	52511	55.22	None
FGSM	82.01	40834	42.57	Low
PIA	22.26	9473	9.96	High
PGD	62.17	26459	27.82	Medium



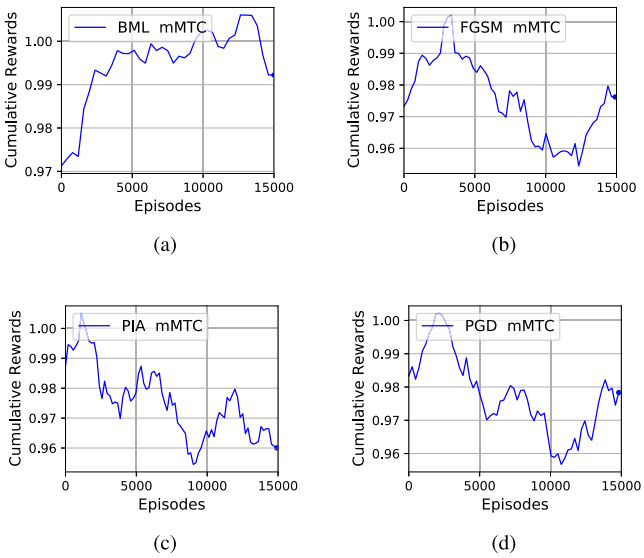
(a) (b)

Fig. 4. Impact of adversarial attacks (FGSM, PIA, PGD) on O-RAN xApp performance. (a) Packet Drop Rate % per iterations. (b) Latency and outage of uRLLC devices under different traffic load.



(a) (b) (c) (d)

Fig. 5. eMBB Slice: Reward accumulation trajectories for eMBB users under different attack scenarios versus baseline (no-attack). PIA (Policy Infiltration Attack) shows the most severe impact, followed by PGD and FGSM, with the baseline achieving the highest reward. (a) Baseline Model (BML). (b) After FGSM attack. (c) Policy Infiltration Attack (PIA). (d) After PGD attack.



(a) (b) (c) (d)

Fig. 6. mMTC Slice: Cumulative rewards for mMTC slice under adversarial attacks and baseline scenario, illustrating the impact on IoT services. (a) Baseline Model(BML). (b) After FGSM attack. (c) Policy Infiltration Attack(PIA). (d) After PGD attack.

services. The performance evaluation in Table VII suggests that the 'No Attack' scenario showcases the DRL-driven resource allocation model's impressive performance, with a high packet delivery rate, successful RBG allocation, and the lowest packet loss rate as depicted in Fig. 4(a). Fig. 4(b) illustrates that the PIA attack in the 'After Attack' scenario results in a 77.74% reduction in packet delivery rates, indicating a significant decline in QoS. The figure also shows that successful PRB allocations decreased, resulting in a 40% decrease in data transmission rate and an additional latency of up to 2.5 milliseconds across various traffic loads (a 150% increase to τ). The high effects of PIA might be associated with its precise targeting method, which directly influences decision-making processes inside the resource allocation model. The findings emphasize the considerable damage caused by the PIA on the QoS and QoE of uRLLC users, indicating serious threats to their safety and operational efficiency.

Besides, Fig. 4(b) also shows that the PGD attack is more damaging than the FGSM attack. The PGD attack leads to a 20% increase in packet drops and up to 0.4 milliseconds of delay for uRLLC devices. The packet drops indicate that the PGD attack is more successful at influencing DRL-based resource allocation models via iterative gradient-based perturbations compared to the FGSM attack, which relies on a single perturbation approach. Our attacks significantly impact low-latency services. The figure also indicates that uRLLC services encounter higher latency and more frequent outages with rising traffic demand. Adversarial attacks use increased network congestion and interference to amplify delays and errors in resource distribution. As a result, the performance and reliability of uRLLC devices are compromised, making these systems more susceptible to malicious tampering in the face of increased network demands. Figs. 5, 6, and 7 demonstrate that the PIA attack outperforms the PGD and FGSM attacks, resulting in a more significant decrease in reward. This comparison is based on the outcomes observed under various traffic conditions (sparse density, medium density, and heavy density of vehicles) in the eMBB, mMTC, and uRLLC contexts.

In highly dynamic V2X scenarios, our results reveal an intriguing phenomenon: adversarial attacks show reduced effectiveness while unexpectedly decreasing latency. Fig. 4(b) shows the latency difference between the baseline model and PIA narrowing from 2.5 ms at low traffic to 1.5 ms at high traffic (30–40 vehicles/km^2). However, this apparent improvement masks severe performance degradation, with Table VII showing PDR dropping from 82.58% to 22.26% under PIA in challenging scenarios. Conversely, in stable environments, attacks are more effective, with Table VI demonstrating a 40% reduction

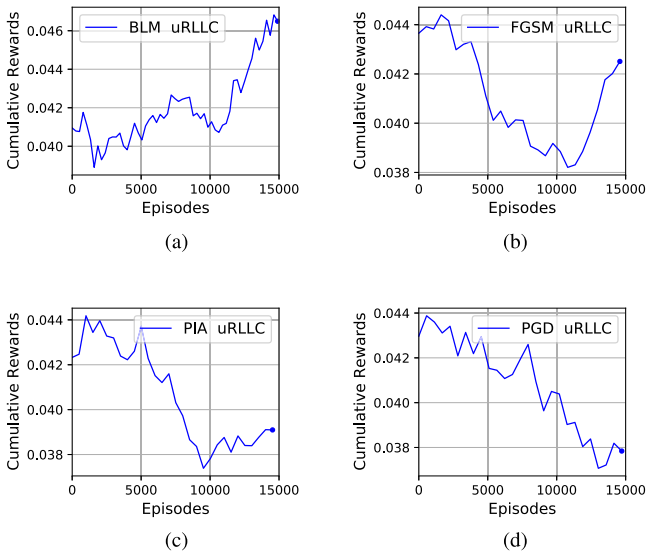


Fig. 7. uRLLC Slice: Cumulative rewards reflect user satisfaction pattern with the agent's resource allocation decisions. Declining cumulative rewards due to adversarial attacks exposes vulnerabilities, hinting at potential disruptions in critical uRLLC applications like remote driving and industrial automation. (a) Baseline Model (BML). (b) After FGSM attack. (c) Policy Infiltration Attack (PIA). (d) After PGD attack.

TABLE VIII
UE DATA RATE COMPARISON BEFORE AND AFTER THE ATTACK

Scenario	[Ours-attacks case] ϵ=0.01				Study [16] ϵ=0.04, ϵ=0.05		
	No Attack	FGSM	PIA	PGD	No Attack	FGSM	PGD
10 UEs	0.8	0.7	0.6	0.2	1.6	0.7	0.8
20 UEs	1.7	1.6	1.2	0.9	1.7	0.7	0.9
30 UEs	2.5	2.2	1.8	1.9	2.0	0.8	1.0
40 UEs	3.4	3.0	2.4	2.5	2.4	1.0	1.2

Data rate in this evaluation is measured by Mbps.

in uRLLC slice data rate under PIA. These findings highlight how adversarial attacks can degrade the DRL-based model's performance and increase latency for allocation decisions while compromising overall network reliability.

E. Attack Capability Comparison on Data Rate With Literature

In a thorough review, we compared our system performance to that of [16], which closely aligned with our work. Table VIII shows that both two techniques have comparable attack performance on user data rate. However, our attack focuses on the DRL interaction phase rather than CNN-based signal categorization. Furthermore, the previous study used a bigger perturbation magnitude ϵ (0.04/0.05 versus 0.01), emphasizing the increased cost of implementing perturbation in their study. The observed differences in data rates under the 'No Attack' condition, compared to prior research, stem from various factors within the experimental setup. These factors encompass variations in the resource allocation model, characteristics of the training dataset, and the network conditions studied.

It is essential to take into account the complexity of the algorithm's learning behavior, which is an iterative optimization method used in our analysis. The differences in data rates come from variations in the convergence speed of the DRL-based PPO algorithm during training. The significant decrease in user data rates, particularly in PGD circumstances, suggests potential disruptions in O-RAN network connections and higher latency. As shown in Table VIII, the PIA attack reduces the UE data rate by 1 Mbps from 3.4 Mbps in the baseline model "No attack" when 40 UE queries are made concurrently. This drop increases the probability of packet loss, resulting in reduced service quality, potentially leading to severe consequences such as lost connections for applications sensitive to delay. The reductions in resource allocation shown in the PIA-based attacks also reveal concerns about poor O-RAN capability under adversarial attacks.

VI. CONCLUSION

This work presents three methods (FGSM, PIA, PGD) of attacking the DRL-based resource allocation models in intelligent O-RAN for vehicular networks by altering signal power or resource requests, which leads to incorrect allocations and reduced data rates. FGSM focuses on less computing approaches, such as generating sequential adversarial perturbations, while PIA and PGD prioritize policy imitation, requiring more memory resources for storing and analyzing policy changes. Among the three, PIA can cause severe degradation in connectivity QoS with the greatest damage in uRLLC, resulting in data rate deterioration between 29% and 40%. PIA also precipitated a staggering 77.74% reduction in packet delivery rate, underscoring a severe QoS deterioration. During attack times, uRLLC services experience an increase in latency and outage, reaching 150% and 30%, respectively.

Based on our findings, we propose four countermeasures to enhance model resilience: (1) adversarial training with adversarial examples (e.g., manipulated CQI reports, falsified GPS coordinates, simulated signal fluctuations). (2) robust DRL model–implementing input bounds (e.g., CQI values, locations, signal strength) and using historical data for anomaly detection–with defensive distillation or signal processing hardening. (3) radio positioning-based vehicles locations verification (e.g., triangulation, multi-array antennas) to independently confirm UE position reports. (4) expanding to heterogeneous networks (e.g., space-air-ground integrated for IoV, multi-agent systems) which introduce more challenges for creating adversarial attacks and defenses.

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 Commun. Mag.*, vol. 59, no. 10, pp. 21–27, Oct. 2021.
[2] L. Bonati, M. Polese, S. D'Oro, S. Basagni, and T. Melodia, "OpenRAN gym: AI/ML development, data collection, and testing for O-RAN on PAWR platforms," *Comput. Netw.*, vol. 220, 2023, Art. no. 109502.
[3] M. Polese et al., "Empowering the 6G cellular architecture with open RAN," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 2, pp. 245–262, Feb. 2024.
[4] P. H. Masur, J. H. Reed, and N. K. Tripathi, "Artificial intelligence in open-radio access network," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 37, no. 9, pp. 6–15, Sep. 2022.
[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 Commun.*, vol. 30, no. 1,, pp. 70–77, Feb. 2023.
[6] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding O-RAN: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 1376–1411, Secondquarter 2023.

[7] R. Kaliski, S.-M. Cheng, and C.-F. Hung, "Supporting 6G mission-critical services on O-RAN," *IEEE Internet Things Mag.*, vol. 6, no. 3, pp. 32–37, Sep. 2023.

[8] O. Sunay et al., "SD-RAN: ONF's software-defined RAN platform consistent with the O-RAN architecture," Open Network Foundation, White Paper, 2020. [Online]. Available: https://opennetworking.org/wp-content/uploads/2020/08/SD-RAN-v2.0.pdf

[9] C.-H. Lai, L.-H. Shen, and K.-T. Feng, "Intelligent load balancing and resource allocation in O-RAN:A multi-agent multi-armed bandit approach," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, 2023, pp. 1–6.

[10] Y. Cui, X. Yang, P. He, D. Wu, and R. Wang, "O-RAN slicing for multi-service resource allocation in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 73, no. 7, pp. 9272–9283, Jul. 2024.

[11] A. Masaracchia, V. Sharma, M. Fahim, O. A. Dobre, and T. Q. Duong, "Digital twin for open RAN: Toward intelligent and resilient 6G radio access networks," *IEEE Commun. Mag.*, vol. 61, no. 11, pp. 112–118, Nov. 2023.

[12] J. Qiu, L. Du, Y. Chen, Z. Tian, X. Du, and M. Guizani, "Artificial intelligence security in 5G networks: Adversarial examples for estimating a travel time task," *IEEE Veh. Technol. Mag.*, vol. 15, no. 3, pp. 95–100, Sep. 2020.

[13] S. P. Jason and S. Boswell, "Zero trust architecture for evolving radio access networks," Ericsson, Tech. Rep., Nov. 2023. [Online]. Available: https://www.ericsson.com/4ae256/assets/local/reports-papers/further-insights/doc/open-ran-security.pdf

[14] O-RAN Working Group 11, "O-RAN security threat modeling and remediation analysis 4.0," WG11.O-RAN-Threat-Model-v04.00, 2023. [Online]. Available: https://orandownloadsweb.azurewebsites.net/specifications

[15] C. Benzaïd and T. Taleb, "AI for beyond 5G networks: A cyber-security defense or offense enabler?," *IEEE Netw.*, vol. 34, no. 6, pp. 104–147, Nov./Dec. 2020.

[16] N. N. Sapavath, B. Kim, K. Chowdhury, and V. K. Shah, "Experimental study of adversarial attacks on ML-Based xApps in O-RAN," in *Proc. IEEE Glob. Commun. Conf.*, 2023, pp. 6352–6357.

[17] E. Habler et al., "Adversarial machine learning threat analysis and remediation in open radio access network," 2023. [Online]. Available: https://arxiv.org/pdf/2201.06093.pdf

[18] D. Adesina, C. Hsieh, Y. Sagduyu, and L. Qian, "Adversarial machine learning in wireless communications using RF data: A review," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 1, pp. 77–100, Firstquarter 2023.

[19] J. Groen et al., "Implementing and evaluating security in O-RAN: Interfaces, intelligence, and platforms," *IEEE Netw. Mag.*, 2024, doi: 10.1109/MNET.2024.3434419.

[20] M. Sadeghi and E. G. Larsson, "Adversarial attacks on deep-learning based radio signal classification," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 213–216, Feb. 2019.

[21] D. Ke, X. Wang, K. Huang, H. Wang, and Z. Huang, "Minimum power adversarial attacks in communication signal modulation classification with deep learning," *Cogn. Computation*, vol. 15, no. 2, pp. 580–589, 2023.

[22] Z. Lu, C. Zhong, and M. C. Gursoy, "Dynamic channel access and power control in wireless interference networks via multi-agent deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 1588–1601, Feb. 2022.

[23] B. R. Manoj, M. Sadeghi, and E. G. Larsson, "Downlink power allocation in massive MIMO via deep learning: Adversarial attacks and training," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 2, pp. 707–719, Jun. 2022.

[24] S. Soltani, M. Shojafar, A. Brighente, M. Conti, and R. Tafazolli, "Poisoning bearer context migration in O-RAN 5G network," *IEEE Wireless Commun. Lett.*, vol. 12, no. 3, pp. 401–405, Mar. 2023.

[25] Y. E. Sagduyu, T. Erpek, and Y. Shi, "Adversarial machine learning for 5G communications security," in *Game Theory and Machine Learning for Cyber Security*. Hoboken, NJ, USA: Wiley-IEEE Press, 2021, pp. 270–288.

[26] A. Bahramali, M. Nasr, A. Houmansadr, D. Goeckel, and D. Towsley, "Robust adversarial attacks against DNN-Based wireless communication systems," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2021, pp. 126–140.

[27] F. Wang, M. C. Gursoy, and S. Velipasalar, "Adversarial reinforcement learning in dynamic channel access and power control," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2021, pp. 1–6.

[28] Y. A. Ergu, V.-L. Nguyen, R.-H. Hwang, Y.-D. Lin, C.-Y. Cho, and H.-K. Yang, "Unmasking vulnerabilities: Adversarial attacks against DRL-based resource allocation in O-RAN," in *Proc. IEEE Int. Conf. Commun.*, 2024, pp. 2378–2383.

[29] R. Bellman, "A Markovian decision process," *J. Math. Mechan.*, vol. 6, pp. 679–684, 1957.

[30] P. Sroka, Ł. Kułacz, S. Janji, M. Dryjański, and A. Kliks, "Policy-based traffic steering and load balancing in o-ran-based vehicle-to-network communications," *IEEE Trans. Veh. Technol.*, vol. 73, no. 7, pp. 9356–9369, Jul. 2024.

[31] H. Zhang, H. Zhou, and M. Erol-Kantarci, "Team learning-based resource allocation for open radio access network (O-RAN)," in *Proc. IEEE Int. Conf. Commun.*, 2022, pp. 4938–4943.

[32] S.-P. Yeh, S. Bhattacharya, R. Sharma, and H. Moustafa, "Deep learning for intelligent and automated network slicing in 5G open RAN deployment," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 64–70, 2024.

[33] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "ColO-RAN: Developing machine learning-based xApps for OpenRAN closed-loop control on programmable experimental platforms," *IEEE Trans. Mobile Comput.*, vol. 22, no. 10, pp. 5787–5800, Oct. 2023.

[34] Quad Critical and Emerging Technology Working Group, "Open ran security report," May 2023. [Online]. Available: https://www.ntia.gov/sites/default/files/publications/open_ran_security_report_full_report_0.pdf

[35] A. Chiejina, B. Kim, K. Chowhdury, and V. K. Shah, "System-level analysis of adversarial attacks and defenses on intelligence in O-RAN based cellular networks," in *Proc. 17th ACM Conf. Secur. Privacy Wireless Mobile Netw.*, 2024, pp. 237–247.

[36] H. Jiang, H. Chang, S. Mukherjee, and J. Van der Merwe, "OZTrust: An O-RAN zero-trust security system," in *2023 IEEE Conf. Netw. Function Virtualization Softw. Defined Netw.*, 2023, pp. 129–134.

[37] O-RAN Next Generation Research Group (nGRG), "Rr-2023-04," O-RAN Alliance, Tech. Rep., Dec. 2023, Quantum security considerations for future RAN architectures [Online]. Available: https://mediastorage.o-ran.org/ngrg-rr/nGRG-RR-2023-04-Quantum-Security-v1_1.pdf

[38] G. Apruzzese, R. Vladimirov, A. Tastemirova, and P. Laskov, "Wild networks: Exposure of 5G network infrastructures to adversarial examples," *IEEE Trans. Netw. Service Manag.*, vol. 19, no. 4, pp. 5312–5332, Dec. 2022.

[39] V. Sharma, I. You, and N. Guizani, "Security of 5G-V2X: Technologies, standardization, and research directions," *IEEE Netw.*, vol. 34, no. 5, pp. 306–314, Sep./Oct. 2020.

[40] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," 2015, *arXiv:1412.6572*.

[41] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," 2019, *arXiv:1706.06083*.