Trustworthy and Efficient Deep Reinforcement Learning-Driven Physical-Layer Security for 6G Networks

Alex Pierron[1], Joaquin Garcia-Alfaro[1], Jose Rubio-Hernan[1], Michel Barbeau[2], Luca De Cicco[3]

[1] SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France
[2] Carleton University, Ottawa, Canada
[3] Politecnico di Bari, Bari, Italy

alex.pierron@telecom-sudparis.eu

Résumé

Cet article court présente nos recherches sur le contrôle par Deep Reinforcement Learning (DRL) de surfaces intelligentes reconfigurables (RIS) pour la sécurité physique des futurs réseaux 6G. Nous mentionnons d'abord notre résultat sur l'intégration de l'équité dans un contrôleur DRL-RIS multi-utilisateur, puis situons nos travaux en cours sur les backdoors en DRL à partir d'une littérature fondatrice encore peu appliquée aux systèmes physiques. Nous soulignons enfin l'intérêt de notre environnement de simulation CTRL_RIS comme ressource open source reproductible pour relier résultats actuels et pistes futures sur l'équité, la sécurité et la robustesse des contrôleurs DRL-RIS pour les réseaux 6G.

Mots-clés

Intelligence Artificielle Physique, Apprentissage par Renforcement, 6G, Cybersécurité, Sécurité de la Couche Physique, Surfaces Intelligentes Reconfigurables

Abstract

This short paper presents our research on Deep Reinforcement Learning (DRL) control of reconfigurable intelligent surfaces (RIS) for the physical security of future 6G networks. We first mention our results on integrating fairness into a multi-user DRL-RIS controller, then situate our ongoing work on DRL backdoors based on foundational literature that has yet to be widely applied to physical systems. Finally, we highlight the relevance of our CTRL_RIS simulation environment as a reproducible open-source resource for linking current results and future avenues of research on fairness, security, and robustness in DRL-RIS controllers for 6G networks.

Keywords

Physical AI, Reinforcement Learning, 6G, Cybersecurity, Physical-Layer Security, Reconfigurable Intelligent Surfaces

1 Introduction

6G systems are expected to increasingly rely on Reconfigurable Intelligent Surfaces (RIS), in passive, active, or hybrid forms, to shape propagation conditions in difficult environments [1, 2]. Yet RIS remain embedded, hardware-constrained devices: large arrays of low-cost elements must be controlled in real time with limited on-board resources, and the joint optimization of phase shifts, beamforming, and physical-layer security quickly becomes difficult for classical iterative methods in dense settings [3, 4, 5]. This motivates a Physical AI approach where a policy trained by Deep Reinforcement Learning (DRL) [6, 7] directly controls the RIS, adapts to radio and security conditions, and makes fast decisions under physically grounded channel and control constraints, while also opening a new attack surface [8, 4, 9, 10].

Our recent work addressed one side of this problem, namely fairness in DRL-driven RIS-assisted physical-layer security [11]. We showed that maximizing aggregate performance is not sufficient in multi-user settings, where unfair allocations may deprive weaker users from service. Using the Jain Fairness Index [12] and the broader notions of the *price of fairness* developed in [13], we showed a misalignment between the classical Sum Secrecy Rate (SSR) used generally in multi-user telecommunications systems [14] and the behaviour obtained once the neural network is trained. This contribution provides a first step toward trustworthy DRL-driven RIS and a concrete baseline for future robustness studies. The second block concerns backdoors

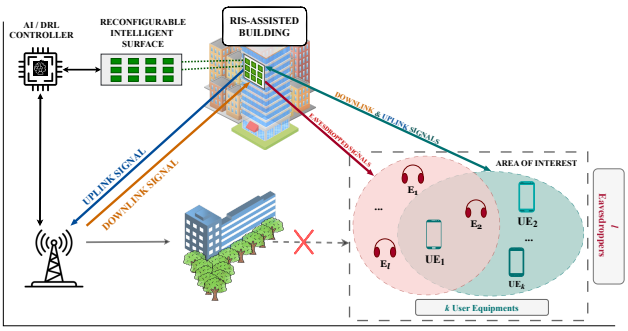

Figure 1: General DRL-driven RIS-assisted secure communication setup for 6G networks.

in DRL. While backdoor attacks are now well established

in supervised learning [15, 16], their adaptation to reinforcement learning is more recent and especially relevant for controllers deployed in long-lived cyber-physical systems. In the RIS setting, poisoning can affect training observations, rewards or environmental dynamics. The attack surface is also shaped by hardware realities such as imperfect channel estimation, finite-resolution phase control, reconfiguration latency, and RF impairments [14, 9]. This short paper positions our research along three complementary axes: fairness-aware RIS control, a concise taxonomy of DRL backdoors, and an open-source experimental basis through the *CTRL_RIS* DRL environment.[1] We provide a reproducible Python environment for secure RIS beamforming, configurable multi-user scenarios with eavesdroppers, and DRL baselines for fairness and robustness experiments.

2 Fairness vs. Backdoor-Aware RIS

Figure 1 summarizes the secure wireless environment considered in our prior fairness study [11], with a RIS model kept consistent with physically grounded formulations [14]. The main result of that paper is simple but important regarding alignment: fairness is not only a side metric but a structural requirement for reward engineering for DRL-driven RIS systems.If a reward only encourages global performance, the agent tends to privilege the user with the most favourable channels; fairness-aware shaping is needed to maintain acceptable service across users while keeping secrecy objectives meaningful [9, 11].

This observation naturally connects to trustworthiness. A poisoned DRL policy does not need to produce a spectacular failure to be harmful. An attacker may instead cause a subtle but targeted degradation of service, for instance by selectively hurting disadvantaged users, biasing resource allocation, or weakening the secrecy-fairness trade-off. In a RIS controller, such effects are further filtered by hardware limits: coarse phase updates, stale channel estimates, or impaired RF chains can hide small policy deviations while still moving the system toward an unfavourable operating point. Fairness therefore becomes both a design objective and a possible indicator of malicious behaviour.

3 Representative DRL Backdoors

Recent DRL backdoor literature assumes two main dimensions [17, 18, 19]: the attack loop and the adversarial access. *Inner-loop* attacks poison the training stream step by step, typically through state and reward manipulations; *outer-loop* attacks operate at the trajectory or episode level and can exploit richer information. A second dimension concerns the attacker interface: poisoning may alter the observed state, the reward signal, or part of the environment itself. Table 1 condenses three selected representative techniques from the current DRL backdoor literature [17, 18, 19]. TroJDRL [17] first showed that targeted malicious behaviour can be implanted through poisoned training interactions. BadRL [18] reduces the amount of poisoning required, making the attack more stealthy. SleeperNets [19] broadens the threat model through trajectory-level poisoning. However, this literature still mostly relies on benchmark-style RL environments rather than on Physical AI systems with sensing, actuation, channel, and hardware constraints. For RIS controllers, this distinction matters: a backdoor must remain effective despite quantized phase shifts, estimation noise, delayed reconfiguration, and constrained embedded control, while these same limitations can amplify the impact of small poisoned inputs once the policy is deployed [14, 9]. The literature thus provides a strong vocabulary, but not a complete answer for dynamic physical systems. The RIS case gives a concrete physical interpretation. A recent study on *pilot backdoor attacks* against DRL-empowered RIS control shows how an adversary-controlled IRS can contaminate channel state information and implant malicious behaviour in a radio controller [10]. This result is particularly relevant for our agenda because it bridges the DRL-backdoor literature and programmable wireless environments without requiring unrealistic assumptions about the attack outcome. In practice, the most plausible triggers are those aligned with the hardware and signal chain itself, such as biased pilots, corrupted Channel State Information (CSI), or malicious phase updates, rather than arbitrary perturbations. In such settings, compromised behaviour may affect secrecy, fairness, or both, which reinforces the need for evaluation protocols that go beyond average utility and remain compatible with physically grounded RIS models [14]. It is our primary goal to find effective detection and mitigation techniques to prevent problematic behaviours caused by these potential backdoors in the DRL controller.

Table 1: Three representative DRL backdoor techniques.

Method	Loop	Access	Effect
TroJDRL [17]	Inner	State/reward	Targeted policy shift
BadRL [18]	Inner	Sparse poisoning	Stealthier attack
SleeperNets [19]	Outer	Trajectories	Universal behavior

4 Conclusion

Our research addresses the challenge of building trustworthy Physical AI for future 6G networks. For DRL-driven RIS control, fairness, security, and robustness must be considered jointly. First contribution has established that fairness-aware reward design is necessary for secure multiuser RIS control [11]. We now extend this perspective to adversarial training and backdoor-aware threat modeling for DRL, using our open-source *CTRL_RIS* [2] DRL environment to evaluate poisoned training, triggered behaviour and robustness under physically grounded constraints.

Acknowledgments

The work was funded by the French National Research Agency under the France 2030 ANR program "PEPR Networks of the Future" (NF-HiSec ANR-22-PEFT-0009).

[1] https://github.com/alex-pierron/CTRL_RIS

[2] https://github.com/alex-pierron/CTRL_RIS

References

[1] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable Intelligent Surfaces: Principles and Opportunities," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021.

[2] ETSI ISG RIS, "Reconfigurable Intelligent Surfaces (RIS); Use Cases, Deployment Scenarios and Requirements," Tech. Rep. V1.2.1, European Telecommunications Standards Institute (ETSI), February 2025.

[3] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. De Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2450–2525, 2020.

[4] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser miso systems exploiting deep reinforcement learning," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1839–1850, 2020.

[5] Y. Feng, Q. Hu, K. Qu, W. Yang, Y. Zheng, and K. Chen, "Reconfigurable intelligent surfaces: Design, implementation, and practical demonstration," *Electromagnetic Science*, vol. 1, no. 2, pp. 1–21, 2023.

[6] R. S. Sutton and A. G. Barto, *Reinforcement learning: an introduction*. Adaptive computation and machine learning, Cambridge, Mass: MIT Press, 1998.

[7] X. Wang, S. Wang, X. Liang, D. Zhao, J. Huang, X. Xu, B. Dai, and Q. Miao, "Deep reinforcement learning: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 4, pp. 5064–5078, 2022.

[8] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 375–388, 2020.

[9] Z. Peng, Z. Zhang, L. Kong, C. Pan, L. Li, and J. Wang, "Deep reinforcement learning for RIS-aided multiuser full-duplex secure communications with hardware impairments," *IEEE Internet of Things Journal*, vol. 9, no. 21, pp. 21121–21135, 2022.

[10] Y. Huang, H.-M. Wang, Z. Wang, and W. Liu, "Pilot backdoor attack against deep reinforcement learning empowered intelligent reflection surface for smart radio," *IEEE Transactions on Wireless Communications*, 2025.

[11] A. Pierron, M. Barbeau, L. De Cicco, J. Rubio-Hernan, and J. Garcia-Alfaro, "A fairness-aware strategy for b5g physical-layer security leveraging reconfigurable intelligent surfaces," in *Foundation and Practice of Security 2025*, Springer, 2025.

[12] R. K. Jain, D.-M. W. Chiu, W. R. Hawe, *et al.*, "A quantitative measure of fairness and discrimination," tech. rep., Digital Equipment Corporation, 1984.

[13] D. Bertsimas, V. F. Farias, and N. Trichakis, "The price of fairness," *Operations Research*, vol. 59, no. 6, pp. 1380–1393, 2011.

[14] E. Björnson, Ö. T. Demir, *et al.*, *Introduction to multiple antenna communications and reconfigurable surfaces*. Now Publishers, Inc., 2024.

[15] Y. Li, Y. Jiang, Z. Li, and S.-T. Xia, "Backdoor learning: A survey," *IEEE transactions on neural networks and learning systems*, vol. 35, no. 1, pp. 5–22, 2022.

[16] S. Zhang, Y. Pan, Q. Liu, Z. Yan, K.-K. R. Choo, and G. Wang, "Backdoor attacks and defenses targeting multi-domain ai models: A comprehensive review," *ACM Computing Surveys*, vol. 57, no. 4, pp. 1–35, 2024.

[17] P. Kiourti, K. Wardega, S. Jha, and W. Li, "TrojDRL: Evaluation of Backdoor Attacks on Deep Reinforcement Learning," in *2020 57th ACM/IEEE Design Automation Conference (DAC)*, pp. 1–6, July 2020. ISSN: 0738-100X.

[18] J. Cui, Y. Han, Y. Ma, J. Jiao, and J. Zhang, "BadRL: Sparse Targeted Backdoor Attack against Reinforcement Learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, pp. 11687–11694, Mar. 2024.

[19] E. Rathbun, C. Amato, and A. Oprea, "SleeperNets: Universal Backdoor Poisoning Attacks Against Reinforcement Learning Agents," *Advances in Neural Information Processing Systems*, vol. 37, pp. 111994–112024, Dec. 2024.