RAN-Aware Adversarial Training for Robust Near-RT RIC Handover Control

Vaskar Chakma Wooyeol Choi

School of Computer Science and Engineering, Chung-Ang University, Seoul, Republic of Korea

vaskar@cau.ac.kr wchoi@cau.ac.kr

Abstract

The integration of artificial intelligence (AI) into Radio Access Networks (RAN) enables adaptive handover control but introduces vulnerabilities to adversarial attacks. This paper presents a RAN-constrained adversarial threat model and develops an adversarial robust AI-RAN framework that improves handover reliability under physically plausible attacks. Unlike unconstrained adversarial models in computer vision, our approach enforces temporal smoothness, measurement bounds, and mobility continuity constraints reflective of real wireless environments. Evaluation on 3GPP-compliant Urban Macro scenarios demonstrates that adversarial training reduces handover failure rates by 4.8% points and improves throughput by 7.7% under 2.5 dB attacks, while maintaining real-time inference feasibility in Near-RT RIC deployments.

1. INTRODUCTION

Cellular networks are transformed by the Open Radio Access Network (O-RAN) design, which separates RAN com- ponents and introduces AI into the Near-RT RIC [1]. To optimize handover management, AI-driven control applications (xApps) evaluate radio metrics like RSRP, CQI, and mobility attributes to improve cell selection decisions [2]. According to recent research, O-RAN systems' resource allocation, traffic steering, and handover control can all be affected by adversarial machine learning (AML) attacks [3].

Unlike adversarial computer vision research with unconstrained perturbations, wireless settings involve physical constraints such as temporal smoothness, measurement bounds, and mobility continuity, which are beyond the capabilities of current threat models [4]. By including physically realistic disturbances that resemble actual wireless channels, the proposed RAN- constrained adversarial threat model fills this gap. Our physically grounded adversarial threat model and offline training framework enable deployment in latency-sensitive Near-RT RIC scenarios.

2. SYSTEM MODEL

Fig. 1 illustrates our framework. The RAN observes the state vector at time t as given by $X_T = RSRP_t(s), RSRP_t(n_1), \dots, RSRP_t(n_k), CQI_t, v_t, L_t$, where s denotes the serving cell, n_k neighboring cells, v_t UE velocity, and L_t cell load. Near-RT RIC processes observation sequences $X_t = \{X_{t-T+1}, \dots, X_t\}$ through a bi-directional LSTM model $f_\theta(\cdot)$ to generate handover decision $\hat{y}_t = f_\theta(X_t)$.

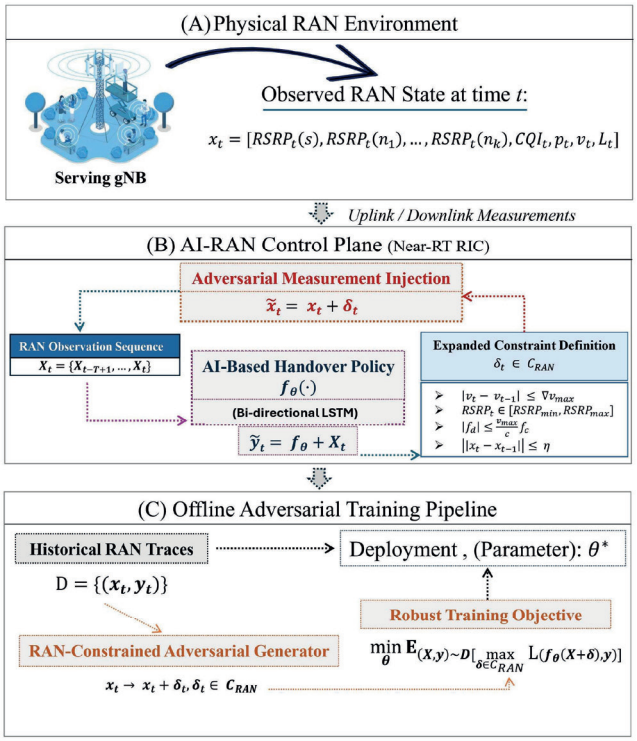

Fig. 1. A RAN-constrained framework that uses measurements and controlled adversarial inputs to train robust models.

We define adversarial perturbed measurements as $\tilde{X}_t = X_t + \delta_t$, where perturbation $\delta_t \in C_{RAN}$ satisfies physical constraints: measurement bounds $RSRP_t \in [-140, -44]$ dBm (3GPP TS 38.215), mobility continuity $|v_t - v_{t-1}| \le \Delta v_{max}$, and temporal smoothness $||X_t - X_{t-1}|| \le \eta$. The attack is sparse this is because perturbations activate only when the serving-cell margin satisfies $RSRP_s(t) -$

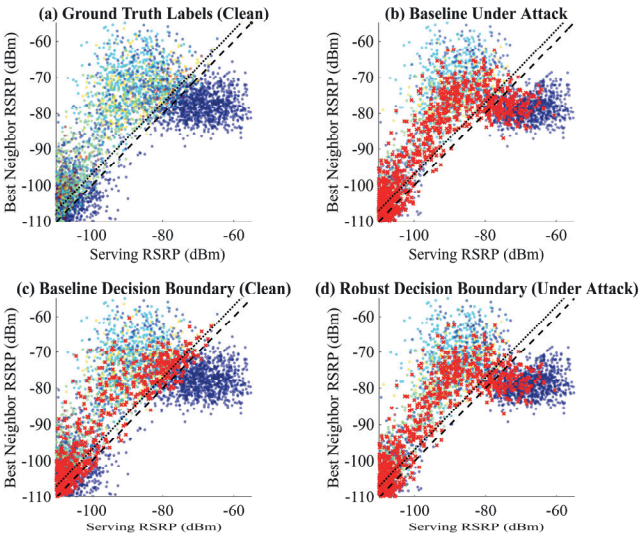

Fig. 2. Serving-neighbor RSRP scatter plots showing clean labels, adversarial effects on the baseline model, and the comparison of baseline and robust decision boundaries.

$max_k\, RSRP_{n_k}(t) \leq \nabla HO = \varepsilon$, targeting only the vulnerable handover moments. These constraints ensure attacks remain stealthy and physically plausible, distinguishing our model from unconstrained threat scenarios.

3. RESULTS AND ANALYSIS

We simulate a 3GPP Urban Macro cellular network with 6 neighboring cells (ISD = 500 m) and UE mobility at 2-30 m/s, using 3GPP TR 38.901 UMa NLOS path loss with 6 dB shadowing at $f_c = 3.5\,GHz$, a -104 dBm noise floor, and 3 dB handover hysteresis. We train in a bi-directional LSTM model (64 hidden units, T = 15) on N = 6000 samples using the Adam optimizer, with attacks varying from ε = 0 to 5 dB.

The clean ground-truth cell distribution in serving vs. best-neighbor RSRP space is shown in Fig. 2(a). The baseline model produces 2208 misclassified handover decisions under a 2.5 dB attack; these are clustered at the hysteresis border, where the margin is most vulnerable, and are indicated by red × markers in Fig. 2(b). Compared to the baseline, the resilient model retains a narrower and more organized decision boundary under the same attack, as seen in Fig. 2.

Fig. 3 shows that at ε = 2.5 dB the handover failure rate increases from 20% (clean) to 37.3% under attack. With adversarial training, this is reduced to 32.5%, giving a 4.8 percentage-point improvement. A similar pattern appears in throughput: the baseline drops from 9.3 Mbps to 8.1 Mbps, while the robust model maintains 8.7 Mbps. Shannon capacity analysis provides context: at a mean SINR of about −1.2 dB, the theoretical limit is $log_2\left(1 + 10^{-\frac{1.2}{10}}\right) \times 20 \approx$ 16.3 Mbps, while the observed 9.3 Mbps clean throughput reflects handover failures and protocol overhead.

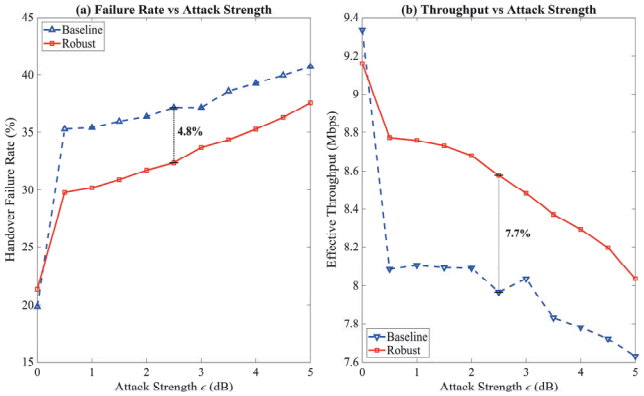

Fig. 3. Impact of attack strength ϵ on handover failure rate and effective throughput, comparing baseline and robust models.

4. CONCLUSION

This paper introduces a RAN-constrained adversarial robustness framework for AI-driven handover control in O-RAN architectures. Through offline multi-epsilon curriculum adversarial training, we achieve a 4.8% reduction in handover failure rates and a 7.7% throughput improvement under 2.5 dB attacks, while maintaining real-time inference feasibility. Future research will investigate adaptive defenses, and certified robustness bounds for reliable 6G network automation.

ACKNOWLEDGEMENT

REFERENCES

[1] C. Feng, H. H. Yang, K. Guo, W. Xia, C. Liu, and T. Q. Quek, “AI-RAN: The pathway to future wireless networks,” Journal of Information and Intelligence, p. S2949715926000016, Jan. 2026.

[2] K. Sthankiya, N. Saeed, G. Mcsorley, M. Jaber, and R. G. Clegg, “A Survey on AI-Driven Energy Optimization in Terrestrial Next Generation Radio Access Networks,” IEEE Access, vol. 12, pp. 157540–157555, 2024.

[3] Y. Abera Ergu and V.-L. Nguyen, “RADAR: Robust DRL-Based Re- source Allocation Against Adversarial Attacks in Intelligent O-RAN,” IEEE Transactions on Green Communications and Networking, vol. 9, pp. 2305–2318, Dec. 2025.

[4] A. Guesmi, M. A. Hanif, B. Ouni, and M. Shafique, “Physical Adversarial Attacks for Camera-Based Smart Systems: Current Trends, Categorization, Applications, Research Challenges, and Future Outlook,” IEEE Access, vol. 11, pp. 109617–109668, 2023.