CICI: IPAAI: REPAIRT: Securing xApps in Open RANs with Reliable and Principled AI Red-Teaming

PI: Francesco Restuccia, Northeastern University

Co-PI: Flavio Esposito, Saint Louis University

https://www.nsf.gov/awardsearch/show-award/?AWD_ID=2530896

Northeastern University
Institute for Intelligent Networked Systems

SAINT LOUIS UNIVERSITY

The Open Radio Access Network (O-RAN) paradigm:

O-RAN splits the RAN into three main components: Radio Unit (RU), Distributed Unit (DU), and Central Unit (CU) that operate under the supervision of the Near Real-Time RAN Intelligent Controller (Near-RT-RIC) and use third-party applications (xApps) for data-driven network control. Third-party developers create and monetize xApps through shared marketplaces.

Need for Integrity, Provenance, and Authenticity (IPA):

- AI is key for efficient O-RAN management using xApps
- This enables adversarial attacks aiming at altering network operations through evasion/poisoning attacks
- IPA is needed to secure O-RAN deployments

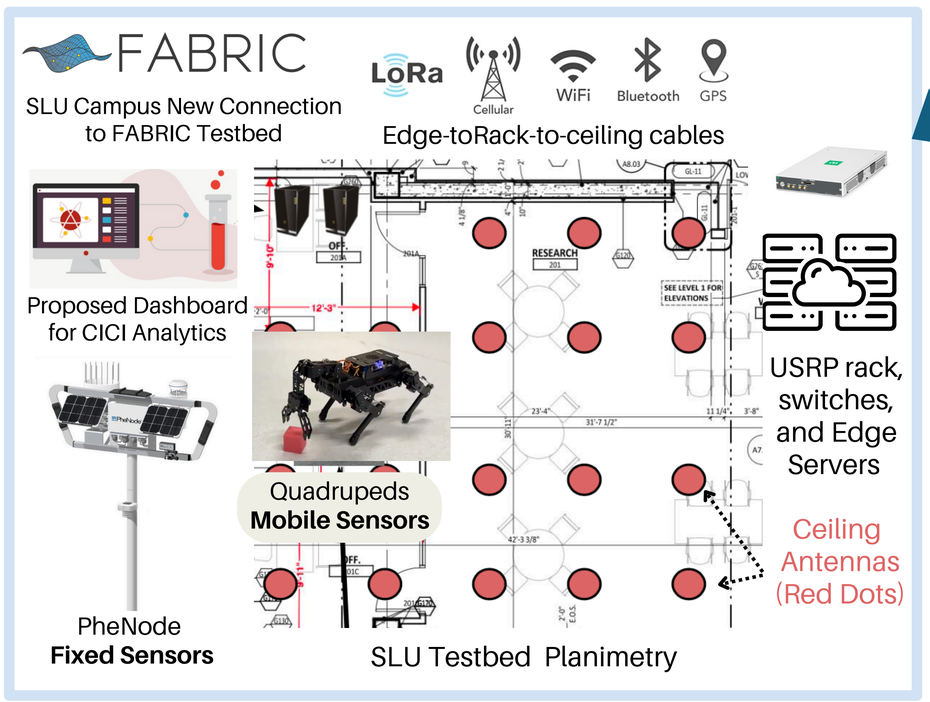
Improve IPA of Wireless Data and AI Models in O-RAN by:

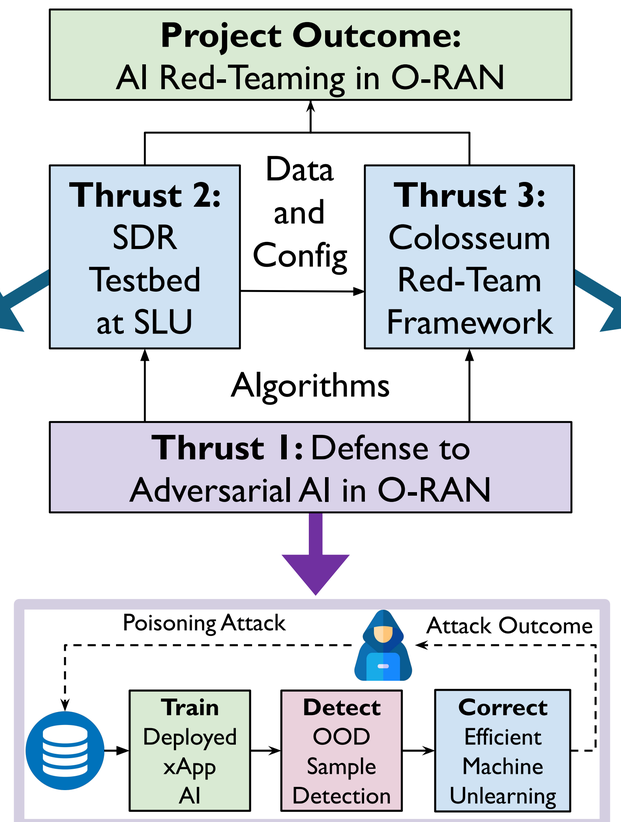
- Expanding existing cyberinfrastructures for realistic data collection
- Enabling AI red-teaming in O-RAN to study stealthy threat vectors
- Enhancing data integrity/provenance incorporating traceability
- Sharing datasets for AI research under open-source license

The REPAIRT red-teaming effort will catalyze reproducible AML research in O-RAN, fulfilling the IPAAI program objective of making AI training data and results transparent, authentic, and broadly available

Technical Approach:

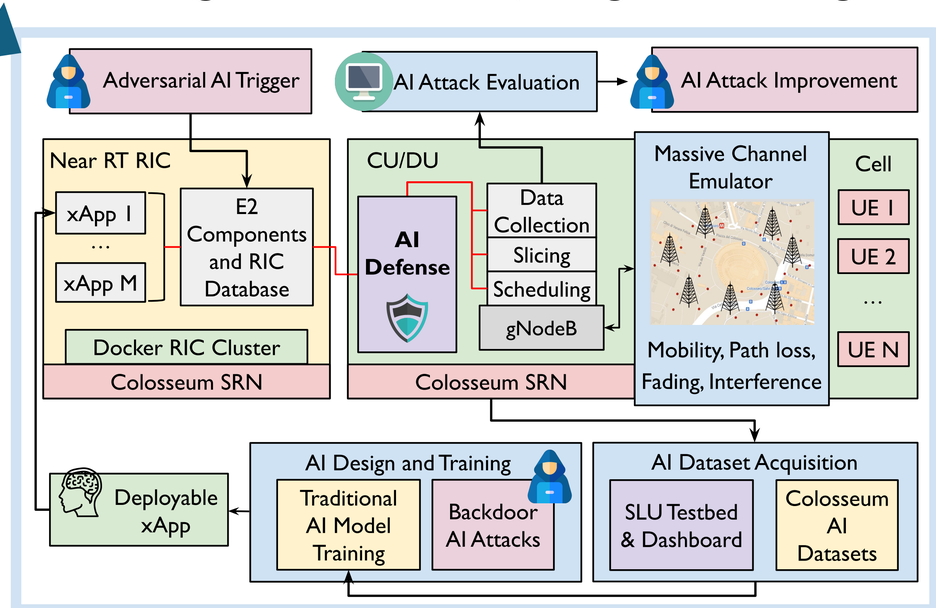
Develop hardware/software cyberinfrastructure for AI red-teaming of O-RAN xApps at scale.





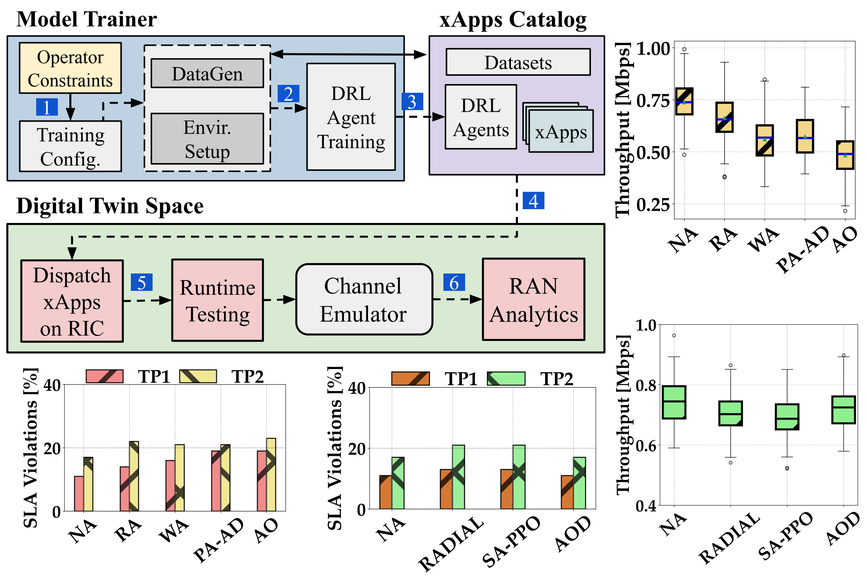
Unique Properties of the Scientific Domain:

O-RANs operate in near-real-time, handling dynamic key performance metrics under strict latency/reliability constraints. This requires real-time fine-grained detection/mitigation strategies.



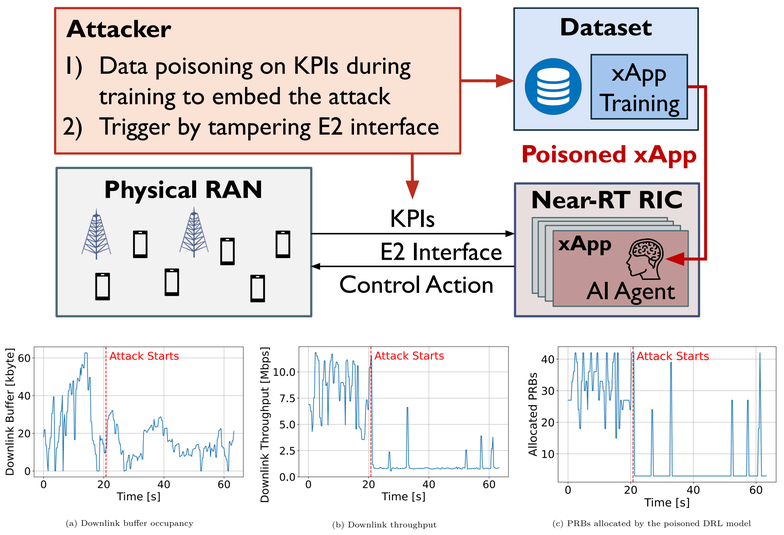
Research Results:

Inference-time Evasion Attack



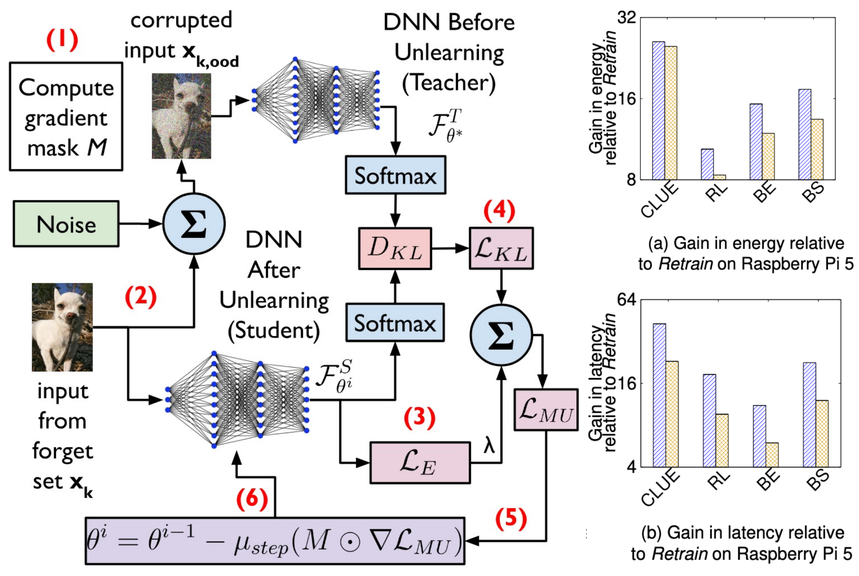
T. Hassan, F. Meneghello, and F. Restuccia, "AdvO-RAN: Adversarial Deep Reinforcement Learning in AI-Driven Open Radio Access Networks," ACM MobiHoc, Houston, Texas, USA, 2025.

Training-time Poisoning Attack



A. Lacava, S. Maxenti, L. Bonati, S. D'Oro, A. Oprea, T. Melodia, and F. Restuccia, "How to Poison an xApp: Dissecting Backdoor Attacks to Deep Reinforcement Learning in Open Radio Access Networks," Computer Networks, 2025.

Machine Unlearning-based Defense



S. Sayyed, N. Bastian, M. De Lucia, A. Swami, and F. Restuccia, "CLUE: Bringing Machine Unlearning to Mobile Devices," IEEE/CVF WACV, Tucson, Arizona, USA, 2026.