# RadioShield: A PHY-Layer Defense Framework for RRC Signaling Storms in O-RAN

Farah Abed Zadeh*, Bartlomiej Siniarski†, Shen Wang‡, Madhusanka Liyanage§
*†‡§Network Softwarization and Security Labs (NetsLab), School of Computer Science, University College Dublin, Ireland
Email: *farah.abedzadeh@ucdconnect.ie, †bartlomiej.siniarski@ucd.ie, ‡shen.wang@ucd.ie, §madhusanka@ucd.ie

***Abstract*—Open Radio Access Network (O-RAN) disaggregation introduces architectural flexibility but significantly expands the cellular attack surface. Specifically, Radio Resource Control (RRC) signaling storms exploit these functional splits, severely amplifying resource exhaustion across distributed RAN components. State-of-the-art defenses typically deploy security rApps or xApps at the core or upper network layers; however, these approaches are inherently reactive, detecting anomalies only after radio resources and control-plane entities are already saturated. By relying on aggregated counters and volumetric thresholds, they overlook the immediate physical contention and disruption occurring at the radio edge. To address these limitations, we propose RadioShield, a novel, multi-tier defense framework that shifts the detection scope from upper-layer monitoring to the PHY/MAC layer at the Open Distributed Unit (O-DU). By leveraging radio telemetry, RadioShield enables early detection through a distributed application (dApp), in close proximity to malicious User Equipment (UEs), allowing the network to identify and halt signaling storm traffic microseconds after it originates. To validate this framework, we utilized a physical O-RAN testbed with realistic UE reboot and handover storms, demonstrating over 99% accuracy in both binary and multi-class detection. Computational profiling confirms microsecond-scale inference and reduced resource overhead, proving that dApp-based early detection outperforms traditional, upper-layer network monitoring in both computational efficiency and threat mitigation speed.**

***Index Terms*—O-RAN security, RRC Signaling Storms, Control-plane traffic, Intrusion Detection Systems, Machine Learning, Radio telemetry, PHY/MAC-layer detection**

## I. Introduction

The shift toward Open Radio Access Networks (O-RAN) is redefining how mobile networks are built and operated by introducing disaggregation, virtualization, and open interfaces across the RAN stack. However, this architectural transformation also expands the threat surface, exposing critical interfaces and control functions to new vectors of exploitation. Among these threats, signaling storm attacks represent a critical availability risk. The severity of this threat is recognized by the O-RAN Alliance, which explicitly identifies signaling overload and control-plane abuse as high-priority security risks [1].

This research was funded by the ROBUST-6G Project through the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe Research and Innovation Program (Grant number: 101139068), and CONNECT phase 2 project that has received funding from Science Foundation Ireland under grant no. 13/RC/2077_P2.
**ISBN 978-3-903176-82-9 ©2026 IFIP**

Signaling storms exploit the stateful nature of the control plane, leading to disproportionate processing overhead, increased latency, and service disruption for legitimate users. Adversaries can trigger excessive Radio Resource Control (RRC) state transitions to overwhelm network resources with relatively low-rate traffic. In O-RAN, this impact is amplified by functional disaggregation of the base station; a single control-plane trigger can cascade into multiple signaling exchanges across RAN components, potentially saturating centralized control resources [2]. These attacks threaten not only network availability but also the reliability of intelligent RAN control applications, making timely and accurate detection a fundamental requirement for secure O-RAN operation.

Existing approaches predominantly rely on core-network counters, control-plane message volumes, or xApp/rApp-level KPIs collected after signaling procedures have already propagated through the RAN stack [3]–[6]. As a result, they detect the *effects* of the storm through the impacted network area rather than its earliest manifestation at the radio edge. This causes the RAN stack and the core network to have already suffered severe resource starvation and service disruption. This limits both **response timeliness** and **attack interpretability**: UE reboot storms and handover storms can appear similarly anomalous at upper layers because both ultimately generate bursts of RRC, F1AP, and NGAP activity.

This paper addresses that gap by introducing **RadioShield**, a novel, multi-tier intrusion detection architecture integrated into the O-RAN stack. We posit that RRC signaling storms generate distinct, detectable footprints at the Physical (PHY) and Medium Access Control (MAC) layers well before they cause degradation in the upper control plane. A key factor of this architecture is the design of a **distributed application (dApp)** [7], which is a localized microservice deployed directly at the Open Distributed Unit (O-DU) to perform low-latency analytics, monitoring, or control functions in close proximity to radio resources. Rather than relying solely on network-level inspection, RadioShield dApp uses radio telemetry (e.g., SINR variance, block error rates, and resource block utilization) to support earlier localization of the attack source, stopping malicious User Equipment (UEs) at the radio edge.

### A. Our Contributions

The main contributions of this work are as follows:

- **Early Detection Security Framework:** This work introduces a novel O-DU-level detection framework implemented via a dApp, demonstrating that PHY/MAC telemetry can identify signaling-storm activity earlier at the radio edge, before aggregated upper-layer symptoms become visible to xApp/rApp or core-side monitors.
- **Radio-Level Attack Identification:** Using radio telemetry, we demonstrate the scheduler- and radio-side impact of control-plane abuse, preserving distinguishable signatures of different attack vectors. This allows separation in multiclass detection, whereas upper-layer network features largely collapse into a generic volumetric anomaly.
- **Deployment on a physical O-RAN testbed:** We execute realistic UE reboot and handover storm scenarios on a fully functional 5G O-RAN testbed. Unlike prior works that rely on simulations, we capture the true processing overhead and physical dynamics of disaggregated components. To facilitate further research in O-RAN security, the radio-level and time-aggregated datasets are available [8].
- **Cross-Layer Comparative Validation:** We quantify the predictive and computational tradeoffs of our radio-centric approach against traditional upper-layer network monitoring. Results confirm that dApp-based radio telemetry achieves over 99% accuracy with microsecond-scale inference latency, validating that physical-layer monitoring outperforms legacy approaches and enables the proactive mitigation of signaling storms long before network resources are exhausted.

## II. Background and Related Work

### A. 5G O-RAN Architecture

The O-RAN architecture represents a fundamental shift from traditional monolithic RAN deployments toward a disaggregated, virtualized, and programmable radio access ecosystem, as standardized by the O-RAN Alliance [16]. In O-RAN, base station functionalities are decomposed into three logical entities: the Open Radio Unit (O-RU), Open Distributed Unit (O-DU), and Open Central Unit (O-CU), interconnected through standardized and open interfaces. This separation enables multi-vendor interoperability, flexible deployment, and cloud-native scalability, while also expanding the system's attack surface.

Beyond functional disaggregation, O-RAN introduces intelligence as a first-class architectural component through two logically separated RAN Intelligent Controllers (RICs): Near-RT RIC and the Non-RT RIC. The Near-RT RIC operates on timescales ranging from tens of milliseconds to one second and hosts xApps responsible for near-real-time control and optimization functions, such as mobility management, interference mitigation, and anomaly detection. In contrast, the Non-RT RIC operates at longer timescales and hosts rApps that perform policy generation, network-wide optimization, and AI/ML model training and lifecycle management.

The Service Management and Orchestration (SMO) framework coordinates these components and facilitates data collection, model training, and policy enforcement across the O-RAN ecosystem. Standardized interfaces such as E2, A1, and O1 enable telemetry exchange, control actions, and management signaling between RAN elements and intelligent controllers. While these interfaces enable unprecedented observability, they also serve as the foundation for our proposed detection mechanism. Crucially, the O-DU exposes granular telemetry regarding the PHY/MAC layers, such as scheduling decisions, modulation schemes, and interference levels, which are traditionally abstracted away from core-level monitoring.

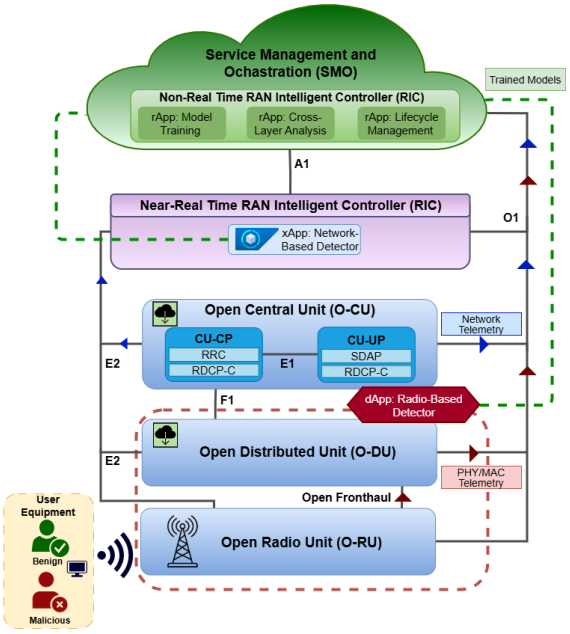


**Fig. 1:** The RadioShield multi-tier security architecture for RRC signalling storm in O-RAN

AI/ML capabilities are deeply embedded into the O-RAN control loop, enabling data-driven decision-making across RAN layers. Telemetry exposed through E2 service models includes key performance measurements (KPMs) spanning physical, MAC, and RRC layers, allowing intelligent applications to reason about network behavior in near real time. This architectural design positions O-RAN as a promising platform for adaptive security mechanisms, particularly those that leverage lower-layer radio intelligence to secure the control plane.

### B. RRC Signaling Storm Background

RRC is the primary 5G control-plane protocol, responsible for connection management, mobility, and radio bearer configuration. In 5G NR, the RRC state machine comprises three states: RRC Idle, RRC Inactive, and RRC Connected. State transitions are triggered by data arrival, inactivity timers, or mobility events. As depicted in Fig. 2, standard RRC connection establishment requires a multi-step handshake across the disaggregated RAN, progressing from the initial RRC SetUp Request up to the RRC Setup Complete.

A signaling storm is a control-plane availability attack where an adversary triggers excessive state transitions to

**TABLE I:** Comparison of Existing Signaling Storm Research and Proposed Contributions

| Reference | Experimental Setting | Observation Layer | Detector Deployment | Attack Execution Strategy | Radio-Level Features | Multiclass Support | Dataset |
|---|---|---|---|---|---|---|---|
| Ettiane *et al.* [3], [9] | Simulated | Control Plane | Core Network / O-CU | Synthetic Modeling | ✗ | ✗ | ✗ |
| Pavloski [10] | Simulated | Control Plane | Core Network | Traffic & Timer Synchronization | ✗ | ✗ | ✗ |
| Abdelrahman *et al.* [11] | Simulation | Packet / User Plane | Core Network | Network-Unfriendly Mobiles | ✗ | ✗ | ✗ |
| Park *et al.* [4] | Emulated Core | Core Network | Core Network | NAS/Registration Flooding | ✗ | ✗ | ✗ |
| Feng *et al.* [12] | Operational Core | Core Network | Core Network | Artificial Counter Injection | ✗ | ✗ | ✗ |
| Nguyen *et al.* [5], [13] | Emulation / OAI | Control Plane | O-CU / Near-RT RIC | Protocol Modification (Msg3) | ✗ | ✗ | ✗ |
| Hoffmann *et al.* [14] | O-RAN Emulation | Near-RT RIC | Near-RT RIC (xApp) | Registration Request Flooding | ✗ | ✗ | ✗ |
| Mayhoub *et al.* [15] | O-RAN Emulation | Near-RT RIC | Near-RT RIC (xApp) | O-RU Reboot (Forced Handover) | ✗ | ✗ | ✗ |
| **Proposed Work** | **Real O-RAN Testbed** | **PHY / MAC & Network** | **O-DU (dApp) & Near-RT RIC (xApp)** | **Real UE Reboot & Handover** | ✓ | ✓ | ✓ |

exhaust network resources. Unlike volumetric DDoS attacks, signaling storms exploit protocol logic, allowing low-rate, protocol-compliant traffic to induce significant processing overhead across RAN and control-plane functions.

Adversaries exploit these RRC procedures through several primary vectors. A common mechanism is UE-induced transitions, in which compromised UEs or malicious botnets synchronize reboots or toggle connectivity, forcing repeated NAS and RRC registration sequences. Another threat arises from protocol exploitation, where malicious devices flood the network with partial connection requests—such as repeatedly transmitting the RRC Setup Request shown in Fig. 2—without ever completing the procedure. This behavior artificially occupies control-plane state resources at both the O-DU and O-CU.

In O-RAN, the disaggregation of the base station significantly amplifies the impact of these attacks [2]. Because RRC functionality resides in the O-CU control plane (CU-CP), a single malicious trigger propagates as cascaded signaling exchanges across open interfaces. These interfaces create additional opportunities for message reordering or delayed control message release, further destabilizing the control loop. Another critical vulnerability arises from virtualization: the software-defined nature of O-DUs and O-CUs increases susceptibility to software-level resource exhaustion. Moreover, functional disaggregation requires more internal signaling exchanges to complete a single RRC procedure compared to legacy monolithic architectures, resulting in signaling multiplication.

Most existing discussions of RRC signaling storms focus on protocol behavior, control-plane load, or architectural vulnerabilities. However, much less attention has been given to how such attacks manifest at the radio level, despite the fact that all signaling ultimately induces PHY/MAC activity. When an adversary initiates a storm, it physically manifests as a surge in RAN synchronization, a rapid increase of data latency and throughput, and distinct variations in signal quality due to contention. Consequently, these attacks leave a physical footprint that can be detected before the signaling flood saturates the upper layers.

### C. Existing Defenses and Their Limitations

Existing signaling-storm defenses can be broadly grouped by the layer at which they observe the attack, with emphasis on extracted detection signals and execution environment.

*1) Core-Network and Upper-Layer Network Analysis:* The majority of existing research detects signaling storms by monitoring aggregated counters at the Core Network (CN) or O-CU. Early studies relied on analytical modeling and

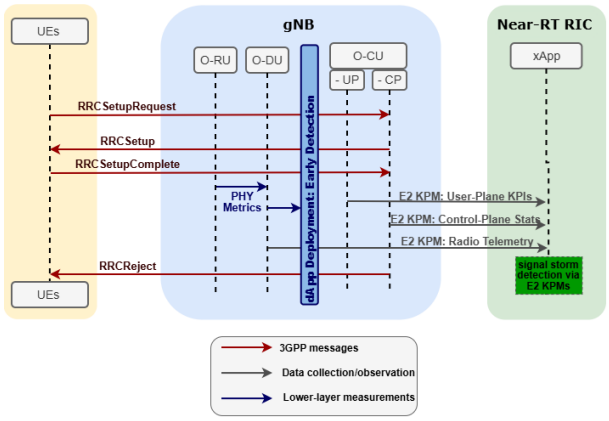


**Fig. 2:** 5G NR RRC signaling flow in an O-RAN architecture with E2 reporting for signaling storm detection.

discrete-event simulations. Ettiane *et al.* [3], [9] analyzed RRC state transitions and connection requests in simulated 5G environments, using threshold-based monitoring of signaling counters. Similarly, Pavloski [10] demonstrates how low-rate, repetitive service requests synchronized with RRC inactivity timers can induce frequent state promotions and demotions, while Abdelrahman *et al.* [11] modeled "network-unfriendly" mobiles using packet-level statistics in simulated environments.

More recent works have focused on the 5G Core. Park *et al.* [4] emulated signaling DDoS attacks by configuring large numbers of emulated UEs to repeatedly trigger NAS-level procedures such as registration and PDU session establishment and release, thereby overloading core network functions. Feng *et al.* [12] generated artificial storm scenarios at the counter level and utilized Network Data Analytics Function (NWDAF) indicators for detection. While effective for protecting the core, these methods suffer from detection latency; they identify the attack only after it has consumed RAN resources and reached the core, failing to prevent radio-level degradation.

*2) RAN-Centric and O-RAN Approaches:* With the advent of O-RAN, research has shifted toward leveraging RAN-specific telemetry. Nguyen *et al.* [5], [13] moved closer to realistic execution by modifying UE behavior in an O-RAN–compatible OpenAirInterface (OAI) testbed to transmit partial RRC connection requests (e.g., Msg3 flooding). However, their detection logic primarily utilized aggregated control statistics. Hoffmann *et al.* [14] and Mayhoub *et al.* [15] utilized O-RAN RIC telemetry to detect registration floods and forced handover storms. Although these works emulate O-RAN environments, they typically focus on KPMs such as

"number of connected UEs" or "handover success rate." These metrics are essentially lagging indicators of an ongoing storm, rather than the immediate physical footprints of the malicious signaling itself.

*3) Discussion and Limitations:* Table I summarizes these works and highlights persistent limitations. First, most existing approaches remain *reactive*: they detect the storm only after repeated access, setup, or handover procedures have already saturated the control plane and consumed O-DU/O-CU resources. In practice, xApp/rApp-only and core-side indicators are often lagging symptoms of an ongoing attack rather than its earliest manifestation. Second, upper-layer counters provide limited *class separability*. UE reboot storms and handover storms can both appear as bursts of RRC-, F1AP-, or mobility-related activity when viewed only through aggregate signaling statistics, making it difficult to distinguish the underlying attack mechanism.

### *D. Our Threat Model and Scope*

To contextualize the proposed RadioShield framework, we define a formal threat model outlining the adversary's capabilities, targeted assets, and the specific scope of our defense mechanism. This model is grounded in the classical Confidentiality–Integrity–Availability (CIA) security paradigm [17], with primary emphasis on availability, as RRC signaling storms are control-plane availability attacks.

**Assets and Security Objectives:** The primary assets under protection in our model include: (i) RRC state machine resources residing at the O-CU, (ii) scheduling and PHY/MAC processing capacity at the O-DU, and (iii) the operational stability of the Near-RT RIC control loop. The security objective is to preserve control-plane availability and prevent cascading signaling amplification across disaggregated O-RAN components.

**Adversary Capabilities:** To account for the diverse attack surface of the O-RAN architecture, we consider two distinct adversary profiles. First, we assume an **external adversary** operating at the radio edge. This attacker controls a botnet of compromised UEs to emulate legitimate UE behavior. They possess valid cryptographic credentials and can initiate standard, protocol-compliant NAS and RRC procedures.

Second, we consider an **infrastructure adversary** capable of localized disruption within the virtualized RAN. Given the software-defined nature of O-RAN, this attacker is assumed to have influenced the availability of a specific virtualized component (e.g., an O-DU), either through operational misconfiguration or by compromising the underlying host infrastructure.

**Attack Surface and Vectors:** Within these adversary profiles, we focus on protocol-compliant control-plane abuse that repeatedly triggers costly RRC procedures via two primary strategies:

1) *UE Reboot / Registration Storms:* Synchronized, repetitive initiation of complete RRC establishment and release procedures to overload the O-CU control plane.
2) *Handover Storms:* Artificial triggering of continuous cell-to-cell mobility procedures to overwhelm the Near-RT RIC and O-DU/O-CU scheduling queues. This scenario also captures realistic environmental scenarios where benign but destructive group handover storms are caused by extreme, synchronized UE mobility (e.g., high-speed trains).

Unlike volumetric flooding attacks, these behaviors exploit legitimate signaling logic to amplify processing overhead with relatively low traffic rates, making them difficult to distinguish using only coarse upper-layer counters.

**Scope of Defense:**

RadioShield is designed to detect control-plane signaling amplification at the earliest observable layer. Compared with xApp/rApp-only or core-side defenses that monitor higher-layer indicators such as connected-UE counts, registration activity, or RRC counters, this O-DU vantage point captures the immediate radio-layer footprint of signaling abuse, including abnormal modulation schemes, retransmission failures, throughput starvation, and signal degradation. This enables lower-latency detection and clearer discrimination between reboot-driven and handover-driven storm behaviors. RadioShield therefore complements existing upper-layer defenses through a lightweight O-DU dApp for multiclass storm detection, validated on a physical O-RAN testbed using realistic reboot and handover storm scenarios.

## III. RadioShield: dApp-based Detection Framework

To address the inherent limitations of reactive, upper-layer monitoring, we propose **RadioShield**, a multi-tier intrusion detection framework natively integrated into the O-RAN architecture. RadioShield shifts the primary detection horizon from the core network to the radio edge, leveraging the physical footprint of control-plane attacks to enable early detection.

As illustrated in Fig. 1, the RadioShield architecture orchestrates several interdependent workflows to extract multi-layer telemetry, perform central model training and analysis, and enforce decentralized real-time inference leveraging both O-DU dApps and Near-RT RIC xApps.

Rather than relying on a single vantage point, RadioShield utilizes a multi-vantage point telemetry collection strategy. Radio data originates at the O-RU and is transmitted to the O-DU via the Open Fronthaul interface. Concurrently, network telemetry is generated at the O-CU, comprising metrics from both CU-UP and CU-CP protocols. This heterogeneous data is continuously streamed upstream via the standardized O1 and E2 interfaces to the Non-RT RIC within the SMO.

Central orchestration occurs within the SMO, where a specialized rApp for Model Training consumes the gathered telemetry from the data stream to train location-aware detection models. Specifically, the rApp trains dual models: a lightweight, high-speed model optimized for the unique physical footprint observable at the O-DU, and a secondary model trained on upper-layer network traffic. Once validated, these models are deployed downstream via the O1 and A1 interfaces to their respective inference nodes.

The core novelty of RadioShield is its hierarchical detection strategy, which minimizes detection latency while maximizing accuracy through a two-tiered implementation:

**Tier 1: Early Detection at the Edge (O-DU dApp).** The first line of defense is deployed at the O-DU as a dApp. Operating on timescales under 10 milliseconds, the dApp ingests the real-time PHY/MAC telemetry streams. Because signaling storms induce scheduling failures and hardware jitter before control-plane saturation occurs, the dApp acts as an early-warning system. By running lightweight inference locally at the O-DU, RadioShield identifies the radio footprint of a storm near-instantaneously, allowing for localized rate-limiting of suspected rogue UEs before the traffic propagates to higher layers.

**Tier 2: Network Detection and Verification (Near-RT RIC xApp).** The second line of defense resides in the Near-RT RIC as an xApp, operating on a 10 ms to 1-second control loop. By ingesting network-layer telemetry from the O-CU, this centralized xApp performs upper-layer anomaly detection using aggregated control-plane indicators. Acting as a redundant verification layer, it provides cross-layer validation of events detected at the O-DU, enhancing overall architectural robustness. Through E2 reporting, the xApp maintains a global view of behavior across multiple RAN components, enabling it to identify anomalous traffic coming from localized failures or misconfigurations. The resulting network-level inconsistencies remain structurally visible to the upper-layer detector.

At the orchestration layer, the framework integrates with the inherent capabilities of the Non-RT RIC. Within this broader data exchange ecosystem, standard rApps can be leveraged to support long-term security operations, such as cross-layer forensic analysis and machine learning lifecycle management. By correlating physical-layer anomalies at the edge with logical state-machine exhaustion at the O-CU, these overarching functions can continuously refine the deployed security models. This orchestration ensures automated validation and retraining downstream, systematically improving the system's ability to distinguish malicious signaling storms from benign flash-crowd events over time.

Through this multi-tiered architecture, RadioShield achieves the latency benefits of edge-level PHY/MAC detection while retaining the contextual accuracy of network monitoring.

## IV. Experimental Methodology

### A. 5G O-RAN Testbed Setup

We conduct our experiments on a physical 5G O-RAN testbed built on the OpenAirInterface (OAI) platform, as illustrated in Fig. 3. The testbed comprises a 5G CN, an O-CU, two O-DUs, and two O-RUs. Specifically, a Benetel 5G RAN550 O-RU is connected to its corresponding DU using functional Split 7.2a, representing a vendor-grade deployment, while a USRP B200 O-RU operates under Split 8 to enable flexible experimentation at lower protocol layers.

RadioShield's dApp consumes O-DU-visible PHY/MAC telemetry rather than vendor-specific O-RU internals. Accordingly, the framework is not inherently tied to a particular O-RU vendor or functional split; rather, its portability depends on whether equivalent lower-layer radio indicators remain observable at the O-DU.

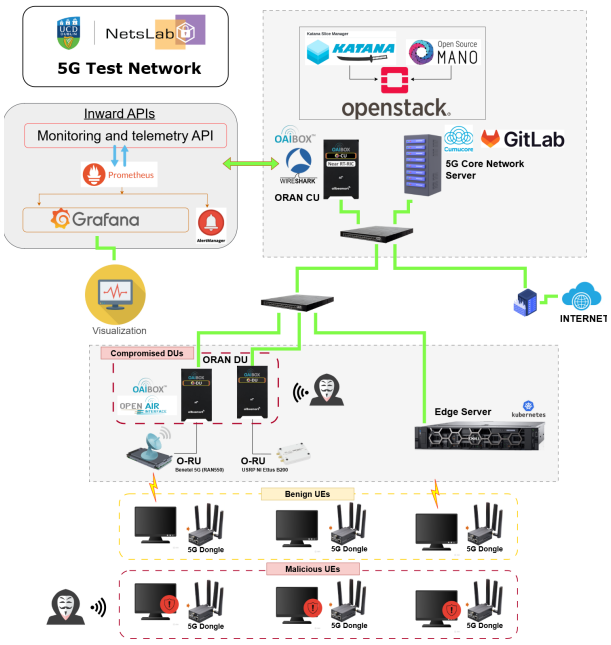


**Fig. 3:** Overview of the 5G O-RAN testbed, illustrating the network components and the origination points of the executed attack vectors.

An edge server, emulating the O-Cloud environment, is deployed at the network edge and co-locates key O-RAN components, including the Near-RT RIC. This edge infrastructure supports both near-real-time analytics and model training workflows. Real user equipment (UEs) are instantiated using commercial laptops equipped with 5G Quectel RM500Q-GL modems, which connect to the O-RUs and generate both benign and attack traffic. This setup enables end-to-end experimentation using real radio transmissions and protocol interactions, rather than simulated or emulated traffic.

### B. Attack Execution Vectors

Having established the theoretical threat model in Section II-D, we operationalized these attack vectors within our physical testbed. Unlike prior studies that rely on analytical thresholds or software-based traffic generators, our methodology enforces realistic control-plane exhaustion by programmatically driving physical hardware. Both attack scenarios were executed over sustained 30-minute observation windows to capture steady-state storm dynamics and ensure statistical significance in the resulting telemetry.

To execute the asynchronous **UE Reboot scenario**, three devices were connected to the testbed, each equipped with Quectel RM500Q-GL modems. Rather than performing manual restarts, we developed OS-level automation scripts (via Python) that programmatically toggled the *cellular network interface* using low-level system calls.

Each UE underwent repeated connect–disconnect cycles, where the cellular interface was brought down and reinitialized via ip link commands, thereby triggering a full RRC release and subsequent re-establishment sequence. The three UEs are configured with base disconnect/reconnect cycles of 5, 7, and 10 seconds, respectively. using controlled jitter injection to avoid deterministic synchronization artifacts while maintaining coordinated storm conditions. This design ensured repeated NAS registration, RRC Setup, and security negotiation procedures across all compromised UEs.

The second scenario targeted mobility amplification through controlled **O-DU availability manipulation.** Using the OAIBox interface, we programmatically toggled the operational state of one O-DU instance at 1–2 minute intervals over a 30-minute window. This forced all actively connected UEs to migrate to the alternate O-DU/O-RU pair via standard handover procedures.

Unlike the reboot scenario, where connections are gracefully terminated, this attack vector triggered concentrated bursts of F1AP and RRC reconfiguration signaling, forcing the target O-DU to rapidly ingest and allocate resources for multiple migrating UEs simultaneously, generating intense, concentrated spikes in MAC-layer scheduling latency and control-plane handover signaling.

### C. Multi-Layer Dataset Construction

*1) Radio Telemetry Dataset:* Using the OAIBox interface, we extracted cell-level O-DU telemetry to construct the Radio Telemetry Dataset. This consisted of 9,589 samples, where each row represents one second of aggregated PHY/MAC layer behavior. Unlike upper-layer logs that only reflect logical protocol states, this dataset captures the immediate physical stress on the air interface. The 24-feature set exposes the physical footprint through key metrics:

- **Signal Quality & Interference:** Metrics such as Signal-to-Noise Ratio (SINR), Reference Signal Received Power (RSRP), and Channel Quality Indicator (CQI), which distinguish genuine coverage gaps from storm-induced contention.
- **Throughput & Modulation:** Downlink/Uplink throughput and Modulation and Coding Scheme (MCS) indices. Rapid scheduling failures during a storm force unique MCS degradation patterns.
- **Power Utilization:** Maximum UE Power Capacity (PCMAX) and Power Headroom, reflecting the scheduler's struggle to accommodate power surges during mass synchronization.

*2) Network Telemetry Dataset:* To establish a traditional, upper-layer baseline for comparison, we simultaneously captured network traffic at the O-CU via Wireshark. Raw packet captures were aggregated into fixed 500 ms time windows to capture the temporal dynamics of the signaling storms. This yielded a behavior-oriented dataset of 9,437 windows with 22 features, characterized by:

- **Volumetric Intensity:** Packet counts, byte counts, and packet length statistics (mean/std).

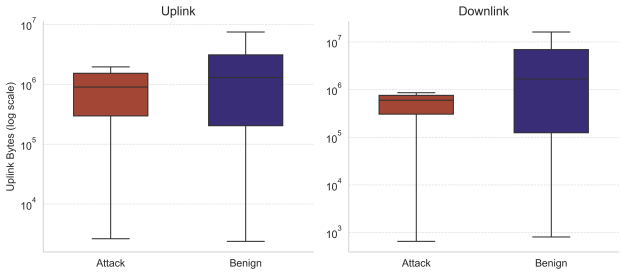


**Fig. 4:** Distribution of cumulative uplink and downlink bytes under benign and attack conditions (log scale)

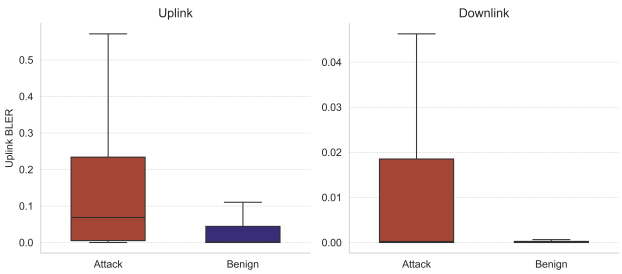


**Fig. 5:** Distribution of uplink and downlink Block Error Rate (BLER) under benign and attack conditions

- **Protocol Composition:** Per-window counts of SCTP, NGAP, F1AP, and E2AP packets.
- **UE Dynamics:** Counts of unique UE identifiers (AMF, RAN, CU, DU IDs) showcasing active device churn.
- **Temporal Burstiness:** Mean and maximum inter-arrival times between packets.

The released dataset and associated artifacts are publicly available on Kaggle [8].

### D. IDS Model Selection and Training

We deploy a diverse set of supervised learning models that represent distinct architectures, learning paradigms, and computational characteristics. Specifically, we consider Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), k-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGBoost), and a Deep Neural Network (DNN).

All deployed models were trained using a fixed set of simple hyperparameters to evaluate the robustness and generalizability across heterogeneous O-RAN observation layers. LR employs L2 regularization with class balancing, DT and RF use balanced class weights to mitigate data imbalance, and RF is configured with 300 estimators to ensure stable ensemble behavior. XGBoost is trained using a binary logistic objective and multiclass softprob objective with controlled depth and learning rate to balance expressiveness and generalization. The DNN follows a lightweight fully connected architecture with two hidden layers, ReLU activations, dropout regularization, and early stopping to prevent overfitting.

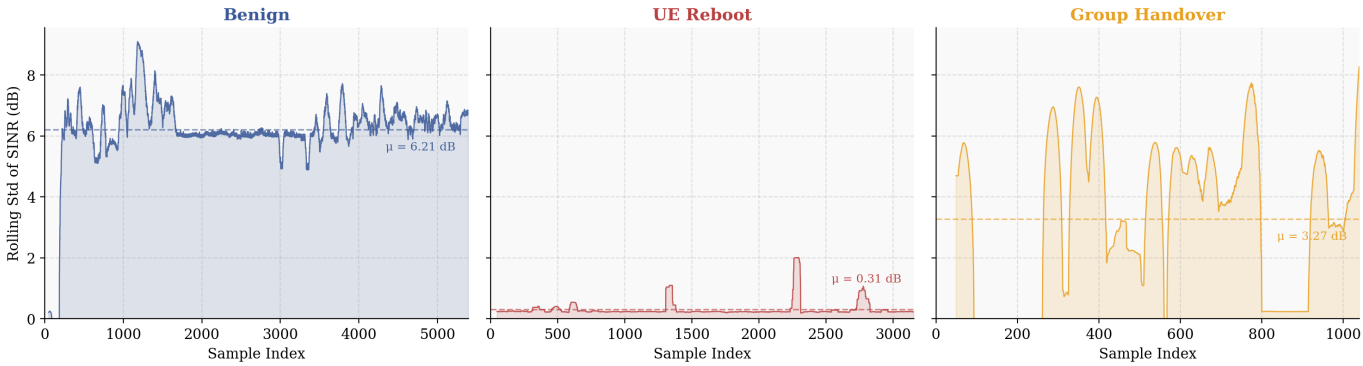


**Fig. 6:** Rolling standard deviation of Signal-to-Interference-Plus-Noise Ratio (SINR) across behavioral classes.

**TABLE II:** Detection performance of binary and multiclass models evaluated on the O-DU Radio Telemetry Dataset.

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
|---|---|---|---|---|
| **Binary Objective** | | | | |
| XGBoost | 99.43 | 99.17 | 99.52 | 99.35 |
| RF | 99.22 | 99.28 | 98.93 | 99.10 |
| DT | 98.07 | 97.74 | 97.85 | 97.80 |
| DNN | 95.62 | 95.87 | 94.04 | 94.95 |
| KNN | 93.53 | 94.97 | 89.99 | 92.41 |
| LR | 91.66 | 92.49 | 88.08 | 90.23 |
| **Multiclass Objective** | | | | |
| XGBoost | 99.53 | 99.33 | 99.20 | 99.27 |
| RF | 98.80 | 98.88 | 97.65 | 98.24 |
| DT | 97.81 | 97.32 | 96.48 | 96.89 |
| DNN | 93.33 | 89.54 | 89.52 | 89.52 |
| KNN | 90.88 | 90.19 | 83.10 | 85.82 |
| LR | 82.69 | 74.84 | 79.77 | 76.41 |

**TABLE III:** Detection performance of binary and multiclass models evaluated on the Upper-Layer Network Dataset.

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
|---|---|---|---|---|
| **Binary Objective** | | | | |
| XGBoost | 97.88 | 97.67 | 97.17 | 97.42 |
| RF | 96.93 | 96.63 | 95.88 | 96.25 |
| DT | 95.71 | 94.5 | 95.11 | 94.8 |
| DNN | 95.87 | 95.81 | 94.08 | 94.94 |
| KNN | 95.02 | 95.84 | 91.89 | 93.82 |
| LR | 90.2 | 90 | 85.71 | 87.8 |
| **Multiclass Objective** | | | | |
| XGBoost | 55.67 | 35.82 | 33.96 | 31.04 |
| RF | 57.47 | 33.37 | 34.44 | 30.58 |
| DT | 47.93 | 35.71 | 35.89 | 35.73 |
| DNN | 58.69 | 25.16 | 33.27 | 24.76 |
| KNN | 55.03 | 37.67 | 34.46 | 32.53 |
| LR | 58.85 | 61.30 | 33.56 | 25.27 |

## V. Results and Discussion

### A. Signaling Storm Detection using PHY/MAC Layer Features

To validate the RadioShield framework, we analyzed cell-level O-DU telemetry and observed that signaling storms produce PHY/MAC degradation because the attack is injected at the air interface and therefore perturbs radio access and scheduling before only aggregated upper-layer symptoms become visible.

For instance, Fig. 4 shows a reduction in cumulative uplink and downlink bytes during attack periods. This occurs because repeated RRC setup, re-establishment, and handover procedures consume radio resources in short-lived control-heavy exchanges, interrupting stable user-plane service. As a result, fewer resources remain for sustained payload transfer, causing less traffic to accumulate within each observation window.

A similar trend is observed for the Block Error Rate (BLER). As shown in Fig. 5, both uplink and downlink BLER shift upward and become more dispersed during attack periods. This implies that the attack injects many transmissions in fragile phases, such as initial access, re-access, incomplete link adaptation, uncertain timing/power alignment, and bursty scheduler churn. These present the conditions under which uplink and downlink decoding are less stable, and HARQ retransmissions are more likely.

Beyond binary detection, the temporal dynamics of radio metrics provide distinct signatures for classifying specific attack vectors. Fig. 6 displays that the rolling standard deviation of SINR changes markedly across classes, reflecting the changes in the set and state of active radio links. Benign traffic exhibits the highest variance ($\mu \approx 6$ dB) due to the natural channel diversity of active UEs. In contrast, reboot storms suppress sustained link diversity ($\mu \approx 0.31$ dB) due to devices repeatedly disappearing and reappearing, suppressing stable user-plane activity. Handover storms repeatedly move UEs across radio contexts, producing sharp, bursty swings in the aggregate SINR distribution ($\mu \approx 3.27$ dB). These results show how PHY/MAC telemetry improves class separability by capturing the radio-side consequences of how the attack is executed, not just the delayed control-plane volume it generates.

### B. IDS Performance Evaluation Using Radio Telemetry

We evaluate six machine learning models on the Radio Telemetry Dataset, with binary and multiclass metrics presented in Table II. The results reveal that tree-based algorithms demonstrate clear superiority in capturing the physical footprints of RRC signaling storms. XGBoost achieves the highest performance in both classification objectives, maintaining robust performance in terms of accuracy and F1-score.

LR accuracy degrades significantly from 91.66% in binary detection to just 82.69% in the multiclass objective. Furthermore, the DNN model yields comparatively lower performance across both objectives. This shows that while previously discussed metrics exhibited clear visual separability

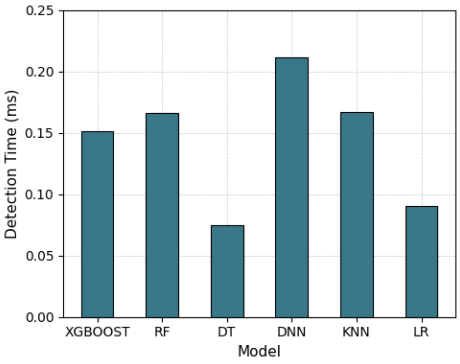


**Fig. 7:** Detection time of the binary models for a single inference at the PHY/MAC layer

between benign and malicious traffic, the broader PHY/MAC feature space contains highly overlapping, non-linear distributions. This renders simple rule-based heuristics ineffective and explains why models, such as KNN and LR, struggle to capture the attack footprint.

Beyond predictive accuracy, deploying an IDS as an early-warning dApp mandates strict adherence to sub-millisecond control loops to alert the overarching network before congestion precipitates. Fig. 7 illustrates the inference delay distribution for each evaluated model. While deep learning models inherently introduce longer inference delays, tree-based models like XGBoost and DT execute in a fraction of the time. XGBoost achieves microsecond-scale inference latency while delivering state-of-the-art accuracy. This optimal balance of speed and precision ensures that signaling storms can be detected and mitigated at the physical layer long before control-plane saturation propagates to the O-CU.

### *C. IDS Performance Evaluation Using Upper-Layer Network Metrics*

To validate the superiority of our radio-centric approach, we evaluated the same six machine learning models using the upper-layer network dataset. As detailed in Table III, the binary classification models demonstrate capable performance, with XGBoost again leading. Because both UE reboot and group handover storms ultimately cause massive traffic surges, network-based volumetric features can successfully identify the presence of a generalized anomaly. Furthermore, this upper-layer detection is inherently reactive, detecting a storm after the control plane is already saturated, whereas radio telemetry evaluates immediate physical state conditions.

The critical limitation of traditional upper-layer network detection is exposed when the models attempt the multiclass objective. As shown in the lower half of Table III, performance across all algorithms collapses. This catastrophic degradation highlights the fundamental ”blind spot” of upper-layer monitoring: at the network level, both UE Reboot and Group Handover storms manifest identically as a volumetric flood of NAS, NGAP, and RRC connection messages. The time-windowed dataset lacks the granular physical context required to separate them.

Comparing these results with the radio telemetry evaluation emphatically validates our core hypothesis. While upper-layer network monitors can only detect a generic storm after resources are already saturated, the PHY/MAC layer retains the immediate and unique physical signatures of the attack's root cause and origin.

### *D. Computational Profiling and Efficiency*

To compare deployment across the O-RAN architecture, we conducted a comprehensive computational profiling of the evaluated models, contrasting their suitability for xApp and dApp integration. By isolating the training and inference routines from background processes and overall system activity, we measured execution latency, CPU effort (defined as the product of mean CPU utilization and execution time), memory footprint, and inference energy per 1000 detections.

As summarized in Table IV, processing radio-level telemetry is inherently more resource-efficient than analyzing upper-layer network traffic. Across nearly all algorithms, models trained on the Radio-Based dataset exhibit lower memory footprints and reduced CPU effort. For instance, XGBoost requires only $653.01$ MB of average memory under radio telemetry, compared to $861.44$ MB for the time-based dataset. This efficiency stems from the compact yet information-rich nature of the radio telemetry features, which reduces redundancy and computational overhead compared to the more extensive network representations.

Furthermore, Fig. 7 reveals that model size does not strictly dictate inference speed. Parametric models like the DNN maintain a compact size ($0.106$ MB) but incur the highest detection latency and sustained computational intensity. This confirms that deep neural architectures are unnecessarily resource-intensive for this structured tabular task. Conversely, tree-based models produce dataset-dependent structures that may be larger (e.g., Random Forest at $10.489$ MB) but execute significantly faster.

Ultimately, XGBoost achieves the most favorable balance of detection performance and computational efficiency. Under the Radio-Based dataset, it requires just $1.31$ seconds of training time, $43.56$ %·s of CPU effort, and $25.10$ mWh per $1000$ inferences, while maintaining a lightweight $1.387$ MB footprint. This combination of moderate latency, controlled energy consumption, and minimal storage makes the radio-level XGBoost detector highly suitable for containerized dApp deployment within the O-DU, enabling continuous monitoring for RRC signaling storms without monopolizing critical radio management resources.

## VI. Conclusion and Future Work

This paper introduced **RadioShield**, a novel PHY/MAC-layer defense framework for mitigating RRC signaling storms in disaggregated O-RAN environments. Unlike reactive upper-layer approaches, RadioShield shifts detection to the radio edge via a lightweight O-DU dApp, identifying the physical footprints of control-plane attacks well before upper-layer saturation occurs. Using a physical 5G O-RAN testbed

**TABLE IV:** Computational profiling of binary detection models across the datasets computed from different O-RAN layers

| Model | Training Time (s) | CPU Effort (%·s) | Average Memory Usage (MB) | Model Size (MB) | Inference Energy (mWh / 1000 inf) |
|---|---|---|---|---|---|
| **Radio-Based Traffic** | | | | | |
| LR | 0.39 | 46.21 | 541.72 | 0.002 | 23.34 |
| DT | 0.31 | 8.36 | 546.56 | 0.025 | 22.93 |
| RF | 4.43 | 90.03 | 554.69 | 10.489 | 27.80 |
| KNN | 0.13 | 18.36 | 561.09 | 1.114 | 31.58 |
| XGBoost | 1.31 | 43.56 | 653.01 | 1.387 | 25.10 |
| DNN | 34.48 | 671.01 | 823.38 | 0.106 | 30.63 |
| **Time-Based Traffic** | | | | | |
| LR | 0.80 | 93.02 | 852.94 | 0.002 | 8.313 |
| DT | 0.38 | 11.01 | 852.79 | 0.040 | 7.235 |
| RF | 4.57 | 96.26 | 857.31 | 12.442 | 8.535 |
| KNN | 0.18 | 18.19 | 860.10 | 1.154 | 9.911 |
| XGBoost | 1.43 | 44.94 | 861.44 | 1.665 | 8.042 |
| DNN | 16.74 | 336.50 | 855.90 | 0.106 | 10.078 |

with realistic UE reboot and handover storms, we demonstrated that radio-level telemetry offers superior discriminative power. Tree-based models, notably XGBoost, achieved over 99% accuracy in both binary and multiclass scenarios with microsecond-scale inference latency. Furthermore, computational profiling confirmed that RadioShield dApps significantly reduce memory and energy overhead in comparison to traditional xApps, proving the practical feasibility of embedding proactive, sub-millisecond security directly within the radio control path.

Future work will focus on validating the framework under higher-density and more operationally realistic conditions. Ongoing work is extending the current platform to a larger deployment with more UEs and additional RAN resources. This will enable validation of scalability, robustness, and service continuity under both adversarial signaling storms and legitimate high-density events. We are also integrating a mitigation xApp that operates on dApp alerts to support operator-defined response and resource-control policies, moving RadioShield from early detection toward closed-loop protection, including configurable action duration, safeguards for legitimate demand surges, and tunable policies that balance security enforcement with service continuity.

## REFERENCES


[1] "O-ran security threat modeling and risk assessment," O-RAN Alliance, Tech. Rep. O-RAN.WG11.TR.Threat-Modeling-R005-v08.00, 2026, working Group 11 (Security). [Online]. Available: https://specifications.o-ran.org/specifications

[2] A. Tabiban, H. A. Alameddine, M. A. Salahuddin, and R. Boutaba, "Signaling storm in O-RAN: Challenges and research opportunities," *IEEE Communications Magazine*, vol. 62, no. 6, pp. 58–64, 2023.

[3] R. Ettiane, A. Chaoub, and R. Elkouch, "Toward securing the control plane of 5g mobile networks against dos threats: Attack scenarios and promising solutions," *Journal of Information Security and Applications*, vol. 61, p. 102943, 2021.

[4] S. Park *et al.*, "Machine learning based signaling ddos detection system for 5g stand alone core network," *Applied Sciences*, vol. 12, no. 23, p. 12456, 2022.

[5] D. K. Nguyen, R. E. Malki, and F. Rebecchi, "Rrc signaling storm detection in o-ran," in *Proceedings of the IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2025.

[6] M. Hoffmann, C. Ide, and H. Ploennigs, "Signaling storm detection in iiot network based on the open ran architecture," *arXiv preprint arXiv:2302.08239*, 2023.

[7] A. Lacava, L. Bonati, N. Mohamadi, R. Gangula, F. Kaltenberger, P. Johari, S. D'Oro, F. Cuomo, M. Polese, and T. Melodia, "dapps: Enabling real-time ai-based open ran control," *Computer Networks*, vol. 269, p. 111342, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128625003093

[8] F. A. Zadeh, B. Siniarski, S. Wang, and M. Liyanage, "Rrc signalling storm dataset (5g o-ran testbed)," 2026. [Online]. Available: https://www.kaggle.com/dsv/15899347

[9] R. Ettiane and R. E. Kouch, "Mitigating denial of service signaling threats in 5g mobile networks," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 2, 2021.

[10] M. Pavloski, "Detecting and mitigating storm attacks in mobile access to the cloud," in *Proceedings of the IEEE International Conference on Fog Computing (ICFC)*. IEEE, 2019.

[11] O. H. Abdelrahman and E. Gelenbe, "Detecting network-unfriendly mobiles with the random neural network," *Probability in the Engineering and Informational Sciences*, vol. 30, no. 4, pp. 514–531, 2016.

[12] D. Feng *et al.*, "Research of deep learning and adaptive threshold-based signaling storm prediction and top cause tracking," *IEEE Access*, vol. 11, pp. 120 603–120 611, 2023.

[13] D. K. Nguyen *et al.*, "Beyond static thresholds: Adaptive rrc signaling storm detection with extreme value theory," in *Proceedings of the International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*. IEEE, 2025.

[14] M. Hoffmann and P. Kryszkiewicz, "Signaling storm detection in iiot network based on the open ran architecture," in *IEEE INFOCOM 2023 Workshops*. IEEE, 2023.

[15] S. Mayhoub *et al.*, "A new sub-use case for signaling storm attack in open ran and an ml-based detection approach," in *Proceedings of the IEEE Conference on Standards for Communications and Networking (CSCN)*. IEEE, 2024.

[16] "O-ran architecture description," O-RAN Alliance, Tech. Rep. O-RAN.WG1.TS.OAD-R005-v16.00, 2026. [Online]. Available: https://specifications.o-ran.org/specifications

[17] C. M. Osazuwa, "Confidentiality, integrity, and availability in network systems: A review of related literature," *International Journal of Innovative Science and Research Technology*, 2024.