XCOL: xApp-based Framework for Detecting Layer-2 DoS Attacks in 5G NR

Arpit Tripathi[1], Abdulla Ovais[1], Bheemarjuna Reddy Tamma[1], Abhishek Thakur[2]
[1]IIT Hyderabad, India [2]IDRBT Hyderabad, India
{cs20resch1006, cs24mtech12014, tbr}@iith.ac.in, abhishekt@idrbt.ac.in

***Abstract*—The growing adoption of 5G in industrial deployments expands the attack surface for adversarial exploitation. Fake Base Stations (FBSs), which lure benign UEs in a fake attach procedure, and Malicious User Equipment (MUEs), which simultaneously flood legitimate gNBs, can be used to launch combined Denial-of-Service (DoS) attacks. Mobile networks lack the capability to detect Layer-2 DoS attacks originating from these FBSs and MUEs. In this work, we propose an O-RAN-compliant XCOL framework that uses a passive air-interface sniffer (SNI5GECT) to collect Layer-2 telemetry alongside standard E2 KPM telemetry from the O-DU, and deploys an xApp on the Near-RT RIC to detect combined FBS- and MUE-induced attacks. The xApp employs sequence models (BiLSTM-Attention, Conv1D) over per-UE RACH event traces, leveraging three engineered feature families: Temporal Burstiness, RAR suppression ratio, and RACH Completion Ratio. We show that the proposed XCOL achieves an F1 score of 1.0 for gNodeB-level MUE flood detection. At the UE level, it attains an F1 score of 0.678 with 100% precision for MUE detection and an F1 score of 0.903 for FBS-victim identification, outperforming threshold baselines while satisfying Near-RT RIC latency requirements.**

***Index Terms*—5G mobile communication, Open RAN, Base stations, Denial-of-service attack, Attack detection**

I. Introduction

One of the most common attacks on 5G network infrastructure used in industrial and mission-critical use cases is a DoS attack originating from FBS and MUE [1]. An FBS is defined by two distinct characteristics: (i) actively monitoring the radio environment to spoof System Information Blocks (SIBs) broadcast by legitimate gNodeBs and (ii) leveraging these intercepted parameters to lure legitimate UEs within its vicinity through a higher Reference Signal Received Power (RSRP) or through advantageous physical proximity to the target UEs [2]. An MUE launches a flooding attack against a legitimate gNodeB [3] by exhausting the Physical Random Access Channel (PRACH) occasions allocated for legitimate UE access. Both attack vectors fundamentally disrupt service availability by either denying or delaying UE registration. A few recent works on FBS-induced DoS attacks include Random Access Channel (RACH) drops [2], RRCSetupRequest drops, and Tracking Area Update (TAU) Reject attacks [4].

While considerable research has focused on mitigating these DoS attacks, modern detection strategies increasingly leverage the disaggregated O-RAN architecture, where xApps and rApps operating within the Near-RT and Non-RT RICs interact with the O-CU/O-DU/O-RU via the E2 interface [5]. This enables fine-grained control over resource allocation and scheduling, providing mechanisms for detecting and countering network threats.

Existing O-RAN-based threat detection work can be classified into three groups. The first group operates at the upper layer of the protocol stack: [5] and [6] detect RRC/NAS level signaling anomalies but do not observe Layer-2 RACH behavior, so FBS and MUE-induced attacks are invisible to them. The second group relies on UE-side observation: threshold-based detection by Arpit et al. [2] and Purification et al. [7] uses UE-reported timing and retry counts, assuming UE is trustworthy and reachable. This assumption fails when UE is the victim or is compromised. The third group targets MUE signaling floods [8] and [3] but does not address FBS attacks. None of these works jointly monitors and detects combined FBS and MUE attacks in 5G NR.

In order to address this research gap, we propose an **x**App-based detection mechanism for **CO**mbined **L**ayer-2 (XCOL) attacks in an O-RAN-based 5G NR system. The key contributions are as follows:

- **Combined Attack Model:** We formalize a joint FBS+MUE attack model for 5G NR, and measure its impact on RACH message failures and radio contentions.
- **Detection Framework:** We design a detection pipeline that collects passive Layer-2 RACH telemetry to jointly detect FBS victims and MUEs, benchmarking session-aggregated ML models (XGBoost, MLP) and temporal sequence models (BiLSTM-Attention, Conv1D) under Near-RT RIC latency constraints.
- **`nr-attack`:** An ns-3 extension for simulating 5G NR RAN attacks, covering RACH procedures, FBS/MUE attack behavior, UE mobility and association, and attack-impact metrics.

II. System & Attack Model

This section formalizes the system model, RACH model, and the attack model.

A. System Model

We consider a 5G network with a set of UEs within a bounded region $\mathcal{A}$. $U = \{u_1, \ldots, u_N\}$ represents randomly distributed UEs (static or mobile) in the region $\mathcal{A}$. $G = \{g_1, \ldots, g_K\}$ represents gNodeBs that are deployed uniformly in that region $\mathcal{A}$, each broadcasting system info like MIB and SIBs. The UEs attach to the nearest gNodeB in the region.

B. 3GPP RACH Model

The 4-step RACH procedure of UE with a legitimate (benign) gNodeB completes in $\tau_{\text{normal}} = \tau_{\text{RAR}} + \tau_{\text{Msg3}} + \tau_{\text{Msg4}}$, where τ_{RAR} denotes the modeled time taken for the UE to send Random Access Request (aka Msg1) and receive a RA Response (RAR aka Msg2) from the base station, τ_{Msg3} denotes the modeled scheduling interval from receiving Msg2 to transmitting Msg3, and τ_{Msg4} denotes the interval from transmitting Msg3 to receiving the contention resolution message (Msg4). If Msg4 is not received, the UE waits until the contention resolution timer expires, $\tau_{\text{Msg4_TO}} = T_{\text{CR}}$, before declaring RACH failure.

TABLE I: RACH parameter vector $\boldsymbol{\theta}_{\text{RACH}}$ (Ref Specs)

Sym.	IE	Spec.	Permissible Values
μ	SCS	38.211	$\{0, 1, 2, 3\}$
W_{RA}	`ra-ResponseWindow`	38.331	$\{1, 2, 4, 8, 10, 20, 40, 80\}$
T_{CR}	`ra-CRTimer`	38.321	$\{8, \ldots, 64\}$
$P_{\max}$	`preambleTxMax`	38.321	$\{3, \ldots, 200\}$
BI	Backoff Interval	38.321	$\{5, \ldots, 960\}$ ms
T_{311}	T311	38.331	$\{1, 3, 5, 10, 15, 20, 30\}$ sec
k_2	UL offset	38.214	$\{1, \ldots, 32\}$
ΔP	`pwrRamp`	38.321	$\{0, 2, 4, 6\}$ dB

3GPP defines the RACH procedure and its timing constraints. We parameterize all RACH timings through a vector:

$$\boldsymbol{\theta}_{\text{RACH}} = \big(\mu,\ W_{\text{RA}},\ T_{\text{CR}},\ P_{\max},\ \text{BI},\ T_{311},\ k_2,\ \Delta P\big), \tag{1}$$

whose elements capture the principal 3GPP-configured parameters governing NR random access, including RACH-ConfigCommon parameters, RAR-associated parameters, and UL scheduling and timing parameters relevant to the attachment procedure. The attach delay scales with numerology (see public repo[1] for the full parameter table and other relevant notations). The NR slot duration is given by:

$$T_{\text{slot}}(\mu) = \frac{1}{2^{\mu}}\ \text{ms} \tag{2}$$

yielding 1.0 ms at $\mu = 0$ (15 kHz SCS), 0.5 ms at $\mu = 1$ (30 kHz), and so on.

C. Attack Model

The attack components consist of FBSs and MUEs (static or mobile) capable of launching Layer-2 DoS attacks in 5G NR. $F = \{f_1, \ldots, f_{n_F}\}$ represents randomly distributed FBSs in region $\mathcal{A}$, each capable of sniffing nearest gNodeB parameters like MIB and SIBs. Each FBS lures legitimate UEs through *proximity advantage*. The FBSs could be static or mobile which is at the attacker's discretion. A set of $M = \{m_1, \ldots, m_{n_M}\}$ MUEs are randomly placed in the region $\mathcal{A}$. We now characterize the two attack vectors.

1) DoS Attack by FBS: Dropping of RACH Messages from legitimate UEs by FBS

We extend the RACH drop attack proposed in [2]. In this scenario, we assume that the FBS drops Msg1 of legitimate UEs with probability $p_{\text{drop}}^{\text{Msg1}} = 0.5$ (representing a balanced attacker strategy), and the UE waits $\tau_{\text{Msg2_TO}}$ for a RAR. But the UE does not receive the RAR because Msg1 was dropped. Alternatively, if it does not drop, the FBS sends a crafted Msg2. The UE responds with Msg3. The FBS then always drops Msg3 to force the attack, and the victim UE waits for $\tau_{\text{Msg4_TO}}$ for Msg4 before declaring the RACH attempt failed. The UE retries until $P_{\max}$ is exhausted.

2) DoS attacks by MUEs on legitimate gNodeBs

MUEs trigger multiple spurious RACH requests to legitimate gNodeBs, resulting in wasted RACH occasion slots. The additional scheduling delay for benign UEs is modeled as a function of the contention ratio $\rho(g_k) = n_M(g_k)/|U_G(g_k)|$. Here, $n_M(g_k)$ denotes the number of MUEs targeting g_k, while $|U_G(g_k)|$ denotes the number of legitimate UEs associated with g_k. For each retry, Msg1 may collide with MUE preambles with probability $p_{\text{drop}} = 1 - e^{-\rho}$. The UE succeeds after multiple retries with an increased attach delay.

3) Combined attacks: FBS+MUE

The RACH drop rate at gNodeB g_k equals to the per-attempt collision probability:

$$p_{\text{drop}} = \frac{\#\ \text{RACH Drops}}{\#\ \text{RACH Attempts}} = 1 - e^{-\rho}. \tag{3}$$

Since $\rho(g_k) = n_M(g_k)/|U_G(g_k)|$, diverting UEs via an FBS reduces the denominator while n_M remains fixed:

$$\begin{aligned} |U_G^{\text{combined}}| < |U_G^{\text{MUE-only}}| &\Rightarrow \rho^{\text{combined}} > \rho^{\text{MUE-only}} \\ &\Rightarrow p_{\text{drop}}^{\text{combined}} > p_{\text{drop}}^{\text{MUE-only}}. \end{aligned} \tag{4}$$

For $n_M = 1$ and 10 benign UEs, $\rho = 0.1$ and $p_{\text{drop}} = 9.5\%$. If an FBS diverts 3 UEs, $|U_G| = 7$, giving $\rho = 0.143$ and $p_{\text{drop}} = 13.3\%$, a 40% relative increase. Thus, the FBS amplifies MUE contention not by flooding PRACH, but by reducing the legitimate UE population, increasing ρ and contention for the remaining UEs.

Having defined the attack model, we now present the proposed XCOL detection framework.

III. XCOL Attack Detection Framework

This section presents XCOL (Fig. 1), a joint-attack-detection framework. It combines passive air-interface telemetry from SNI5GECT [9] with E2-sourced RAN KPMs to detect FBS and MUE attacks. The key components are:

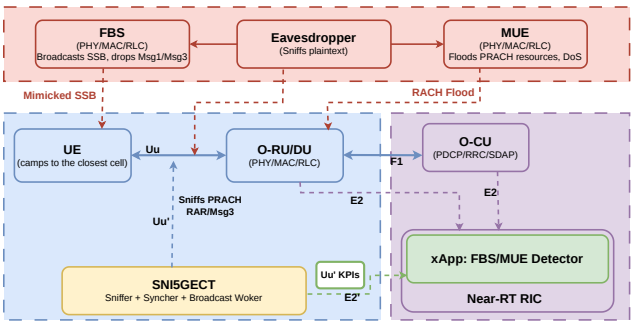


Fig. 1: XCOL Framework

[1]https://github.com/arpit376/nr-attack

*1) **SNI5GECT: Telemetry Collection***

The XCOL framework, as shown in Fig 1, uses the SNI5GECT sniffer [9] to listen to the Uu air interface of 5G NR. It is a passive sniffer that decodes both uplink and downlink messages related to the Synchronization and Random Access Procedure on the Uu air interface, including those between benign UEs and the FBS.

SNI5GECT extracts per-UE, per-attempt RACH KPIs, including Msg1 transmission counts, Msg2 reception/timeout status, incomplete Msg3-Msg4 cycles, contention resolution timer expiries, and $P_{\max}$ exhaustion events. These logs are converted into custom KPIs and then forwarded to the Near-RT RIC as a structured stream via the custom E2' interface. The SNI5GECT module does not report any flags or labels.

*2) **E2 KPIs (via E2SM-KPM) from the O-DU***

We define the RACH drop rate at a gNodeB as the number of dropped RACH messages divided by the total number of RACH attempts at that cell. The legitimate gNodeB reports cell-level RACH statistics to the Near-RT RIC over the E2 interface, including RACH drop rate, active sessions, and RACH contention statistics per cell. These measurements can be used by an xApp to detect MUE flooding attacks.

*3) **xApp: Joint FBS/MUE Detector***

The proposed xApp jointly detects the presence of FBSs, MUEs, and victim UEs by analyzing RACH behavior. Instead of independent telemetry paths, our proposed mechanism correlates UE-level and gNodeB-level observations within a single inference pipeline. The xApp combines E2' and E2 KPIs within a single inference pipeline. RACH attempts are grouped into per-UE sessions using temporary identifiers (e.g., preamble ID, RA-RNTI, timing context), enabling the reconstruction of Msg1–Msg3 event sequences even before contention resolution.

For each UE, the xApp computes: (i) RACH completion ratio (CR): ratio of successful Msg4 transmissions to Msg1 attempts, $CR = |Msg4|/|Msg1|$ and (ii) temporal burstiness (TB) (Msg1 inter-arrival statistics) of RACH attempts. At the gNodeB level, aggregated features include the density of failed UEs. Detection is performed jointly across the UE and gNodeB domains. At the cell level, each gNodeB is classified as NORMAL or FBS based on aggregated RACH behavior. A gNodeB exhibiting persistently low RACH CRs across multiple UEs, combined with stable Timing Advance (TA) patterns, is flagged as a candidate FBS. At the UE level, each UE is classified as BENIGN, MUE, or FBS_VICTIM using the temporal feature sequence.

The xApp processes telemetry through three stages. *Feature Extraction:* per-UE RACH CR and TB stats are computed every 500 ms windows, and then joined with E2 cell-level context (RACH drop rate, failure probability). *Joint Detection:* Conv1D and BiLSTM-Attention models classify each UE's per-window feature time series, capturing temporal patterns that static threshold-based schemes miss, such as an MUE flooding preamble faster than the standard backoff interval across consecutive RACH occasions and an FBS victim exhausting all $P_{\max}$ retries due to repeated RAR or Msg4 timeouts. Thresholds on individual features achieve zero per-class F1 (Table IV), as both attack types share low CR, whereas temporal ordering discriminates them. Output: each UE is classified as BENIGN, MUE, or FBS_VICTIM. Detected MUEs trigger a higher BI at the affected gNB, throttling RACH contention. Detected FBS victims can be redirected to other cells via RRC reconfiguration.

IV. Experimental Setup and Data Collection

This section presents the simulation environment, data collection pipeline, and experimental setup for evaluating the XCOL framework.

A. Implementation of Combined Attack Model in NS-3

We implement the `nr-attack` module as an extension for ns-3.46, which is explained below, along with its constituent modules. The module does not modify PHY-layer parameters or model PHY-layer collisions; instead, it adds attack-specific RACH behavior and evaluation logic around the existing 5G-NR simulation framework. The details are listed in Table II:

TABLE II: Components of NS-3 `nr-attack` Module

Category	Components
Model	`UeAssociationModel` (distance-based cell selection); `RachProcedureSimulator` (4-step RACH with FBS/MUE effects, Sec. IV-A1); `AttackCoordinator` (evaluation loop); `AttackMobilityModel` (Random Waypoint + T311).
Attack Entities	`FbsNode` (`NrGNB` + Msg1/Msg3 dropping); `MueNode` (`NrUE` + simultaneous RACH); `RachFloodApplication` (RACH attack procedures).
Helper	`NrAttackHelper`: configure → create nodes → mobility → configure attackers → applications → start → metrics.
Examples	`coordinated-attack-example.cc`: coordinated FBS/MUE attack scenario.

1) RACH Procedure Simulation

The `RachProcedureSimulator` class generates event-level RACH logs for every UE at every evaluation instance. The simulator computes distance-based association for all N UEs across the combined set $G \cup F$ and identifies impacted gNodeBs (G_{impacted}) from MUE target assignments.

a) Motivation to build the Simulator

FBS and MUE attacks are observable only through the four-step RACH exchange, but 5G-LENA [10] provides a contention-free RACH. 5G-LENA does not model MUE-induced preamble collisions, RAR scheduling, or message drops, and lacks per-message event granularity and standard-compliant timeouts. We therefore implement a dedicated RACH timing model that derives all delays directly from the parameters in 3GPP TS 38.321 and TS 38.331, following the numerology-aware, slot-based design principle used internally by 5G-LENA.

b) Justification for periodic cell reselection

UEs in RRC_IDLE state autonomously perform cell reselection and execute the four-step RACH to reach RRC_CONNECTED. This is consistent with mMTC/IoT periodic uplink reporting (TR 37.868) and is modeled by

Bello et al. [11] as a four-state semi-Markov chain, where each cycle begins with a fresh RACH from RRC_IDLE. Our one-sec evaluation interval with N=100 UEs represents a dense IoT deployment (e.g. smart factory) generating ∼100 independent access attempts, each selecting the geographically closest cell from $G \cup F$.

c) Implementation

A new structure `NrRachConfig` encapsulates $\boldsymbol{\theta}_{\text{RACH}}$ and exposes derived timing through inline methods, implementing the RACH model (Section II-B). The three RACH simulation paths: FBS, impacted gNodeB, and clean gNodeB, all query `NrRachConfig` for every delay, timeout, and backoff value, ensuring that a change in numerology or any RACH-ConfigCommon parameter propagates consistently to all simulation outputs. Each path writes per-gNodeB statistics: arrival count, drop count, contention ratio, failure probability, cumulative drop rates, RRC delay, and flood detection flag.

B. Dataset Preparation

We evaluate four network configurations (Baseline, FBS-only, MUE-only, and Combined), each repeated across 30 independent random seeds. The gNodeBs and the FBS are static. All benign UEs and MUEs follow the Random Waypoint mobility model. Legitimate gNodeB positions are computed by a deterministic grid formula and are identical across 30 seeds. The FBSs (static) positions are chosen randomly. Benign UE and MUE parameters, including initial positions, waypoint destinations, speeds, and pause durations, are drawn from ns-3's RNG stream and reseeded per run. This allows the motion to be random across seeds. Refer to Fig. 2 for simulation parameters and a snapshot of the network topology. Each run produces two telemetry streams. The complete set of metrics is available in the supporting GitHub repository.

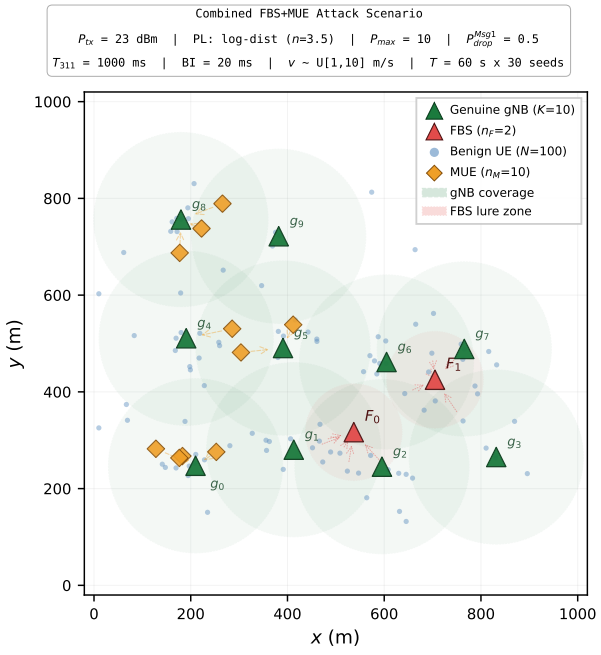


Fig. 2: 3 topology and NS-3 simulation parameters

C. Data Processing Pipeline

A post-processing step performs three operations on the telemetry streams:

(i) **Label separation.** To rule out label leakage, labels are separated from observable telemetry during preprocessing and are never used as model inputs. The gNodeB-side columns that encode ground truth or simulator internals, including `gnb_event`, `drop_reason`, and `flood_detected`, are removed from the feature set and retained only as labels for training and evaluation purposes. The dataset is split by simulation seeds to prevent run overlap across partitions, using seeds 1–21 for training, 22–26 for validation, and 27–30 for testing.

(ii) Per-UE windowed features. UE-side RACH events are grouped into 500 ms windows per (UE, gNodeB) pair using *exact* event matching to distinguish `RACH_MSG2_RECV` (RAR received) from `RACH_MSG2_TIMEOUT` (RAR suppressed). The resulting features are used to provide separability across all three classes:

- `msg1_burst_score`: captures temporal burstiness (TB) of Msg1 transmissions, separating MUE from {BENIGN, FBS_VICTIM}.
- `rar_suppression_ratio`: measures Msg2 absence, separating FBS_VICTIM from {BENIGN, MUE}.
- `rach_completion_ratio (CR)`: ratio of successful RACH completions, separating {MUE, FBS_VICTIM} from BENIGN. $CR = |Msg4|/|Msg1|$

These windowed sequences are fed directly to the sequence models.

(iii) Session aggregation. For tabular baselines, all *active* windows (those with at least one RACH event) are aggregated into a single row per UE per simulation. The sum, max, mean, and standard deviation of each feature are calculated. Silent windows are filtered before aggregation, preventing dilution of the concentrated failure signatures of FBS victims and the burst patterns of MUEs.

D. FBS/MUE Attack Detection

Experiments are conducted on a dual Intel Xeon Platinum 8480+ server (224 logical CPUs) with an NVIDIA RTX 5000 Ada GPU (32 GB VRAM).

For cell-level flood detection, we compare two unsupervised baselines with two tree-based classifiers: a fixed-threshold detector on the raw `rach_drop_rate` E2 KPI, and Isolation Forest, which flags cells whose feature vectors are separable from normal-traffic samples. These are compared against Random Forest and XGBoost, which provide strong predictive performance with sub-15 ms inference latency suitable for Near-RT RIC deployment.

For UE-level detection, *tabular models* operate on session-aggregated features, with one row per UE and only active windows retained. To address class imbalance, we apply SMOTE [12] and evaluate three models: (i) Random Forest, an ensemble baseline with feature-importance estimates; (ii) Gradient Boosting, offering strong predictive performance with sub-15 ms inference latency suitable for Near-RT RIC

deployment; and (iii) MLP, which captures non-linear feature interactions beyond tree-based boundaries. *Sequence models* instead use per-window feature vectors as ordered time series, capturing patterns across multiple windows. BiLSTM-Attention captures long-range dependencies through learned attention, Conv1D detects short local burst patterns with low latency, and LSTM-CNN combines temporal and local pattern extraction.

Three feature families drive 3-class separation. RACH CR separates attack victims (CR ≈ 0) from benign (CR ≈ 1). TB: Msg1 inter-arrival time (IAT) statistics separate MUE (burst IAT ≈ 1 ms, violating 3GPP backoff) from benign/FBS (IAT >20 ms); *RAR suppression* (`msg2_timeout`/`msg1_count`) separates FBS victims (≈ 0.5) from MUE/benign (≈ 0).

V. Performance Results

We divide the results into two parts. The first presents the impact of the individual and combined FBS+MUE attacks, while the second presents the attack-detection results of the xApp.

A. Attack Impact

In Fig. 3, we show the impact of FBS and MUE attacks on the RACH procedure. In a normal scenario, RACH attempts are 100% successful (since benign UEs face negligible contention from other UEs). In the FBS-only scenario, 12.2% of RACH messages are dropped due to a DoS attack induced by the FBS. In the MUE-only scenario, 8.9% of RACH requests succeed only after several retries (which aligns with our MUE attack model). In the combined FBS+MUE scenario, we observe that approximately 12.3% of RACH requests failed due to DoS attacks while 8.8% succeeded after several retries, resulting in delayed access for the victim UEs.

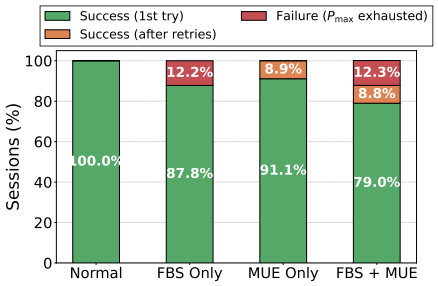


Fig. 3: Success of RACH attempts in normal and attack scenarios.

Fig. 4 shows that most gNodeBs exhibit high RACH drop rates in the combined FBS+MUE scenario. The FBS diverts benign UEs away from legitimate cells, reducing the contender pool $|U_G(g_k)|$ while the MUE density $n_M(g_k)$ remains fixed. This raises $\rho(g_k) = n_M(g_k)/|U_G(g_k)|$ at the remaining gNodeBs. A minority of cells show a slight decrease, as FBS victims that exhaust retries reposition to a different gNodeB, diluting $\rho(g_k)$ at that cell. In the FBS-only scenario, RACH drops are not observed at legitimate gNodeBs; failures occur exclusively at the FBS and are therefore invisible to the legitimate RAN. A controlled single-gNodeB experiment (K=1, N=100, n_F=2, n_M=10, all UEs and MUEs mobile under Random Waypoint, 30 seeds/configuration) eliminates cross-cell redistribution, producing a monotonic drop-rate progression: $0\% \rightarrow 0\% \rightarrow 10\% \rightarrow 18.2\%$ across Baseline, FBS-only, MUE-only, and Combined, confirming FBS-induced diversion as the amplification mechanism.

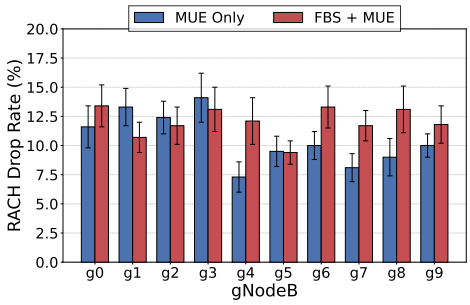


Fig. 4: RACH drop rate for different base stations in two scenarios.

TABLE III: Cell-level MUE flood detection (E2 KPMs, normal vs. flood)

Model	Acc	Prec	Rec	F1	Inference (ms)
Thresh.	0.986	1.000	0.952	0.975	0.01
Isolation Forest	0.981	0.960	0.976	0.968	65.6
RF	**1.000**	**1.000**	**1.000**	**1.000**	22.5
XGB	**1.000**	**1.000**	**1.000**	**1.000**	0.6

B. Attack Detection xApp

This section presents the results of cell-level MUE flood detection using E2 KPMs and UE-level FBS+MUE detection using E2 and E2' KPIs. We then compare the benefits of combining the two telemetry sources.

1) Aggregate RACH Floods from MUE

E2SM-KPM RACH measurements are cell-aggregate (`NRCellDU`), not per-UE. This means that the preambles do not carry UE identity until contention resolution. An MUE that floods without completing the procedure remains anonymous to the gNB. Identifying the MUE will be addressed in Section V-B2. The cell-level flood signature is unambiguous, and the appropriate response is collective RACH back-off via E2SM-RC. Table III confirms this: a threshold achieves F1 = 0.975 whereas XGBoost reaches 1.0 in 0.6 ms.

2) UE-Level FBS+MUE Detection (SNI5GECT telemetry)

Since an FBS operates outside the legitimate RAN and produces no E2 telemetry, its victims can only be identified from passive Uu-interface observation via SNI5GECT. We evaluate tabular models on session-aggregated features and temporal sequence models (BiLSTM-Attention, Conv1D, LSTM-CNN) on per-window feature sequences. Table IV reports the results on the same test set; tabular models use session-aggregated features, whereas sequence models use per-window feature sequences.

3) Comparing Attack Detection based on Attack Scenarios

Table V reports BiLSTM-Attn detection performance for each attack scenario independently. Under the combined FBS+MUE attack, the model achieves MUE F1 = 0.678 and

TABLE IV: UE-Level (Benign / MUE / FBS Victim)

Type	Model	MUE Prec.	MUE F1	FBS Prec.	FBS F1	Inf Time
Tabular	RF+SMOTE	0.964	0.495	0.709	0.663	43 ms
	GB+SMOTE	0.879	0.483	0.836	0.713	14 ms
	MLP+SMOTE	0.872	0.491	0.772	0.722	1.0 ms
Seq.	BiLSTM-Attn	**1.000**	**0.678**	0.864	**0.922**	168 ms
	Conv1D	**1.000**	**0.678**	0.871	0.922	72 ms
	LSTM-CNN	**1.000**	**0.678**	0.870	0.916	101 ms

FBS F1 = 0.903, closely matching the aggregate performance and confirming robustness when both attack vectors co-exist (XCOL). MUE detection peaks in the MUE-only scenario (F1 = 0.741), where burst patterns are unambiguous.

TABLE V: Per-Scenario Detection Performance (BiLSTM-Attn)

Scenario	MUE F1	FBS F1
MUE only (n_M=10)	0.741	—
FBS only (n_F=2)	—	0.918
Combined (n_F=2, n_M=10)	**0.678**	**0.903**

The key findings are:

(i) Sequence models significantly outperform tabular approaches: BiLSTM-Attention achieves MUE F1 = 0.678 and FBS F1 = 0.903, compared to the best tabular model (MLP+SMOTE) at MUE F1 = 0.491 and FBS F1 = 0.722. This demonstrates that temporal structure (burst timing for MUEs and failure-reposition sequences for FBS victims) provides information that is substantially lost when aggregating per-UE temporal windows into session-level features.

(ii) MUE precision of 1.0 and F1 of 0.678 indicate conservative detection: every flagged MUE is genuine, but recall is limited by temporal sparsity. Each MUE burst lasts only ∼9 ms within a one sec reporting window, leaving ∼991 ms without burst activity and thereby diluting its temporal signature at low loads. Despite missed detections, the operational impact remains bounded. Undetected MUEs increase contention ratio (ρ) at target gNodeBs, but the per-gNB RACH drop rate remains below 15% even under full MUE flooding (n_M=10), while contention-induced delay increases by only ∼2.5 ms over the clean baseline. Crucially, MUE precision of 1.0 means that no legitimate UE is classified as an MUE, avoiding false MUE alerts and the associated risk of incorrectly applying RACH throttling.

(iii) FBS recall reaches 98.9% (BiLSTM-Attention), meaning virtually no victim goes undetected. The 86.4% precision indicates that a small number of benign UEs experiencing transient congestion-induced failures are misclassified as victims.

(iv) Threshold and unsupervised baselines achieve zero F1 for the attack classes, indicating that the evaluated UE-level 3-class task benefits substantially from supervised or sequence-based ML.

(v) All model latencies fall within the Near-RT RIC's 10 ms–1 sec control-loop budget.

C. Limitations

Our evaluation uses ns-3 simulations under controlled RACH assumptions. This isolates the effects of FBS and MUE attacks but does not capture hardware non-idealities, RF noise, transport-layer jitter, or the overhead associated with live O-RAN integration. The results should therefore be considered a proof of concept, with validation on a real or emulated O-RAN testbed left for future work.

VI. Conclusions

This work proposed XCOL, a cross-layer detection framework that combines cell-level E2 KPIs with per-UE E2' KPIs to detect and distinguish coordinated FBS+MUE Layer-2 DoS attacks. To evaluate XCOL, we developed `nr-attack`, an ns-3 module for modeling coordinated FBS/MUE attacks and their effects on 5G RACH procedures. The simulation results show that cell-level MUE flooding can be reliably identified using E2SM-KPM, while per-UE discrimination among Benign, MUE, and FBS entities benefits substantially from temporal sequence information. Three complementary indicators emerge from the results: TB captures the bursty behavior of MUEs, RAR suppression captures FBS-induced failures, and CR separates attack from benign conditions. In the combined-attack scenario, the sequence model achieves MUE F1=0.678 (precision=1.0) and FBS F1=0.903, while cell-level MUE detection achieves F1=1.0.

Acknowledgments

This work is partially supported by the ISEA Phase-III project titled 'Security in Distributed Wireless Networks', funded by MeitY, Govt. of India, and by the DST-FIST grant.

References

[1] M. Gong, Z. Wei, S. Chen, W. Yu, and F. Wang, "A survey of existing attacks on 5g sa," *Computer Networks*, p. 111545, 2025.

[2] A. Tripathi, A. Rajput, A. K. Subudhi, K. Kondepu, A. Thakur, and B. R. Tamma, "Denial of service attacks targeting layer 2 in 5g ran," in *Proc of IEEE Future Networks World Forum (FNWF)*. IEEE, 2025.

[3] B. Zhang, P. Zeinaty, N. Limam, and R. Boutaba, "Mitigating signaling storms in 5g with blockchain-assisted 5gaka," in *Proc. of International Conference on Network and Service Management (CNSM)*. IEEE, 2023.

[4] K. S. Mubasshir, I. Karim, and E. Bertino, "Gotta detect'em all: Fake base station and multi-step attack detection in cellular networks," in *Proc. of USENIX Security Symposium*, 2025.

[5] H. Wen, P. A. Porras, V. Yegneswaran, A. Gehani, and Z. Lin, "5g-spector: An o-ran compliant layer-3 cellular attack detection service." in *Proc. of Network and Distributed System Security (NDSS) Symposium*, 2024.

[6] A. Scalingi, S. D'Oro, F. Restuccia, T. Melodia, and D. Giustiniano, "Det-ran: Data-driven cross-layer real-time attack detection in 5g open rans," in *Proc. of IEEE Conference on Computer Communications (ICC)*. IEEE, 2024.

[7] S. Purification, S. Wuthier, J. Kim, J. Kim, and S.-Y. Chang, "Fake base station detection and blacklisting," in *Proc. of International Conference on Computer Communications and Networks (ICCCN)*. IEEE, 2024.

[8] W. Azariah, Y.-Q. Chen, Z.-X. You, R.-G. Cheng, S.-T. Sheu, and B. Chen, "Jamming attacks on the random access channel in 5g and b5g networks," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2026.

[9] S. Luo, M. Garbelini, S. Chattopadhyay, and J. Zhou, "{SNI5GECT}: A practical approach to inject {aNRchy} into 5g {NR}," in *Proc. of USENIX Security Symposium*, 2025.

[10] N. Patriciello, S. Lagen, B. Bojovic, and L. Giupponi, "An e2e simulator for 5g nr networks," *Simulation Modelling Practice and Theory*, vol. 96, 2019.

[11] H. Bello, J. Xiaoping, R. Yue, C. Min, and I. Huseynov, "Energy-delay evaluation and optimization for NB-IoT PSM with periodic uplink reporting," *IEEE Access*, vol. 7, 2019.

[12] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *JAIR*, vol. 16, 2002.