Beyond RSRP: UE-Side Detection of Multi-Stage Fake Base Station Attacks through NAS/RRC Layer in 5G Networks*

Chaeyeon You, Vincent Abella, Hoseok Kwon, I Wayan Adi Juliawan Pawana, Bonam Kim, and Ilsun You†

Kookmin University, Seoul, Republic of Korea
{dbcodus326, vincent, hoseok1997, adijuliawan, kimbona, isyou}@kookmin.ac.kr

Abstract

Fake base station attacks threaten mobile networks by enabling MSAs such as DoS, location tracking, and bidding-down. Since many commercial devices remain vulnerable even in 4G and 5G systems, on-device detection has become crucial. Traditional FBS detection relies on radio-layer metrics, but this method fails against sophisticated MSAs where attackers subtly manipulate signals or protocol behavior. To address this limitation, this study proposes a network-layer detection approach that analyzes RRC and NAS protocol sequences. We construct a self-generated rooted dataset capturing real FBS scenarios and compare detection models trained on this dataset with those trained on conventional, RSRP-based data. The results demonstrate that RRC/NAS based analysis significantly improves detection accuracy and reduces false positives, highlighting its effectiveness as a foundation for advanced on-device FBS detection systems.

Keywords: 5G security, fake base station detection, NAS and RRC analysis, machine learning anomaly detection

1 Introduction

A fake base station (FBS) is an attack mechanism that exploits the early access phase of mobile networks to lure a user device (UE) into connecting to an adversary controlled base station, after which various protocol manipulation attacks can be carried out [1, 2]. Because the attacker intervenes before security procedures are fully activated in the 4G/5G system architecture [3], FBS attacks can trigger not only simple international mobile subscriber identity (IMSI) collection but also multi-stage threats such as denial of service (DoS), user tracking, and security-level downgrades [4, 5]. These attacks remain practical due to the continued presence of vulnerable devices in commercial networks and the relatively low cost of building FBS equipment [1].

Network side countermeasures, such as upgrading base station hardware or deploying security patches, are often difficult to apply promptly due to technical, operational, and financial constraints [6]. In environments where full trust in network operators cannot be assumed, such as public venues, roaming scenarios, or adversarial regions, relying solely on network side defenses becomes impractical. Consequently, on-device FBS detection has emerged as a practical and immediately deployable defense mechanism that does not require modifications to network infrastructure [4].

*Proceedings of the 9th International Conference on Mobile Internet Security (MobiSec'25), Article No. 12, December 16-18, 2025, Sapporo , Japan.

†Corresponding author

Many existing UE-based detection approaches rely on physical-layer indicators available on non-rooted Android devices, particularly signal-strength metrics such as reference signal received power (RSRP), reference signal received quality (RSRQ), and received signal strength indicator (RSSI) [7]. However, signal strength can fluctuate significantly under normal conditions, leading to frequent false positives, and it becomes ineffective when an attacker intentionally adjusts the FBS transmit power to match normal levels [8]. Furthermore, advanced FBS and multi-step attacks manifest anomalous message patterns, irregular state transitions, and unusual message combinations at the RRC and NAS layers characteristics that cannot be captured through RSRP alone [4, 9]. These observations suggest the necessity of incorporating higher-layer protocol sequence features to achieve reliable detection performance.

To address these challenges, this study takes a threefold approach. First, we analyze previously reported multi-step FBS attacks and reproduce representative scenarios in a controlled environment, using prior work on multi-step attack (MSA) taxonomies and detection as a baseline [2, 4]. Second, we construct a detection dataset that includes not only RSRP values but also radio resource control (RRC) and non-access stratum (NAS) protocol sequences, following the dataset construction principles in [10]. Finally, we compare the performance of an RSRP-only detection strategy with a combined approach that integrates RSRP, RRC, and NAS information. The overarching goal of this work is to design an FBS detection method that operates directly on the device without requiring support or modifications from the network side, while demonstrating improved detection accuracy and reduced false positives compared to traditional RSRP-based methods.

2 Background

2.1 5G NR RRC/NAS Overview and Malicious Handover Flow

In 5G New Radio (NR) access networks, the RRC and NAS layers manage device attachment, registration, mobility, and security procedures [11, 12]. In the `RRC_CONNECTED` state, the UE exchanges data with the serving cell and measures the signal strength of surrounding cells for mobility management.

The proposed attack scenario begins with the UE in the `RRC_CONNECTED` state, attached to a legitimate digital unit (DU). The FBS transmits signals with higher strength to be detected by the UE. The UE identifies the FBS signal as the strongest adjacent cell and transmits a `Measurement Report` to the DU [11]. Based on the reported signal quality, the DU makes a `Handover Decision` and transmits an `RRC Connection Reconfiguration` message to the UE to initiate handover to the FBS. The UE synchronizes with the FBS and completes the procedure by transmitting an `RRC Connection Reconfiguration Complete` message.

Upon successful handover, the FBS operates as a legitimate base station and initiates signaling procedures. At the NAS layer, the FBS may transmit an `Identity Request` to collect the IMSI, send `Registration Reject` or `TAU Reject` messages to cause a DoS, or issue a `Security Mode Command(NEA0)` to downgrade security. At the RRC layer, the FBS may terminate the connection using `RRC Release` or disrupt communication through repeated `RRC Reconfiguration` messages [12, 3].

2.2 Role of RSRP and Its Limitations in FBS Detection

Most UE-side FBS detection approaches rely on layer 1 radio measurements such as RSRP and RSRQ, which can be collected without root privileges [7]. However, RSRP merely reflects

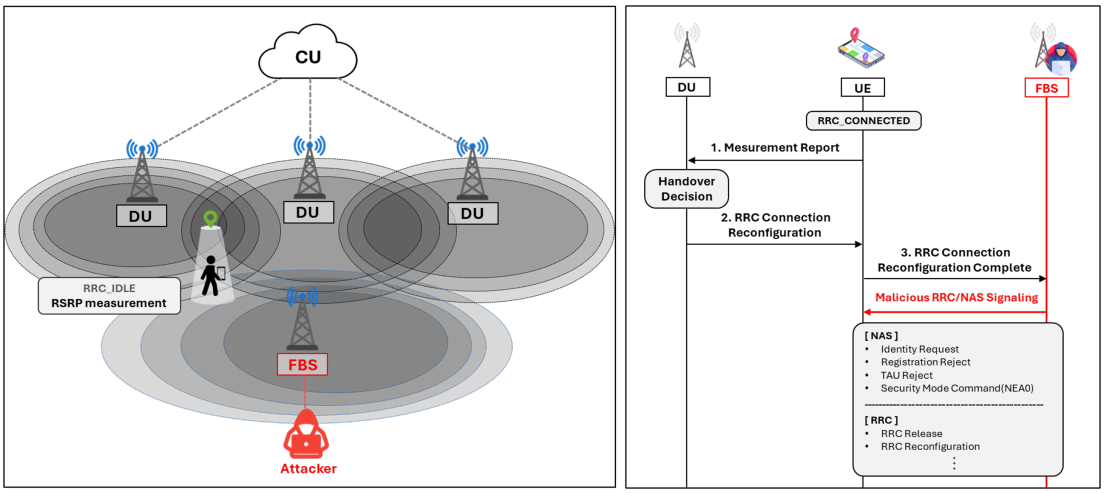

Figure 1: Overview of the FBS attack scenario.

received signal strength and can easily be manipulated by adjusting the transmit power of a rogue base station. Moreover, layer 1 metrics do not capture protocol-level manipulations such as injecting registration reject messages, forcing identity request responses, or altering security mode command parameters [8, 4]. Consequently, RSRP-based methods fail to detect abnormal state transitions or higher-layer signaling manipulations that are central to modern attacks [2].

2.3 Fake Base Stations and Multi-Step Attacks (MSAs)

Recent studies show that FBS attacks have evolved beyond simple IMSI harvesting into multi-phase procedures that manipulate both RRC and NAS signaling [2, 4, 5]. These multi-step attacks (MSAs) deliberately modify multiple protocol stages in sequence to drive the UE into an abnormal or vulnerable state. Rather than relying on a single message anomaly, MSAs combine several RRC/NAS procedures to achieve their objectives. Representative MSA types discussed in recent literature include:

- Identity request injection for IMSI exposure, where a NAS identity request is injected at a non-standard point in the procedure to force the UE to transmit its unprotected IMSI [5].
- Registration or tracking area update (TAU) reject manipulation, using abnormal cause values to push the UE into limited service or deregistration states, leading to extended DoS or forced fallback to legacy radio access technologies (RATs) [4].
- Security mode command downgrade, where the attacker alters security negotiation and forces the UE to adopt null encryption (NEA0) or weaker algorithms, exposing subsequent signaling to interception and modification [3, 2].
- Abnormal RRC release or reconfiguration, transmitting `RRCRelease` or `RRCReconfiguration` at unnatural timing to prematurely terminate connections or repeatedly trigger reselection, enabling continuous IMSI exposure or repeated DoS cycles [11, 4].

These MSAs interfere with (i) the cell selection phase, (ii) NAS registration and authentication procedures, and (iii) security and session management stages [12, 3]. Because they span multiple protocol layers and phases, they cannot be detected using single-metric layer 1 indicators such as RSRP. In contrast, RRC/NAS sequences reveal clear abnormal patterns in message ordering, cause values, and procedure invocations [4, 9].

3 Data and Methodology

3.1 Dataset Description

The dataset utilized in this study was collected from a custom 4G LTE testbed environment. The testbed was constructed using Open5GS [13] as the core network, srsRAN [14] as the software-defined eNodeB, and USRP B210 hardware as the RF front-end. We executed various attack scenarios, including IMSI catching and Authentication Reject attacks, to generate malicious traffic alongside normal baseline traffic, following procedures described in recent studies [4, 5, 10]. The collected data is categorized into three distinct datasets based on the protocol layer and accessibility.

- **RSRP Dataset (Physical Layer):** This dataset consists of physical signal measurements, such as RSRP and RSRQ. These metrics were collected using standard Android APIs on non-rooted devices, representing the most accessible features for user-side detection as explored in prior works [7, 8].
- **RRC Dataset (Layer 3 Control):** This dataset captures RRC messages, which manage connection establishment and mobility as defined in 3GPP TS 38.331 [11]. Extracted from cellular diagnostic logs (requiring root privileges), it includes System Information Blocks (SIBs) and Measurement Reports, which are critical for identifying cell spoofing mechanisms [9].
- **NAS Dataset (Layer 3 Non-Access Stratum):** This dataset focuses on the NAS protocol, which handles user authentication and security [12, 3]. It contains high-level signaling messages such as Identity Requests and Authentication Rejects, providing direct evidence of attacker intent [1, 2].

3.2 Data Preprocessing and Feature Selection

Before training the detection models, we performed data preprocessing and labeling to distinguish between legitimate and malicious traffic.

3.2.1 Data Labeling

We adopted a binary labeling scheme for the supervised learning tasks. Legitimate base station traffic was assigned the label **0 (Normal)**, while malicious traffic originating from false base stations was assigned the label **1 (Attack)**. For the Physical layer dataset (RSRP), labeling was performed based on the Physical Cell ID (PCI) observed during the experiments, adopting a strategy similar to Nakarmi et al. [7]. Signals originating from legitimate PCIs (1 and 2) were labeled as Normal (0), while signals from attacker PCIs (3 and 4) were labeled as Attack (1).

In the case of RRC and NAS datasets, data collection was conducted separately for normal and attack scenarios. Consequently, all packets captured during attack experiments were labeled as Attack (1), whereas those acquired during standard operations were labeled as Normal (0). These datasets were subsequently merged to construct the final training set.

3.2.2 Feature Selection

To mitigate high dimensionality and overfitting, we extracted only critical features relevant to FBS characteristics. The selected features focus on signal strength anomalies in the Physical layer and procedural irregularities in the RRC and NAS layers. Table 1 summarizes the selected features.

3.3 Model Architecture and Training

This study analyzes the time-series evolution of signal strength and the order of protocol messages for FBS detection. We selected three deep learning architectures to process these sequential data inputs.

- **Long Short-Term Memory (LSTM):** LSTM is a recurrent neural network that addresses the vanishing gradient problem. We used this model to capture sequential dependencies in RSRP logs and NAS/RRC signaling, allowing for the distinction between normal mobility and attack sequences.
- **Gated Recurrent Unit (GRU):** GRU is a variant of LSTM with a simplified gating mechanism. We selected GRU to reduce computational cost on UE resources while maintaining the capability to model sequential data.
- **One-dimensional Convolutional Neural Network (1D-CNN):** 1D-CNN utilizes convolution kernels to extract local features from sequential input. We employed this model to identify localized anomalies within the signal and protocol stream, such as RSRP fluctuations or specific signaling patterns.

Table 1: Selected Features for False Base Station (FBS) Detection

Layer	Feature Name	Description & Role in Detection
Physical	RSRP	Signal strength (dBm). FBS typically transmits higher power to lure victims.
	RSRQ	Signal quality (dB). High RSRP with low RSRQ indicates potential interference or spoofing.
	Physical Cell ID (PCI)	Identifier for the physical cell. Anomalous or rapidly changing PCIs suggest FBS.
	EARFCN	Frequency channel number. Checks if the frequency band matches the legitimate carrier.
RRC	Tracking Area Code (TAC)	Location area code. FBS uses distinct TACs to trigger Tracking Area Update (TAU) procedures.
	Cell Identity	Unique identifier within the PLMN. Used to verify if the cell belongs to the legitimate network.
	Establishment Cause	Reason for connection request (e.g., *mo-Signalling*). FBS may force specific causes.
	System Information (SIB)	Broadcast information (e.g., SIB1, SIB2). FBS often broadcasts malformed or missing SIBs.
	Release Cause	Reason for connection release. FBS may use causes like *loadBalancingTAURequired* to drop users.
NAS	Mobile Identity Type	Distinguishes between IMSI and TMSI. FBS requests IMSI (permanent ID) instead of TMSI.
	IMSI / TMSI	The actual identity values. Plaintext IMSI transmission is a strong indicator of an IMSI Catcher.
	EMM Cause	Error codes (e.g., #15 *No Suitable Cells*). Critical for identifying DoS attacks.
	Security Header Type	Indicates if the message is integrity protected or sent in plain NAS (Null ciphering).
	Ciphering Algorithm	Encryption algorithm used (e.g., EEA0). FBS forces null encryption (EEA0) to eavesdrop.
	Message Auth Code (MAC)	Integrity check code. Mismatches indicate potential tampering or authentication failure.

Note: Features are categorized into Physical, RRC, and NAS layers based on the LTE protocol stack.

4 Experimental Results and Analysis

In this study, we evaluated the performance of three deep learning models—LSTM, GRU, and 1D-CNN—to validate the proposed detection method. The experiments were conducted in three scenarios based on dataset characteristics: RSRP (baseline), RRC, and NAS. We used accuracy, precision, recall, and F1-score as performance metrics.

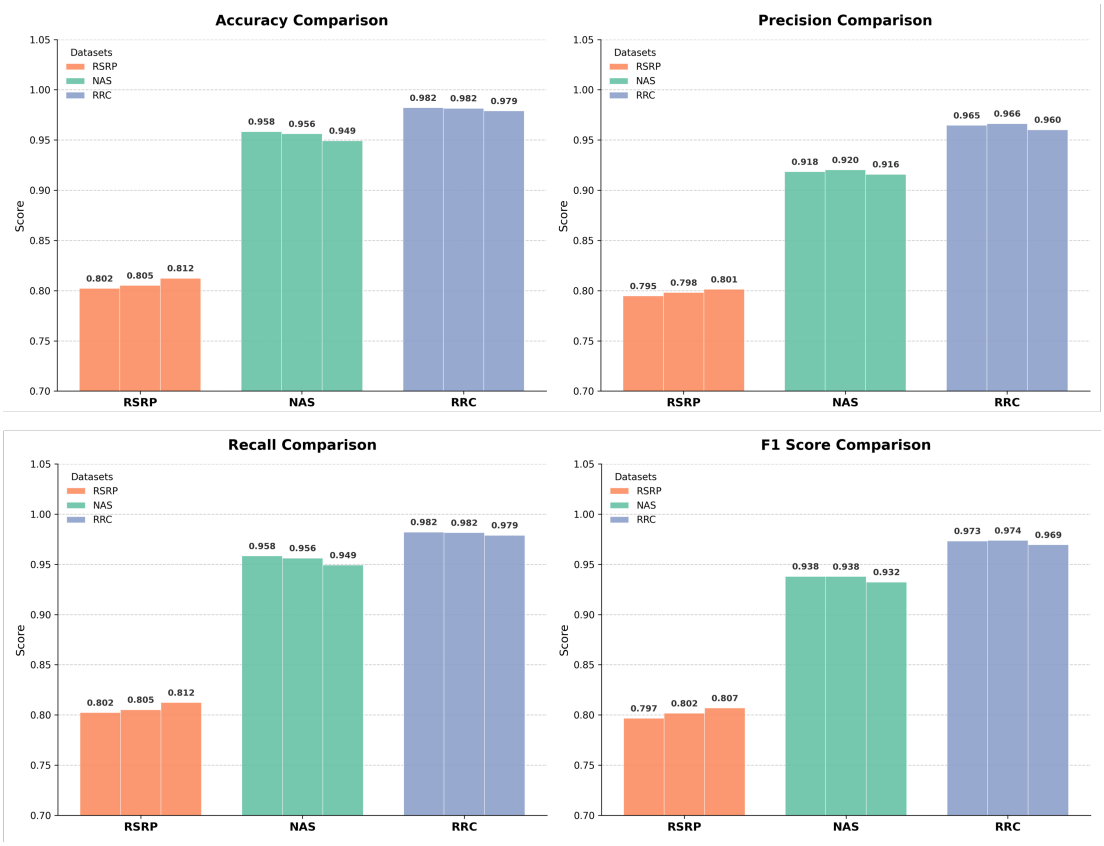

Figure 2: Performance comparison of detection models.

4.1 Performance Comparison by Dataset

The experimental results confirmed a significant gap in detection performance depending on the protocol layer information used for training.

4.1.1 Limitations of RSRP-based Detection (Layer 1)

The RSRP-only dataset, set as the baseline, recorded accuracies between 0.8022 and 0.8123, showing the lowest performance among the three datasets. Although the 1D-CNN model showed a relatively high value of 0.8123, the overall F1-score remained around 0.80. This empirically demonstrates that signal strength information from the physical layer (layer 1) alone is limited

in clearly distinguishing between normal signal fluctuations and intelligent FBS attacks. In other words, signal strength information is insufficient as a standalone detection indicator and serves only as an auxiliary tool.

4.1.2 Effectiveness of RRC and NAS Features

Models trained on the RRC dataset showed accuracies of approximately 0.949 to 0.958, proving a dramatic performance improvement compared to RSRP. This implies that the procedural information of RRC messages acts as a key feature for attack detection.

The most notable results appeared in the NAS dataset. The NAS dataset recorded the highest performance across all models. Specifically, the LSTM model achieved an accuracy of 0.9821 and an F1-score of 0.9732, showing the best performance in the experimental group. GRU and 1D-CNN also maintained high accuracies of 0.9815 and 0.9789, respectively. This strongly suggests that message sequences in the NAS layer (for example, identity request and authentication reject) contain the most certain and unique signatures for identifying FBS attacks.

4.2 Model Stability Analysis

In terms of model architecture, recurrent neural network (RNN)-based models, which have strengths in processing time-series data, showed excellent results.

- LSTM and GRU: These two models recorded accuracies of 0.9821 and 0.9815, respectively, on the NAS dataset, showing slightly better performance than 1D-CNN (0.9789). This is analyzed to be because the order and context of message occurrence are important due to the characteristics of the NAS protocol, and the gating mechanisms of LSTM and GRU effectively learned these sequential dependencies.
- 1D-CNN: 1D-CNN showed the highest accuracy (0.8123) on the RSRP dataset, confirming its strength in extracting local features or processing lower-dimensional data.

5 Conclusion

This study was initiated to counter fake base station attacks that threaten reliable mobile security environments. While many existing detection studies have relied on accessible layer 1 signal metrics like RSRP, this research sought to demonstrate that such approaches have fundamental limitations in defending against sophisticated MSAs.

To this end, we comparatively analyzed layer 1 RSRP data and layer 3 RRC/NAS protocol data, which requires root access, using the same machine-learning and deep-learning models. Experimental results showed that layer 1 data was effective in identifying the physical presence of an FBS, but could be easily neutralized if the attacker spoofs the signal strength to mimic a legitimate base station. Furthermore, it had fundamental limitations in detecting actual attack behaviors such as IMSI theft or downgrades. In contrast, layer 3 RRC/NAS data detected these behavior-based MSA signatures with high accuracy, providing empirical justification for the necessity of RRC/NAS-based data.

Furthermore, these findings hold significant implications for the e-business domain, where mobile connectivity serves as a critical infrastructure for digital transactions and user authentication. By securing the cellular layer against sophisticated surveillance and disruption, the proposed method enhances the integrity of mobile commerce platforms, ultimately fostering the user trust necessary for the sustainable expansion of e-business services.

6 Acknowledgement

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No.2024-00437252, Development of anti-sniffing technology for mobile communication and AirGap environments).

References

[1] CableLabs. False base station or imsi catcher: What you need to know, 2019.

[2] H. Park et al. Smdfbs: Specification-based misbehavior detection for false base stations. *Sensors*, 23(23):9504, 2023.

[3] 3gpp ts 33.501: Security architecture and procedures for 5g system (release 18). Technical report, 3rd Generation Partnership Project (3GPP), 2024.

[4] Kazi Samin Mubasshir, Imtiaz Karim, and Elisa Bertino. Gotta detect 'em all: Fake base station and multi-step attack detection in cellular networks. In *USENIX Security Symposium*, 2025. To appear; preprint available at arXiv:2401.04958.

[5] T. Tucker, N. Marlin, et al. Detecting imsi-catchers by characterizing identity-exposing behavior. In *NDSS Symposium*, 2025.

[6] Ericsson AB. 5g nr radio access architecture and security. *Ericsson Technology Review*, 2022.

[7] P. K. Nakarmi et al. Applying machine learning on rsrp-based features for false base station detection. *arXiv preprint*, 2022.

[8] P. K. Nakarmi. Murat: Multi-rat false base station detector. Technical report, arXiv / Technical Report, 2021.

[9] H. Wen, P. Porras, V. Yegneswaran, A. Gehani, and Z. Lin. 5g-spector: An o-ran compliant layer-3 cellular attack detection service. In *NDSS Symposium*, 2024.

[10] Purdue CS / POWDER and collaborators. Fake base station and multi-step attack detection in user-side traces (dataset and detection). arXiv / project page, 2025.

[11] Nr; radio resource control (rrc); protocol specification (3gpp ts 38.331). Technical report, 3rd Generation Partnership Project (3GPP), 2024. Release 17.4.0.

[12] Non-access-stratum (nas) protocol for evolved packet system (eps) (3gpp ts 24.301). Technical report, 3rd Generation Partnership Project (3GPP), 2024. Release 17.6.0.

[13] Open5GS Project. Open5GS: Open source 5g core and epc. https://open5gs.org/, 2019.

[14] srsRAN Project. srsRAN: 4g/5g software radio suite. https://www.srsran.com/, 2020.