SigDetect: Collaborative Endpoint-based Signal Injection Attack Detection based on Channel Frequency Response

Yingjing Wu
Kahlert School of Computing
University of Utah
Salt Lake City, Utah
wyj1993@cs.utah.edu

Dustin Maas
Kahlert School of Computing
University of Utah
Salt Lake City, Utah
dmaas@cs.utah.edu

Jacobus Van der Merwe
Kahlert School of Computing
University of Utah
Salt Lake City, Utah
kobus@cs.utah.edu

***Abstract*— Unencrypted broadcast data in cellular networks is vulnerable to signal injection attacks that are capable of revealing sensitive information and disrupting critical services. Existing detection methods struggle with such attacks, especially for attacks with low transmission power. This paper introduces `SigDetect`, a collaborative anomaly detection system that leverages complex channel frequency response (CFR) measurements and machine learning to detect signal injection attacks reliably. Extensive evaluations demonstrate that individual `SigDetect` detectors outperform a Received Signal Strength (RSS)-based method by 32.8% in outdoor experiments with stationary radios and 26.7% in indoor experiments where one of the radios is in motion. `SigDetect` also outperforms a CFR approach while eliminating the need to adjust thresholds to adapt to different wireless environments. Finally, `SigDetect`'s collaborative approach, in which neighboring network endpoints aid in detection, improves detection accuracy from 90.2% to 97.2% while reducing the false alarm from rate 11.2% to 0.7% and the missed detection rate from 8.3% to 4.9% in an indoor environment without mobile endpoints. These results suggest that `SigDetect` offers a promising solution for protecting cellular networks against low-power signal injection attacks.**

***Index Terms*—signal injection attack, anomaly detection, cellular network, channel frequency response**

I. Introduction

In Long-Term Evolution (LTE) and 5G networks, broadcast messages from base stations (BSs) to endpoints without authentication are vulnerable to attacks. Traditional spoofing attacks, like the well-known fake base station (FBS) attack, exploit this vulnerability by overpowering all of the transmissions from the legitimate base station, requiring significant sustained transmit powers and receiver sensitivity, or proximity to the target endpoints [27], [29]. This requirement and the accompanying large deltas in received power at the victim endpoints makes the attacks detectable [15], [18], [24]. A more recently proposed category of signal injection attacks, known as signal overshadowing attacks, including SigOver [38] and its variations, such as AdaptOver, LTrack and SigUnder [8], [16], [23], can launch a variety of attacks (e.g., denial of service (DoS), network downgrade, and coarse-grained tracking) using much lower sustained power. They accomplish this by only overshadowing specific broadcast subframes rather than overpowering most or all legitimate transmissions.

Security enhancements in 5G have mitigated signal injection attack vectors, such as those based on international mobile subscriber identity (IMSI) and system information blocks (SIB). By employing the subscription concealed identifier (SUCI), which encrypts the IMSI, and implementing "on-demand" transmission for SIBs, these measures can prevent the messages from being easily intercepted and misused. However, eradicating signal injection threats in cellular networks remains challenging; messages for the master information block (MIB), SIB1, and periodically broadcast SIBs remain susceptible to attacks. Besides, LTE systems are still in use, and a complete transition to more secure systems will take time. Moreover, the growing reliance on commercial networks in military applications [4], [26], [32] further highlights the need for more robust protections from targeted physical layer attacks.

Reliably detecting low-power signal injection attacks in cellular networks is challenging because their effects on physical layer measurements are harder to discern from those caused by natural channel fluctuations due to the dynamic nature of the cellular environment (e.g., the motion of objects in the area of the base station and endpoints, or the motion of the endpoints themselves). Traditional approaches may not effectively address this challenge. For instance, methods relying on variation in received signal strength (RSS) [6], [10], [37] may fall short in reliably distinguishing between normal fluctuations and actual attacks since typical RSS fluctuations (due to resource scheduling as well as well as channel variability) can far exceed those caused by low-power injection attacks. Indeed, we have measured RSS fluctuations of up to 7.5 dB in a commercial network. Also, the tight time and frequency synchronization necessary to successfully overshadow legitimate transmissions with signal injection attacks makes methods that rely on detecting carrier frequency

This work was supported by the NSF through the following grants: #2232463 and #2346555.

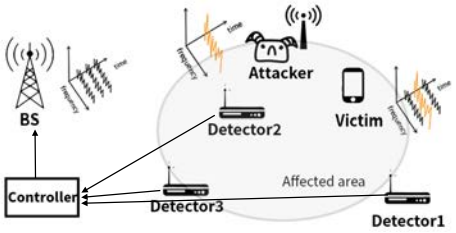


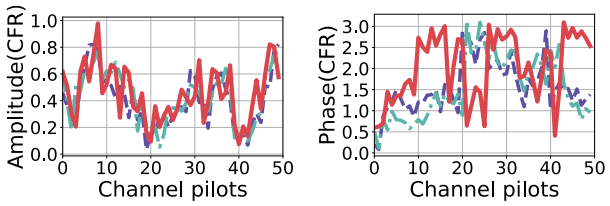
Fig. 1: SigDetect uses channel measurements made at one or more detectors to identify signal injection attacks.

offset (CFO) changes less effective [12], [35].

To address the threats of this new category of more surreptitious physical layer attacks, we propose SigDetect: a collaboration-based anomaly detection system that uses one or more endpoints attached to and/or monitoring the base station of the victim UE to detect overshadowing attacks (illustrated in Figure 1). The detectors use a one-class support vector machine (OCSVM) trained on a carefully curated set of features based on multiple channel characteristics derived from the complex CFR estimates used for equalization in cellular networks (rather than focusing on the perturbations of, e.g., CFR amplitude or CFO alone).

We imagine specialized endpoints for this scenario, e.g., military handsets designed for use on commercial networks, but with SigDetect built in, and potentially another radio access technology to use as a backhaul to raise the alarm. Of course, software-defined radio (SDR) -based detectors could be deployed in the vicinity of a commercial network to perform detection, perhaps attaching to the same network or also using some other out-of-band communications link for raising alarms (the approach we use for our evaluation). SigDetect relies on the nature of the wireless medium, i.e., that endpoints attached to (or in the coverage area of) the same network that does not fall victim to an injection attack may still be affected by it to a detectable degree. As such, these endpoints would be capable of raising the alarm. In addition, since the effects of an attack will range from subtle to significant depending on the relative positions of the attacker and a given endpoint (and the environment), leveraging multiple detectors improves the system's accuracy. Note that we focus solely on robust signal injection attack detection in this work and leave mitigation approaches and defenses for future work.

Compared to RSS-based methods, SigDetect achieves approximately 30% higher accuracy and at least 50% lower false negative rates in indoor scenarios with both stationary and mobile endpoints. SigDetect performs better than the most comparable CFR-based approach (to the best of our knowledge, nobody has employed CFR measurements for injection attack detection in cellular networks, but they have for similar attacks in WiFi networks, which also uses orthogonal frequency domain multiplexing (OFDM)) without the need for threshold or parameter adjustments across different wireless environments. Finally, We show that a collaborative decision-making strategy significantly improves detection accuracy in indoor scenarios with stationary endpoints, increasing average accuracy from 90.2% for a single detector, to 97.2% with multiple detectors.



(a) Outdoor CFR amplitude (b) Outdoor CFR phase

Fig. 2: The impact of SigOver on the amplitude and phase of CFR in outdoor scenario for an affected subframe (red) and two normal subframes (green and blue) ahead of it.

The rest of the paper is organized as follows. Section II presents some preliminaries and our motivations. Section IV describes the design of SigDetect. A series of experimental results and analysis are shown in Section V. We review the related work in Section VI.

II. Background and Motivation

A. LTE/5G Physical Layer Measurements

Modulation Scheme and Frame Structure. LTE and 5G use the orthogonal frequency-division multiple access (OFDMA) modulation scheme to schedule multiple users across time and frequency resources. In this paradigm, the fundamental unit of resources is one sub-carrier for one OFDM symbol duration or a single resource element (RE). Base stations (BSs) schedule downlink and uplink REs for endpoints (a.k.a, user equipment (UEs)) in a structured radio frame system, where each radio frame is 10 ms long and contains 10 subframes, each lasting 1 ms. Each subframe contains a number of OFDM symbols. In 5G, the OFDM symbol duration and sub-carrier spacing can vary, but these parameters are fixed in LTE, where each subframe contains 14 OFDM symbols. In both cases, resources are organized into several logical and physical channels serving various needs, e.g., synchronization, broadcasting system information, paging, transferring user-plane data, etc.

Channel Measurements and Equalization. The multipath nature of the wireless channel will cause a receiver to experience multiple time-shifted and attenuated versions of the transmitted signal. In order to compensate for such distortions, OFDM uses a subset of resource elements to send known pilot symbols that the receiver can use to estimate the effects of the channel and compensate for them in order to demodulate the other symbols correctly. The effects of the channel can be modeled as a channel impulse response (CIR) or its Fourier pair, the channel frequency response (CFR). The effects of the channel can be represented as $y = h * x + n$ or $Y = HX + N$, where y and its Fourier transform Y represents the received signal, x and its Fourier transform X represents the transmitted signal, h and its Fourier transform H represent the CIR and CFR, and n and its Fourier transform N represent noise. Since OFDM/OFDMA modulates symbols in the frequency domain, the CFR is estimated for each OFDM symbol and used to correct the effects of the channel.

B. Signal Overshadowing Attack

In order to successfully overshadow specific OFDMA resources, the attacker must first achieve accurate time and frequency synchronization with the legitimate BS. After that; the attacker crafts malicious subframes by altering the frequency domain data symbols in the original messages while preserving the pilot symbols used for channel estimation at the victim UE, then transmitting a time domain signal that precisely overlaps with the desired subframe. The precise synchronization and manipulation ensure that the victim can decode the altered messages correctly. The attacker then transmits the counterfeit subframe from a position and at a power level that ensures it will be at least 3 dB higher than the legitimate subframe at the victim UE. Figure 1 illustrates the attack in the presence of `SigDetect` detectors. If the attacker is well synchronized with the legitimate BS and knows the precise relative distances between itself, the BS, and the victim, it can further improve the time synchronization of the legitimate and malicious signals arriving at the victim UE. However, `SigDetect` can potentially leverage the fact that the time synchronization is not likely to be as good at the detectors in this circumstance because they occupy different relative positions to the BS and attacker.

C. Motivation

The need for a CFR-based detection method like SigDetect arises from the limitations of detectors relying solely on simpler channel metrics to combat signal injection attacks like SigOver. While RSSI measurements appear attractive for detector development, their fluctuations alone may not suffice for reliable detection. Our experiments using an Ettus B210 SDR and srsRAN across various scenarios found significant RSSI fluctuations, up to 7.5 dB between consecutive subframes. These pose challenges for reliable detection with RSSI alone, especially in the presence of more minor amplitude changes induced by SigOver attacks. Additionally, RSSI measurements are influenced by the allocation of radio resources in a subframe, which is particularly noticeable in cells with wider bandwidth.

Hence, a detector capable of leveraging more discriminative measurements is essential. Channel estimates offer a promising avenue due to their calculation for every subframe in LTE/5G and richer data source compared to RSSI. Furthermore, channel estimates are susceptible to attacks transmitting the same pilots from different locations due to the multipath nature of the channel, resulting in significant differences in the channels observed by receivers. While channel estimates have been used to detect spoofing attacks in other wireless networks, their effectiveness in detecting low-power injection attacks on LTE/5G networks remains largely unexplored.

Our work comprehensively explores the effectiveness of leveraging channel estimate perturbations to detect these attacks. By examining various channel response attributes and their reactions to signal injection attacks under different channel conditions, we aim to devise an effective and resilient detection mechanism. For instance, we highlight the potential of CFR phase estimates, particularly in scenarios where amplitude alone may not suffice for detection effectiveness. This exploration aims to address the challenges outlined by SigOver and establish a robust CFR-based detection framework.

III. Threat Model

The attacker and victims camp on the same BS and periodically monitor the same broadcast channels for, e.g., SIBs or paging messages. We assume an adversary capable of mounting SigOver attacks, i.e., one that is able to: (1) obtain and maintain accurate synchronization in time and frequency with the legitimate BS using a GPS disciplined oscillator (GPSDO) or another sufficiently accurate reference; (2) eavesdrop on legitimate unencrypted broadcast signals; and (3) craft and transmit malicious signals that precisely overlap specific subframes and arrive at the target UEs with enough power ($\geq$ 3 dB more than the legitimate signal for that period) to successfully "overshadow" the legitimate transmission. As we've already noted, an attacker may also have information about the location of the BS and the target that it can use more precisely time-align the arrival of the malicious and legitimate signals at the UE, but the attacker will not able perform this time alignment perfectly for all endpoints in the network simultaneously! In our evaluation, the BSs and the endpoints under attack are all single-input single-output (SISO) SDR-based systems that use software to implement the necessary LTE functions, but the methods generalize well to multiple-input multiple-output (MIMO) endpoints (we would expect detection performance to improve with the addition CFR estimates that come with MIMO BSs and endpoints).

IV. SigDetect Design

A. Overview

`SigDetect` uses one or more endpoints deployed within range of a cellular network to serve as detectors. These detectors monitor channel estimates and raise an alarm to a central control node when they detect a subframe with anomalous channel features. The control node aggregates detector outputs and decides an attack has occurred if enough of the detectors agree. This is illustrated in Figure 1.

Individual Detector Architecture. Figure 3 shows the architecture used at a single detector. The detector constantly generates CFR estimates for relevant subframes to demodulate downlink transmissions. A Feature Extraction module processes channel estimates from these subframes for each radio frame and extracts the desired features (described in Section IV-B) as a sample. The Classification module then applies an OCSVM model (described in Section IV-C) to the sample in order to determine if it is anomalous. An alert is raised to the central control node if an anomaly is detected. Otherwise, the normal sample is placed in a queue that accumulates samples for a predetermined time and uses them to update the model periodically. Note that we assume no attack occurs before we obtain the first model.

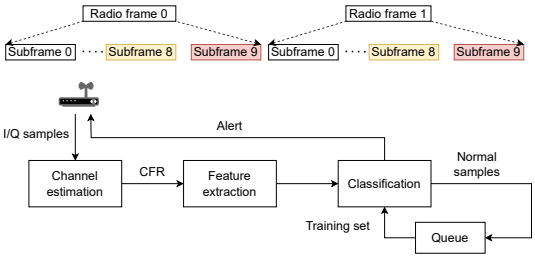


Fig. 3: SigDetect's Architecture

Central Control Node. The central control node aggregates the alerts from the original detectors and implements a decision-making policy in order to decide to take action. In its current design, it simply takes the majority vote to decide whether or not an attack was present for a given time period. Increasing its sophistication and capabilities is currently planned as future work, which we discuss in Section VII.

Deployment Challenges. In Figure 1, we illustrate the fact that some detectors may be far enough from the attacker that the perturbations to its CFR estimates are negligible. These effects may become particularly pronounced for detectors positioned much closer to the base station than they are to the attacker. Other channel effects like shadowing or fading will affect performance as well. These circumstances, coupled with the fact that we assume no precise knowledge of the attacker's location (other than it must be in the range of the base station and its targets), require consideration when deploying detectors. Of course, these challenges will exist for any non-SigDetect approach that relies on the perturbations of physical layer measurements caused by such attacks. However, in our outdoor evaluation of SigDetect, we deploy the attacker, detectors, and victim endpoints at considerable distances given the limited transmit power available, and the detectors are able to consistently identify attacks that lead to signal strength increases down to 0.5 dB, showing that they should perform well in a single cell sector provided they are placed in a way that offers reasonable coverage of the area of interest.

B. Feature Selection

Potential CFR Features. Motivated by the observations in the previous section and a study of previous CFR-based approaches in spoofing detection, we explore a large set of potential features for our detector. These are listed in Table I[1]. In the table, h_i, H_i, and A_i represent the CIR, CFR and the CFR amplitude of frame i respectively. The operation V is defined as: $V([x_1, x_2, x_3, ..., x_n]) = \sum_{i=2}^{n}(|x_i - x_{i-1}|)$. These features describe the statistical properties of CFR variation and include both direct and indirect indicators reflecting CFR changes. They fall into three groups: F_A, F_B, and F_C. The F_A group includes features from f_1 to f_7 that are related to the amplitude change of CFR in two contiguous subframes. F_B includes two features that are associated with performing IFFT operations on the CFR. F_C includes five features related to the phase difference of CFR estimates. Note that, to remove frequent phase flips that occur near $\pm\pi$, we flip the phase below 0 to the upper side for f_{10}, f_{11}, and f_{13}. The CFR estimates themselves represent a high-dimensional feature whose dimension is proportional to bandwidth. Since high-dimensional features increase computational complexity and may cause overfitting, we extract features from the CFR that reflect emerging path components composed of attack channels for dimensionality reduction.

[1]We merged three correlated features into f_2.

Combining Features for Improved Detection. The robust detection of signal injection attacks within cellular networks presents a significant challenge due to the inherent dynamism of channel characteristics. Relying solely on a single feature space proves insufficient to offer a clear separation between anomalies and normal samples. The distribution of normal samples often exhibits complexity beyond simple unimodal distribution, with significant portions exhibiting bimodal, trimodal (as observed in the distribution of f_6 in Figure 4), or even more intricate patterns. These dispersed samples constitute an average of 12.5% in our outdoor experiments (ranging from 5% to 18%) and can arise from dynamic environmental factors like moving reflectors. While attacks causing significant fluctuations beyond the natural range are readily identifiable, the worst-case scenario arises when minor attack disturbances lie within the natural fluctuation range, especially in areas with weak attacker signals. In these scenarios, where samples corresponding to an attack overlap with natural variations in single feature spaces (such as f_5 and f_6 in Figure 4), defining clear anomaly boundaries becomes a significant challenge. This leads us to explore multi-feature integration and select a non-linear classifier.

TABLE I: Features List

	Feature expression	Description
$\mathbf{f_1}$	$\overline{A_2} - \overline{A_1}$	The average CFR amplitude difference
f_2	$\|H_2 - H_1\|^2$	Squared ℓ_2 of CFR difference [21], [36]
	$\|h_2 - h_1\|^2$	Squared ℓ_2 of CIR difference [20], [25]
	$\overline{\|H_2 - H_1\|}$	The average amplitude of CFR difference
f_3	$\text{pcorr}(A_2, A_1)$	Pearson correlation of CFR amplitude [38]
f_4	$\text{scorr}(A_2, A_1)$	Spearman correlation of CFR amplitude
$\mathbf{f_5}$	$V(A_1) - V(A_2)$	Volatility difference of CFR amplitude
$\mathbf{f_6}$	$\|A_1 - A_2\|$	ℓ_2 of CFR amplitude difference [13]
f_7	$\|A_1\| - \|A_2\|$	ℓ_2 difference of CFR amplitude
f_8	$\max(\|h_2\|^2) - \max(\|h_1\|^2)$	Power difference of CIR main component
f_9	$\max(\|\text{IFFT}(H_2 - H_1)\|^2)$	Power of the main component of CIR
$\mathbf{f_{10}}$	$V(\angle(H_2 - H_1))$	Volatility of the phase of CFR difference
f_{11}	$V(\angle H_2) - V(\angle H_1)$	Volatility difference of CFR phase
f_{12}	$\text{scorr}(\angle H_2, \angle H_1)$	Spearman cross correlation of CFR phase
f_{13}	$\|\angle H_2 - \angle H_1\|$	ℓ_2 of the absolute CFR phase difference
f_{14}	$\|\angle H_2\| - \|\angle H_1\|$	ℓ_2 difference of the CFR phase

The feature set chosen (after the evaluation that follows) for SigDetect is denoted as c_1 and comprises of the features f_1, f_5, f_6, and f_{10}.

C. Classifier Selection

We require a classifier capable of distinguishing normal samples from anomalous ones without the need for training data from prior attacks since it is impractical to create a set of

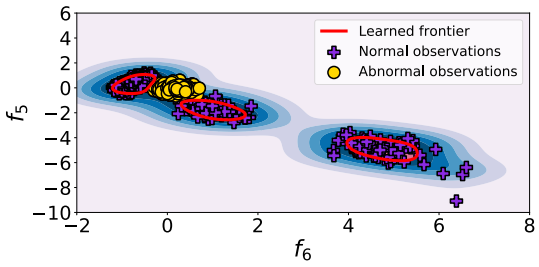


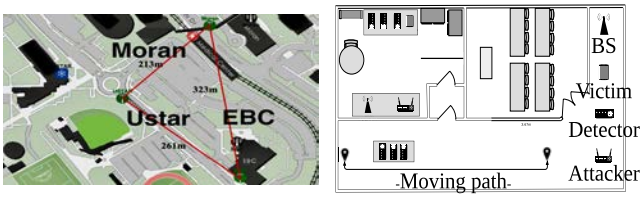
Fig. 4: Anomaly detection with OCSVM. The samples of f_5 and f_6 were gathered in a time window of 300 ms in our outdoor experiments on link MUE (see Section V) and used to train the frontier after normalization. Neither of these features alone make a good detector.

training data that covers the multitude of changing parameters that contribute to the effect an overshadowing attack may have on the features we derive from CFR estimates, e.g., positions of BSs and endpoints, the dispersiveness of the channels involved, etc. Furthermore, labeling data as anomalous simply because it was gathered during a period of attack may add noise to the model, since some detectors may not be impacted by the attack. For these reasons, we choose an unsupervised learning approach.

Why One-class SVM. Classification-based anomaly detection methods train a classifier on what are assumed to be "normal" samples and decide whether a new sample is anomalous if it is not similar enough to the training data. One-class support vector machines (OCSVM) and auto-encoders (AE) [11] are common one-class classification-based anomaly detection methods. Considering the feature space we have selected is not large, and we plan to keep the number of training samples small to more efficiently re-train the classifier, we choose OCSVM [28].

An alternative would be to use nearest-neighbor-based anomaly detection, which classifies new samples based on their distances to nearest neighbors, but this family of classifiers is known for high computational demands, as determining the nearest neighbors for a sample involves calculations involving the entire training dataset. Furthermore, KNN-based approaches are very sensitive to how distance is measured and the selection of 'k,' the number of neighbors to consider. These factors can significantly affect its performance and accuracy.

Hyperparameters. To address the challenge of nonlinear data separability in OCSVM, we choose a radial basis function (RBF) kernel, which projects the data points into a higher-dimensional space where the separation between classes is more likely, leading to improved classification performance. The kernel coefficient γ in an RBF kernel controls the spread of the Gaussian function—a larger value of γ results in a narrower spread of the Gaussian kernel. To achieve a tight boundary around the normal clusters, we set $\gamma = 0.9$. In terms of the parameter ν, it is an upper bound on the fraction of anomalies and a lower bound on the fraction of support vectors in the training set. It essentially controls the trade-off between the model's sensitivity to outliers and the decision



(a) Outdoor Deployment (b) Indoor Deployment

Fig. 5: Experiments Deployment

boundary's flexibility. However, it does not directly adjust the false alarm rate in the same way a threshold on a decision function or probability would. As we know, there are no outliers in the training set, and considering the training sample size, we set ν to be 0.01. Figure 4 is an example of how OCSVM generates boundaries for our outdoor data with these settings. The boundaries of the normal clusters are represented by the red circles.

Evaluation metrics. Our classifier can exhibit two types of error: (1) wrongfully identifying an attacked frame as normal (called false negative (FN)) or (2) wrongfully identifying a legitimate frame as anomalous (called false positive (FP)). As with any detector, there is a fundamental trade-off between the rates of these two errors (FNR, FPR), and the goal is to find a balance between them while minimizing both as much as possible. Since we leave that action taken after detection for future work, we try for a close balance.

V. Evaluation

A. Experiments Setup

*1) **Outdoor Setup**:* To assess the effectiveness of `SigDetect` in a typical cellular network setting, we performed outdoor experiments using three Ettus USRP B210 SDRs, referred to as Ustar, EBC, and Moran, as shown in Figure 5(a). Each of these units was connected to an Intel NUC compute node running Ubuntu 18.04. The Moran node is located atop a small hill, EBC on a rooftop, and Ustar is 20 ft above ground level. A train station and a construction site for a parking lot lie between Moran and the other two nodes, obstructing Moran's line of sight to them. Ustar and EBC have clear visibility to each other, with a road (and intermittent traffic) between them. The radios are separated by hundreds of meters (precise distances in the figure). These nodes are used in three distinct deployment arrangements, notated using their initials arranged in order of UE, attacker, and BS (MUE, UME, and MEU).

To examine the impact of the attack signal's arrival time on detection results, precise control over the signal's transmission timing is achieved using a White Rabbit (WR) synchronization system [3]. This system provides all three nodes with highly precise 10 MHz and pulse-per-second (PPS) reference signals, allowing for sub-nanosecond synchronization of transmissions and the ability to finely adjust the transmission (and reception) time of the attack signals. We used this capability to analyze the effect of timing discrepancies between the legitimate and attack signals on detector performance. We

note that an attacker or detector will not have this level of synchronization in a real-world scenario, and our detector does not leverage it to improve performance. It was used only to understand the effects of specific timing offsets in an effort to improve the detector. All outdoor experiments are conducted in LTE band 78 at a center frequency of 3460 MHz.

Data Collection. We collect CFR estimates and extract features for the attacked subframe (subframe 9) and the preceding subframe (subframe 8; no attack present) during every LTE frame. We gather a comprehensive dataset of normal and anomalous radio signals across a range of attack scenarios, including different attack transmission gains, cell bandwidths, and attack transmission timing offsets. We repeat this exercise for the three different deployment arrangements (MUE, UME, and MEU).

*2) **Indoor Setup**:* Figure 5(b) illustrates our indoor experiment setup. We use two USRP X310 SDRs [2], placed on a desk, with one functioning as the attacker and the other as the BS, both connected to an external Octoclock [1] for synchronization. (Again, the detector does not use this synchronization to improve performance. In this scenario, it is used to give the attacker a similar advantage that it might get from using a GPSDO in an outdoor scenario with a real BS, e.g., the ability to get tighter frequency synchronization to the BS than two undisciplined SDRs are capable of.) The attacker is positioned close to the BS to facilitate easy adjustments of its transmission power. Four B210 SDRs are set up on a table two meters away; the unit at the far right serves as the victim, while the others act as LOS detectors. Three additional B210s are placed on a cart outside the lab, obstructed by walls and doors, in order to act as NLOS detectors. All of these nodes are connected to Intel NUCs with an Ubuntu OS. In order to simulate mobility, we move the cart back and forth parallel to the wall, within the coverage areas of both the BS and the attacker. All indoor experiments are conducted in LTE band 78 at a downlink center frequency of 3580 MHz with 5 MHz bandwidth.

Data Collection Once again, we collect CFR estimates and extract features for subframes 8 and 9 of every LTE frame. Since the detectors in this experiment don't have access to the OctoClock references, we use LTE system frame numbers and subframe numbers to verify the time-aligned transmission/reception of normal and attacked subframes. We gather data for different attack transmission gains and with stationary or mobile endpoints, gathering around 4000 samples.

*3) **Bridging the Gap between Reality and Experiments**:* When conducting a preliminary analysis of our indoor experiments, we noticed that they exhibited RSSI fluctuations that diverged from those typically seen in commercial cellular networks. Specifically, unlike the considerable variation observed in real-world deployments, most samples in this part of the evaluation are more stable, showing variations below 1 dB, with only a few reaching around 2 dB. This is in contrast to an actual network where, with the same bandwidth, RSSI variations can surpass 8 dB, and the distribution of RSSI

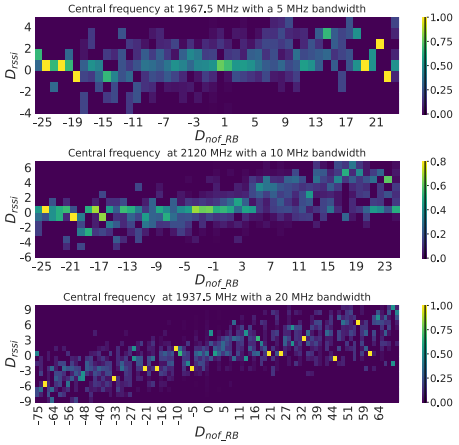


Fig. 6: The observed pattern in RSSI variation between the subframes targeted for attack (paging subframes) and the preceding subframes across commercial cells of varying bandwidths. Subframes with more RBs allocated for data transmission generally exhibit higher RSSI values, with a substantial difference in the payload sizes of consecutive subframes leading to larger absolute D_{rssi} values.

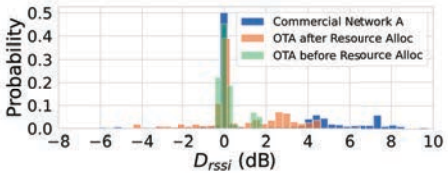


Fig. 7: Variations in RSSI data gathered from a commercial network compared to those obtained from our BS for the same UE.

generally exhibits more variance.

Understanding the Discrepancy. Given that RSSI computation includes the energy across the whole bandwidth, we hypothesized that cellular network RSSI changes depend not only on environmental factors but also due to payload changes - specifically, bursty changes in the number of occupied resource blocks (RBs) from one subframe to the next. (Resource blocks are an aggregation of the resource elements described in Section II-A). To verify this hypothesis, we designed experiments to investigate the potential relationship between RSSI changes (D_{rssi}) and the RB changes (D_{nof_RB}). We used a B210 SDR to measure RSSI across three different commercial networks at various central frequencies (2120 MHz for 10 MHz bandwidth, 1967.5 MHz for 5 MHz, and 1937.5 MHz for 20 MHz). We acquired the RBs allocation by monitoring downlink traffic and decoding the downlink control information with Falcon [9]. The heat maps in Figure 6 illustrate that as the variation in RBs allocated between two subframes grows, there is a corresponding rise in the variability of RSSI. The variance increases with higher bandwidths.

Solution. To replicate realistic RSSI fluctuations in a controlled testing environment, we monitored a commercial network cell (Network A, 1967.5 MHz, 5 MHz) using Falcon, recording 3 seconds of RSSI readings and RBs allocation data. During the experiments, we allocated the same amount of resource blocks in a cyclic manner using the scheduler_metric.cc file in an earlier version of srsRAN (prior

to release 19). Figure 7 demonstrates that allocating the radio resource following the pattern from a real network makes RSSI variations in our indoor experiments more closely match those of a real network.

B. Performance

1) ***Baseline Methods****:* CSITE [13] was used for source authentication in OFDM-based Wi-Fi systems. CSITE utilizes KNN on 'Time Gain Factor' (TGD), which is the Euclidean distance of the CFR amplitude (feature f_6) between two frames, adjusted by a "following coefficient", which increases the Euclidean distance for frames that are temporally distant. The average TGD between the frame and its k nearest neighbors determines a "Degree of Following" (DoF), and the classification is based on this DoF. Any sample with a DoF larger than the i percentile of the DoFs of the most recently accepted k frames in the slicing window is regarded as anomalous. According to the paper, we configure CSITE with a slicing window of 40 frames, the time gain factor of the "following coefficient" is set to one, and $k = 5$. And based on the fairness consideration (see Section V-B3), we set $i = 90$.

Moreover, our study includes a comparative analysis between `SigDetect` and an approach based on LOF, which utilizes feature f_1 and is denoted as LOF_AMP. The LOF algorithm's implementation utilized the scikit-learn library [5], leveraging its novelty detection features and adjusting the contamination parameter to 0.1. For indoor experiments, we also compare `SigDetect` to an RSSI-based approach. The confidence level of the threshold was set at the 90th percentile. We also show the performance for `SigDetect` using all the listed features (noted as `SigDetect`(c_2)). Except for CSITE, all the other methods were consistently implemented with a time window of 300 frames.

TABLE II: Performance Comparison of Single Detector Capabilities

Method	Metrics		
	Accuracy	FPR	FNR
Outdoor LOS&NLOS Stationary			
`SigDetect`	**88.7**	**9.5**	**13.1**
`SigDetect`(c_2)	82.7	32.7	1.8
CSITE	87.3	22.2	3.2
LOF_AMP	59.7	8.9	71.6
Indoor LOS&NLOS Stationary			
`SigDetect`	**90.2**	**11.2**	**8.3**
CSITE	86.3	22.3	5.1
LOF_AMP	79.9	10.5	29.5
RSSI	57.4	15.6	69.6
Indoor NLOS Moving			
`SigDetect`	**81.2**	**12.3**	**25.4**
CSITE	74.6	23.8	26.9
LOF_AMP	76.5	10.6	30.7
RSSI	54.5	14.4	76.6

2) ***Performance Comparison****:* We compare the performance of `SigDetect`'s detector with other detection approaches across three distinct scenarios: outdoor stationary, indoor stationary, and indoor NLOS moving. The performance evaluation aimed to minimize false alarms while maximizing the detection of anomalous frames. This evaluation criterion was chosen because a low miss detection rate does not guarantee that the classifier effectively distinguished anomalous samples since SigOver attack signals are typically low-power, potentially leading to overlapping distributions of normal and anomalous samples. Based on the criterion, `SigDetect` consistently outperforms the others across all scenarios, as indicated in Table II.

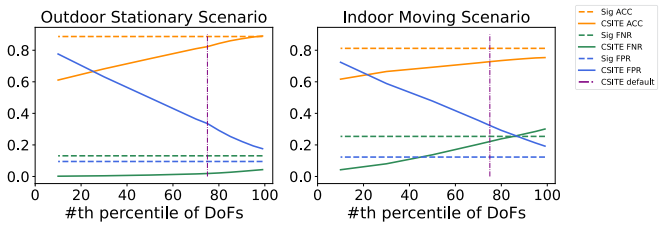


Fig. 8: Impact of Threshold Settings in CSITE on Accuracy, FPR, and FNR Relative to SigDetect. The area to the left of the purple line represents the threshold setting area recommended by CSITE, yet our data indicates an alternate optimal threshold zone.

In the outdoor scenario, the LOF_Amp method exhibits the worst performance, with an FPR of 8.9% but an FNR of 71.6%. The lower FPR results from the contamination factor, while the high FNR indicates that the method faces challenges in effectively identifying anomalies. This can be attributed to the method's limited capacity to differentiate between overlapping normal and abnormal samples. However, there is a significant improvement in its performance in indoor experiments. This enhancement results from averaging the performance across multiple detectors, some of which may detect stronger signals from the attack, thereby increasing the method's accuracy. On the other hand, `SigDetect`(c_2), despite achieving relatively high accuracy, incurs a substantial FPR of 32.7%. Including multiple features introduces irrelevant or noisy data that does not aid in distinguishing between normal and abnormal frames, potentially leading to confusion within the OCSVM model and imprecise boundary definition. CSITE faces similar challenges in outdoor and indoor stationary scenarios with high false alarms and low miss detection rates. We explore the reason in detail in Section V-B3. For moving scenarios, the efficacy of all the approaches generally decreases because the movement increases channel variability, complicating the distinguishing between normal and anomalous frames.

Run Time: Our implementation trains the model using 300 samples (corresponding to the time window size) on an i8-Core E5-2630 "Haswell" processor within 6ms. This facilitates real-time processing, ensuring the model is updated with new data and ready for the classification and collection of new samples in the next radio frame (10ms).

3) ***Robustness Consideration****:* From Table II, we notice a trend for CSITE in stationary scenarios: it maintains a low FNR but suffers from a high FPR. Further examination of the normal and abnormal samples processed by CSITE shows a

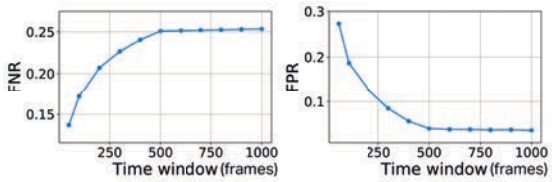


Fig. 9: Impact of Time Window

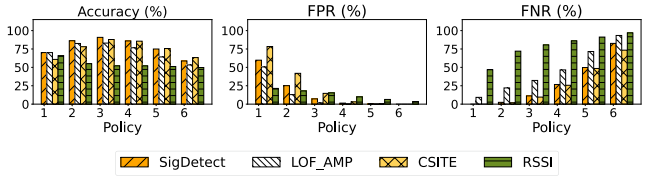


Fig. 10: Collaborative Performance in Indoor Stationary Scenario

clear separation between them, which explains the low FNR and also indicates that the high FPR is from an inappropriate threshold setting. CSITE determines its threshold by the ith percentile of DoFs from the most recently accepted k normal samples. According to the paper, the initial value of i is recommended to be 75, with adjustments to lower values contingent on channel instability. Our findings suggest that in stationary conditions, an increase in i can decrease the FPR without significantly impacting the FNR, as depicted in Figure 8. Besides, we find that increasing the value of k for threshold determination could further improve CSITE's detection efficacy.

While fine-tuning the threshold offers the potential for CSITE to outperform `SigDetect`, setting the optimal threshold necessitates intimate knowledge of the underlying data distribution, potentially limiting its generalizability. In the settings shown in Figure 8, increasing the 'i' value initially improves performance. Nevertheless, as the threshold increases, especially in the moving scenario, a clear tradeoff between FPR and FNR becomes evident. The tradeoff stems from the overlap of normal and abnormal samples. A faster movement can increase normal data variability, exacerbating this overlap. Consequently, selecting an appropriate threshold that fits all scenarios proves challenging. In contrast, `SigDetect` avoids the need for scenario-specific parameter adjustments and consistently delivers solid performance, showing its robustness.

*4) **Impact of Time Window Length**:* The curve in Figure 9 illustrates how the FPR and FNR change as the time window length varies. It can be observed from the curve that a short time window length results in a high FPR. This could be attributed to the inadequacy of the window length in capturing the changes in the channel. As the time window length extends, there appears to be a stabilization in both FPR and FNR. We chose a 300-sample window for balance between FPR and FNR.

*5) **Collaborative Performance**:* To understand the impact of employing multiple detectors on the detection, we evaluated indoor collaborative performance, considering both stationary and walking-speed movement scenarios. Six detectors were used in stationary conditions, including three LOS and three NLOS detectors. For moving scenarios, three moving

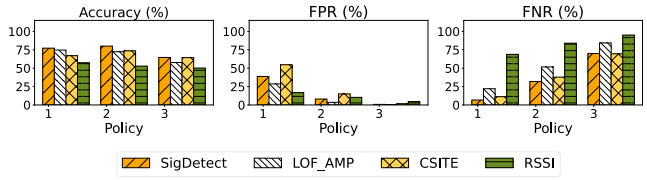


Fig. 11: Collaborative Performance in Indoor Moving Scenario

NLOS detectors were deployed. The decision-making policy for determining an attack is based on whether n detectors concurrently recognize it.

Figures 10 and 11 display the collaborative detection effectiveness under various policies. In both stationary and moving contexts, except for the RSSI method, the detection performance of other methods initially improves as the number of detectors (n) required to confirm an attack increases. However, performance declines as n continues to rise. In stationary conditions, policy 3 improves `SigDetect`'s detection accuracy from 90.2% for a single detector to 96.9% with an FPR of only 5.6% and an FNR of 0.5%. Policy 4 further enhances detection accuracy to 97.2%, with an FPR of just 0.7% and an FNR of 4.9%. Conversely, the performance of the RSSI-based method deteriorates as n increases. While the RSSI approach sees a benefit when any detector identifies an attack, its effectiveness diminishes if the decision-making requires the consensus of more detectors. We did not see an obvious benefit from the collaboration in the moving scenario. This could be attributed to the insufficient number of mobile detection devices we employed, coupled with the fact that the impact of attacks on different devices does not significantly vary in our mobile settings.

VI. Related Work

Numerous studies in the wireless communication field focus on detecting impersonation attacks from unauthorized transmitters using channel information during communication processes. RSS-based schemes have been fully studied in [6], [7], [10], [30], [37], [39], [40], having their specialty in network type, classification methods, detection and localization, and stationary and moving scenarios. Channel response-based methods have the potential to provide location-specific details about a communication channel compared to RSS-based techniques, which makes them valuable for identifying the genuine radio source either directly or indirectly through radio source localization. CIR-based detection approaches have been extensively studied in previous research, including works such as [19], [20], [25], [31]. Similarly, CFR-based schemes that use hypothesis testing and fingerprint techniques have also been explored in literature, such as [13], [14], [17], [33], [36]. Furthermore, channel response-based schemes have been investigated for different moving scenarios, as shown in [22]. In addition, there have been attempts to incorporate deep-learning and reinforcement-learning techniques in channel response-based attack detection algorithms. For instance, [34] provides a deep learning-based fingerprinting scheme for indoor localization. In [36], Q-learning was employed

to find the optimal threshold for spoofing detection in a dynamic environment by iteratively testing different values. The majority of the studies mentioned above do not specifically target cellular networks and are primarily focused on indoor environments. Moreover, they mainly concentrate on detecting spoofing attacks rather than signal-overshadowing attacks.

VII. Conclusion

Signal overshadowing attacks, which can be perpetrated at low cost with commodity hardware, represent a significant threat to the security of LTE/5G networks, especially with the growing reliance on commercial networks for mission-critical applications. However, methods for reliably detecting such attacks on cellular networks do not exist in the literature. This paper presents `SigDetect`, a machine learning-based collaborative solution capable of detecting low-power attacks like SigOver by monitoring channel response perturbations on endpoints camped on the same cell. We evaluate `SigDetect` against SigOver in multiple cellular scenarios (indoor and outdoor; with or without endpoint mobility), showing that it performs better than other common detection methods (RSSI and CFR-based) that have generally been applied to less surreptitious attacks.

While `SigDetect` performs well in detecting SigOver, there is always room for improvement. For example, the control node could offer feedback to the detectors to improve their performance in scenarios where they disagree, mitigate false alarms by checking in with the endpoints after a suspected attack, or potentially use the CFR estimates of all the detectors to estimate the direction the attack came from. We leave these improvements for future work.

References

[1] Ettus. octoclock. https://www.ettus.com/all-products/octoclock/.

[2] Ettus. usrp x300/x310 spec sheet. https://www.ettus.com/content/files/X300_X310_.

[3] White rabbit. https://ohwr.org/project/white-rabbit/.

[4] DOD Announces $600 Million for 5G Experimentation and Testing at Five Installations. https://www.defense.gov/News/Releases/Release/Article/2376743/dod-announces-600-million-for-5g-experimentation-and-testing-at-five-installati/, October 2020.

[5] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

[6] Yingying Chen, Wade Trappe, and Richard P Martin. Detecting and localizing wireless spoofing attacks. In *2007 4th Annual IEEE Communications Society Conference on sensor, mesh and ad hoc communications and networks*, pages 193–202. IEEE, 2007.

[7] Yingying Chen, Jie Yang, Wade Trappe, and Richard P Martin. Detecting and localizing identity-based attacks in wireless and sensor networks. *IEEE Transactions on Vehicular Technology*, 59(5):2418–2434, 2010.

[8] Simon Erni, Patrick Leu, Martin Kotuliak, Marc Röschlin, and Srdjan Čapkun. Adaptover: Adaptive overshadowing of lte signals. *arXiv preprint arXiv:2106.05039*, 2021.

[9] Robert Falkenberg and Christian Wietfeld. FALCON: An accurate real-time monitor for client-based mobile network data analytics. In *2019 IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, Hawaii, USA, December 2019. IEEE.

[10] Daniel B Faria and David R Cheriton. Detecting identity-based attacks in wireless networks using signalprints. In *Proceedings of the 5th ACM workshop on Wireless security*, pages 43–52, 2006.

[11] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[12] Weikun Hou, Xianbin Wang, Jean-Yves Chouinard, and Ahmed Refaey. Physical layer authentication for mobile systems with time-varying carrier frequency offsets. *IEEE Transactions on Communications*, 62(5):1658–1667, 2014.

[13] Zhiping Jiang, Jizhong Zhao, Xiang-Yang Li, Jinsong Han, and Wei Xi. Rejecting the attack: Source authentication for wi-fi management frames using csi information. In *2013 Proceedings IEEE INFOCOM*, pages 2544–2552. IEEE, 2013.

[14] Zhiping Jiang, Kun Zhao, Rui Li, Jizhong Zhao, and Junzhao Du. Phyalert: identity spoofing attack detection and prevention for a wireless edge network. *Journal of Cloud Computing*, 9(1):1–13, 2020.

[15] Leyli Karaçay, Zeki Bilgin, Ayşe Bilge Gündüz, Pinar Çomak, Emrah Tomur, Elif Ustundag Soykan, Utku Gülen, and Ferhat Karakoç. A network-based positioning method to locate false base stations. *IEEE Access*, 9:111368–111382, 2021.

[16] Martin Kotuliak, Simon Erni, Patrick Leu, Marc Roeschlin, and Srdjan Capkun. Ltrack: Stealthy tracking of mobile phones in lte. In *31st USENIX Security Symposium (USENIX 2022)*, 2022.

[17] Zang Li, Wenyuan Xu, Rob Miller, and Wade Trappe. Securing wireless systems via lower layer enforcements. In *Proceedings of the 5th ACM workshop on Wireless security*, pages 33–42, 2006.

[18] Zhenhua Li, Weiwei Wang, Christo Wilson, Jian Chen, Chen Qian, Taeho Jung, Lan Zhang, Kebin Liu, Xiangyang Li, and Yunhao Liu. Fbs-radar: Uncovering fake base stations at scale in the wild. In *NDSS*, 2017.

[19] Fiona Jiazi Liu, Xianbin Wang, and Serguei L Primak. A two dimensional quantization algorithm for cir-based physical layer authentication. In *2013 IEEE International Conference on Communications (ICC)*, pages 4724–4728. IEEE, 2013.

[20] Fiona Jiazi Liu, Xianbin Wang, and Helen Tang. Robust physical layer authentication using inherent properties of channel impulse response. In *2011-MILCOM 2011 Military Communications Conference*, pages 538–542. IEEE, 2011.

[21] Hongbo Liu, Yan Wang, Jian Liu, Jie Yang, and Yingying Chen. Practical user authentication leveraging channel state information (csi). In *Proceedings of the 9th ACM symposium on Information, computer and communications security*, pages 389–400, 2014.

[22] Hongbo Liu, Yan Wang, Jian Liu, Jie Yang, Yingying Chen, and H Vincent Poor. Authenticating users through fine-grained channel information. *IEEE Transactions on Mobile Computing*, 17(2):251–264, 2017.

[23] Norbert Ludant and Guevara Noubir. Sigunder: a stealthy 5g low power attack and defenses. In *Proceedings of the 14th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pages 250–260, 2021.

[24] Peter Ney, Ian Smith, Gabriel Cadamuro, and Tadayoshi Kohno. Seaglass: Enabling city-wide imsi-catcher detection. *Proc. Priv. Enhancing Technol.*, 2017(3):39, 2017.

[25] Neal Patwari and Sneha K Kasera. Robust location distinction using temporal link signatures. In *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, pages 111–122, 2007.

[26] "PAWR Project Office (PPO)". New $2.7M PAWR Project Funded by the U.S. Department of Defense Will Test AI-Driven Spectrum Sharing Optimization In a 5G-NR Network. https://advancedwireless.org/new-2-7m-pawr-project-funded-by-the-us-department-of-defense-will-test-ai-driven-spectrum-sharing-optimization-in-a-5g-nr-network/.

[27] David Rupprecht, Katharina Kohls, Thorsten Holz, and Christina Pöpper. Breaking lte on layer two. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1121–1136. IEEE, 2019.

[28] Bernhard Schölkopf, Robert C Williamson, Alex Smola, John Shawe-Taylor, and John Platt. Support vector method for novelty detection. *Advances in neural information processing systems*, 12, 1999.

[29] Altaf Shaik, Ravishankar Borgaonkar, Shinjo Park, and Jean-Pierre Seifert. On the impact of rogue base stations in 4g/lte self organizing networks. In *Proceedings of the 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, pages 75–86, 2018.

[30] Yong Sheng, Keren Tan, Guanling Chen, David Kotz, and Andrew Campbell. Detecting 802.11 mac layer spoofing using received signal

strength. In *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*, pages 1768–1776. IEEE, 2008.

[31] Jitendra K Tugnait and Hyosung Kim. A channel-based hypothesis testing approach to enhance user authentication in wireless networks. In *2010 Second International Conference on COMmunication Systems and NETworks (COMSNETS 2010)*, pages 1–9. IEEE, 2010.

[32] United States Department of Defense. Department of Defense (DoD) 5G Strategy (U). https://www.cto.mil/wp-content/uploads/2020/05/DoD_5G_Strategy_May_2020.pdf, May 2020.

[33] Ning Wang, Weiwei Li, Ting Jiang, and Shichao Lv. Physical layer spoofing detection based on sparse signal processing and fuzzy recognition. *IET Signal Processing*, 11(5):640–646, 2017.

[34] Xuyu Wang, Lingjun Gao, Shiwen Mao, and Santosh Pandey. Deepfi: Deep learning for indoor fingerprinting using channel state information. In *2015 IEEE wireless communications and networking conference (WCNC)*, pages 1666–1671. IEEE, 2015.

[35] Jianliang Wu, Yuhong Nan, Vireshwar Kumar, Mathias Payer, and Dongyan Xu. {BlueShield}: Detecting spoofing attacks in bluetooth low energy networks. In *23rd International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2020)*, pages 397–411, 2020.

[36] Liang Xiao, Yan Li, Guoan Han, Guolong Liu, and Weihua Zhuang. Phy-layer spoofing detection with reinforcement learning in wireless networks. *IEEE Transactions on Vehicular Technology*, 65(12):10037–10047, 2016.

[37] Wenqing Yan, Sam Hylamia, Thiemo Voigt, and Christian Rohner. Phy-ids: A physical-layer spoofing attack detection system for wearable devices. In *Proceedings of the 6th ACM Workshop on Wearable Systems and Applications*, pages 1–6, 2020.

[38] Hojoon Yang, Sangwook Bae, Mincheol Son, Hongil Kim, Song Min Kim, and Yongdae Kim. Hiding in plain signal: Physical signal overshadowing attack on {LTE}. In *28th {USENIX} Security Symposium ({USENIX} Security 19)*, pages 55–72, 2019.

[39] Jie Yang, Yingying Chen, and Wade Trappe. Detecting spoofing attacks in mobile wireless environments. In *2009 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pages 1–9. IEEE, 2009.

[40] Jie Yang, Yingying Chen, Wade Trappe, and Jerry Cheng. Detection and localization of multiple spoofing attackers in wireless networks. *IEEE Transactions on Parallel and Distributed systems*, 24(1):44–58, 2012.