Physical Layer-Based Device Fingerprinting For Wireless Security: From Theory To Practice

Junqing Zhang, *Senior Member, IEEE*, Francesco Ardizzon, *Member, IEEE*,
Mattia Piana, *Graduate Student Member, IEEE*, Guanxiong Shen, and Stefano Tomasin, *Senior Member, IEEE*

***Abstract*—The identification of the devices from which a message is received is part of security mechanisms to ensure authentication in wireless communications. Conventional authentication approaches are cryptography-based, which, however, are usually computationally expensive and not adequate in the Internet of Things (IoT), where devices tend to be low-cost and with limited resources. This paper provides a comprehensive survey of physical layer-based device fingerprinting, which is an emerging device authentication for wireless security. In particular, this article focuses on hardware impairment-based identity authentication and channel features-based authentication. They are passive techniques that are readily applicable to legacy IoT devices. Their intrinsic hardware and channel features, algorithm design methodologies, application scenarios, and key research questions are extensively reviewed here. The remaining research challenges are discussed, and future work is suggested that can further enhance the physical layer-based device fingerprinting.**

***Index Terms*—Channel state information, deep learning, device authentication, hardware impairments, Internet of Things, machine learning, radio frequency fingerprint, and wireless security.**

I. Introduction

The Internet of Things (IoT) is expected to significantly impact our lifestyles. According to IoT Analytics, the number of connected devices reached to 18.8 billion in 2024, an increase of 13% from 2023 [1]. These massively connected IoT devices have transformed our everyday lives with exciting applications such as smart homes, smart cities, connected healthcare, industry 4.0, etc. Wireless communications are preferred to connect these devices seamlessly. There have been many techniques for IoT, including WiFi (IEEE 802.11), ZigBee (IEEE 802.15.4), long range (LoRa), Bluetooth low energy (BLE), and narrowband IoT (NB-IoT), to name but a few [2].

This revolution requires security at all levels. Security is quite a broad topic, involving confidentiality, integrity, availability, authentication, etc. [3], [4]. This article will focus on device authentication, which is the first important step for network security. The receiver verifies the legitimacy of the received signal by checking specific features in the same signal. Our current computer and communications networks are protected by cryptography-based approaches, including both symmetric encryption, such as advanced encryption standard (AES), and public-key cryptography (PKC) such as Rivest-Shamir-Adleman (RSA). In particular, authentication is performed using a cryptographic challenge-response protocol based on symmetric encryption or PKC.

However, cryptographic solutions may not be applicable to IoT devices. Symmetric encryption requires a key pre-shared, whose refresh turns to be challenging for IoT [5]. PKC requires computationally expensive algorithms, which often have severe power and computational limitations [3], hence they are unsuitable for IoT devices. In addition, on the eve of quantum computing, PKC may be compromised due to the exponential increase in the computational power of attackers [6]. Due to the above limitations, there is a lack of competent IoT security solutions, and there have been many notorious security threats to IoT devices [4].

This background is driving the development of lightweight, yet secure technologies for the IoT. Regarding device authentication, the two most promising non-cryptographic approaches are physical layer-based device fingerprinting [7], which includes hardware impairments-based radio frequency fingerprint identification (RFFI) [8] and channel-based authentication [9]. In detail,

- RFFI uses unique hardware impairments as the device identifier. Due to the imperfect manufacturing process, the nominal values of hardware components slightly deviate from their specification. These hardware impairments are unique and stable, which can be exploited as device fingerprints.

Manuscript received xxx; revised xxx; accepted xxx. Date of publication xxx; date of current version xxx. The work of J. Zhang was supported in part by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant ID EP/Y037197/1 and in part by the UK Royal Society Research Grants RGS\R1\231435. The work of J. Zhang and S. Tomasin was also supported by the EU COST Action CA22168 - Physical layer security for trustworthy and resilient 6G systems (6G-PHYSEC). The work of M. Piana was funded by the European Commission through the Horizon Europe/JU SNS project ROBUST-6G (Grant Agreement no. 101139068). The work of G. Shen was supported in part by the National Natural Science Foundation of China under Grant 62401138. The work of S. Tomasin was supported by the project ISP5G+ (CUP D33C22001300002), which is part of the SERICS program (PE00000014) under the NRRP MUR program funded by the EU-NGEU. The review of this paper was coordinated by xxx. *(Corresponding author: Junqing Zhang.)*

J. Zhang is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GJ, United Kingdom. (email: junqing.zhang@liverpool.ac.uk)

F. Ardizzon, M. Piana, and S. Tomasin are with the Department of Information Engineering, University of Padova, Padova, Italy. (emails: francesco.ardizzon@unipd.it; mattia.piana@phd.unipd.it; stefano.tomasin@unipd.it)

G. Shen is with the School of Cyber Science and Engineering, Southeast University, China. (email: gxshen@seu.edu.cn)

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier xxx

- Channel-based authentication exploits the channel characteristics through which the signal propagates to identify the source (or, better, its location) at the receiver, taking advantage of the fact that signals transmitted by devices at different locations travel through different channels (i.e., different delays and attenuations for each path). Thus, the propagation environment, rather than the transmitting device characteristics, and the relative position between transmitter and receiver, guarantee the authenticity of the transmitter.

A. Existing Surveys

Here we provide a review of the recent tutorials and survey papers published on similar topics [10]–[19].

1) Existing Surveys on Device Fingerprinting: Paper [10] provides a tutorial on fingerprinting at different layers including physical, MAC, and network layers. We will focus on the physical layer techniques and significantly extend [10] by summarizing the recent advances in the area as deep learning has brought several exciting improvements.

Paper [11] focuses on device behavior fingerprinting, which is related to not only communication networks-based fingerprints but also in-device fingerprints, e.g., resource usage, software signatures, etc. Moreover, it is not solely concerned with security issues; it also encompasses a significant amount of fault detection content. The authors only briefly introduce the availability of the physical layer device fingerprinting technique but do not provide sufficient details on the latest studies and state-of-the-art schemes.

Survey [12] examined device fingerprinting techniques for resource-constrained IoT applications. While traffic and impairment-based approaches were considered, the survey did not include wireless channel-based methods. From the perspective of identification algorithms, although the deep learning techniques were mentioned, they were limited to a conceptual overview, with insufficient in-depth profiling of state-of-the-art deep learning-based fingerprinting algorithms.

The work in [13] surveyed numerous available device fingerprints, which span the entire cyber-physical system and encompass various characteristics, including thermal, optical, chemical, magnetic, and electrical aspects. However, it only briefly introduced the physical layer device fingerprinting problem and did not discuss the latest technologies. Furthermore, the authors focused on feature selection, while the introduction to the latest authentication algorithms is missing.

2) Existing Surveys on Physical Layer Security and Authentication: Surveys [14], [15] provide comprehensive coverage of physical layer authentication (PLA) techniques, with both passive and active approaches. Our survey will focus on the passive approaches as they can be readily applied to our pervasive IoT devices. Additionally, while the authors already considered the use of machine learning (ML) techniques, the coverage of the literature on ML solutions for device authentication is only partial, as the use of ML has become popular only in recent years.

When looking at physical layer solutions, many techniques require models of specific channels they work on. Existing surveys, such as [16]–[18] cover physical layer authentication techniques tailored for specific application domains. In particular, [16] considers device fingerprinting for global navigation satellite systems (GNSS) antispoofing. Both crypto and physical layer solutions are considered in [17], but only for satellite Internet. Illi *et al.* focus instead on physical layer security solutions and the IoT [18].

Finally, the survey [19] reviews both physical layer authentication and secure transmission, and it mainly focuses on channel-based authentication. We will delve into device fingerprinting by covering both hardware impairments-based and channel-based approaches.

3) Summary: A common shortfall in all existing papers is the absence (or very limited coverage) of experimental results and their derivation, which are crucial for assessing the merits and fostering the implementation of new security approaches. Several new techniques have appeared in recent years that are not covered by those surveys, e.g., generative AI for authentication, reconfigurable wireless environments, e.g., with reflective intelligent surfaces (RISs) and drones for challenge-response authentication at the physical layer, etc. Lastly, fingerprinting and authentication have been investigated in several domains, including different frequency bands and applications (IoT, mobile sixth generation (6G), WiFi, ...) for radio transmissions, but also in underwater acoustic communications (UWAC). An extensive survey of such domains and their peculiarities for fingerprinting/authentication is missing.

B. Survey Aims

As summarized in Table I, this paper complements and extends the published surveys with a comprehensive review of the physical layer-based fingerprinting for wireless security. We will review the design principles of both RFFI and channel-based authentication. We will also compare these two approaches and discuss their integration for more secure authentication mechanisms. Among the most promising and recent advances in these areas, we mention the availability of new technologies (such as RIS), the use of new transmission bands that fostered related technologies such as integrated communication and sensing, the experimentation (thus with higher technology readiness level) of physical-layer security mechanisms, and the use of ML techniques to secure transmissions by merging information coming from different communication layers. As unique features of our survey paper, we cover topics from theoretical development to practical implementation and share our experiences and insights on the design considerations of practical implementation. Thus, while looking at a specific domain, it will still provide a general framework to discuss solutions across different domains.

C. Survey Structure

Section II gives an overview of physical layer-based device fingerprint, which is further categorized into two techniques. The rest of the survey is comprised of three parts. The first part will cover the first technique, which is hardware impairments-based authentication, i.e., RFFI. The second part will describe channel-based authentication.

TABLE I
COMPARISON WITH EXISTING SURVEYS. ✗, ❍, AND ✓MEAN THE TOPIC IS NOT COVERED, PARTIALLY COVERED, AND EXTENSIVELY COVERED.

Ref	Year	ML	Exp.	Domains	New Tech.
[10]	2015	✗	✗	Wireless Networks	✗
[11]	2021	✓	✗	IoT	✗
[12]	2022	✓	✗	IoT	✗
[13]	2023	✓	✗	Cyber-Physical System	✗
[14]	2020	❍	✗	Wireless Networks	✗
[15]	2020	❍	✗	Wireless Networks	✗
[16]	2021	✓	✓	GNSS	✗
[17]	2023	✓	✗	Satellite Internet	✗
[18]	2024	✓	✗	IoT	✗
[19]	2024	✓	✗	Wireless Networks	✗
This	2025	✓	✓	IoT/6G/UWANs	✓

The first part is on RFFI and spans Sections III to VI. In particular, Section III presents the RFFI tasks, while Section IV models the hardware impairments for both transmitter and receiver. The algorithm design for deep learning-based RFFI is explained in Section V. For the practical implementation of RFFI, Section VI describes the key research topics, publicly available datasets, and the investigated scenarios. Section VII explains the experimental methodologies for RFFI.

The second part is on channel-based authentication and spans Sections VIII to XI. In particular, Section VIII introduces the definition and the approaches used for channel-based (CB)-PLA and Section IX is devoted to an overview of the channel features exploited for CB-PLA. An in-depth delve into the methodologies used for CB-authentication, including both statistical and ML approaches, is provided in Section X. Lastly, Section XI provides an overview of CB-authentication datasets publicly available and existing applications.

The third part provides an overview of challenges and future research activities discussed in Section XII. The main conclusions are reported in Section XIII.

The abbreviations used in this paper can be found in Table II.

II. DEVICE FINGERPRINTING AT THE PHYSICAL LAYER

The authentication on the basis of the signals exchanged at the physical layer comprises security mechanisms that can be classified as hardware fingerprinting or CB authentication techniques, which provide lightweight security mechanisms particularly useful in the IoT. As shown in Fig. 1, we will consider a system involving K transmitting IoT devices and a receiver. The IoT transmitter sends packets, which are captured by the receiver. Based on the received signals, the receiver aims to authenticate the transmitter based on its intrinsic hardware impairments and random channel features.

1) Transmitter: For each transmitter, the modulated signal, $x(t)$, passes to the transmitter chain, including the mixer, oscillator, and power amplifier [8], [20]. These hardware components are not perfect due to the variation in the manufacturing process, and their specifications deviate slightly from their nominal values. Their effects are collectively represented by $\mathcal{F}(\cdot)$. The radio frequency (RF) signal at the transmitter becomes $s(t) = \mathcal{F}(x(t))$.

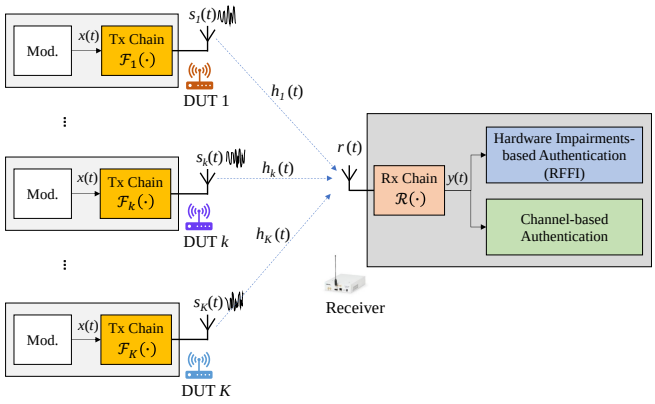


Fig. 1. Physical layer-based device fingerprinting system overview.

2) Channel: The RF signal is emitted into the wireless channel, which experiences multipath channel effects, denoted as $h(\tau, t)$, resulting in the received signal as

$$r(t) = h(\tau, t) * \mathcal{F}(x(t)), \tag{1}$$

where $*$ denotes the convolution operation. Note that many IoT devices are mobile; hence the channel impulse response $h(\tau, t)$ is time-varying.

3) Receiver: The receiver captures the received signal $r(t)$, which is passed to the receiver chain, including the mixer and oscillator too [8]. The receiver hardware components are not perfect either, and their effects are represented by $\mathcal{R}(\cdot)$. Considering all the above processes, the received signal $y(t)$ can be mathematically written as

$$y(t) = \mathcal{R}\Big(h(\tau, t) * \mathcal{F}_k(x(t))\Big) + n(t), \tag{2}$$

where $n(t)$ is the additive white Gaussian noise (AWGN).

A. Device Fingerprinting

As can observed from (2), the received signal, $y(t)$, involves both the hardware impairments and channel features, which can be exploited for device authentication.

1) Hardware Impairments-Based Authentication: Due to the manufacturing process, the hardware components are not perfect. Hence, hardware components are subject to impairments, such as mixer imbalance, oscillator imperfection, and power amplifier non-linearities [8], [20]. These impairments are minute and do not affect the communication functionalities because they can be compensated for by the receiver. These features are unique and can be used as device identifiers. RFFI protocols extract the hardware impairments embedded in the signal and infer its corresponding device identity.

2) Channel Based Authentication: The channel over which the transmitted signal travels is characterized by reflections, scattering, attenuations, as well as angular / time / Doppler features, and the position of the transmitter and the receiver. CB authentication uses this information from the channel to identify the sender of the message (and the channel over which the signal is going). A basic assumption is that devices are slowly moving and the environment is slowly changing; thus,

TABLE II
LIST OF ABBREVIATIONS

Abbreviation	Definition	Abbreviation	Definition
6G	sixth generation	LRT	likelihood ratio test
AE	autoencoder	LSTM	long-short term memory
AML	adversarial machine learning	LT	likelihood test
AoA	angle-of-arrival	LTE	long-term evolution
AoD	angle-of-departure	MD	misdetection
AWGN	additive white Gaussian noise	MIMO	multiple-input multiple-output
BAAE	Bahdanau attention autoencoder	ML	machine learning
BLE	Bluetooth low energy	NLOS	non-line-of-sight
CB	channel-based	NN	neural network
CFO	carrier frequency offset	OC-SVM	one-class support vector machine
CFR	channel frequency response	OFDM	orthogonal frequency-division multiplexing
CIR	channel impulse response	pdf	probability density function
CNN	convolutional neural network	PDP	power-delay profile
COTS	commercial off-the-shelf	PLA	physical layer authentication
CR	challenge-response	QuaDRiGa	quasideterministic radio channel generator
CSI	channel state information	RFF	radio-frequency fingerprint
DRL	deep reinforcement learning	RFFI	radio frequency fingerprint identification
DT	decision tree	RIS	reflective intelligent surface
DUT	devices under test	RL	reinforcement learning
EL	enseamble learning	RMS	root-mean square
FA	false alarm	RNN	recurrent neural network
FFT	fast Fourier transform	RSS	received signal strength
GAN	generative adversarial network	SCM	normalized sample covariance matrix
GLRT	generalized likelihood-ratio test	SDR	software-defined radio
GNN	graph NN	SNR	signal-to-noise ratio
GNSS	global navigation satellite systems	SVM	support vector machine
GPR	Gaussian process regression	TDOA	time difference of arrival
GPS	global positioning system	TLE	two-line element
GRU	gated recurrent unit	TOA	time of arrival
IoT	Internet of Things	USRP	universal software radio peripheral
KF	Kalman filter	UWAC	underwater acoustic communications
KNN	K-nearest neighbors	UWB	Ultra-Wideband
LEO	low Earth orbit	V2V	vehicle to vehicle
LLM	large language models	VAE	variational autoencoder
LoRa	long range	VANET	vehicular ad-hoc network
LOS	line-of-sight	VLC	visible-light communications

the authentication mechanism checks if different transmissions experience similar propagation channels. Other approaches are also discussed in the following, where the channel can change fast, but its consistent evolution over time provides the authentication feature.

III. RADIO FREQUENCY FINGERPRINT IDENTIFICATION

Deep learning has transformed many areas thanks to its powerful automatic feature extraction capability, which has also significantly enhanced RFFI. To the best knowledge of the authors, the work in [21] is the first paper applying deep learning to RFFI. Specifically, convolutional neural network (CNN) and multilayer perceptron (MLP) are used to classify LoRa devices. After that, deep learning has attracted massive interest in the RFFI area. Many deep learning approaches, such as CNN [22]–[24], recurrent neural network (RNN) including long-short term memory (LSTM) [22], [24] and gated recurrent unit (GRU) [24], Transformer [24], etc, have demonstrated significant impact, which can alleviate the difficulties of manual feature engineering.

Depending on whether there are rogue devices involved, RFFI can be categorized into closed-set classification, open-set recognition, and anomaly detection [25], whose implementations are illustrated in Fig. 2. A deep learning-based RFFI protocol involves two stages, namely training and inference. A deep learning model will be trained using a training dataset, $\mathcal{D}_{\text{train}}$, and the trained deep learning model will be used for inference in the second stage.

A. Closed-Set RFFI Classification

As shown in Fig. 2(a), there are K legitimate transmitters, a.k.a. devices under test (DUT), to be identified, and no rogue device is considered in the closed-set RFFI classification. The devices in the training and inference stages remain the same, hence the name "closed-set" comes from. The approach will predict the identity of the DUT.

Close-set RFFI classification is probably the most studied scenario in RFFI, which is a multi-class classification problem. Hence, deep learning is perfect for such tasks. A training dataset, $\mathcal{D}_{\text{train}} = \{(y_i, \ell_i)\}_{i=1}^{NK}$, will be constructed, where ℓ_i is the device label of the collected i-th packet and N is the number of packets collected for each DUT. The number of packets from each DUT should be kept the same, to ensure

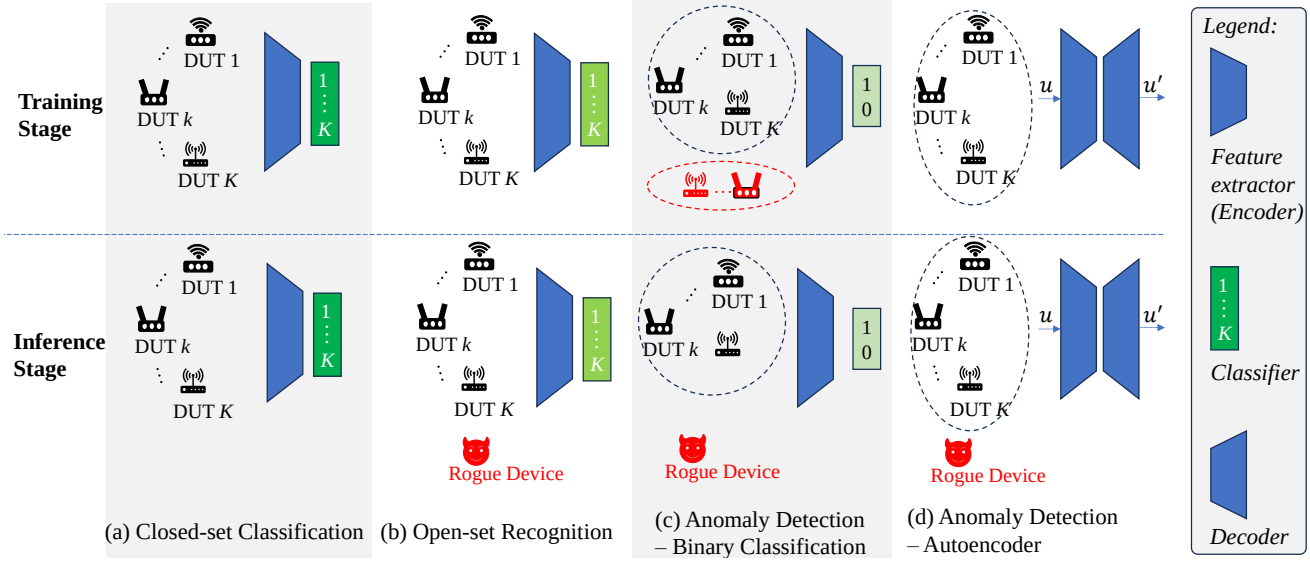


Fig. 2. Deep learning-based RFFI tasks.

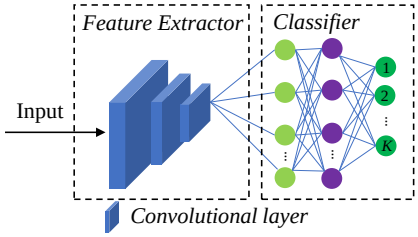


Fig. 3. Illustration of a CNN architecture.

a balanced dataset. A deep learning model can be partitioned into a feature extractor and a classifier. A CNN architecture is given as an example in Fig. 3. The feature extraction includes convolutional layers and pooling layers. The classifier is composed of a few fully connected layers, and the last layer has K neurons corresponding to K classes.

In the training stage, the parameters Θ of the deep learning model f are optimized as

$$\Theta = \arg\min_{\Theta} \sum_{(y_i,\ell_i)\in\mathcal{D}_{\text{train}}} \mathcal{L}(f(y_i;\Theta),\ell_i), \tag{3}$$

where $\mathcal{L}(\cdot)$ is the loss function, e.g., the cross-entropy loss.

In the inference stage, the softmax is used as the activation function, then the last fully connected layer of the classifier will return a list of probabilities $\boldsymbol{p} = (p_1, p_2, ..., p_K)$, where p_k represents the probability of the k-th device, given as

$$p_k = \frac{e^{z_k}}{\sum_{i=1}^{K} e^{z_i}}, \tag{4}$$

where $\boldsymbol{z} = (z_1, z_2, ..., z_K)$ is the output of the layer before the softmax activation. The final prediction is obtained by locating the maximum probability, given as

$$\hat{\ell} = \arg\max_{k}(\boldsymbol{p}). \tag{5}$$

B. Open-Set Recognition

Under closed-set classification, rogue devices will be classified as the legitimate DUT with the closest features, which is not acceptable as attackers will be admitted. Therefore, open-set recognition is required.

In open-set recognition, there are K legitimate DUTs and rogue devices, as illustrated in Fig. 2.b. Because rogue devices do not appear in the training stage, it is named "open-set". We need to first detect whether the DUT is legitimate or rogue, then further classify the index for legitimate DUTs. Open-set recognition can be addressed by the deep learning-based approaches with an adjustment to the softmax activation function. Hence, the training and inference stages will be similar to the closed-set classification. The training dataset can be constructed in the same way as the closed-set classification.

Gritsenko *et al.* leveraged the output probabilities of softmax given in (4) for open-set recognition [26]. Specifically, when the signal is from an unseen rogue device, the confidence level of the neural network prediction is low, hence, the output probability will be smaller than a pre-calculated threshold. In contrast, when it is from a legitimate DUT, the neural network can predict as it does in the closed-set classification.

Hanna *et al.* adopted a new activation function, the OpenMax [25]. The activation vector z prior to softmax is extended to $K+1$ outputs, given as

$$z'_k = \begin{cases} z_k\omega_k, & k \in \{1, ..., K\} \\ \sum_{k=1}^{K} z_k(1-\omega_k), & k = K+1 \end{cases} \tag{6}$$

where ω_k is a confidence parameter of the sample belonging to k-th class[1] and the additional $K+1$ output refers to the rogue devices. The vector z'_k is then fed into the softmax function, and the prediction can be obtained using (5). Different from [26] only leveraging the softmax output probabilities, this work exploits the entire activation vector, which is more robust.

[1]Please refer to [25] for the detailed calculation.

Open-set recognition can also be tackled by non-deep learning-based methods. Shen *et al.* designed a K-nearest neighbors (KNN)-based method [23]. They created a radio-frequency fingerprint (RFF) database that stores a few RFF features for each legitimate DUT. In the inference stage, RFF features will be extracted from the input signal and compared with the features in the database. The attacker is not registered beforehand, hence their features are largely different, which can be detected via a large feature distance. In contrast, the legitimate devices can be identified because there will be a matching feature in the database.

C. Anomaly Detection

There are K legitimate DUTs and rogue devices involved in the inference stage. Different from open-set recognition, anomaly detection only detects whether the DUT is legitimate or rogue. Because it is not practical to assume attackers are cooperative, hence, they are not available in the training stage.

Anomaly detection can be achieved by binary classification. As shown in Fig. 2.c, the K legitimate DUTs are treated as one class (label 1). A few other DUTs will be used to represent rogue devices, which serve as the other class (label 0). The system design will be similar to the closed-set classification, but the number of classes reduced to two. However, in the inference stage, when the rogue device appears, it is supposed to be classified as label 0.

Autoencoder (AE) is a popular unsupervised deep learning architecture for anomaly detection [25]. An AE-based RFFI approach is portrayed in Fig. 2.d. In the training stage, similar to the binary classification approach, the K DUTs are treated as a single class. But differently, there is no other device required. AE consists of an encoder and a decoder. The encoder first compresses the input, u, to a latent feature; the decoder will then try to reconstruct the input signal from the latent feature and output u'. The mean square error (MSE) between u and u' is typically used as the reconstruction error. The training process will learn the features of the training data and reduce the MSE. In the inference stage, if the signal is from the legitimate DUT, the trained AE can reconstruct the input, and a low MSE will be returned. Otherwise, when the signal is from a rogue device, the MSE will be higher than a threshold, indicating an outlier is detected.

IV. Hardware Impairments for RFFI

Due to the variations in the manufacturing processes, the hardware components of the radio devices will not be perfect. Their specifications will deviate from their nominal values slightly, which are referred to as RF hardware impairments. This section will provide the key parts for the modelling of transmitter and receiver impairments. The detailed mathematical derivation can be found in [8].

A. Transmitter Impairments

The architecture of a direct conversion transmitter is portrayed in Fig. 4. Their overall effects are represented as $\mathcal{F}(\cdot)$ in Section II while their individual effects will be modelled in this section.

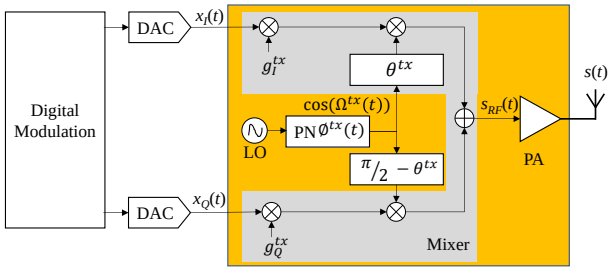


Fig. 4. Transmitter impairment model.

The local oscillator (LO) will produce a sinusoidal waveform with a specific carrier frequency. The output frequency is not stable but is subject to temperature and ageing. When the system's nominal carrier frequency is f_c^0, the real frequency is $f^{tx} = f_c^0 + \Delta f^{tx}$, where Δf^{tx} is the offset. Besides the carrier frequency offset (CFO), the LO is also subject to phase noise (PN), $\phi^{tx}(t)$. Considering all the LO imperfections, the carrier phase of the transmitter can be written as

$$\begin{aligned}\Omega^{tx}(t) &= 2\pi f_c^{tx} t + \phi^{tx}(t) \\ &= 2\pi f_c^0 t + 2\pi \Delta f^{tx} t + \phi^{tx}(t).\end{aligned} \tag{7}$$

The mixer will then mix the baseband signal with the carrier wave. However, the mixer is also subject to gain and phase imbalance. Specifically, g_I^{tx} and g_Q^{tx} represent the gain of in-phase (I) and quadrature (Q) branches, respectively; θ_{tx} denotes the phase imbalance. Due to the existence of gain and phase imbalance, the RF band signal then becomes [27]

$$\begin{aligned}s_{RF}(t) =& g_I^{tx} x_I(t)\cos(\Omega^{tx} + \theta^{tx}) - g_Q^{tx} x_Q(t)\sin(\Omega^{tx} - \theta^{tx}), \\ =& \Re\{s_{BB}(t) e^{j\Omega^{tx}}\},\end{aligned} \tag{8}$$

where $x_I(t)$ and $x_Q(t)$ are the baseband data at the I and Q branches, respectively, and

$$s_{BB}(t) = g_I^{tx} x_I(t) e^{j\theta^{tx}} + j g_Q^{tx} x_Q(t) e^{-j\theta^{tx}}. \tag{9}$$

The RF signal then undergoes the power amplifier, which introduces additional nonlinearities. A power amplifier in a narrowband system is usually modelled with memoryless nonlinear effects, including amplitude/amplitude (AM/AM) and amplitude/phase (AM/PM) characteristics [28]. There are several behavioural models, such as the Saleh, Rapp, and Ghorbani models, etc. [28].

After passing through a power amplifier, the signal becomes

$$\begin{aligned}s(t) &= A(|s_{BB}(t)|) e^{j(\angle s_{BB}(t) + \Omega^{tx}(t) + \Phi(|s_{BB}(t)|))}, \\ &= s'(t) e^{j\Omega^{tx}(t)},\end{aligned} \tag{10}$$

where $\angle s_{BB}(t)$ is the angle of the baseband signal and

$$s'(t) = A(|s_{BB}(t)|) e^{j(\angle s_{BB}(t) + \Phi(|s_{BB}(t)|))}. \tag{11}$$

B. Receiver Impairments

Similarly, the receiver will also have RF impairments. Fig. 5 depicts the receiver architecture and its impairments, i.e., receiver LO imperfection and mixer imbalance. Their overall

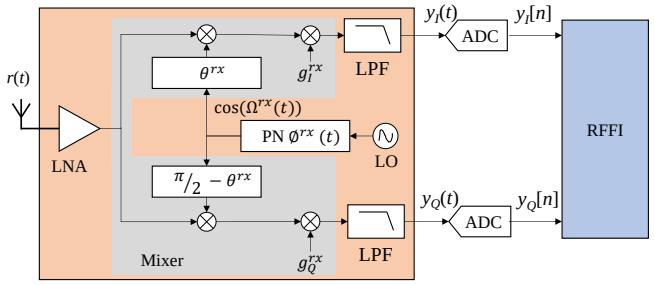


Fig. 5. Receiver impairments.

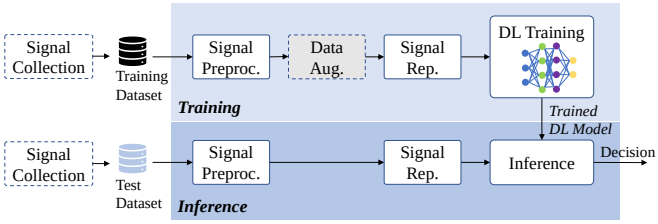


Fig. 6. Deep learning-based RFFI algorithm design.

effects are denoted as $\mathcal{R}(\cdot)$ in Section II. In this section, we will model their individual effects.

The LO at the receiver is also subject to frequency offset, Δf^{rx}, and phase noise, $\phi^{rx}(t)$. The receiver carrier can then be written as

$$\begin{aligned}\Omega^{rx}(t) &= 2\pi f_c^{rx} t + \phi^{rx}(t) \\ &= 2\pi f_c^0 t + 2\pi \Delta f^{rx} t + \phi^{rx}(t). \end{aligned} \tag{12}$$

The receiver uses a mixer to mix the received signal, which will downconvert the signal from the RF band to the baseband. Similarly to the transmitter mixer, the receiver mixer also has gain imbalance (g_I^{rx} and g_Q^{rx}) and phase imbalance (θ^{rx}).

Considering the effects of the LO imperfection and mixer imbalance, the receiver's carrier wave becomes

$$C^{rx}(t) = K_1^{rx} e^{-j\Omega^{rx}(t)} + K_2^{rx} e^{j\Omega^{rx}(t)}, \tag{13}$$

where $K_1^{rx} = (g_I^{rx} e^{-j\theta^{rx}} + g_Q^{rx} e^{j\theta^{rx}})/2$ and $K_2^{rx} = (g_I^{rx} e^{j\theta^{rx}} - g_Q^{rx} e^{-j\theta^{rx}})/2$.

The RF signal captured by the receiver can be written as

$$r(t) = (h(\tau, t) * s)(t) = (h(\tau, t) * s')(t) e^{j\Omega^{tx}(t)}. \tag{14}$$

After the downconversion (by the LO and mixer) and low-pass filter, the received signal at the baseband becomes

$$\begin{aligned} y(t) &= r(t) C^{rx}(t) \\ &= K_1^{rx} h(\tau, t) * s'(t) e^{j\Delta\Omega} + K_2^{rx} (h(\tau, t) * s')^*(t) e^{-j\Delta\Omega}, \end{aligned} \tag{15}$$

where $\Delta\Omega = 2\pi(\Delta f^{tx} - \Delta f^{rx})t + \phi^{tx}(t) - \phi^{rx}(t)$, $\Delta f = \Delta f^{tx} - \Delta f^{rx}$ is the commonly known CFO.

The baseband signal $y(t)$ in (15) possesses all the RF impairments of both the transmitter and receiver. The analogue signal is sampled by the analogue-to-digital converter (ADC), which produces a digital sequence, $y[n]$, and is used for RFFI. The transmitter impairments are the unique hardware features that RFFI explores. Regarding the receiver impairments, when the same receiver is used for collecting training and test datasets, the effects brought by receiver impairments are consistent and can be ignored. However, when different receivers are used, they will indeed affect RFFI performance, which will be reviewed in Section VI-C.

V. Deep Learning-based RFFI Algorithm Design

The deep learning-based RFFI algorithm design is shown in Fig. 6, including dataset collection, signal preprocessing, data augmentation, signal representation, and deep learning model.

- Training Stage: Once a training dataset is created, the sampled signals are processed by signal preprocessing (Section V-B) and then converted to a proper signal representation (Section V-D). An additional data augmentation approach is usually adopted to enhance the dataset diversity (Section V-C). The samples are then input into a deep learning model for training, which will produce a trained deep learning model when completed.
- Inference Stage: The signal undergoes the same signal preprocessing and signal representation algorithms, then is input to the trained deep learning model.

The deep learning training can usually be done offline, while the inference should be done in real-time in practice, even though many papers do it offline for evaluation purposes.

A. Signal Collection

The signal data collection is essential to collect IQ samples and build up dedicated datasets. The readers can also opt to use public datasets, which will be summarized in Section VI.

As introduced in [29], most RF fingerprinting studies utilize software-defined radio (SDR) as the wireless receiver to capture IQ samples for fingerprint extraction [22]. In addition, some WiFi-focused studies have explored the extraction of RF fingerprints from channel state information (CSI) provided by commercial network interface cards (NICs) or system-on-chips (SoCs), such as the Intel 5300 NIC [30], Atheros CSI tool [31], [32], Nexmon CSI tool [33], [34], and ESP32 CSI tool [35], [36]. More details can be found in Section VII.

B. Preprocessing

Signal preprocessing involves power normalization and CFO compensation. Power normalization can be done by normalizing the signal power with respect to the root-mean square (RMS) value of the power.

CFO represents the difference between the carrier frequencies of the transmitter and receiver, as embedded in (**??**). CFO has been widely adopted in the literature. For example, the work in [37] designed an accurate algorithm to estimate CFO from CSI in WiFi, as CFO is twisted with frame detection delay, sample frequency offset, and time of flight. CFO is also used together with other parameters to classify WiFi devices in [38] and ZigBee devices in [39].

However, CFO is subject to temperature drift. In a seven-month study carried out in [22], it is revealed that CFO is not suitable as a unique and stable feature. Specifically, the instantaneous CFO varies quickly when the device is

powered on due to the emitted heat of the device. While the instantaneous CFO is varying, the work in [22] also found that the CFO mean value remains relatively stable over the seven-month test period, which is used as an auxiliary approach to calibrate the prediction of the deep learning model.

In summary, due to the time-varying nature of CFO, it is suggested to carry out CFO compensation to preprocess the sampled signals, especially for low-cost IoT devices.

C. Data Augmentation

Data augmentation is used to augment the training dataset in a simulation manner. It is very time-consuming and labour-intensive to collect a comprehensive training dataset using experiments. In contrast, data augmentation can generate many artificial samples by adding channel and noise effects, which can significantly reduce the data collection overhead [40].

Specifically, the original training dataset can be constructed by sampling high-quality signals, $\{y_i\}$, which can be achieved by placing the DUT and receivers apart with a relatively short distance (e.g., less than 1 meter). We can then augment $\{y_i\}$ by emulating channel and noise as

$$y_i'(t) = (y_i * h'(\tau, t))(t)s + n'(t), \tag{16}$$

where $h'(\tau, t)$ is the multipath channel and $n'(t)$ is the AWGN noise, both generated by a simulation model.

In particular, the multipath channel modelling involves both the power-delay profile (PDP) and Doppler shift [23]. The PDP describes the attenuation gains of each channel tap. For example, the exponential PDP can be mathematically given as

$$P(m) = \frac{1}{\tau_d} e^{-mT_s/\tau_d}, m = 0, 1, ..., m_{\max}, \tag{17}$$

where τ_d is the RMS delay spread, $m_{\max}$ is the index of the last tap, and T_s is the sampling interval. Regarding the Doppler shift, it describes how the channel gain changes over time, with common models such as the Jakes model. By incorporating as many PDP and Doppler shift models as possible, data augmentation can significantly enhance the comprehensiveness of the training dataset.

The channel modelling can be achieved by employing the fading channel realization in Matlab [41]. The `comm.RayleighChannel` and `comm.RicianChannel` functions provide abundant interfaces to configure PDP and Doppler shift. Furthermore, the Wi-Fi channel models are also available in Matlab [42], with PDP pre-configured.

Besides the channel effect, AWGN can be added to emulate scenarios with different signal-to-noise ratio (SNR) levels.

D. Signal Representation

The signal captured by the receiver is always in the *time domain* initially, which is named IQ samples in the literature, as shown in (1). Utilizing raw IQ samples is applicable across any wireless protocol. However, as shown in (1), it is a time convolution between the hardware features and the channel, which makes it difficult to separate them in the time domain. Hence, IQ samples tend to be less effective for channel-robust RFFI, as evidenced in [43].

Frequency domain signal is popularly employed, which can be simply obtained by applying fast Fourier transform (FFT) operations to the time domain signal [22]. The channel effect can be separated from frequency domain signals more easily compared to the time domain counterpart. The *time-frequency domain* spectrogram is a widely employed signal representation in RFFI research [22], [44]. This can be obtained by applying a short-time Fourier transform (STFT) to the time domain signal. The time-domain IQ samples, frequency-domain FFT coefficients, and time-frequency domain spectrogram of LoRa preambles are exemplified in Fig. 7.

In addition to these domain transform methods, there are also other specially designed signal representations. For instance, Peng *et al.* post-process the constellation figures, generating image-like differential constellation trace figures (DCTFs) [45]. The authors in [46] subtract the ideal signals from the received ones, creating error signals as neural network inputs. Other available signal representations include bi-spectrum [47], and Hilbert-Huang spectrum [48], etc.

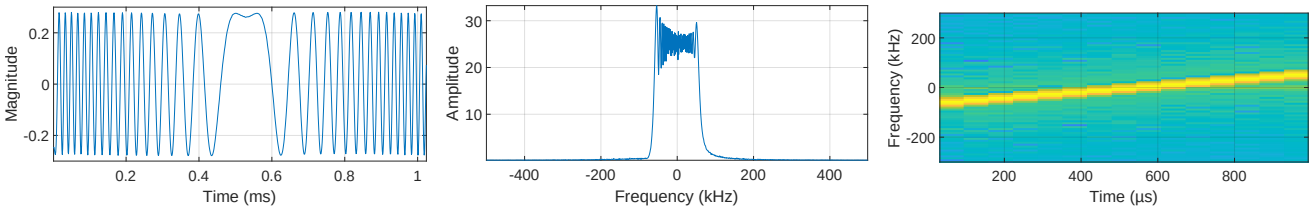
Aside from the signal representations derived from the steady-state portion of signals, some studies focus on extracting RF fingerprints from the transients that occur when transmitters are powered on or off [49]. However, this approach requires high-end receivers capable of operating at high sampling rates, which can significantly increase the cost of system deployment.

E. Deep Learning Model

The deep learning models are capable of extracting unique features from the input signal representations and subsequently predicting the device identity. The design of neural networks should take into account the employed signal representation, as illustrated in the following example. Image-like representations, such as spectrograms [44], [50], DCTF [45], and Hilbert-Huang spectrum [48], [51], are suitable for processing with CNNs, while time-domain IQ samples are suitable for processing with 1D CNNs or specially designed complex-valued neural networks [52]. Some studies also utilize MLP to process frequency-domain spectrum [21], [22]. As the captured radio signals exhibit temporal dependencies, sequence models can be employed for RFFI tasks as well. Recent studies have investigated the application of RNN, LSTM, GRU, and the latest Transformer models [24].

VI. RFFI Key Research Topics, Public Datasets and Applications

As shown in (2), RFFI performance is affected by channel and noise effects as well as receiver impairments. Therefore, this section reviews the RFFI research activities related to these three areas, namely channel effects elimination, noise mitigation, and receiver distortion mitigation. In addition, public datasets are critical to the development of deep learning-based RFFI techniques. As summarized in Table III, there are some RFFI datasets shared by the community. We provide a list of available datasets to evaluate the above three research topics. Finally, we review the RFFI literature in terms of their application techniques, including WiFi, ZigBee, LoRa, cellular, and Iridium satellites.



(a) (b) (c)

Fig. 7. Signal representation: (a) Time domain signals (I branch), (b) FFT coefficients, (c) spectrogram.

TABLE III
SUMMARY OF PUBLICLY AVAILABLE RFFI DATASETS

Wireless Technology	Dataset	Paper	# DUT	Receiver	Brief Summary	Evaluation Purpose		
						Channel	Noise	Receiver
WiFi	[53]	[54]	up to 150	USRP B210, X310 and N210 (up to 32)	WiFi signals collected from different days and positions.	✓		✓
WiFi	[55]	[56]	19	Xilinx Zynq ZC706 board + FMCOMMS5 ADI daughter board	WiFi signals collected from different days and positions (indoor, outdoor, car park).	✓		
WiFi	[57]	[58]	10	USRP X310	WiFi signals collected from different days and positions (indoor, static, mobile).	✓		
ZigBee	[59]	[60]	60	USRP N210	Outdoor-NLOS, outdoor-LOS, indoor-NLOS, and indoor-NLOS	✓		
LoRa	[61]	[23]	60	USRP N210	LoRa preambles. Signals from different channels available	✓	✓	
LoRa	[62]	[24]	10	USRP N210	LoRa preambles with different spreading factors			✓
LoRa	[63]	[64]	10	USRP N210 (3), B210 (2), B200 mini (2), Pluto (2), RTL (9)	LoRa preambles. Signals from different receivers and channels available	✓	✓	✓
LoRa	[65]	[66]	60	USRP N210, B210, B200, B200 mini, Pluto, RTL	LoRa preambles. Signals from different receivers and channels available	✓		
LoRa	[67]	[44]	100	USRP N210	Signals from indoor and outdoor testbeds, and different days	✓		
LoRa	[68]	[69]	25	USRP B210	Signals from different days, distances, locations and receivers	✓		
Bluetooth	[70]	[71]	10	USRP X300	Signals collected from different locations and days	✓		
UWB	[72]	[73]	13	COST UWB board	Signals collected from different locations and days	✓		
LTE	[74]	[75]	8	USRP N210	Signals collected from different locations and days	✓		
Satellite	[76]	[77]	66	USRP X310	Signals from 66 Iridium satellites		✓	
Satellite	[78]	[79]	66	USRP N210	Signals from 66 Iridium satellites		✓	

A. Channel Effects Elimination

1) Research Activities: The received RF signals are affected not only by the transmitter hardware impairments but also by the wireless channel. In particular, the movement or relocation of wireless transmitters can result in fluctuations in the characteristics of received RF signals, which may subsequently interfere with the RFF extraction process.

The negative impacts of wireless channels have been revealed in numerous recent RFFI studies. The authors in [52] and [54] conducted comprehensive experiments to assess the effects of wireless channels on RFFI. Their findings indicate that channel variations can lead to significant performance degradation when fingerprinting WiFi signals. Similar results are also found in fingerprinting wireless signals with narrower bandwidths, such as LoRa [23], [69] and ZigBee [80]. These studies experimentally demonstrate that the inevitable channel effects can degrade the RFFI performance, presenting a significant challenge that must be addressed.

Recent studies attempt to mitigate the channel effects through two categories of approaches: signal processing and deep learning algorithms. The former category often employs expertise and prior knowledge in wireless communication to design signal processing algorithms to manually separate channel distortions from the received RF signal, constructing channel-robust features for identification purposes. For example, the authors in [23] propose to mitigate the multipath effects in the time-frequency domain by dividing neighbouring columns in a spectrogram, thereby generating a channel-independent signal representation. This method has been demonstrated to be effective in identifying LoRa devices,

exhibiting excellent channel-independent properties. The work in [81] designs a channel-robust WiFi RFF named difference of the logarithm of the spectrum (DoLoS), which is based on the fact that the long training symbols (LTSs) and short training symbols (STSs) in a WiFi packet share similar channel frequency response (CFR). In addition, some studies employ a channel estimation module to approximately measure the channel impulse response (CIR) or CFR, subsequently utilizing the result to perform channel equalization, thereby largely eliminating the multipath effect [52], [82]–[84]. However, the channel equalization process can inevitably eliminate some hardware features while eliminating the multipath effect, resulting in the degradation of identification accuracy.

Deep learning algorithms are also employed to mitigate channel effects. More specifically, these approaches aim to enhance the capacity of neural networks, enabling them to extract channel-independent features automatically. The most prevalent method is data augmentation, which has been introduced in Section V-C. In summary, it utilizes a wireless channel simulator to synthesize a greater number of signals exhibiting various channel effects during the neural network training process. This can expand the distribution of training data effectively, aligning it with the test phase to prevent potential performance degradation. The benefits of data augmentation have been illustrated in numerous recent studies [23], [40], [44], [83], [85], and gradually evolved into a standard procedure in designing RFFI systems. In addition to data augmentation, several studies have attempted to use the latest deep learning methods, such as transfer learning [86], domain adaptation [87], or disentangled representation learning [80] to improve the system's robustness to channel variations.

2) Available Datasets: As discussed, the channel effects, i.e., multipath and Doppler effects, have a considerable impact on the performance of RFFI. It is, therefore, essential to design effective mitigation algorithms. A number of datasets are available for this purpose, containing signals collected under a variety of positions and channel conditions [23], [44], [52], [54], [60]. Shen *et al.* release a public dataset consisting of 60 LoRa devices. The signal collection is carried out at six locations, including both line-of-sight (LOS) and non-line-of-sight (NLOS) channel conditions [23]. The authors in [44] and [69] also perform experiments in outdoor environments, and the datasets are made available. Shi *et al.* collected signals from 60 ZigBee devices, including both indoor/outdoor and LOS/NLOS scenarios [60].

In addition to the narrowband LoRa and ZigBee datasets listed above, there are also publicly available datasets within the research community that contain wideband WiFi and long-term evolution (LTE) signals. Hanna *et al.* collected a large-scale WiFi dataset, consisting of 174 transmitters and 41 universal software radio peripheral (USRP) receivers [54]. The experiment is conducted four times, spanning a mont,h within a grid indoor testbed. Additionally, Al-Shawabka *et al.* present a large-scale WiFi dataset collected in a grid testbed of 6,000 square ft [52]. The transmitters are 13 USRP N210 and seven USRP X310, while a USRP N210 receiver is used for signal capture. In addition to WiFi, the authors in [75] also provide an LTE dataset which consists of signals collected from seven mobile phones. It should be noted that wideband signals are more susceptible to channel variations than narrowband ones.

B. Noise Mitigation

1) Research Activities: The propagation of wireless signals over long distances can result in severe attenuation, which in turn leads to a low SNR condition at the receiver. Given that transmitter hardware impairments are often minute, RFFs are probably submerged in noise. It is therefore difficult to accurately extract them for identification. As wireless systems often operate in RF conditions where the SNR is less than 20 dB, it is necessary to explore RFFI solutions that are robust to noise contamination.

Denoising can be leveraged to improve the system's robustness.Wang *et al.* demonstrate that smooth filtering is effective in combating noise contamination [88], and Xing *et al.* conclude that stacking multiple identical symbols is also effective in reducing noise [89]. The authors in [90] reveal that converting the PHY waveform to a logarithmic power spectral density can improve identification accuracy in low-SNR environments. Although these manual denoising algorithms are experimentally shown to be effective against noise interference, whether the RFFs are unintentionally eliminated remains unclear. In addition, there are also studies utilizing multiple observations to improve low-SNR performance, such as merging the identification results of multiple receivers [64], [91], [92] or multiple data packets [24].

Apart from these, some studies have attempted to enhance the ability of deep learning models to process low SNR signals. The authors in [24], [93] evaluate different data augmentation strategies, concluding that adding artificial noise to mini-batches during training, i.e., online augmentation, can lead to the most significant improvement. Some studies improve the low-SNR performance by using specially designed neural networks. For example, the authors in [94] utilize the dynamic shrinkage learning network, which can integrate denoising capabilities into deep learning models.

2) Available Datasets: RFFs caused by hardware impairments are often faint, and their effective extraction at low SNR conditions is challenging. The authors in [77] and [79] present datasets collected from IRIDIUM satellites, which are particularly suitable for low-SNR RFFI research. The IRIDIUM satellites operate in low Earth orbit, at an altitude of approximately 780 kilometers. This results in severe propagation attenuation and makes the signal extremely weak at the ground receiver. In addition to satellites, the LoRa dataset in [64] also contains low-SNR signals. Specifically, the transceiver distances are up to 30 meters, and the signal SNRs are clearly labelled, ranging from 10 dB to 50 dB. Similarly, the authors in [69] collect LoRa signals at various distances as well, i.e., ranging from 5 m to 25 m. Despite the low-quality RF signals provided by the above-introduced datasets, an alternative and efficient method is to add artificial Gaussian noise to high-quality signals, thereby synthesizing low-SNR conditions [24], [44].

C. Receiver Distortion Mitigation

1) Research Activities: While RFFI aims to exploit the transmitter's unique hardware impairments for identification, the receiver impairments will also affect the received signal, as shown in (2). Most RFFI studies assume the same receiver is used during the training and inference stages, thus, the receiver effect can be neglected. However, this assumption is not always valid, as the transmitter is frequently served by multiple receivers in practical wireless systems. The authors in [8] build simulation models to evaluate the effect of receiver impairments. The simulation results demonstrate that the identification accuracy decreases by up to 20% when the IQ imbalances of the receivers are different between training and test. The performance degradation caused by receiver effects is experimentally validated in [54], [64], [69], [95].

To overcome the receiver effect, the authors in [54], [95] recommend including as many receivers as possible in the training stage, with the aim of improving the model's generalization ability. Moreover, the work in [64] proposes a receiver-agnostic training scheme that employs a gradient reversal layer to direct the deep learning model to learn receiver-independent features. The proposed algorithm was evaluated using 20 SDR receivers, demonstrating excellent generalization ability. The authors in [96] applied the concept of generative adversarial network (GAN) to learn receiver-independent features, resulting in an improvement of 20% in accuracy. Furthermore, recent studies demonstrate that transfer learning and fine-tuning strategies can adapt classifiers trained on one receiver to perform effectively on others [97].

2) Available Datasets: As discussed in Section VI-C, the RFFs are not only affected by the transmitter but also by the receiver chain. It is therefore necessary to have datasets containing signals collected by multiple receivers to explore effective receiver-agnostic RFFI solutions. Shen *et al.* present a dataset comprising 20 SDR receivers of varying types, spanning from low-end RTL-SDR to high-end USRP N210 [64]. This dataset was specifically created for the purpose of evaluating receiver effects, and the signals collected at various positions are available. While other LoRa datasets also contain multiple receivers [66], [69], they are incompatible with [64] in terms of number and types of receivers. With regard to WiFi protocols, the WiSig dataset comprises 41 USRP receivers, which are suitable for use in research into mitigating WiFi receiver distortion [54].

D. Applications

There have been many deep learning-based RFFI papers published in the last few years, with applications in WiFi, ZigBee, LoRa, LTE, and satellite communications.

1) WiFi: Recent studies have attempted to apply RFFI to secure WiFi systems, ranging from IEEE 802.11b to IEEE 802.11ax standards [52], [54], [81], [84], [89], [98]. Li *et al.* design a fractal dimension estimation method to extract features from direct-sequence spread spectrum (DSSS) IEEE 802.11b signals, and use support vector machine (SVM) or KNN for identification [89]. For wideband orthogonal frequency-division multiplexing (OFDM) signals, the authors in [52] and [54] evaluated the channel effects, showing that the multipath effects can significantly degrade the identification performance. To alleviate this problem, Xing *et al.* design a DoLoS algorithm, extracting channel-robust features from IEEE 802.11 OFDM signals as the neural network input [81]. The most significant challenge for WiFi RFFI systems is the design of effective algorithms for eliminating channel effects.

2) LoRa: A considerable amount of research has been conducted for designing LoRa RFFI systems [21]–[24], [44], [64], [69]. To the best knowledge of the authors, [21] is the first work attempting to use RFFI to identify LoRa transmitters, which carries out experiments in a transceiver distance of up to 100 meters and achieves 59% to 99% accuracy. Furthermore, Shen *et al.* conducted a series of research aimed at developing practical and robust LoRa RFFI systems [22]–[24], [64], [66], and released all the datasets and codes to the public. In [22], different signal representations are studied, i.e., IQ samples, FFT coefficients, and spectrogram, concluding that spectrogram is the most appropriate for LoRa signals because of their frequency-changing property. The work in [23] designed the channel-independent spectrogram to mitigate the channel effects and a three-stage protocol for open-set identification. Afterwards, [24] aims at improving identification accuracy in low-SNR environments, which can be achieved by online augmentation and merging predictions derived from multiple LoRa packets. Finally, the studies in [64] and [66] explore receiver-agnostic and federated RFFI protocols, respectively. In addition to these, the researchers in [44] and [69] also carried out in-the-wild experiments to validate RFFI performance. As a low-power, long-range communication technology, the most significant challenge for LoRa RFFI is to design effective algorithms to combat noise contamination.

3) ZigBee: RFFI is also utilized to authenticate ZigBee/IEEE 802.15.4 devices [39], [45], [46], [80], [99], [100]. The authors in [100] construct multiple discriminant analysis (MDA) classifiers to identify ZigBee devices. With the development of deep learning, Merchant *et al.* propose inputting error signals, i.e., the difference between received and ideal signals, into CNNs for the identification task [46]. Peng *et al.* design the image-like DCTF feature as the CNN input as well [39], [45]. The authors in [99] and [80] adopt more advanced deep learning algorithms, using multisampling convolutional neural network (MSCNN) and disentangled representation (DR) learning to construct RFFI systems.

4) BLE: The work in [101] studied using BLE hardware features to track mobile devices. The authors designed a non-deep learning-based approach by computing the Mahalanobis distance to track registered devices. With comprehensive experiments over 17 mobile devices, involving smartphones, laptops, etc, they revealed it is viable, although sometimes unreliable, to use BLE hardware fingerprints to track mobile devices. The same group later used a CFO obfuscation strategy by modifying CFO [102] to prevent such a tracking attack.

Regarding deep learning approaches, Jagannath *et al.* designed an embedding-assisted attentional framework and achieved a significant reduction of the memory usage and trainable neural network model complexity [71]. Yuan *et al.* designed a denosing AE, with a CNN as the backbone, to

improve the classification performance under low SNR [103]. The authors achieved over 75% accuracy over 10 dB SNR for 18 BLE devices.

5) Ultra-Wideband: Ultra-Wideband (UWB) is usually used for high-resolution localization with a precision at the centimeter level. This is enabled at a cost of high bandwidth, e.g., 500 MHz. To the best knowledge of the authors, the work in [73] is the only one studied RFFI for UWB. Due to the high bandwidth, it will be challenging using SDR for capturing UWB signals. Instead, the authors employed commercial off-the-shelf (COTS) UWB devices to collect CIR measurements and converted them to spectrograms. They designed a Vision Transformer-based deep learning model and achieved over 99% classification accuracy.

6) Cellular Communications: Most of the recent RFFI research has focused on identifying devices operating in the unlicensed industrial, scientific, and medical (ISM) bands, and there are only a few studies have investigated its application in cellular systems [75], [104]–[107], e.g., GSM, 4G LTE, 5G NR. Zhuang *et al.* utilize RFFI technology to identify GSM base stations to detect fake base station (FBS) crimes, which is achieved by extracting modulation errors and statistical features as unique RFFs [104]. The authors in [105] propose an RFFI method by applying wavelet decomposition to the 4G LTE demodulation reference signal (DMRS). Yin *et al.* capture the transient-on, transient-off, and modulation segment of LTE physical layer random access channel (PRACH) preambles, converting them to DCTF representations separately and designing a multi-channel CNN for identification [106]. Peng *et al.* combines wavelet transform (WT) coefficient graphs and differential spectrum to extract RFFs from LTE signals and conduct experimental evaluations using commercial phones and SDRs [75]. Finally, the authors in [107] create a 5G RFFI system involving four base stations, which demonstrates that RFFI is effective in securing 5G networks as well.

7) Satellite Communications: Recent studies have applied the RFFI technique to satellite identification [77], [79], [108]. For instance, the authors in [108] utilize it to detect global positioning system (GPS) spoofing attacks. Specifically, they use multivariate normal distribution (MVN) models to extract features from the captured IQ samples and set a threshold to detect the spoofed GPS signals. Oligeri *et al.* targets to identify the low Earth orbit (LEO) IRIDIUM satellites, which was originally developed by Motorola in the last century [77]. The authors observe and collect IQ samples from 66 satellites, and then train a CNN for identification. The results show that the accuracy is above 80%. Smailes *et al.* also conducts extensive research in fingerprinting IRIDIUM satellites [79].

VII. RFFI Experimental Methodologies

While it is fundamental to carry out experimental evaluation for RFFI to assess its performance in practical scenarios, there is a lack of explanation of the experimental methodologies in the literature. This section aims to bridge the gap.

Building a testbed is mandatory to carry out an experimental evaluation, therefore this section will cover necessary information for building a testbed, including DUT and receiver. Then, we will discuss the requirements for the dataset collection.

A. DUT

There are several DUT options, which include IoT development kits, consumer electronics, and SDR platforms.

IoT kits are probably the most commonly used devices. Some examples are given below:

- WiFi: ESP32 [109]
- BLE: ESP32 and Nordic Semiconductor nRF52840 dongles [103]
- LoRa: Pycom LoPy4 & FiPy (discounted), mbed shield, and Dragino shield [23].

The vendors usually provide example code snippets for transmitting and receiving. Their transmission behavior can be customized, e.g., transmission intervals, transmit power, etc.

It is desirable to demonstrate that RFFI can work with **consumer electronics**. We exemplify the following consumer electronics DUTs:

- WiFi: WiFi dongles [110] and Nexus 5 smartphones [111]
- BLE: Smartphone, laptop, Apple Watch etc [101]
- LTE: Smartphone [106], [111]

Compared to IoT kits, it is relatively difficult to control transmission parameters accurately. The transmissions can instead be triggered by, e.g., running the ping command for WiFi [110]. For BLE, the device will transmit advertising packets periodically, which can be leveraged [101]. On the other hand, there is also research reported by using a third-party firmware, nexmon, to inject I/Q imbalance into the baseband signal of smartphones [111].

SDR platforms are also employed as DUTs for RFFI. For example, the authors in [115] used 10 HackRF One SDRs and the work in [121] used 20 USRP SDRs. Using SDRs can provide full hardware control. For example, the hardware impairments are reconfigured in [122]. However, their cost is usually higher than IoT kits and consumer electronics.

B. Receiver

The receiver plays an important role in RFFI, which will perform signal collection and then feed the collected signals for classification. In deep learning-based RFFI, there are two categories of signals for deep learning input, namely I/Q samples and CSI.

1) Platform for I/Q Samples Collection: Most of the RFFI works rely on I/Q samples, which can be captured by SDR platforms. There are different ways to access data from SDR, including Matlab, GNURadio, Python-based libraries, and PicoScenes, as summarized in Table IV.

While SDR can capture the I/Q samples, a signal analysis program is required to decode and interpret the data samples. For example, signal synchronization algorithms are required to locate the starting point of the collected packets. MAC address decoding is required for WiFi and Bluetooth to ensure that the captured packets are sent from the target DUT, because there are numerous WiFi and Bluetooth transmissions over the air. Such functions can be either achieved by custom-built codes or available third-party solutions such as Matlab toolboxes, GNURadio implementations (see Table IV). Regarding PicoScenes, it is a middleware specifically created for WiFi.

TABLE IV
SUMMARY OF SDR PLATFORMS AND THEIR APPLICATIONS IN RFFI

Software	Supported SDR Platform	Wireless Technology	Representative RFFI Papers
Matlab [112]	ADALM-PLUTO SDR, RTL-SDR, USRP SDR and Xilinx® Zynq®-Based Radio	WiFi (Matlab WLAN toolbox)	Not reported
		LoRa (custom code)	USRP N210 [113]
		BLE (Matlab Bluetooth toolbox)	USRP N210 [103]
GNURadio	All SDR platforms support GNURadio	WiFi IEEE 802.11a/g/p [114]	HackRF One [115], USRP N210 [111]
		IEEE 802.15.4 [116]	Not reported
Python [117]	ADALM-PLUTO SDR, RTL-SDR, USRP SDR, HackRF One and BladeRF	Custom code	Not reported
PicoScenes [118]	USRP SDR and HackRF One	WiFi IEEE 802.11a/g/n/ac/ax/be	USRP N210 [110]

TABLE V
SUMMARY OF CSI TOOLS AND THEIR APPLICATIONS IN RFFI

CSI Tool	Supported Amendaments and Chipsets/SDR	Representative RFFI Papers
Intel 5300 CSI tool [30]	IEEE 802.11n for IWL5300	[37], [119], [120]
Atheros CSI tool [31], [32]	IEEE 802.11n	Not reported
Nexmon CSI tool [33], [34]	IEEE 802.11a/g/n/ac for Broadcom WiFi Chips	Not reported
ESP32 CSI tool [35], [36]	IEEE 802.11n	Not reported
PicoScenes [118]	IEEE 802.11a/g/n/ac/ax for USRP SDR and HackRF One, AX210/AX200, IEEE 802.11n for QCA9300 and IWL5300	AX210 [120]
SDR + Matlab WLAN Toolbox	IEEE 802.11a/g/n/ac/ax for ADALM-PLUTO SDR, RTL-SDR, USRP SDR and Xilinx® Zynq®-Based Radio	Xilinx® Zynq®-Based Radio [109]

2) Platform for WiFi CSI Collection: CSI can represent fine-grained channel information. While most of the WiFi chipsets do not provide the CSI, there are a few exceptions, as summarized in Table V. Intel 5300 CSI tool [30] is probably the most widely used as it is the first CSI tool. However, it can only report channel matrices for 30 subcarrier groups, i.e., every 2/4 subcarriers at 20/40 MHz. Nexmon CSI tool can support up to 80 MHz and return estimated CSI for all the subcarriers, which can significantly increase the extracted information. PicoScenes can support the latest WiFi 6 with up to 160 MHz bandwidth for AX210/AX200. Besides, the MATLAB WLAN toolbox can also provide CSI.

C. Requirement of Dataset Collection

As indicated in [52], the training and test datasets are collected on two different days, and channel conditions are similar, but their deep learning-based RFFI cannot work at all [52, Fig. 11]. This reveals that even slight variations in the channel, noise, or hardware impairments will result in significant consequences. In practice applications, the test datasets are highly likely collected on different days from the training datasets. Hence, we need to design a robust RFFI algorithm. In order to demonstrate the robustness of RFFI algorithms, it is always necessary to have training and test datasets collected from different days.

It is important to avoid overfitting in RFFI. For example, when evaluating RFFI against channel variations, the training and test datasets should not be collected from the same environment or environments with similar channel conditions. In particular, as many different channel scenarios as possible should be covered, e.g., LOS & NLOS, static & mobile, indoor & outdoor, etc.

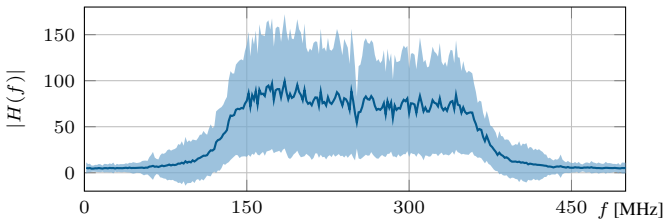
VIII. CHANNEL-BASED AUTHENTICATION

CB PLA verifiers leverage the effects of the communication channel for authentication; thus, in this case, the nature of the communication channel itself enables authentication.

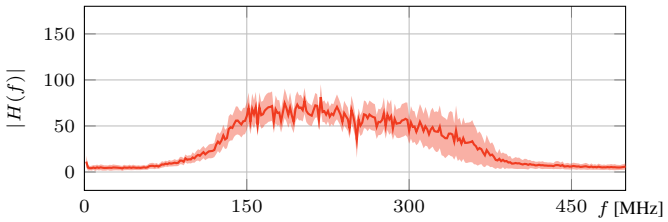
In particular, a channel measurement, typically called *channel feature*, is selected. As an example, Fig. 8 depicts the absolute value of the CFR measured using Wi-PoS, an UWB hardware platform, with carrier frequency $6.489\,\mathrm{GHz}$ and bandwidth of $499.2\,\mathrm{MHz}$, place at the fourth floor of the iGent Tower and in Portus Ganda, both located in Ghent, Belgium [123]. In particular, we report mean and 1σ bounds computed over 300 measurements. Indeed, while the traces collected in the same environments are related, there are significant changes when the devices are collected in different locations. For instance, the indoor environment is associated with a much higher standard deviation, e.g., due to multipath, than the outdoor environment. This highlights that, indeed, we can exploit traces like these, and thus the channel, to authenticate the devices, as a signal sent by a spoofer placed in a different environment will induce a significantly different CFR. The review of the channel features to be used for authentication purposes is reported in Section IX.

Concerning the authentication mechanism, three main approaches have been studied: the tag-based, channel variation, and challenge-response (CR) approaches.

- The *tag-based approach* assumes that the channel features do not change over time and the impersonating attacker is in another location than the legitimate transmitter, enabling the verifier to distinguish between the two transmitting locations by processing the received signal.
- The *CR approach* still considers static channels but it also assumes that the verifier can modify the propagation environment and predict the resulting channel features. Thus, authentication is performed by introducing a random



(a) Indoor scenario.



(b) Outdoor (open area) scenario.

Fig. 8. Examples of CFR, mean and 1σ bounds for an indoor and an outdoor scenario.

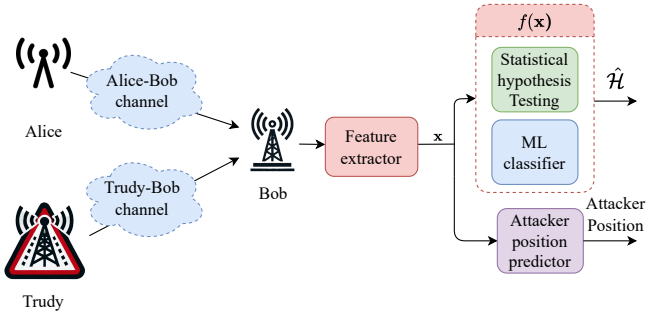


Fig. 9. High-level representation of the CB-authentication scheme.

modification (challenge) and comparing the predicted channel features with those obtained from the received signal.

- The *channel variation approach* is suitable in scenarios where the channel changes, as it includes a prediction of the channel features from previous observations and a comparison between the predicted features and those estimated from the currently received message.

Note that the latter two approaches assume instead that the channel changes, but variations are in part known by the verifier: either change can be predicted (e.g., channel variation) or controlled (e.g., CR-authentication) by the verifier.

Fig. 9 provides a general scheme for CB PLA, in which first the verifier Bob extracts the vector feature $\boldsymbol{x}$ from the channel of the received signal and then classifies it as legitimate or not by means of a function $f(\cdot)$. In particular, the classification variable is $\hat{\mathcal{H}} = \hat{\mathcal{H}}_0$ when the received message is considered authentic and $\hat{\mathcal{H}} = \hat{\mathcal{H}}_1$ otherwise. In case the signal is marked as malicious, it may be possible to localize the attacker (see Section X-E).

The following two sections will describe the features and methodologies used for PLA. Table VI summarizes the main publications on CB-authentication. It provides a categorization based on the application domain (radio communications on air, UWAC, and visible-light communications (VLC) on air), the decision methodologies (statistical and ML), and the channel features used for authentication.

A. Tag-based PLA

A typical tag-based PLA protocol includes two phases:

- *Identification Association*: the legitimate transmitting device (Alice) sends a pilot sequence to the verifier device (Bob), which *estimates or learns* channel characteristics/behavior, which we call *tag*. The pilot transmission is assumed to be authenticated because either we are sure that the attacker is not transmitting or an authentication mechanism using a pre-shared secret between Alice and Bob is implemented. This phase is performed only once and is not repeated for each transmitted packet.
- *Identification Verification*: Bob receives a message containing pilot symbols, estimates the channel, and compares such an estimate with the tag obtained in the identification association phase. If the two are *compatible*, the message is considered authentic, otherwise, it is discarded as fake. This phase is performed for each message transmission, after the identification association.

Tag-based PLA is subject to two issues: a) the channel may change over time due to the mobility of either the user or surrounding objects, and b) the channel estimate is affected by noise and receiver impairments (e.g., synchronization issues). Both make the definition of *compatibility* between the channels estimated in the two phases problematic. To cope with these issues, two main research paths have been taken. About the first issue, we note that tag-based PLA is well suited for static channels, while some modifications are needed to make it work under channel variations.

The first research path has looked for channel characteristics that are more robust against both issues to be used for authentication. For example, the number of taps of the CIR is less time-variant than the full CIR. Similarly, focusing on the received power and dropping the channel phase information in narrowband transmissions provides a robust approach against synchronization errors.

The second research path focuses on the methodology to assess the message authenticity, taking into account the impairments (interference, noise, channel variations) of the estimated tag. A first set of solutions is obtained by framing tag-based PLA as a hypothesis testing problem between two hypotheses (the received signal is legitimate or fake) and resorting to statistical tests for its solution: this however, requires the knowledge of the tag statistics. A second set of solutions is obtained by considering tag-based PLA as a classification problem to be solved using ML models: this approach requires a dataset of tag realizations. Solutions mixing the two approaches are also possible, e.g., using a raw statistical approach as the initial test, then a more refined ML test.

In a general tag-based PLA approach, two hypotheses are considered, $\mathcal{H}_0$ and $\mathcal{H}_1$, corresponding to the legitimate and the under-attack case, respectively. To assess the authenticity

of the signals, the verifier uses the tag verification function $f(\boldsymbol{x})$ and computes the decision via thresholding, i.e., deciding

$$D: \quad \hat{\mathcal{H}} = \begin{cases} \hat{\mathcal{H}}_0 & f(\boldsymbol{x}) \geq \lambda, \\ \hat{\mathcal{H}}_1 & f(\boldsymbol{x}) < \lambda, \end{cases} \tag{18}$$

where λ is a threshold chosen by the verifier, e.g., to minimize the misdetection (MD) for a target false alarm (FA) probability.

When using a statistical approach, the function $f(\cdot)$ is derived analytically, typically exploiting the statistics of the $\boldsymbol{x}$ under $\mathcal{H}_0$ and, eventually, $\mathcal{H}_1$. More details on this approach are provided in Section X-A. On the other hand, in the ML domain, we aim at classifying the observed tag $\boldsymbol{x}$ into the two classes of legitimate and attack messages. We still perform the test with (18), where now $f(\cdot)$ represents an ML model, trained with a dataset of labelled tags, where the label indicates the class to which the tag belongs.

Comparisons between the statistical and ML paradigms have been reported in [124]–[126]. By properly designing the ML model and its training, it is possible to achieve the same performance as the statistical approaches. In general, however, the choice between the two approaches is dictated by the knowledge of statistics or the availability of datasets.

A relevant distinction, common to solutions of both statistical and ML domains, concerns the knowledge about the attacker. In particular, we distinguish

- *Binary Classification*: the verifier Bob knows the distribution of both Alice and Trudy (statistical domain) or has a labeled dataset with tags belonging to both Alice and Trudy (ML domain). We remark that in the literature, binary classification is also referred to as two-class classification, as it exploits information (i.e., distribution and/or labeled data) of both Alice and Trudy.
- *Artificial Dataset*: the verifier has a dataset with tags belonging to Alice but makes some assumptions about Trudy and generates an artificial dataset of Trudy's tags.
- *One-Class Classification*: the verifier knows only the distribution of Alice (statistical domain) or has a dataset with tags belonging only to Alice. Especially in ML contexts, one-class classification is also referred to as an anomaly/outlier detection task.

Remark: it is worth pointing out that, differently from the binary class case, in the artificial dataset case, the knowledge about the attacker is only partial, and thus such knowledge does not allow the legitimate party to build a fully reliable Trudy dataset. For instance, we know a region where Trudy may be, but we do not know the exact position; we can then build a dataset with observations from the whole region.

B. Challenge-Response PLA

Modern communication systems enable a partial modification of the electromagnetic propagation environment. For example, a RIS can be used in a PLA context and be controlled by the verifier to steer impinging signals in desired directions. Another example is obtained when the verifier is a moving device, e.g., a drone, that can modify the channel from the transmitter by changing its position. CR-PLA is a mechanism that leverages such control of the electromagnetic environment to strengthen authentication [127]. We define the different conditions of the channel induced by the behaviour of the verifier (e.g., the RIS configuration or the drone position) as *channel configurations*.

As shown in Fig. 10, the CR-PLA procedure includes two phases, similar to tag-based PLA, but with different contents:

- *Identification Association:* First, the channel features from the device to be identified are estimated by the verifier *for different channel configurations*, namely Channel $1, \ldots, M$. The estimation of the first phase enables the receiver to obtain estimates of the channel also for configurations that have not been explored in this phase, through interpolation algorithms. We must ensure that in this phase the transmission is legitimate (thus no spoofing attack is possible), as it happens in the identification association phase of tag-based PLA.
- *Identification Verification:* The identification verification phase is split into two steps, challenge and response.
  - *Challenge*: The verifier selects at random a channel configuration before transmission of the message.
  - *Response*: The message is transmitted and the verifier estimates the resulting channel from the received signals. Lastly, the verifier compares the estimated channel with the channel predicted for the selected channel configuration, according to the information acquired in the first phase.

Note that an attacker to be successful must transmit its signals through the legitimate channel as modified by the receiver (e.g., through the RIS rather than directly to the receiver) or it must know the instantaneous channel configuration and shape its attack accordingly (see [128] for an in-depth analysis).

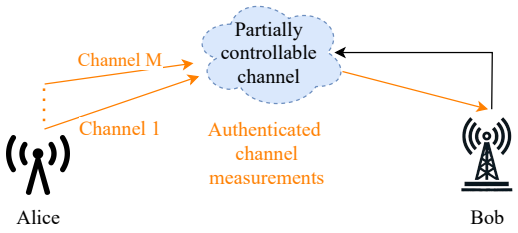
CR-PLA introduces additional randomness to the channel conditions (with respect to tag-based PLA), thus achieving higher robustness against attacks. The optimization of the defence strategy (choice of the random channel configuration in the Challenge phase) and of the attack strategies has been investigated in [129], [130].

C. Channel Variation PLA

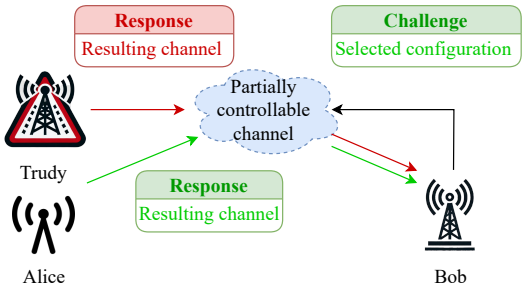
Channel variation PLA is an authentication mechanism specifically designed for time-varying channels. A typical channel variation protocol works as follows:

1) The legitimate transmitter sends N packets to the verifier, which estimates the channel features for each packet (*identification association* step).
2) From the estimated features, the verifier predicts the channel features for future transmissions.
3) Upon reception of the packet to be verified, the verifier checks the consistency of the predicted and the measured features (*identification verification* step).

As in the tag-based PLA, the main research directions investigate both the choice of robust features and the design of predictors providing a good trade-off between FA and MD probabilities. The predictor can be implemented both via statistical methods or, more recently, via ML.



(a) Identification association in CR-PLA.



(b) Identification verification in CR-PLA.

Fig. 10. Challenge-response paradigm scheme.

In formulas, given the previous channel observations $[\boldsymbol{x}_{-N}, \boldsymbol{x}_{-N+1}, \ldots, \boldsymbol{x}_{-1}]$, the prediction function is denoted by $g(\cdot)$ and it provides

$$\boldsymbol{y}_0 = g\left([\boldsymbol{x}_{-N}, \boldsymbol{x}_{-N+1}, \ldots, \boldsymbol{x}_{-1}]\right) , \tag{19}$$

and the consistency check for the current observation $\boldsymbol{x}_0$ is

$$\hat{\mathcal{H}} = \begin{cases} \hat{\mathcal{H}}_0 & \text{if } \|\boldsymbol{y}_0 - \boldsymbol{x}_0\| < \lambda, \\ \hat{\mathcal{H}}_1 & \text{if } \|\boldsymbol{y}_0 - \boldsymbol{x}_0\| \geq \lambda. \end{cases} \tag{20}$$

Variations to (20) include tests where no prediction is performed but the threshold is dynamically updated, e.g., as in [131].

We remark that differently from tag-based PLA, channel variation PLA is typically considered as a single-sided hypothesis testing or a one-class classification problem, without any knowledge of the attacker's behavior.

D. Channel-Based Attacker Localization

When dealing with PLA attacks, it is important that the receiver first detects the attack and then possibly localizes the attacker. We will also see that some features used for verification are also used for localization. Thus, when implementing CB-PLA protocols, on top of the authentication, it may be possible to add an attacker localization step at a relatively low cost. Thus ,during the rest of the survey, we will also include solutions for attacker localization.

IX. Channel Features For CB-Authentication

In this Section, we describe the main channel features that are used in CB-PLA. Although most of the works consider specific features, it may be possible to extend their procedures to other features. Moreover, some works do not focus on specific channel features in their schemes but rather aim to design solutions that work for any feature selected by the user.

For each channel feature, we will also include a discussion on techniques used to localize the attacker, if any.

A. Channel Impulse and Frequency Response

One of the most popular channel features for PLA is the entire CIR. Apart from channel estimation techniques that are common to all approaches, it is possible to distinguish between two main paradigms in their use. In the first approach, raw CIR is used for authentication, which has been proposed in different contexts, e.g., 6G [143], UWAC [147], and more recently in VLC [152]. In the second approach, preprocessing is performed on the estimated CIR, and a new metric is extracted. Often, such a metric is the result of the comparison between the measured CIR and a database of previously collected (and trusted) responses. Examples of this approach are for the UWAC context, the Froebenious norm in [150], the time reversal-based metric in [148], the maximum and minimum correlation amplitude in [151], while in the radio context, both the Euclidean distance and the Pearson correlation coefficient in [149]. A related approach for WiFi was also proposed in [132] where a channel feature is obtained by comparing the measured CIR with a locally generated replica.

The CIR/CFR is also used in CR-PLA. Some of them refer to the CIR over the multiple-input multiple-output (MIMO) channels. In this case, by increasing the number of antennas in either or both the transmitter and the receiver, we obtain a richer description of the channel that typically improves the accuracy of the authentication procedure. In particular, CR-PLA has been applied to scenarios with RISs [128], where the verifier controls the RIS, and the channels between the devices and the RIS are MIMO. Another context where CR-PLA has been applied, still on the CIR/CFR of the channel is in communications with drones, where the channel variations are due to the movement of the drones, not necessarily equipped with multiple antennas [136], [137]. Similar approaches of CR-PLA have been proposed in other frameworks. For example, [138] proposes a scheme in which the CIR is used to hide both the challenge and the response of the attacker using wiretap coding techniques; an extension of this approach includes the use of artificial noise [139].

An alternative to CIR is CFR, which can be easily obtained from the CIR by FFT, or is immediately available in communication systems operating in the frequency domain, e.g., the OFDM. Examples of application domains where CFR is used for CB-PLA are IoT/industrial IoT [160]–[164], cellular systems [166], [167], or WiFi [153]–[157]. Other works exploit a dataset of previously collected CFRs to derive a metric, e.g., in [162] the authors compare the performance when using the whole CFR matrix as input rather than the difference between a reference channel matrix and a measured one. Several works exploit a database of previously collected CFRs to predict the current one, which is then compared to the measured CFR to verify authenticity, e.g., [159]. The CFR in an OFDM system can also be interpreted as an image and then used to feed deep learning models [165].

TABLE VI
LITERATURE CLASSIFICATION OF CB-PLA MECHANISMS.

Feature list	Radio (Air)		Acoustic Underwater		Visible Light (Air)
	Statistical	ML	Statistical	ML	Statistical
CIR	[128]–[130], [132]–[140]	[141]–[146]	[147]–[149]	[150], [151]	[152]
CFR	[153]–[155]	[156]–[168]			
RSS or SNR	[135], [137], [169]–[172]	[131], [141]–[143], [173]–[181]		[182]	[183]
Channel Statistics	[184], [185]	[143]	[186], [187]	[188]–[190]	
Time Measurements	[191], [192]	[141]	[193]		
AoA	[175], [194]–[198]	[175], [176], [199], [200]	[193], [201]		
Doppler Frequency	[169], [178], [202], [203]	[142], [173], [177]			

Although CIR and CFR provide a complete description of the channel, resulting in a very detailed channel feature for PLA, their estimate is subject to several limitations that either reduce the accuracy of the authentication or require additional processing. Indeed, the main problem is related to synchronization, which may differ upon reception of different messages. Synchronization errors lead to time and phase shifts in the baseband equivalent CIR, and correspondingly to phase changes in CFR. Moreover, the estimate of each channel tap is affected by estimation noise that may significantly change the resulting CIR and CFR.

B. Received Power

To address the issues of CIR and CFR features, it is possible to use *channel parameters* either directly estimated from the received signal or extracted from the estimated CIR and CFR, which are robust to impairments. The first case is given by the received power, which is a channel parameter basically unaffected by small synchronization errors. Many receivers already compute the received power, for example, for signal quality assessment, range, or power control; thus, no extra effort is required to obtain the input of PLA mechanisms. Note that the received power is also denoted as received signal strength (RSS), while the estimated SNR is analogous to the received power, apart from a normalization of the noise power.

Excluding the fading component, the received power is related to the path loss and thus the relative distance between the transmitter and Bob. Moreover, in a context where Alice-Bob's distance is known, Bob can compute the reference power to be compared against the measured one for authentication purposes. Examples of this approach include the use of the norm between the measured and the expected RSS [131], SNR [175], received power [169], and the detection of anomalous path gains via ML [199], eventually also including dynamic scenarios [174].

Thus, many works propose the use of this already-available feature for PLA purposes in different contexts such as WiFi, vehicular ad-hoc network (VANET), UWAC, and VLC, as now discussed. For instance, in WiFi, the variations of RSS due to the movement of devices have been exploited in [171] to determine whether two pairing devices are in physical proximity to each other, thus authenticating their proximity. In [141], an reinforcement learning (RL) mechanism is used in the VANET context to dynamically adjust the authentication test threshold based on the previous results, including false alarm rates and authentication policy parameters.

Concerning the UWAC context, the normalized sample covariance matrix (SCM), i.e., the power covariance measured at different receivers, is used in [182] to estimate the transmitter position, and later for channel variation PLA.

In the VLC context, a first study of CB authentication is [183], where the attacker transmits when the legitimate transmitter is idle; then LOS direct current channel gain is used as an authentication feature. Then, [152] applies the CR-PLA technique to VLC. In this case, a RIS that operates on visible light signals is configured randomly to enable the receiver photodetector to authenticate the transmitter.

Attacker Localization: The RSS can be used further for attacker localization [179]–[181], typically exploiting ML, as we will detail later in Section X-E. However, RSS is known to be not adequate for localization as RSS-based localization methods may be vulnerable to beamforming attacks [204].

C. Channel Statistics

Beyond synchronization and estimation noise issues, in some contexts, the use of CIR and CFR becomes problematic by fast channel variations. In this case, it is better to use as a channel feature the *statistics of the channel* rather than its instantaneous realization.

Concerning the industrial IoT context, [177] uses the mean and variance of the subcarrier amplitude, carrier phase, and carrier frequency computed over the whole CIR.

A wide range of features has been tested in [185] in the UWB context, with results suggesting skew and kurtosis of the

tap's magnitude, maximum peak-to-earlier peak ratio (MPEP), and the peak-to-average power ratio (PAPR) to be the most promising features.

This approach is also popular in the UWAC context. In particular, in [186], while looking for features that are stable over time but not over space, several channel features have been investigated. The study concluded that the best performance was achieved using the number of channel taps, the average tap power, relative RMS delay, and smoothed received power. Such features also proved their effectiveness in the following works in authentication [188], [189]. In a dynamic UWAC context, the power-weighted arrival delay has been proposed to track the evolution of the channel; therefore, anomalous changes were associated with the start of an attack [187]. The work was extended in [190], where the new feature was integrated with those derived in [186].

D. Time Measurements

In many scenarios, power-related features are not usable, as they are either too predictable by the attacker or too variable to be used for CB authentication, e.g., due to fading. An alternative is offered by time-based features, such as time of arrival (TOA), time difference of arrival (TDOA), or even the estimated transmitter clock bias. Still, it is worth pointing out that, while effective, the use of the TDOA requires the presence of multiple synchronized receivers. The TDOA is used in the satellite context in [191] and in industrial/UWB communication in [192], while the TOA has been used in the UWAC in [193], along with the angle-of-arrival (AoA).

E. Angle of Arrival

For a receiver equipped with multiple antennas, the AoA is another channel parameter to be used as a channel feature for PLA. Indeed, the AoA is related to the transmitter position; thus, exploiting, for instance, a prior knowledge of the legitimate transmitter position, it is possible to discriminate between the legitimate transmitter and the spoofer just by looking at the AoA.

The AoA has been used in UWAC in both [193] and [201] and in [194] to profile the client's WiFi network. For IoT authentication, AoA has been exploited in [199].

In [200], the authors exploit the massive-MIMO geometrical channel to extract an image of angle and delay and then adopt ML techniques to authenticate the transmitters.

Attacker Localization: It is also possible to use the AoA of the received signal to localize the attackers [205]. SecureAngle [197] is a framework to estimate the signal's AoA and create AoA-based signatures to identify the legitimate users. If a user's signal does not belong to the authorized signatures pool, it gets rejected and localized by using the AoAs of the direct path of its signal, estimated by multiple access points. Pilot spoofing attacks are addressed in [195] and [196]: in such attacks, a spoofer corrupts the initial channel estimation phase by sending the same pilot sequence as the legitimate users at the same time. In particular, [195] employs an uplink and downlink training phase to detect and localize an attacker using the spatial spectra on the received signals and exploiting the reciprocity of uplink and downlink channels in time-division-duplex (TDD) systems. Still, due to the duration of the training phases, the method is vulnerable to environment changes (e.g., a moving attacker) [196]. Thus, [196] proposes an uplink joint detection and localization of an attacker via sequential Bayesian inference (i.e., by considering the time correlation on the estimated quantities).

F. Doppler Frequency

The use of the Doppler frequency shift is particularly popular in the satellite communication context. Indeed, analogously to the power-based approaches, a receiver that knows the position and velocity of both itself and the satellites can compute the relative velocity and thus the Doppler shift. Such an estimate can then be compared to the measured Doppler for PLA for authentication purposes. This approach has been used, for instance, in [202], where the receiver computes its position and velocity via GNSS, and in [203], where the authors tackle the problem of inter-satellite link authentication.

Often, the Doppler frequency shift is used in pair with the RSS,e.g., [169], [173].

Attacker Localization: Doppler frequency is also used to localize the attacker. The scenario considered in [206] sees a vehicle to vehicle (V2V) communication system attacked by a fixed or mobile terminal that is spoofing a GNSS signal. The vehicles use commercial GNSS receivers to measure the spoofer signal Doppler frequency. Next, the vehicles share their local measurements with the others, and by combining them, they localize the attacker. Note that, as all the vehicles are locked on the same spoofing signal, no additional synchronization among them is required.

G. CB-Authentication With RISs

If a verifier-controlled RIS is available in the network, specific solutions can be implemented.

In [184], a generalized likelihood-ratio test (GLRT) technique is used, but the second-order statistics of both the legitimate and the attack channel are known. In this case, the configuration of the RIS is fixed, and both the direct channel and the channel through the RIS are estimated for the hypothesis testing procedure. In [140], the impact of residual hardware impairments on authentication mechanisms in the presence of a RIS is investigated. In [198], it is proposed to use the AoAs of the direct and cascaded links at Bob and the effective angle-of-departures (AoDs) at the RIS. The sparsity of the direct channel and also the unique double-structured sparsity of the beamspace cascaded channel are exploited as authentication features.

In all these works, the configuration of the RIS, i.e., the setting of the phase of the elements, is fixed and typically optimized to maximize the communication performance. However, as already mentioned in Section VIII-B, the possibility to control the propagation characteristics of the channel with a RIS allows a new mechanism for authentication exploiting the CR approach. In fact, the use of RISs for this purpose was introduced in the first paper of the topic [127], and the security and communication performance were then studied in [128].

Specific attacks and defense strategies (both for the control of the RISs by the verifier and for suitable beamformers to be used by the attacker to increase the chances of success) have been investigated in [129] and [130].

X. Identification Verification Methodologies

In this Section, we describe in detail the identification verification phase of CB-PLA mechanisms, i.e., the part where the verifiers check that the currently received message is authentic. The description is organized into three parts related to statistical approaches, binary classification approaches, and one-class classification approaches. Lastly, we also provide a survey of techniques for localization based on ML models.

A. Statistical Approaches

With statistical approaches, we assume to have available probability density function (pdf) of the channel feature in either or both the legitimate and under-attack conditions. The authentication process is then seen as a hypothesis testing problem, and the test function is obtained from the pdfs.

1) Tag-based Authentication: Concerning the tag-based authentication, referring to binary hypothesis testing, the likelihood ratio test (LRT) is shown (by the Neyman-Pearson theorem) to minimize the missed detection probability for a fixed false alarm [207]. It provides the test function

$$V: \quad f(\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\mathcal{H}_0)}{p(\boldsymbol{x}|\mathcal{H}_1)}\,, \tag{21}$$

where $p(\boldsymbol{x}|\mathcal{H}_i)$ is the pdf of the tag in case $\mathcal{H}_i$ computed in $\boldsymbol{x}$. Such a test has been used in several works, such as [148], [201], [203]. In particular, in [201], the verifier Bob, upon receiving a new message, computes the Mahalanobis distance between the current observation and a database of previously collected AoAs containing both legitimate and non-legitimate samples.

Still, (21) has a major drawback as it requires the verifier to know, or at least assume to know, both legitimate and under-attack tag statistics, which may be a strong assumption in many practical applications.

In a single-sided testing problem, where only the tag statistics in legitimate conditions are known, the likelihood test (LT) is typically employed, which provides the test function

$$f(\boldsymbol{x}) = p(\boldsymbol{x}|\mathcal{H}_0)\,. \tag{22}$$

In the specific case of a Gaussian-distributed vector, e.g., when the measurement is affected by AWGN, (22) becomes

$$f(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{x}'\|\,, \tag{23}$$

where $\boldsymbol{x}'$ is the expected observation, which is used as a reference. Such an approach has been used, for instance, in [191], where the RMS error between the measured and the expected TDOA is thresholded. Still, it is worth noting that the LT is typically sub-optimal with respect to the LRT, but do not assume any knowledge of Trudy's attack statistical distribution. Such an approach has been used in [131], [153]–[155], [170], [171], [183], [186], [193], [202].

Alternative tests to (22) have been considered, for instance, resorting to the Pearson correlation between the different observation sequences [171], [194]. The Pearson correlation factor between the scalar feature sequences x_i, $i = 1, \ldots, n$, and x'_i, $i = 1, \ldots, n$, is

$$r = \frac{\sum_{i=1}^{n} (x_i - \mu(x))\,(x'_i - \mu(x'))}{\sqrt{\sum_{i=1}^{n} (x_i - \mu(x))^2}\sqrt{\sum_{i=1}^{n} (x'_i - \mu(x'))^2}}, \tag{24}$$

where $\mu(x) = \frac{1}{n}\sum_{j=1}^{n} x_j$ and $\mu(x') = \frac{1}{n}\sum_{j=1}^{n} x'_j$. For example, in [171], *witness* external devices check if two devices that should be paired are in proximity by controlling if the RSS are correlated. This approach is based on the assumption that close-by transmitters will yield correlated time series of RSS to a receiver.

Such tests are often integrated with more complex protocols. A common scenario considers a verifier that coordinates several independent receivers or channels and has to aggregate the local decisions to perform authentication. The local decision is often performed considering either (22) or (23), and then the verifier has to design a function to perform the final decision. A major difference between these works is that while some share with the verifier the soft output, others share only a local decision, i.e., the binary output of the combination between (22) and (18). In this context, in [203], 6 channels are considered, and several methods have been investigated to aggregate the local decision, in particular, OR, AND, and majority rule. On the other hand, a distributed test is considered in [186], where soft information is provided by the devices, and the aggregation is performed by weighting the local observation, considering, for instance, the distance between each receiver.

A different approach is proposed in [147] where the authentication is framed as a game, where the legitimate party utility function is a mixture of FA probability, MD probability, and spoofing cost; the legitimate party and the attacker have to choose the test threshold value and the spoofing probability respectively. A similar approach has been proposed in the satellite context in [169].

2) Channel Variation PLA: The most popular statistical method for channel variation PLA involves the use of Kalman filters (KFs) predictors. In detail, considering a characteristic $\boldsymbol{z}_i$ to be tracked (e.g., the user distance or velocity), typically called *state*, its time evolution is modeled as

$$\boldsymbol{z}_i = \boldsymbol{A}_i \boldsymbol{z}_{i-1} + \boldsymbol{w}_i\,, \tag{25}$$

where $\boldsymbol{A}_i$ is the state transition matrix at time-step i, and $\boldsymbol{w}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_i)$ represents the process noise (assumed to be Gaussian). On the other hand, the *measurement*, which for us is the channel feature, and the state are related

$$\boldsymbol{x}_i = \boldsymbol{B}_i \boldsymbol{z}_i + \boldsymbol{r}_i\,, \tag{26}$$

where $\boldsymbol{B}_i$ is the observation matrix and $\boldsymbol{r}_i \sim \mathcal{N}(\boldsymbol{0}, \Sigma_i)$ models the observation noise.

The Kalman filter has two operating modes: *prediction* and *model update*. During the former, it computes the a priori state estimate and its covariance matrix, respectively, as

$$\tilde{\boldsymbol{z}}_{i|i-1} = \boldsymbol{A}_i \tilde{\boldsymbol{z}}_{i-1|i-1} \,, \tag{27a}$$

$$\boldsymbol{P}_{i|i-1} = \boldsymbol{A}_i \boldsymbol{P}_{i-1|i-1} \boldsymbol{A}_i^{\mathrm{T}} \,. \tag{27b}$$

When a new channel feature is provided $\hat{\boldsymbol{x}}_i$, Bob can refine its model as

$$\boldsymbol{e}_i = \hat{\boldsymbol{x}}_i - \boldsymbol{B}_i \tilde{\boldsymbol{z}}_{i|i-1} \tag{27c}$$

$$\boldsymbol{C}_i = \boldsymbol{B}_i \boldsymbol{P}_{i|i-1} \boldsymbol{B}_i^{\mathrm{T}} + \Sigma_i \tag{27d}$$

$$\boldsymbol{G}_i = \boldsymbol{P}_{i|i-1} \boldsymbol{B}_i^{\mathrm{T}} \boldsymbol{C}_i^{-1} \tag{27e}$$

$$\hat{\boldsymbol{z}}_{i|i} = \hat{\boldsymbol{z}}_{i|i-1} + \boldsymbol{G}_i \boldsymbol{e}_i \tag{27f}$$

$$\boldsymbol{P}_{i|i} = (\boldsymbol{I} - \boldsymbol{G}_i \boldsymbol{B}_i) \boldsymbol{P}_{i|i-1} \,, \tag{27g}$$

where $\hat{\boldsymbol{z}}_{i|i}$ and $\boldsymbol{P}_{i|i}$ are the a posteriori state estimate and its covariance, respectively, while $\boldsymbol{G}_i$ is often called *Kalman gain*. Finally, the prediction error $\boldsymbol{e}_i$ is called *innovation* of the Kalman filter and can be used for security purposes. In particular, Bob computes

$$\beta_n = \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{C}_i^{-1} \boldsymbol{e}_i \,, \tag{28}$$

which Bob uses as input for authenticity verification. We remark that, differently from the general model of the Kalman filter, we have no control input. The interested reader may look for a more formal introduction of the KF in [208]. Several variations can be introduced to the KF. For instance, in the so-called extended KF non-linear equation, replace either or both (25) and (26).

The main issue of statistical approaches, such as the KF, is that they require an analytic model. For instance, KFs typically requires an a priori knowledge of the measurement evolution over time and the relation between the measurement and the true state (e.g., the transmitter position), e.g., $\boldsymbol{B}_i$ and $\boldsymbol{A}_i$ used in the (linear) KF in (25) and (26), respectively. However, such models are often hard to obtain in practice, as they depend on both the features themselves and the environment, thus limiting the set of possible input features.

In the UWAC context, in [187], a set of receivers is tracking the power-weighted arrival delay using a KF. The KF so-called *innovation*, which measures the discrepancy between the value predicted by the KF and the observed measurement, is monitored, thus associating high innovation values with the start of a spoofing attack. Such an approach was generalized in [190], replacing the KF with a LSTM neural network (NN). A similar approach has also been considered in [135], [172], targeting the V2X scenario.

As an alternative to the KF approach, in the MIMO radio context, [134] considered a scenario where the CIR evolves following a Gauss-Markov process. The tests evaluate (the norm of) the difference between the current and the previous CIR at different transmitter-receiver pairs, considering both the cases where each pair observes statistically independent channels, and where the observed channels are correlated. Still, we notice that the check is still related to (28), eventually considering the covariance to be a diagonal matrix.

Differently, the work in [209] considers an online adaptive method where the threshold is dynamically adjusted by the verifier, according to its previous statistics. Another adaptive method has been proposed in [132], where the tag symbol setup is dynamically adapted, following a water-filling approach where the power associated with each tag symbol is adjusted to match a predefined detection probability.

B. ML Approaches - Binary Classification

In ML-based solutions, it is assumed to have a dataset of tag samples under one or both the legitimate and attack conditions. The authentication problem is framed as a classification problem, and the test function is replaced by an ML trained with the available dataset(s).

We first consider the binary classification solutions, where two datasets (of legitimate and attack conditions) are available to train the ML model. Note that the dataset of attack samples can also be artificial, i.e., generated by the verifier under suitable assumptions, when no real data are available, as discussed in Section X-C. The availability of the Trudy dataset is related to some assumptions, i.e., expected position, type of device, or resulting channel feature. To the best knowledge of the authors, no work has considered the use of two-class ML techniques for channel variation PLA and CR-PLA. Thus, all results are relative to the tag-based PLA.

In [124] it has been proven that a sufficiently complex NN trained with a sufficiently large dataset, containing both positive (Alice) and negative (Trudy) samples, when used in (18) achieves the same performance as the optimal LRT.

The work in [162] compares the performance of four standard classification algorithms, namely decision tree (DT), SVM, KNN, and enseamble learning (EL), in particular bagged trees (BTs). The last achieved the best performance in both simulations and tests, but at a higher computational cost with respect to the other solutions.

In [144], a CNN is used to extract the relevant features from the channel observation, and then a RNN is expected to learn the spectral dependencies between the extracted features. In [200], the authors propose to use the sparse nature of the channel in a massive MIMO-OFDM communication system to first build an angle-delay image that is efficient for NN training. Then, they exploit you-only-look-once (YOLO), an advanced single-stage object detection network, to capture the angle-delay features from the received signal, and finally, a lightweight neural network to perform the classification.

We remark that IoT is a relevant use case for PLA, where devices have limited computing capabilities; therefore, ML solutions can be particularly useful. For example, in [199], a two-step authentication mechanism for IoT devices in 5G networks is proposed. The first step aims to detect anomalies in the virtual AoA and path gains of all the IoT devices in the cell; the second step provides an authentication mechanism based on ML. In particular, the anomaly detected in the first step compares the number of communications at the MAC layer with those identified at the physical layer. If the anomaly is detected, a ML is used to decide if a single communication (at the MAC layer) comprises more than one path at the

physical layer to confirm the anomaly. The work in [161] exploits the presence of multiple devices at the edge to build a collaborative authenticator. In particular, the CFRs associated with a MIMO system are divided among different devices of an edge network, where a subclassifier makes a first classification, which is interpreted as a vote. Finally, the decision is taken by aggregating the single votes, each associated with a proper weight. Different techniques are proposed in [164], [166] to reduce the computation load, thus allowing power-constrained devices to perform PLA. In [164], a convolutional denoising AE is used to preprocess the raw CSI measurements to reduce the dimension of the features, eliminate noise, and extract key features. After the pre-processing, a weighted KNN algorithm classifies the extracted features and authenticates the transmitters, which reduces the computational overhead compared to other ML approaches. [166] aims at reducing the computation overhead in training the NN, thus proposing three gradient descent algorithms to accelerate training.

C. ML Approaches - Artificial Dataset

When the dataset available for training contains only channel feature samples under legitimate conditions, two approaches are possible. One solution provides that first, an *artificial dataset* of attack channel features is generated, and then the binary classification approach is used for training on the available legitimate and artificial attack datasets. A second solution uses only the available dataset, which is denoted as a one-class classification approach. In this section, we consider the solution based on the artificial dataset, while in Section X-D, we detail the one-class classification approach.

1) Tag-based Authentication: The solution based on the artificial dataset is employed in [160], [177].

A DT is exploited to perform authentication in [177], in the absence of an attack dataset. In particular, they train the classifier using positive unlabeled data, where only positive (i.e., legitimate) data is used, but part of the data is labeled as non-legitimate and left out during this first training. Then, the procedure is repeated, changing positive and unlabeled data until a robust enough classifier has been trained.

An automated labeling strategy is proposed in [160], which comprises an offline and online procedure. They specifically look for clone or Sybil attack data samples, which are later used to train the more refined SVM-based (online) procedure.

2) Channel Variation PLA: The solution based on the artificial dataset is employed in [149], [151], [174], [210].

An extreme learning approach is proposed in [149], where the input contains both previous trusted observations and the observation to be verified; then, the network is trained to check the coherence between the two. An artificial dataset is generated to improve the network classification performance. In particular, the artificial dataset has the same distribution as the legitimate one but contains samples that are uncorrelated with the (previous) legitimate ones.

An artificial dataset is considered in [151], and it contains samples from nodes that are far from the receiver as negative. Then SVM is used to build a classifier. The features, e.g., the maximal time-reverse resonating strength, are then specifically chosen to minimize the impact of the channel time variability.

Using channel measurements relative to a specific location of devices as an authentication feature makes the authentication very scenario-dependent; thus, channel time-varying patterns as scenario-independent features are used in [210] to authenticate devices even in uncalibrated scenarios, including both CIR and CFR as input to the ML model.

The use of a graph NN (GNN) was considered in [174], where the verifier has to decide whether the RSSs measured from several receivers within a frame comes only from Alice or from multiple transmitters, that is, Alice and Trudy. In particular, a two-step approach is proposed. First, a NN checks whether two consecutive transmissions come from the same position. The latter step has to split the received frame sequence into regions associated with the same transmitter. In particular, in the first step, the training dataset collects pairs of consecutive transmissions where i) in the legitimate case, the transmission comes from transmitters in the same position, while ii) in the under-attack case, the pair contains samples coming from different randomly sampled positions.

D. ML Approaches - One-class Classification

We now consider one-class classification solutions, where only a legitimate dataset is used to train the classifier.

1) Tag-based Authentication: The first set of solutions collects anomaly detection models, such as AEs or one-class support vector machine (OC-SVM). For example, a OC-SVM is considered in [145], with both magnitude and IQ samples of CIR measured at different antennas as input. OC-SVM has been also considered UWB [185]. In [165], a NN determines the position of a device (from a set of discrete positions) from the observed CSI. If the confidence metric associated with the predicted position is below a threshold, the device is considered not authentic. Indeed, this authentication mechanism boils down to a region location verification, i.e., it verifies that the device is in a set of pre-approved positions. The combination of an AE for dimensionality reduction and variational autoencoder (VAE) for its generative capabilities has been considered in [146].

Cooperative One-Class Classification: When multiple detectors are available, cooperative CB-PLA can be applied. In fact, multiple verifiers can cooperate in the final decision by merging the collected data (or local decisions) to form a distributed authenticator and improve the security performance.

Two-step approaches are considered in [143], [156], [168], where first clustering-based methods are used to detect the presence of outliers within the dataset. Then the authentication is completed by using SVM [168], an ensemble of AEs [156], or graph learning [143] approaches.

In [188], the authors compared binary NN and one-class AE for UWAC networks. The same techniques have also been considered in [189], where the authentication process is performed in two steps: a first pre-elaboration on each device, with a NN, and then a second central elaboration made on the pre-elaborated data, using a second NN.

2) Channel Variation PLA: A kernel-based prediction method is proposed in [142], and Gaussian process regression (GPR) is considered in [158]. The use of a Bahdanau attention AE predictor is considered to predict the current CFR

in [159]. Compared to a more traditional AE-based predictor, this architecture includes an attention mechanism to take into account the effects of fading and Doppler shift.

A second solution aims at extending the solution developed for the static to the dynamic context, thus assuming the distribution to change slowly over time. For example, in [167], the concatenation of the legitimate features (tested) before and the measurement under test is fed to CNN, whose output represents the probability that the whole input matrix belongs to Alice. OC-SVM is used for classification in the satellite context in [173] but is progressively updated to take into account the evolution of the statistic over time.

A third alternative is to optimize the test threshold. The authentication problem has been formulated as a zero-sum game, where RL, in particular, Q-learning and Dyna-Q, have been used to optimally set the threshold for LT in [131]. In [141], a deep reinforcement learning (DRL) mechanism is used for the VANET context where Bob sets the authentication threshold and Trudy plays with the attack rate.

The tag-based solution in [168] has been extended in [157] to also take into account mobile users by adding a processing block that monitors the temporal correlation between subsequent CSI blocks.

In the context of UWAC, Casari *et al.* investigate the use of OC-SVM and an AE to fuse the local KF's innovations [187], later extended in [190], where a RNN allowed to track features that cannot have been analytically modelled and thus tracked by KF. In [163], a framework considers federated learning among cooperating edge devices: a group of edge devices is selected using a Q-learning-based adaptive search procedure and collaborates to form an authenticator. Unlike previous works, two threats are examined: the presence of external attackers (i.e., regular spoofers) and internal threats, the latter represented by malicious nodes injecting false parameters that might lead to convergence failure or convergence to a wrong model. A similar approach is also proposed in [182], where first, a CNN estimates the transmitter positions from the SCM, and next, a RNN predictor tracks the transmitter position.

E. ML Approaches - Attacker Localization

Various approaches for the location of the attacker have been considered using ML models. In [179], the authors use the K-means clustering algorithm to detect and locate attackers. In particular, after a training phase, they combine RSS readings from multiple access points and divide them into clusters in the signal space. If there are multiple transmitters at the same time (i.e., a legitimate party and one or multiple spoofers), then the resulting point in the signal space will be far from the centroids of the legitimate clusters, and thus an anomaly is detected. After the spoofer is detected, the cluster centroids are used to localize it. The authors experimentally tested their approach on WiFi and Zigbee networks, reaching $P_{\rm md} > 0.95$ with $P_{\rm fa} < 0.05$.

A drawback of [179] is that it cannot localize multiple attackers. Thus, in [180], the authors propose the IDOL (Integrated Detection and Localization) framework, capable of detecting and localizing multiple attackers. In [180], three types of algorithms were tested to locate the attackers: nearest neighbor matching in signal space, probability-based, and multilateration; while in [181], the authors used a discriminant-adaptive neural network to perform the same task.

XI. CB-Authentication Public Datasets and Applications

This Section describes the common methodologies to test the performance of the proposed CB-PLA mechanisms, considering both simulation tools and experimental datasets.

A. Simulation and Experimental Methodologies

The methodology for CB-PLA mirrors the one used for RFFI, described in Section VII. It typically involves the collection of two datasets. The first is used to set up the detector. For instance, in KF-based channel variation authentication, the first dataset is used to adapt the filter parameters during the initial transient (e.g., [182], [187]). In ML-based solutions, such a dataset is instead used to train the detector. The second dataset is used for testing, to evaluate the performance of the trained detector. We remark that, while formally two datasets need to be collected, often only one is collected, which is then split into two. These operations need to be performed carefully. For instance, training on samples that are collected close to each other in time helps the detector to learn the channel stationarity, and thus, we neglect the evolution of the channel over time. On the other hand, as it happens when overfitting, this also makes the detector less robust to variations. Thus, to make the detector more robust to temporal variations, it is advisable to split the dataset randomly.

Finally, it should be noted that when testing CB-authentication solutions, the hardware impairments are typically neglected, implicitly treating them as estimation noise. Indeed, even if costly, a better practice would involve the collection of multiple datasets, each collected with a different transmitter/receiver hardware pair, which would make the detector, trained on the merged dataset, truly device-independent. Alternatively, future works should include an estimation and correction step to correct the hardware impairments or, even better, a joint RFFI & CB-based authentication, which allows the detector to exploit both techniques at the same time, as detailed in Section XII.

B. Simulation Tools and Setups

WiFi: A broad set of simulation tools is used for the WiFi context. A simple model provides independent Gaussian distributed channel taps [134], while other solutions, such as [142], consider generating more realistic CFRs, including an exponential PDP. Other parameters have been set according to the IEEE 802.11a specification. Another alternative is to consider geometric models, e.g., ray tracing tools, as in [154].

Vehicular: Specifically targeting the V2X context, simulations have been performed in [135], [172], with communication parameters adapted from the SAE J2945/1 standard using Matlab. In [135], the authors simulated two traffic scenarios: straight and intersection. In the first, Alice and Bob are driving

straight on the same road, while in the second, Eve follows Alice who is driving crosswise with respect to Bob. On the other hand, [172] considers instead a more abstract model, where the RSS is modeled after a log-normal distribution.

Cellular Wireless: While some papers generate the features via statistical models (e.g., [199]), three simulators are popular in the literature:

- Quasideterministic radio channel generator (QuaDRiGa) channel simulator [211]: it has been used in [158], [165]. In particular, in [158] the simulation includes also the movement between transmitter and receiver, with parameters set to simulate the ground city macrocell in the Berlin survey in Germany (BERLIN UMa).
- WINNER II channel model [212]: it is used for instance in [145] to model a non-line-of-sight scenario, with users moving at different velocities.
- MATLAB 5G toolbox channel: this was used to test the Bahdanau attention autoencoder proposed in [159] operating in the 5G FR1 n78 band.

Underwater Acoustic: Concerning PLA in the UWAC context, the most popular solution involves the Bellhop ray-tracing simulator [213], [214], used for instance in [182], [186]–[188], [190]. Such a tool also includes the description of several environmental parameters, such as sound speed profile and bathymetry. For instance, among others, the San Diego Bay area was considered in [182], [186], [187], [190].

An alternative simulator used in [151] has been described in [215]. Finally, in [193], the performance is evaluated considering both an AWGN channel and a coloured noise channel with and without frequency-dependent path loss, respectively.

Satellite: Different channel models are considered in the satellite communication, including both AWGN [173] and Rician fading channel [202]. Specifically concerning satellite orbit datasets, two-line element (TLE) datasets have been used in both [169], [191]. In the former, the dataset was derived from [216], while in the latter via the Ansys STK [217].

UWB: Concerning UWB, a MATLAB simulation is performed in [192], modeling the UWB signals as the first derivative Gaussian pulses, and the channel is modeled as AWGN. To test the performance, [185] considers both simulation and experimental tests. The simulations have been performed in MATLAB with receivers implemented following the IEEE 802.15.4z standard and channels compliant with the IEEE 802.15.4a standard. In particular, both LOS and NLOS have been included.

C. Experimental Setups and Public Datasets

Table VII collects a list of datasets available online that may be used to develop/test new channel-based authentication techniques, classified by wireless technology and measured channel features. In the remaining part of the section, we discuss the use of simulation data, experiments, and datasets in wireless technology.

WiFi: Several works have provided experimental results on WiFi networks. First, a public dataset containing RSS and channel measured from WiFi access points, placed at increasing distance from a reference transmitter [221].

In [224], a SDR platform for the WiFi PHY layer has been implemented and CB PLA is performed on the CSI, RSS, and frequency offset. In [153], the CSI was considered in a typical indoor scenario with fixed locations of the users.

A dedicated prototype has been developed to test the performance of the AoA-based PLA solution described in [194]. The developed access point has two FPGA platforms, with four radio front ends and four antennas each. The clients are two Soekris boxes, equipped with Atheros IEEE 802.11g radios.

For the proximity-based PLA solution of [171], experiments were performed with ten Nokia N800 Internet Tablets, showing that the proposed solution can reliably detect attackers as close as two meters away from legitimate devices.

Experimental results on the clustering-based approach of CFR for authentication have been reported in [157], [168] where an IEEE 802.11n WiFi network was considered, with two laptops (Lenovo T500 and T61) serving as monitors that collect the wireless packets. A commercial wireless Linksys E2500 access point is the device to be authenticated, transmitting 10 packets/second. For each packet, the CFR relative to 30 subcarriers is extracted with equal spacing among the 56 subcarriers of a 20 MHz channel.

In [156] experimental results are reported with a commercial WiFi device, Huawei TAS-AN00 operating as a station, transmitting at a rate of 100 packets/second in $20\,\mathrm{MHz}$ WiFi the channel on $2.4\,\mathrm{GHz}$.

The experimental results reported in [155] have been performed on the Microsoft Sora SDR, reaching a false positive and false negative ratio of 10^{-3}.

In [176], a WiFi operating at mmWave (60 GHz band) is considered, with reference to the IEEE 801.11ad standard. The considered feature is the SNR trace obtained at the receiver in the sector level sweep (SLS) process, and an ML approach is used to authenticate the message. Talon AD7200 routers and MG360 WiGig USB Adapters are used to perform experiments in a meeting room, achieving a sum of MD and FA probability less than 1%. In [175] experimental results for authentication are presented, based on the dataset of [176].

In [170], experiments are conducted for the authentication based on the verification of SNR series observed at Alice and Bob, through a statistical method. Alice, Bob, and Trudy are implemented on Dell E5400 laptops, which usethe Intel iwl5300 chipset, operating IEEE 802.11g with channel one in the $2.4\,\mathrm{GHz}$ frequency, with a transmission rate of 12 Mbps and transmission power of 15 dBm.

Many works perform dedicated experiments [131], [132], [144], [167], deploying three or more USRPs in an indoor environment mimicking an office or industrial context, with parameters following the standards, e.g., IEEE 802.11a/g and IEEE 802.11n/ac.

The GNN-based solution proposed for the channel variation in [174] in the artificial dataset training framework, exploits the dataset from [220], a publicly available WiFi fingerprint dataset which collects fingerprints collected with 21 devices in an indoor scenario.

IoT & Industrial IoT: As in the WiFi context, many works only perform dedicated experiments, using again USRPs de-

TABLE VII
SUMMARY OF PUBLICLY AVAILABLE DATASETS FOR CB AUTHENTICATION

Wireless Technology	Dataset	Paper	Features	Brief Summary
WiFi, IoT, Industrial IoT	[218]	[146], [160], [162], [164], [177], [210]	CSI	Data collected in industrial environments and open-area sites by NIST
WiFi	[219]	[175], [176]	SNR	SLS SNR traces collected from the communication between AP and clients at mmWave frequencies
	[220]	[174]	RSS	Fingerprints collected in an indoor scenario from 21 different Android devices
	[221]	–	RSS	Measured from two Raspberry PIs, placed at various distances from one another
UWAC	[222]	–	CIR	Experiment performed in Kauai (Hawaii) in 2011.
	[223]	[151]	CIR	Long-range experiment performed in the Mediterranean Sea in 2019.

ployed in an industrial-like environment, e.g., [161], [209], eventually also in a cooperative setting, such as [163].

In [166], the IoT nodes are emulated using USRPs, placed in various positions in an indoor environment.

A popular dataset for testing in the industrial IoT context is [218], described in [225], [226]. For instance, both the *open area test site* and the *automotive assembly plant* scenario datasets were used in [143]. The dataset was also used for training and testing in [146], [160], [162], [164], [177], [210]. Many works use both the NIST dataset and a dataset from dedicated experiments. A first example is [164], where the experimental dataset was collected using three Lenovo X220 laptops (Alice, Bob, and Trudy) placed approximately 2–4 m apart transmitting at $2.4\,\mathrm{GHz}$, using IEEE 802.11n protocol, with 3 transmitting antennas, 2 receiving antennas, and 30 subcarriers using Linux CSI Tool. In [210], the performance of the channel variation-based ML approach was tested using [226] for the static scenario, then dedicated experiments were performed using two USRPs to account for the dynamic one.

Underwater Acoustic: Due to the lack of a standard channel model and the impact of the environmental condition on the measurement, experiments, and proof of concept, often called *sea trials* are common in UWAC studies [186], [188], [189], eventually re-using datasets from previous experiments, such as in [188], where Bragagnolo *et al.* used the Hadera (Israel) dataset from [186]. However, only a few datasets are actually publicly available. For instance, a popular dataset is the KAM11, which was only published with the Watermark simulator [222]. An example of a public dataset is instead the LR19 [223] used, for instance, in [151].

Another alternative is to perform tests in (typically indoor) pools: in [150] an experiment is run by collecting measurements from a $25 \times 6 \times 1.6\,\mathrm{m}^3$ non-anechoic pool where 9 transmitters and 1 receiver were deployed.

Satellite: Since not many works consider PLA for satellite, very few experiments have been reported yet. An exception is [178], where the Abdrabou *et al.* collected real LEO satellite data using the system toolkit.

UWB: Dedicated experiments have been performed to test the performance of the solution proposed in [185]. In particular, two nRF52840-DK boards have been used to implement the legitimate users, while a NUCLEO-Z429 is used for the attacker, all equipped with Qorvo DWM3000 modules. 1000 CIRs have been collected in both static and dynamic scenarios.

XII. CHALLENGES AND FUTURE RESEARCH

A. Challenges for RFFI

Despite significant development in deep learning-driven RFFI technology, numerous challenges remain unresolved. This section elaborates on these challenges and presents a summary based on the most recent studies.

1) Lack of Capacity Evaluation: The term 'capacity' is used to describe the maximum number of wireless devices that can be accurately distinguished by analysis of their RFFs, which is critical for an authentication technique. Most RFFI studies use commercially available wireless transmitters, but most involve fewer than tens or dozens of devices, and few large-scale experiments have been conducted. To the best of the authors' knowledge, the work in [52] presents the experiment with the largest number of wireless devices, up to 10,000. However, this large dataset has not been published, and researchers in this community cannot use it to explore the maximum capacity of the RFFI technique. A few studies have attempted to provide a theoretical analysis of the user capacity in RFFI systems [227], [228]. Nevertheless, achieving an accurate prediction of the user capacity remains a significant challenge, particularly for deep learning-driven RFFI systems. There is still a need for large-scale experimental evaluation and theoretical analysis to assess the capacity of RFFI.

2) Lack of Stability Evaluation: As an identifier used for authentication, the stability of RFFs is critical. However, there are rare studies that systematically investigated the stability of RFFI systems. In particular, the characteristics of RF components can change slightly due to variations in the surrounding environment, such as temperature and humidity, and hardware ageing. The authors in [21], [22] indicate that the oscillator frequency is sensitive to temperature variations and that CFO compensation at the receiver side can improve system stability. However, these studies only focus on the stability of the frequency offset function resulting from the oscillator impairments, while the other hardware characteristics are not investigated. Moreover, the authors in [229] experimentally demonstrate that the RF fingerprints of SDR

transmitters exhibit significant variations when transitioning between on and off states. However, the evaluation of wireless transmitters beyond SDRs remains an open area for exploration. The comprehensive evaluation of the RFF stability and the design of robust feature extraction algorithms represent crucial directions for future research.

3) Lack of Benchmark Datasets: The RFFI research community does not have a benchmark dataset that is as widely used as ImageNet in computer vision. This limits comparisons among studies. In addition, researchers without RF hardware and experience in designing wireless signal acquisition systems cannot efficiently engage in RFFI research. As discussed in Section VI, some studies have released public datasets detailing the collection environments and hardware setup [29], [54], [230]. However, most of these datasets still do not meet benchmark requirements in terms of dataset size, device population, and diversity in channel conditions. The collection and publication of large-scale benchmark datasets of various wireless protocols remains an urgent need in the field of RFFI.

4) Limited Studies on Adversarial Machine Learning Attacks and Defense: Cutting-edge RFFI schemes heavily rely on deep learning. However, recent research in the ML community has revealed that deep learning is vulnerable to adversarial machine learning (AML) attacks, including in the context of wireless systems [231], [232]. Depending on the attack phase, AML can be categorized into backdoor attacks launched in the training stage [233], [234] and adversarial/evasion attacks launched in the inference stage [235]–[239].

The AML attacks can be launched during the model training stage, named backdoor attacks [233], [234]. Zhao *et al.* propose the first backdoor attack on RFFI systems, and evaluate the algorithm on three WiFi datasets and a LoRa dataset. The results demonstrate that the attack can be successfully launched in either the time domain or the time-frequency domain [233]. The authors in [234] further investigate the backdoor attacks against low-earth orbit satellite fingerprinting systems.

The AML attack can be launched during the model inference stage, named adversarial/evasion attacks [235]–[239]. For instance, Ma *et al.* demonstrate that adding perturbations to the deep learning input can interfere with the identification result and can even mislead into a specific identity [237].

More research is required to study the AML as well as the countermeasure. For example, the transmission of perturbation in evasion attacks will experience channel propagation, but the effect is not properly studied yet. Adversarial training and randomized smoothing are used as countermeasures for Wi-Fi sensing [240], but there is no such study for RFFI.

5) Limited Studies on RFF Concealment: While the majority of the research focuses on using RFF for legitimate purposes, i.e, device authentication, it can also be used maliciously, e.g., device tracking in [101]. Hence, it is essential to design RFF concealment approaches.

Abanto-Leon *et al.* added a randomized phase to each subcarrier in a WiFi OFDM system to ensure privacy, when non-linear phase errors are used for RFFI [241]. They proved that when the phase is generated via a random number generator, the approach is robust against statistical attack. Givehchian *et al.* obfuscated the CFO of BLE devices for preventing tracking attacks [102]. They implemented their CFO obfuscation method using TI CC2640 chipsets and carried out a comprehensive experimental evaluation, which demonstrated the feasibility.

These approaches only focus on the phase errors and CFO as hardware fingerprints, and their identification is based on comparing their similarities. However, deep learning RFFI is learning all the available hardware impairments. It is not clear how RFFI will be affected if only one hardware feature is obfuscated as other impairments remain the same.

B. Challenges for Channel-Based Authentication

1) Lack of Scalability: Attacks against CB authentication can be deployed by transmitting from different positions until the features estimated by Bob are similar to those of Alice's transmissions. An alternative is that the attacker precodes the signal before transmission to introduce the features suitable for authentication [9]. At the moment, the search space for an attack is limited, making the attack easy. Indeed, more efforts should be focused on the factors that make the attack harder in a scalable way. Such efforts include the investigation of a) scaling laws for the attack success probability with respect to design parameters such as the number of antennas or the length of pilot signals, b) new approaches such as the CR-PLA that introduce further randomness in the authentication process, c) new bounds on attacks based on physical constraints obtained from specific technologies (i.e., type of antennas used by Bob): a recent example is given by [242] that proved that an effective attack on AoA-based PLA can succeed only under very stringent conditions on the attacker location and hardware capabilities.

2) Lack of Integration: Since the first studies, PLA has been proposed as a security technique to be integrated with other approaches for authentication. However, such an integration has not been thoroughly investigated. A full protocol for PLA that integrates cryptographic approaches, for example, to secure the identification association phase or to be deployed when PLA is under attack, is yet to be investigated. Moreover, integration of CB-PLA with authentication based on wiretap coding is still in its early stages and deserves further investigation. Lastly, integration may also include the use of diversified features for authentication, also coming from different layers of the communication stack: this is also an area that deserves more studies. In this case, ML techniques could be particularly beneficial to capture the relation among the features, but such solutions should be, at the same time, effective in the specific scenario of deployment and robust against adversarial attacks, which leaves many open research points.

3) Lack of Benchmark Datasets: Also for CB-PLA as RFFI, there are not yet well-established datasets to be used for benchmarking different approaches. The difficulty of obtaining such datasets is related to the specific technologies that can be deployed (type of antennas, operating frequencies), the different kinds of environments in which the testbed operates (indoor, outdoor, with different transmit-receive distances), and the need to obtain measurements from several positions at

the same time to assess also the knowledge of the attacker and the statistical relation of his channel to the legitimate channels. In this sense, apart from more extensive data collection and the use of existing simulators that provide spatially consistent channel realization, the adoption of PLA techniques in the standard would encourage a discussion from the community on reference scenarios to be used for benchmarking, thus giving a boost to PLA adoption.

C. Future Directions

1) Generative AI Approaches: Generative AI represents transformative AI technologies to create new content, such as GAN and large language models (LLM). Generative AI has been widely used in securing communication from the physical layer [243], but its application in device fingerprinting is relatively limited.

Generative strategies/architectures include AEs, VAE, diffusion model, etc. They are used to design detectors in the anomaly detection context. They can also be used to generate the training dataset, e.g, VAE is used to generate satellite data [244]. This may allow a binary classification-based detector to have an initial offline training with artificial but realistic data, later refined online. In the context of anomaly detection, generative models may be used to generate an artificial dataset (see Section X-C). Regarding diffusion models, it is used for denoising in RFFI [110].

Recently, LLM have proven their effectiveness in multiple fields, even in the communication context [245]. Still, no solution that exploits LLM has been proposed in the device fingerprinting context. Due to their generalization capabilities and if trained in a multimodal manner, thus taking as input also information concerning, for instance, the environment, LLM may be used to generate high-fidelity artificial datasets, thus leading to even more robust detectors.

On the other hand, generative models may be used by the attacker to design effective attacks, as done in the RFFI context in [246]. In particular, an attacker provided with the legitimate detector (or dataset used for training it) may exploit a generative architecture to generate the attack samples that are most likely to fool the verifiers. Thus, future research directions should also include these attacks into account.

2) Emerging Communication Technologies: While the use of device fingerprinting for securing communication technologies such as WiFi is consolidated, for newer communication technologies, especially in the optical domain, only a few or even no work at all considers device fingerprinting for securing communication. This is the case, for instance, in VLC [247] or even underwater optical communications, where, to the best of the authors' knowledge, very little research has been done. Thus, a research direction may involve the translation of the more consolidated solutions and algorithms into these new technologies.

3) Interplay between RFFI and Channel-based Authentication: RFFI and channel-based authentication represent two distinct but complementary approaches to wireless device authentication. RFFI relies on the unique hardware impairments inherent to individual devices, which are introduced during the component manufacturing process. RFFI system is implemented at the receiver side, which is well-suited for scenarios where low-cost, infrastructure-independent security solutions are required. In contrast, channel-based authentication exploits the unique properties of the wireless channel, which are influenced by the surroundings; thus, it is effective in rich scattering environments, where it is hard for the attacker to predict, replicate, and compensate the attack signal to effectively mimic the legitimate channel features.

The combination of RFFI and channel-based authentication offers a promising solution to enhance wireless security. Hybrid authentication protocols can be designed: RFFI ensures device-level identification based on unique hardware characteristics, while channel-based authentication validates location or monitors channel characteristics within a communication session. In this case, attackers would need to simultaneously replicate both the hardware impairments and the exact channel conditions to bypass the dual-layer protection, significantly increasing the difficulty of attacks.

XIII. Conclusions

This article presented a comprehensive survey on physical layer-based device fingerprinting, focusing on hardware impairment-based identity authentication and channel features-based location authentication. In particular, RFFI exploits unique hardware impairments as devices are identified. Three RFFI tasks, closed-set RFFI classification, open-set RFFI recognition, and anomaly detection, were explained. The hardware impairments of both transmitters and receivers were modelled. A deep learning-based design was described. Three RFFI research topics, channel effects elimination, noise mitigation, and receiver distortion mitigation, were reviewed as they are essential for RFFI design. Finally, the experimental methodologies for RFFI was described. Regarding CB-based authentication, an overview of existing approaches has been provided, including both statistics-based techniques and ML solutions. Several features used for PLA have been introduced and discussed, and implementations have been classified based on the use of simulation or experimental tools. The remaining research challenges for both topics were discussed, and future research directions were suggested to make more robust device fingerprinting approaches.

References

[1] State of IoT 2024: Number of connected IoT devices growing 13% to 18.8 billion globally. Accessed on 7 Dec., 2024. [Online]. Available: https://iot-analytics.com/number-connected-iot-devices

[2] A. Burg, A. Chattopadhyay, and K.-Y. Lam, "Wireless communication and security issues for cyber–physical systems and the Internet-of-things," *Proc. IEEE*, vol. 106, no. 1, pp. 38–60, Jan. 2017.

[3] W. Trappe, R. Howard, and R. S. Moore, "Low-energy security: Limits and opportunities in the Internet of things," *IEEE Signal Process. Mag.*, vol. 13, no. 1, pp. 14–21, Jan. 2015.

[4] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on Internet-scale IoT exploitations," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2702–2733, Mar. 2019.

[5] J. Zhang, G. Li, A. Marshall, A. Hu, and L. Hanzo, "A new frontier for IoT security emerging from three decades of key generation relying on wireless channels," *IEEE Access*, vol. 8, pp. 138 406–138 446, July 2020.

[6] C. Cheng, R. Lu, A. Petzoldt, and T. Takagi, "Securing the Internet of Things in a quantum world," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 116–120, Feb. 2017.

[7] K. Zeng, K. Govindan, and P. Mohapatra, "Non-cryptographic authentication and identification in wireless networks [security and privacy in emerging wireless networks]," *IEEE Wireless Commun.*, vol. 17, no. 5, pp. 56–62, May 2010.

[8] J. Zhang, R. Woods, M. Sandell, M. Valkama, A. Marshall, and J. Cavallaro, "Radio frequency fingerprint identification for narrowband systems, modelling and classification," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 3974–3987, 2021.

[9] P. Baracca, N. Laurenti, and S. Tomasin, "Physical layer authentication over MIMO fading wiretap channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2564–2573, May 2012.

[10] Q. Xu, R. Zheng, W. Saad, and Z. Han, "Device fingerprinting in wireless networks: Challenges and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 94–104, Jan. 2015.

[11] P. M. S. Sánchez, J. M. J. Valero, A. H. Celdrán, G. Bovet, M. G. Pérez, and G. M. Pérez, "A survey on device behavior fingerprinting: Data sources, techniques, application scenarios, and datasets," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 1048–1077, Feb. 2021.

[12] R. R. Chowdhury and P. E. Abas, "A survey on device fingerprinting approach for resource-constraint IoT devices: Comparative study and research challenges," *Internet of Things*, vol. 20, p. 100632, Nov. 2022.

[13] V. Kumar and K. Paul, "Device fingerprinting for cyber-physical systems: A survey," *ACM Comput. Surv.*, vol. 55, no. 14s, July 2023.

[14] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 282–310, Jan. 2020.

[15] L. Bai, L. Zhu, J. Liu, J. Choi, and W. Zhang, "Physical layer authentication in wireless communication networks: A survey," *J. Commun. Inf. Netw.*, vol. 5, no. 3, pp. 237–264, Mar. 2020.

[16] W. Wang, I. Aguilar Sanchez, G. Caparra, A. McKeown, T. Whitworth, and E. S. Lohan, "A survey of spoofer detection techniques via radio frequency fingerprinting with focus on the GNSS pre-correlation sampled data," *Sensors*, vol. 21, no. 9, Sep. 2021.

[17] Y. Zhang, S. Zhao, H. Ji, Y. Zhang, Y. Shen, and X. Jiang, "A survey of secure communications for satellite Internet based on cryptography and physical layer security," *IET Inf. Secur.*, vol. 2023, pp. 1–15, Oct. 2023.

[18] E. Illi, M. Qaraqe, S. Althunibat, A. Alhasanat, M. Alsafasfeh, M. de Ree, G. Mantas, J. Rodriguez, W. Aman, and S. Al-Kuwari, "Physical layer security for authentication, confidentiality, and malicious node detection: A paradigm shift in securing IoT networks," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 1, pp. 347–388, Jan. 2024.

[19] T. M. Hoang, A. Vahid, H. D. Tuan, and L. Hanzo, "Physical layer authentication and security design in the machine learning era," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1830–1860, Feb. 2024.

[20] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, "Wireless physical-layer identification: Modeling and validation," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 9, pp. 2091–2106, Apr. 2016.

[21] P. Robyns, E. Marin, W. Lamotte, P. Quax, D. Singelée, and B. Preneel, "Physical-layer fingerprinting of LoRa devices using supervised and zero-shot learning," in *Proc. ACM Conf. Secur. Privacy Wireless Mobile Netw. (WiSec)*, 2017, pp. 58–63.

[22] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for LoRa using deep learning," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2604–2616, June 2021.

[23] G. Shen, J. Zhang, A. Marshall, and J. R. Cavallaro, "Towards scalable and channel-robust radio frequency fingerprint identification for LoRa," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 774–787, Feb. 2022.

[24] G. Shen, J. Zhang, A. Marshall, M. Valkama, and J. R. Cavallaro, "Toward length-versatile and noise-robust radio frequency fingerprint identification," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 2355 – 2367, Apr. 2023.

[25] S. Hanna, S. Karunaratne, and D. Cabric, "Open set wireless transmitter authorization: Deep learning approaches and dataset considerations," *IEEE Trans. on Cogn. Commun. Netw.*, vol. 7, no. 1, pp. 59–72, Jan. 2021.

[26] A. Gritsenko, Z. Wang, T. Jian, J. Dy, K. Chowdhury, and S. Ioannidis, "Finding a 'new' needle in the haystack: Unseen radio detection in large populations using deep learning," in *Proc. IEEE Int. Symp. Dynamic Spectr. Access Netw. (DySPAN)*, Newark, NJ, USA, 2019, pp. 1–10.

[27] Z. Zhu, X. Huang, M. Caron, and H. Leung, "Blind self-calibration technique for I/Q imbalances and DC-offsets," *IEEE Trans. Circuits Syst. I*, vol. 61, no. 6, pp. 1849–1859, Jun. 2013.

[28] Z. Zhu, H. Leung, and X. Huang, "Challenges in reconfigurable radio transceivers and application of nonlinear signal processing for RF impairment mitigation," *IEEE Circuits Syst. Mag.*, vol. 13, no. 1, pp. 44–65, Jan. 2013.

[29] G. Shen, J. Zhang, and A. Marshall, "Deep learning - powered radio frequency fingerprint identification: Methodology and case study," *IEEE Commun. Mag.*, vol. 61, no. 9, pp. 170–176, Sep. 2023.

[30] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11n traces with channel state information," *ACM SIGCOMM Comp. Commun. Rev.*, vol. 41, no. 1, pp. 53–53, Jan. 2011.

[31] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity WiFi," in *Proc. 21st Annual International Conference on Mobile Computing and Networking (MobiCom)*. New York, NY, USA: ACM, 2015, p. 53–64. [Online]. Available: http://doi.acm.org/10.1145/2789168.2790124

[32] Atheros CSI toolkit. [Online]. Available: https://wands.sg/research/wifi/AtherosCSI/

[33] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets," in *Proc. Int. Workshop Wireless Netw. Testbeds, Exp. Evaluation & Characterization*, 2019, p. 21–28.

[34] M. Schulz, D. Wegemer, and M. Hollick. Nexmon channel state information extractor. [Online]. Available: https://github.com/seemoo-lab/nexmon_csi

[35] S. M. Hernandez and E. Bulut, "Lightweight and Standalone IoT Based WiFi Sensing for Active Repositioning and Mobility," in *Proc. 21st Int. Symp. a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Cork, Ireland, Jun. 2020.

[36] ESP32 CSI toolkit. [Online]. Available: https://stevenmhernandez.github.io/ESP32-CSI-Tool/

[37] J. Hua, H. Sun, Z. Shen, Z. Qian, and S. Zhong, "Accurate and efficient wireless device fingerprinting using channel state information," in *Proc. IEEE INFOCOM*, 2018, pp. 1700–1708.

[38] T. D. Vo-Huu, T. D. Vo-Huu, and G. Noubir, "Fingerprinting Wi-Fi devices using software defined radios," in *Proc. ACM Conf. Secur. Privacy Wireless Mobile Netw. (WiSec)*, 2016, pp. 3–14.

[39] L. Peng, A. Hu, J. Zhang, Y. Jiang, J. Yu, and Y. Yan, "Design of a hybrid RF fingerprint extraction and device classification scheme," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 349–360, Jan. 2018.

[40] N. Soltani, K. Sankhe, J. Dy, S. Ioannidis, and K. Chowdhury, "More is better: Data augmentation for channel-resilient RF fingerprinting," *IEEE Commun. Mag.*, vol. 58, no. 10, pp. 66–72, Oct. 2020.

[41] Fading channels. Accessed on February 5, 2025. [Online]. Available: https://mathworks.com/help/comm/ug/fading-channels.html

[42] WLAN channel models. Accessed on February 5, 2025. [Online]. Available: https://mathworks.com/help/wlan/gs/wlan-channel-models.html

[43] J. Ma, J. Zhang, G. Shen, L. Peng, and A. Marshall, "Towards channel-robust radio frequency fingerprint identification using contrastive learning," in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC)*, 2025.

[44] A. Al-Shawabka, P. Pietraski, S. B Pattar, F. Restuccia, and T. Melodia, "DeepLoRa: Fingerprinting LoRa devices at scale through deep learning and data augmentation," in *Proc. ACM Int. Symp. Mob. Ad Hoc Netw. Comput. (MobiHoc)*, Shanghai, China, Jul. 2021.

[45] L. Peng, J. Zhang, M. Liu, and A. Hu, "Deep learning based RF fingerprint identification using differential constellation trace figure," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1091–1095, Jan. 2019.

[46] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for RF device fingerprinting in cognitive communication networks," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 160–167, Jan. 2018.

[47] L. Ding, S. Wang, F. Wang, and W. Zhang, "Specific emitter identification via convolutional neural networks," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2591–2594, Dec. 2018.

[48] J. Zhang, F. Wang, O. A. Dobre, and Z. Zhong, "Specific emitter identification via Hilbert-Huang transform in single-hop and relaying scenarios," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 6, pp. 1192–1205, Jun. 2016.

[49] J. Hall, M. Barbeau, and E. Kranakis, "Radio frequency fingerprinting for intrusion detection in wireless networks," *IEEE Trans. Depend. Sec. Comput.*, vol. 12, pp. 1–35, 2005.

[50] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for LoRa using spectrogram and CNN," in *Proc. IEEE INFOCOM*, May 2021.

[51] Y. Pan, S. Yang, H. Peng, T. Li, and W. Wang, "Specific emitter identification based on deep residual networks," *IEEE Access*, vol. 7, pp. 54 425–54 434, Apr. 2019.

[52] A. Al-Shawabka, F. Restuccia, S. D'Oro, T. Jian, B. C. Rendon, N. Soltani, J. Dy, S. Ioannidis, K. Chowdhury, and T. Melodia, "Exposing the fingerprint: Dissecting the impact of the wireless channel on radio fingerprinting," in *Proc. IEEE INFOCOM*, 2020, pp. 646–655.

[53] (2022) WiSig: RF fingerprinting dataset. [Online]. Available: https://cores.ee.ucla.edu/downloads/datasets/wisig/

[54] S. Hanna, S. Karunaratne, and D. Cabric, "WiSig: A large-scale WiFi signal dataset for receiver and channel agnostic RF fingerprinting," *IEEE Access*, vol. 10, pp. 22 808–22 818, Feb. 2022.

[55] (2025) DeepCRF TIFS. [Online]. Available: https://github.com/Oriseven/DeepCRF_TIFS

[56] R. Kong and H. Chen, "Deepcrf: Deep learning-enhanced csi-based rf fingerprinting for channel-resilient wifi device identification," *IEEE Trans. Inf. Forensics Security*, vol. 20, pp. 264 – 278, 2025.

[57] L. Xie, L. Peng, and J. Zhang. (2024) Wi-Fi dataset for channel-robust RFFI. [Online]. Available: https://ieee-dataport.org/documents/wi-fi-dataset-channel-robust-rffi

[58] ——, "Towards robust RF fingerprint identification using spectral regrowth and carrier frequency offset," in *Proc. IEEE INFOCOM*, 2025.

[59] J. Shi, L. Peng, H. Fu, and A. Hu, "ZigBee RFF dataset," 2023. [Online]. Available: https://dx.doi.org/10.21227/b4qd-gv36

[60] ——, "Robust RF fingerprint extraction based on cyclic shift characteristic," *IEEE Internet Things J.*, vol. 10, no. 21, pp. 19 218–19 233, Nov. 2023.

[61] G. Shen, J. Zhang, and A. Marshall. (2022) LoRa RFFI dataset. [Online]. Available: https://dx.doi.org/10.21227/qqt4-kz19

[62] ——. (2023) LoRa RFFI dataset with different spreading factors. [Online]. Available: https://dx.doi.org/10.21227/5q6q-c107

[63] ——. (2024) Radio frequency fingerprint LoRa dataset with multiple receivers. [Online]. Available: https://dx.doi.org/10.21227/d6vx-r538

[64] G. Shen, J. Zhang, A. Marshall, R. Woods, J. Cavallaro, and L. Chen, "Towards receiver-agnostic and collaborative radio frequency fingerprint identification," *IEEE Trans. Mobile Comput.*, vol. 23, no. 7, pp. 7618 – 7634, Dec. 2023.

[65] G. Shen and J. Zhang. (2024) LoRa Federated RFFI dataset. [Online]. Available: https://dx.doi.org/10.21227/nkdv-az07

[66] G. Shen, J. Zhang, X. Wang, and S. Mao, "Federated radio frequency fingerprint identification powered by unsupervised contrastive learning," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 9204 – 9215, Sep. 2024.

[67] LoRa radio data. [Online]. Available: https://www.interdigital.com/data_sets/lora-radio-data.

[68] RF fingerprinting (RFFP) datasets. [Online]. Available: https://research.engr.oregonstate.edu/hamdaoui/datasets

[69] A. Elmaghbub and B. Hamdaoui, "LoRa device fingerprinting in the wild: Disclosing RF data-driven fingerprint sensitivity to deployment variability," *IEEE Access*, vol. 9, pp. 142 893–142 909, Oct. 2021.

[70] A. Jagannath and J. Jagannath. (2022) RF-fingerprint-BT-IoT: Real-world frequency hopping Bluetooth dataset from IoT devices for RF fingerprinting. [Online]. Available: https://dx.doi.org/10.21227/364j-6j73

[71] ——, "Embedding-assisted attentional deep learning for real-world RF fingerprinting of Bluetooth," *IEEE Trans. on Cogn. Commun. Netw.*, vol. 9, no. 4, pp. 940–949, 2023.

[72] T. Ardoin and M. Kholghi. (2024) RUFF – Rotating UWB For Fingerprint. [Online]. Available: https://zenodo.org/records/11083153

[73] T. Ardoin, N. Pauli, B. Groß, M. Kholghi, K. Reaz, and G. Wunder, "Tracking UWB devices through radio frequency fingerprinting is possible," in *Proc. Int. Conf. Comput., Netw. and Commun. (ICNC)*, 2025.

[74] L. Peng. (2024) LTE mobile phone PRACH signal. [Online]. Available: https://ieee-dataport.org/documents/lte-mobile-phone-prach-signal

[75] L. Peng, Z. Wu, J. Zhang, M. Liu, H. Fu, and A. Hu, "Hybrid RFF identification for LTE using wavelet coefficient graph and differential spectrum," *IEEE Trans. Veh. Technol.*, Apr. 2024.

[76] G. Oligeri and S. Sciancalepore. (2022) Physical layer data acquisition of IRIDIUM satellites broadcast messages. [Online]. Available: https://data.mendeley.com/datasets/xcxspv8c2r/2

[77] G. Oligeri, S. Sciancalepore, S. Raponi, and R. Di Pietro, "PAST-AI: Physical-layer authentication of satellite transmitters via deep learning," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 274–289, 2023.

[78] J. Smailes, S. Köhler, S. Birnbach, M. Strohmeier, and I. Martinovic. (2023) Dataset for "watch this space: Securing satellite communication through resilient transmitter fingerprinting". [Online]. Available: https://zenodo.org/records/8220494

[79] ——, "Watch this space: Securing satellite communication through resilient transmitter fingerprinting," in *Proc. ACM SIGSAC Conf. on Comput. and Commun. Secur. (CCS)*, 2023, pp. 608–621.

[80] R. Xie, W. Xu, J. Yu, A. Hu, D. W. K. Ng, and A. L. Swindlehurst, "Disentangled representation learning for RF fingerprint extraction under unknown channel statistics," *IEEE Trans. Commun.*, vol. 71, no. 7, pp. 3946–3962, Jul. 2023.

[81] Y. Xing, A. Hu, J. Zhang, L. Peng, and X. Wang, "Design of a channel robust radio frequency fingerprint identification scheme," *IEEE Internet Things J.*, vol. 10, no. 8, pp. 6946–6959, Aug. 2023.

[82] T. Zheng, Z. Sun, and K. Ren, "FID: Function modeling-based data-independent and channel-robust physical-layer identification," in *Proc. IEEE INFOCOM*, 2019, pp. 199–207.

[83] M. Cekic, S. Gopalakrishnan, and U. Madhow, "Wireless fingerprinting via deep learning: The impact of confounding factors," in *Proc. Asilomar Conf. Signals, Syst., and Comput.*, 2021, pp. 677–684.

[84] H. Fu, L. Peng, M. Liu, and A. Hu, "Deep learning based RF fingerprint identification with channel effects mitigation," *IEEE Open J. Commun. Soc.*, Jul. 2023.

[85] M. Piva, G. Maselli, and F. Restuccia, "The tags are alright: Robust large-scale RFID clone detection through federated data-augmented radio fingerprinting," in *Proc. ACM Int. Symposium Mob. Ad Hoc Netw. Comput. (MobiHoc)*, Shanghai, China, Jul. 2021.

[86] T. Tian, Y. Wang, H. Dong, Y. Peng, Y. Lin, G. Gui, and H. Gacanin, "Transfer learning-based radio frequency fingerprint identification using ConvMixer network," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2022, pp. 4722–4727.

[87] R. Pan, H. Chen, H. Chen, and W.-Q. Wang, "Equalization assisted domain adaptation for radio frequency fingerprint identification," *IEEE Wireless Commun. Lett.*, Apr. 2024.

[88] W. Wang and L. Gan, "Radio frequency fingerprinting improved by statistical noise reduction," *IEEE Trans. on Cogn. Commun. Netw.*, vol. 8, no. 3, pp. 1444–1452, Mar. 2022.

[89] Y. Xing, A. Hu, J. Zhang, L. Peng, and G. Li, "On radio frequency fingerprint identification for DSSS systems in low SNR scenarios," *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2326–2329, Nov. 2018.

[90] T. Ohtsuji, T. Takeuchi, T. Soma, and M. Kitsunezuka, "Noise-tolerant, deep-learning-based radio identification with logarithmic power spectrum," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2019, pp. 1–6.

[91] S. Andrews, R. M. Gerdes, and M. Li, "Crowdsourced measurements for device fingerprinting," in *Proc. ACM Conf. Secur. Privacy Wireless Mobile Netw. (WiSec)*, 2019, pp. 72–82.

[92] B. He and F. Wang, "Cooperative specific emitter identification via multiple distorted receivers," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3791–3806, Jun. 2020.

[93] G. Shen, J. Zhang, A. Marshall, M. Valkama, and J. Cavallaro, "Radio frequency fingerprint identification for security in low-cost IoT devices," in *Proc. Asilomar Conf. Signals, Syst., and Comput.*, 2021, pp. 309–313.

[94] W. Wu, S. Hu, D. Lin, and Z. Liu, "DSLN: Securing Internet of Things through RF fingerprint recognition in low-SNR settings," *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3838–3849, Jul. 2021.

[95] K. Merchant and B. Nousain, "Toward receiver-agnostic RF fingerprint verification," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2019, pp. 1–6.

[96] T. Zhao, S. Sarkar, E. Krijestorac, and D. Cabric, "GAN-RXA: A practical scalable solution to receiver-agnostic transmitter fingerprinting," *IEEE Trans. on Cogn. Commun. Netw.*, vol. 10, no. 2, pp. 403–416, 2023.

[97] K. Li, J. Bao, X. Xie, J. Hong, and C. Hua, "Receiver-agnostic radio frequency fingerprint identification for zero-trust wireless networks," *IEEE J. Sel. Areas Commun.*, 2025.

[98] J. A. Gutierrez del Arroyo, B. J. Borghetti, and M. A. Temple, "Fingerprint extraction through distortion reconstruction (FEDR): A CNN-based approach to RF fingerprinting," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 9258–9269, Sep. 2024.

[99] J. Yu, A. Hu, G. Li, and L. Peng, "A robust RF fingerprinting approach using multisampling convolutional neural network," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6786–6799, Apr. 2019.

[100] T. J. Bihl, K. W. Bauer, and M. A. Temple, "Feature selection for RF fingerprinting with multiple discriminant analysis and using ZigBee device emissions," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 8, pp. 1862–1874, Aug. 2016.

[101] H. Givehchian, N. Bhaskar, E. R. Herrera, H. R. L. Soto, C. Dameff, D. Bharadia, and A. Schulman, "Evaluating physical-layer BLE location tracking attacks on mobile devices," in *Proc. IEEE Symp. on Secur. and Privacy (SP)*, 2022, pp. 1690–1704.

[102] H. Givehchian, N. Bhaskar, A. Redding, H. Zhao, A. Schulman, and D. Bharadia, "Practical obfuscation of BLE physical-layer fingerprints on mobile devices," in *Proc. IEEE Symp. on Secur. and Privacy (SP)*, 2024, pp. 2867–2885.

[103] N. Yuan, J. Zhang, Y. Ding, and S. L. Cotton, "Robust radio frequency fingerprint identification for Bluetooth low energy under low snr and channel variations," in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC)*, 2025.

[104] Z. Zhuang, X. Ji, T. Zhang, J. Zhang, W. Xu, Z. Li, and Y. Liu, "FBSleuth: Fake base station forensics via radio frequency fingerprinting," in *Proc. Asia Conf. Comp. Commun. Secur. (ASIA CCS)*. ACM, 2018, p. 261–272.

[105] X. Yang and D. Li, "LED-RFF: LTE DMRS based channel robust radio frequency fingerprint identification scheme," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 1855–1869, Dec. 2023.

[106] P. Yin, L. Peng, G. Shen, J. Zhang, M. Liu, H. Fu, A. Hu, and X. Wang, "Multi-channel CNN-based open-set RF fingerprint identification for LTE devices," *IEEE Trans. on Cogn. Commun. Netw.*, vol. 10, no. 5, pp. 1788–1800, Apr. 2024.

[107] G. Reus-Muns, D. Jaisinghani, K. Sankhe, and K. R. Chowdhury, "Trust in 5G open RANs through machine learning: RF fingerprinting on the POWDER PAWR platform," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*. IEEE, 2020, pp. 1–6.

[108] M. Foruhandeh, A. Z. Mohammed, G. Kildow, P. Berges, and R. Gerdes, "Spotr: GPS spoofing detection via device fingerprinting," in *Proc. ACM Conf. Secur. Privacy Wireless Mobile Netw. (WiSec)*, 2020, pp. 242–253.

[109] R. Kong and H. Chen, "CSI-RFF: Leveraging micro-signals on CSI for RF fingerprinting of commodity WiFi," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 5301 – 5315, 2024.

[110] G. Yin, J. Zhang, Y. Ding, and S. Cotton, "Noise-robust radio frequency fingerprint identification using denoise diffusion model," in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC) Workshop*, 2025.

[111] X. Gu, W. Wu, Y. Zhou, A. Song, M. Yang, Z. Ling, and J. Luo, "CQP-RFFI: Injecting a communication-quality preserving RF fingerprint for Wi-Fi device identification," in *Proc. IEEE/ACM 32nd Int. Symp. Quality of Service (IWQoS)*, 2024, pp. 1–10.

[112] Supported Hardware – Software-Defined Radio. [Online]. Available: https://uk.mathworks.com/help/comm/supported-hardware-software-defined-radio.html

[113] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for LoRa using deep learning," vol. 39, no. 8, pp. 2604–2616, 2021.

[114] IEEE 802.11 a/g/p Transceiver. [Online]. Available: https://github.com/bastibl/gr-ieee802-11

[115] H. Li, K. Gupta, C. Wang, N. Ghose, and B. Wang, "Radionet: Robust deep-learning based radio fingerprinting," in *Proc. IEEE Conf. Commun. and Netw. Secur. (CNS)*, 2022, pp. 190–198.

[116] IEEE 802.15.4 ZigBee Transceiver. [Online]. Available: https://github.com/bastibl/gr-ieee802-15-4

[117] PySDR: A Guide to SDR and DSP using Python. [Online]. Available: https://pysdr.org/index.html

[118] PicoScenes: Enabling Modern Wi-Fi ISAC Research! [Online]. Available: https://ps.zpj.io/

[119] P. Liu, P. Yang, W.-Z. Song, Y. Yan, and X.-Y. Li, "Real-time identification of rogue WiFi connections using environment-independent physical features," in *Proc. IEEE INFOCOM*, 2019, pp. 190–198.

[120] J. Huang, B. Liu, C. Miao, X. Zhang, J. Liu, L. Su, Z. Liu, and Y. Gu, "PhyFinAtt: An undetectable attack framework against phy layer fingerprint-based WiFi authentication," *IEEE Trans. Mobile Comput.*, vol. 23, no. 7, pp. 7753–7770, 2023.

[121] F. Restuccia, S. D'Oro, A. Al-Shawabka, B. C. Rendon, S. Ioannidis, and T. Melodia, "DeepFIR: Channel-robust physical-layer deep learning through adaptive waveform filtering," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 8054–8066, 2021.

[122] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, "ORACLE: Optimized radio classification through convolutional neural networks," in *Proc. IEEE INFOCOM*, Paris, France, 2019, pp. 370–378.

[123] G. Kia, D. Plets, B. Van Herbruggen, J. Fontaine, L. Verloock, E. De Poorter, and J. Talvitie, "UWB CIR data collected in 9 different environments in Ghent, Belgium," 2023. [Online]. Available: https://dx.doi.org/10.21227/kt06-tw72

[124] A. Brighente, F. Formaggio, G. M. Di Nunzio, and S. Tomasin, "Machine learning for in-region location verification in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2490–2502, Nov. 2019.

[125] L. Senigagliesi, M. Baldi, and E. Gambi, "Comparison of statistical and machine learning techniques for physical layer authentication," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1506–1521, Oct. 2021.

[126] F. Ardizzon and S. Tomasin, "Learning the likelihood test with one-class classifiers for physical layer authentication," 2024. [Online]. Available: https://arxiv.org/abs/2210.12494

[127] S. Tomasin, H. Zhang, A. Chorti, and H. V. Poor, "Challenge-response physical layer authentication over partially controllable channels," *IEEE Commun. Mag.*, vol. 60, no. 12, pp. 138–144, Dec. 2022.

[128] S. Tomasin, T. N. M. M. Elwakeel, A. V. Guglielmi, R. Maes, N. Noels, and M. Moeneclaey, "Analysis of challenge-response authentication with reconfigurable intelligent surfaces," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 9494–9507, 2024.

[129] L. Crosara, A. V. Guglielmi, N. Laurenti, and S. Tomasin, "Divergence-minimizing attack against challenge-response authentication with IRSs," in *Proc. IEEE Int. Conf. on Commun. Workshops (ICC Workshops)*, 2024, pp. 1986–1991.

[130] A. V. Guglielmi, L. Crosara, S. Tomasin, and N. Laurenti, "Physical-layer challenge-response authentication with IRS and single-antenna devices," in *Proc. IEEE Int. Conf. on Commun. Workshops (ICC Workshops)*, 2024, pp. 560–565.

[131] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "PHY-layer spoofing detection with reinforcement learning in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 037–10 047, Dec. 2016.

[132] H. Tan, N. Xie, J. Lu, and D. Niyato, "Generalized tag-based physical-layer authentication under frequency selective fading channels," *IEEE Trans. Commun.*, vol. 71, no. 5, pp. 2876–2890, May 2023.

[133] H. Amin, W. Aman, and S. Al-Kuwari, "On the potential of re-configurable intelligent surface (RIS)-assisted physical layer authentication (PLA)," 2024. [Online]. Available: https://arxiv.org/abs/2405.00426

[134] N. Xie, J. Chen, and L. Huang, "Physical-layer authentication using multiple channel-based features," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2356–2366, Jan. 2021.

[135] J. Wang, Y. Shao, Y. Ge, and R. Yu, "Physical-layer authentication based on adaptive Kalman filter for V2X communication," *Veh. Commun.*, vol. 26, p. 100281, Dec. 2020.

[136] F. Mazzo, S. Tomasin, H. Zhang, A. Chorti, and H. V. Poor, "Physical-layer challenge-response authentication for drone networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2023, pp. 3282–3287.

[137] M. Piana, F. Ardizzon, and S. Tomasin, "Challenge-response to authenticate drone communications: A game theoretic approach," 2024. [Online]. Available: https://arxiv.org/abs/2410.00785

[138] D. Shan, K. Zeng, W. Xiang, P. Richardson, and Y. Dong, "PHY-CRAM: Physical layer challenge-response authentication mechanism for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1817–1827, Sep. 2013.

[139] X. Wu, Z. Yang, C. Ling, and X.-G. Xia, "Artificial-noise-aided physical layer phase challenge-response authentication for practical OFDM transmission," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6611–6625, Oct. 2016.

[140] B. Çiçek and H. Alakoca, "Impact of residual hardware impairments on RIS-aided authentication," in *Proc. Virtual Conference on Communications (VCC)*, 2024, pp. 1–6.

[141] X. Lu, L. Xiao, T. Xu, Y. Zhao, Y. Tang, and W. Zhuang, "Reinforcement learning based PHY authentication for VANETs," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3068–3079, Mar. 2020.

[142] H. Fang, X. Wang, and L. Hanzo, "Learning-aided physical layer authentication as an intelligent process," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2260–2273, Mar. 2019.

[143] R. Meng, X. Xu, G. Li, B. Xu, F. Zhu, B. Wang, and P. Zhang, "Multidimensional fingerprints-based multiattacker detection for 6G systems," *IEEE Internet Things J.*, vol. 11, no. 2, pp. 2665–2683, Feb. 2024.

[144] Q. Wang, H. Li, D. Zhao, Z. Chen, S. Ye, and J. Cai, "Deep neural networks for CSI-based authentication," *IEEE Access*, vol. 7, pp. 123 026–123 034, Aug. 2019.

[145] M. Abdrabou and T. A. Gulliver, "Adaptive physical layer authentication using machine learning with antenna diversity," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6604–6614, Oct. 2022.

[146] R. Meng, X. Xu, B. Wang, H. Sun, S. Xia, S. Han, and P. Zhang, "Physical-layer authentication based on hierarchical variational autoencoder for industrial Internet of Things," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2528–2544, Mar. 2023.

[147] Y. Li, L. Xiao, Q. Li, and W. Su, "Spoofing detection games in underwater sensor networks," in *Proc. OCEANS 2015 - MTS/IEEE Washington*, 2015, pp. 1–5.

[148] M. Khalid, R. Zhao, and X. Wang, "Node authentication in underwater acoustic sensor networks using time-reversal," in *Proc. Global Oceans 2020: Singapore – U.S. Gulf Coast*, 2020, pp. 1–4.

[149] N. Wang, T. Jiang, S. Lv, and L. Xiao, "Physical-layer authentication based on extreme learning machine," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1557–1560, Jul. 2017.

[150] L. Xiao, G. Sheng, X. Wan, W. Su, and P. Cheng, "Learning-based PHY-layer authentication for underwater sensor networks," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 60–63, Jan. 2019.

[151] R. Zhao, T. Shi, C. Liu, X. Shen, and O. A. Dobre, "Physical layer authentication without adversary training data in resource-constrained underwater acoustic networks," *IEEE Sensors J.*, vol. 23, no. 22, pp. 28 270–28 281, Nov. 2023.

[152] A. Brighente, S. Xu, S. Soderi, and M. Conti, "Physical layer authentication for distributed RIS (DRIS) enabled VLC systems," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2024, pp. 3340–3345.

[153] L. Xiao, A. Reznik, W. Trappe, C. Ye, Y. Shah, L. Greenstein, and N. Mandayam, "PHY-authentication protocol for spoofing detection in wireless networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, 2010, pp. 1–6.

[154] L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, "Fingerprints in the Ether: Using the physical layer for wireless authentication," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2007, pp. 4646–4651.

[155] M. Liu, A. Mukherjee, Z. Zhang, and X. Liu, "TBAS: Enhancing Wi-Fi authentication by actively eliciting channel state information," in *Proc. Annu. IEEE Int. Conf. on Sens., Commun., and Netw. (SECON)*, 2016, pp. 1–9.

[156] Y. Song, B. Chen, T. Wu, T. Zheng, H. Chen, and J. Wang, "Enhancing packet-level Wi-Fi device authentication protocol leveraging channel state information," *Wireless Commun. Mob. Comput.*, vol. 2021, no. 1, p. 2993019, Jan. 2021.

[157] H. Liu, Y. Wang, J. Liu, J. Yang, Y. Chen, and H. V. Poor, "Authenticating users through fine-grained channel information," *IEEE Trans. Mobile Comput.*, vol. 17, no. 2, pp. 251–264, Feb. 2018.

[158] H.-M. Wang and Q.-Y. Fu, "Channel-prediction-based one-class mobile IoT device authentication," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7731–7745, Oct. 2022.

[159] J. Han, Y. Li, G. Liu, J. Ma, Y. Zhou, H. Fang, and X. Wu, "Model-driven learning for physical layer authentication in dynamic environments," *IEEE Commun. Lett.*, vol. 28, no. 3, pp. 572–576, Mar. 2024.

[160] S. Chen, Z. Pang, H. Wen, K. Yu, T. Zhang, and Y. Lu, "Automated labeling and learning for physical layer authentication against clone node and Sybil attacks in industrial wireless edge networks," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 2041–2051, Mar. 2021.

[161] F. Xie, Z. Pang, H. Wen, W. Lei, and X. Xu, "Weighted voting in physical layer authentication for industrial wireless edge networks," *IEEE Trans. Ind. Informat.*, vol. 18, no. 4, pp. 2796–2806, Apr. 2022.

[162] F. Pan, Z. Pang, H. Wen, M. Luvisotto, M. Xiao, R.-F. Liao, and J. Chen, "Threshold-free physical layer authentication based on machine learning for industrial wireless CPS," *IEEE Trans. Ind. Informat.*, vol. 15, no. 12, pp. 6481–6491, Dec. 2019.

[163] T. Zhang, Y. Huo, Q. Gao, L. Ma, Y. Wu, and R. Li, "Cooperative physical layer authentication with reputation-inspired collaborator selection," *IEEE Internet Things J.*, vol. 10, no. 24, Dec. 2023.

[164] Y. Chen, H. He, S. Liu, Y. Zhang, Y. Li, B. Xing, B. Guo, and L. Chen, "Physical layer authentication for industrial control based on convolutional denoising autoencoder," *IEEE Internet Things J.*, vol. 11, no. 9, May 2023.

[165] S. Wang, K. Huang, X. Xu, Z. Zhong, and Y. Zhou, "CSI-based physical layer authentication via deep learning," *IEEE Wireless Commun. Lett.*, vol. 11, no. 8, pp. 1748–1752, Aug. 2022.

[166] R.-F. Liao, H. Wen, J. Wu, F. Pan, A. Xu, H. Song, F. Xie, Y. Jiang, and M. Cao, "Security enhancement for mobile edge computing through physical layer authentication," *IEEE Access*, vol. 7, pp. 116 390–116 401, Aug. 2019.

[167] X. Qiu, J. Dai, and M. Hayes, "A learning approach for physical layer authentication using adaptive neural network," *IEEE Access*, vol. 8, pp. 26 139–26 149, Feb. 2020.

[168] H. Liu, Y. Wang, J. Liu, J. Yang, and Y. Chen, "Practical user authentication leveraging channel state information (CSI)," in *Proc. ACM Symp. on Inf., Comput. and Commun. Secur. (ASIA CCS)*. ACM, 2014, p. 389–400.

[169] M. Abdrabou and T. A. Gulliver, "Game theoretic spoofing detection for space information networks using physical attributes," *IEEE Trans. Commun.*, vol. 72, no. 7, pp. 3947–3956, Feb. 2024.

[170] K. Zeng, K. Govindan, D. Wu, and P. Mohapatra, "Identity-based attack detection in mobile wireless networks," in *Proc. IEEE INFOCOM*, 2011, pp. 1880–1888.

[171] A. Kalamandeen, A. Scannell, E. de Lara, A. Sheth, and A. LaMarca, "Ensemble: cooperative proximity-based authentication," in *Proc. Int. Conf. on Mobile Syst., Appl., and Services (MobiSys)*, 2010, p. 331–344.

[172] J. Wang, Y. Shao, Y. Wang, Y. Ge, and R. Yu, "Physical layer authentication based on nonlinear Kalman filter for V2X communication," *IEEE Access*, vol. 8, pp. 163 746–163 757, Sept. 2020.

[173] M. Abdrabou and T. A. Gulliver, "Physical layer authentication for satellite communication systems using machine learning," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 2380–2389, Nov. 2022.

[174] D. Romero, T. N. Ha, and P. Gerstoft, "Spoofing attack detection in the physical layer with robustness to user movement," in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC)*, 2024, pp. 1–6.

[175] Y. Jiang, L. Jiao, L. Zhao, and K. Zeng, "Beam pattern fingerprinting with missing features for spoofing attack detection in millimeter-wave networks," in *Proc. ACM Workshop on Wireless Secur. and Mach. Learn. (WiseML)*. ACM, 2022, p. 75–80.

[176] N. Wang, L. Jiao, P. Wang, W. Li, and K. Zeng, "Exploiting beam features for spoofing attack detection in mmWave 60-GHz IEEE 802.11ad networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3321–3335, May 2021.

[177] R. Du, L. Zhen, and Y. Liu, "Physical layer authentication based on integrated semi-supervised learning in wireless networks for dynamic industrial scenarios," *IEEE Trans. Veh. Technol.*, vol. 72, no. 5, pp. 6154–6164, May 2023.

[178] M. Abdrabou and T. A. Gulliver, "Authentication for satellite communication systems using physical characteristics," *IEEE Open J. Veh. Technol.*, vol. 4, pp. 48–60, 2022.

[179] Y. Chen, W. Trappe, and R. P. Martin, "Detecting and localizing wireless spoofing attacks," in *Proc. Annu. IEEE Int. Conf. on Sens., Commun., and Netw. (SECON)*, 2007, pp. 193–202.

[180] J. Yang, Y. Chen, W. Trappe, and J. Cheng, "Detection and localization of multiple spoofing attackers in wireless networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 1, pp. 44–58, Jan. 2012.

[181] Y. Gajbhiye and R. Daruwala, "RSS-based spoofing detection and localization algorithm in IEEE 802.11 wireless networks," in *Proc. Int. Conf. on Commun. and Signal Process. (ICCSP)*. IEEE, 2016, pp. 1642–1645.

[182] G. Ventura, F. Ardizzon, and S. Tomasin, "Authentication by location tracking in underwater acoustic networks," 2024. [Online]. Available: https://arxiv.org/abs/2410.03511

[183] A. Ijaz, M. M. U. Rahman, and O. A. Dobre, "On safeguarding visible light communication systems against attacks by active adversaries," *IEEE Photonics Technol. Lett.*, vol. 32, no. 1, pp. 11–14, Jan. 2020.

[184] J. He, M. Niu, P. Zhang, and C. Qin, "Enhancing PHY-layer authentication in RIS-assisted iot systems with cascaded channel features," *IEEE Internet of Things Journal*, vol. 11, no. 14, pp. 24 984–24 997, 2024.

[185] K. Joo and W. Choi, "Enhancing security of HRP UWB ranging system based on channel characteristic analysis," *IEEE Internet Things J.*, vol. 11, no. 24, pp. 39 794–39 808, 2024.

[186] R. Diamant, P. Casari, and S. Tomasin, "Cooperative authentication in underwater acoustic sensor networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 954–968, Feb. 2019.

[187] P. Casari, F. Ardizzon, and S. Tomasin, "Physical layer authentication in underwater acoustic networks with mobile devices," in *Proc. Int. Conf. on Underwater Netw. & Syst. (WUWNet)*. ACM, 2022.

[188] L. Bragagnolo, F. Ardizzon, N. Laurenti, P. Casari, R. Diamant, and S. Tomasin, "Authentication of underwater acoustic transmissions via machine learning techniques," in *Proc. IEEE Int. Conf. on Microw., Antennas, Commun. and Electron. Syst. (COMCAS)*, 2021, pp. 255–260.

[189] F. Ardizzon, R. Diamant, P. Casari, and S. Tomasin, "Machine learning-based distributed authentication of UWAN nodes with limited shared information," in *Proc. Underwater Commun. and Netw. Conf. (UComms)*, 2022, pp. 1–5.

[190] F. Ardizzon, P. Casari, and S. Tomasin, "A RNN-based approach to physical layer authentication in underwater acoustic networks with mobile devices," *Comput. Netw.*, vol. 243, p. 110311, Apr. 2024.

[191] E. Jedermann, M. Strohmeier, M. Schäfer, J. Schmitt, and V. Lenders, "Orbit-based authentication using TDOA signatures in satellite networks," in *Proc. ACM Conf. Secur. Privacy Wireless Mobile Netw. (WiSec)*. New York, NY, USA: Association for Computing Machinery, 2021, p. 175–180.

[192] S. Gao, Y. Ding, Y. Lu, L. Han, L. Zhou, C. Chen, X. Yu, and X. Huang, "A lightweight fingerprint-based device authentication architecture for wireless industrial automation networks," in *Proc. Int. Conf on Ind. Artif. Intell. (IAI)*, 2019, pp. 1–6.

[193] W. Aman, M. M. U. Rahman, J. Qadir, H. B. Pervaiz, and Q. Ni, "Impersonation detection in line-of-sight underwater acoustic sensor networks," *IEEE Access*, vol. 6, pp. 44 459–44 472, Aug. 2018.

[194] J. Xiong and K. Jamieson, "SecureArray: improving wifi security with fine-grained physical-layer information," in *Proc. Annu. Int. Conf. on Mobile Comput. & Netw. (MobiCom)*, 2013, p. 441–452.

[195] L. Ning, B. Li, C. Zhao, Y. Tao, and X. Wang, "Detection and localization of the eavesdropper in MIMO systems," *IEEE Access*, vol. 8, pp. 94 984–94 993, May 2020.

[196] Y. Tao, X. Wang, B. Li, and C. Zhao, "Pilot spoofing attack detection and localization with mobile eavesdropper," *IEEE Trans. Mobile Comput.*, vol. 22, no. 3, pp. 1688–1701, Mar. 2023.

[197] J. Xiong and K. Jamieson, "SecureAngle: improving wireless security using angle-of-arrival information," in *Proc. ACM SIGCOMM Workshop on Hot Topics in Netw.*, 2010.

[198] A. Bendaimi, A. Abdallah, A. Celik, A. M. Eltawil, and H. Arslan, "How to leverage double-structured sparsity of RIS channels to boost physical-layer authentication," *IEEE Wireless Commun. Letters*, vol. 13, no. 8, pp. 2260–2264, 2024.

[199] N. Wang, L. Jiao, P. Wang, M. Dabaghchian, and K. Zeng, "Efficient identity spoofing attack detection for IoT in mm-Wave and massive MIMO 5G communication," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–6.

[200] N. Gao, Q. Huang, C. Li, S. Jin, and M. Matthaiou, "EsaNet: Environment semantics enabled physical layer authentication," *IEEE Wireless Commun. Lett.*, vol. 13, no. 1, Jan. 2023.

[201] M. Khalid, R. Zhao, and N. Ahmed, "Physical layer authentication in line-of-sight underwater acoustic sensor networks," in *Proc. Global Oceans*, 2020, pp. 1–5.

[202] Q.-Y. Fu, Y.-H. Feng, H.-M. Wang, and P. Liu, "Initial satellite access authentication based on Doppler frequency shift," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 498–502, Mar. 2021.

[203] O. A. Topal and G. Karabulut Kurt, "Physical layer authentication for LEO satellite constellations," in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC)*, 2022, p. 1952–1957.

[204] T. Wang and Y. Yang, "Analysis on perfect location spoofing attacks using beamforming," in *Proc. IEEE INFOCOM*, 2013, pp. 2778–2786.

[205] M. H. Yılmaz and H. Arslan, "A survey: Spoofing attacks in physical layer security," in *Proc. IEEE Local Comput. Netw. Conf. Workshops (LCN Workshops)*, 2015, pp. 812–817.

[206] C. Sanders and Y. Wang, "Localizing spoofing attacks on vehicular GPS using vehicle-to-vehicle communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15 656–15 667, 2020.

[207] J. Neyman and E. S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses," *Phil. Trans. Roy. Soc. Lond. A*, vol. 231, no. 694-706, pp. 289–337, Feb. 1933.

[208] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[209] X. Yin, X. Fang, N. Zhang, P. Yang, X. Sha, and J. Qiu, "Online learning aided adaptive multiple attribute-based physical layer authentication in dynamic environments," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 1106–1116, Feb. 2021.

[210] Q. Wang, Z. Pang, W. Liang, J. Zhang, and K. Wang, "Transferable physical layer authentication based on time-varying patterns toward zero training deployment for mobile IIoT devices," *IEEE Trans. Ind. Informat.*, vol. 20, no. 5, pp. 7675–7685, Feb. 2024.

[211] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, Jun. 2014.

[212] P. Kyösti, J. Meinilä, L. Hentila, X. Zhao, T. Jämsä, C. Schneider, M. Narandzi'c, M. Milojevi'c, A. Hong, J. Ylitalo, V.-M. Holappa, M. Alatossava, R. Bultitude, Y. Jong, and T. Rautiainen, "IST-4-027756 WINNER II D1.1.2 v1.2 WINNER II channel models," *Inf. Soc. Technol*, vol. 11, Feb. 2008.

[213] M. Porter *et al.*, "Bellhop Gaussian beam/finite element beam code," last accessed: Sept. 2024. [Online]. Available: http://oalib.hlsresearch.com/Rays/index.html

[214] N. Morozs, W. Gorma, B. T. Henson, L. Shen, P. D. Mitchell, and Y. V. Zakharov, "Channel modeling for underwater acoustic network simulation," *IEEE Access*, vol. 8, pp. 136 151–136 175, July 2020.

[215] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 701–717, Apr. 2013.

[216] N. (NORAD), "Celestrak: Norad two-line element sets current data," 2019, last access: Oct. 2024. [Online]. Available: https://celestrak.org/NORAD/elements/

[217] Ansys, "Ansys System Tool Kit home Page." [Online]. Available: https://www.ansys.com/products/missions/ansys-stk

[218] R. Candell, "Radio frequency measurements for selected manufacturing and industrial environments," 2016. [Online]. Available: https://data.nist.gov/od/id/mds0139sck

[219] N. Wang, "802.11ad SLS SNR trace-based authentication," 2025. [Online]. Available: https://github.com/wangning8566/SLS-SNR-trace-based-authentication

[220] E. S. Lohan, J. Torres-Sospedra, H. Leppäkoski, P. Richter, Z. Peng, and J. Huerta, "Wi-Fi crowdsourced fingerprinting dataset for indoor positioning," *Data*, vol. 2, no. 4, Oct. 2017.

[221] A. A. S. AlQahtani and T. Alshayeb, "RSSI measurements of beacon frames from Wi-Fi radio waves," 2023. [Online]. Available: https://dx.doi.org/10.21227/2bk3-dw90

[222] P. van Walree, R. Otnes, and T. Jenserud, "Watermark: A realistic benchmark for underwater acoustic modems," in *Proc. Underwater Commun. and Netw. Conf. (UComms)*, 2016, pp. 1–4.

[223] J. Huang and R. Diamant, "Channel impulse responses from Mar. 2019 long range experiment (Mediterranean Sea)," 2019. [Online]. Available: https://dx.doi.org/10.21227/nzgr-ds72

[224] X. Li, J. Liu, B. Ding, Z. Li, H. Wu, and T. Wang, "A SDR-based verification platform for 802.11 PHY layer security authentication," *World Wide Web*, vol. 23, pp. 1011–1034, Jan. 2020.

[225] R. Candell, C. Remley, J. Quimby, D. Novotny, A. Curtin, P. Papazian, G. Koepke, J. Diener, and M. Hany, "Industrial wireless systems: Radio propagation measurements," Jan. 2017.

[226] J. Quimby, R. Candell, C. Remley, D. Novotny, J. Diener, P. Papazian, A. Curtin, and G. Koepke, "NIST channel sounder overview and channel measurements in manufacturing facilities," Nov. 2017.

[227] W. Wang, Z. Sun, K. Ren, and B. Zhu, "User capacity of wireless physical-layer identification: An information-theoretic perspective," in *Proc. IEEE Int. Conf. on Commun. (ICC)*. IEEE, 2016, pp. 1–6.

[228] ——, "User capacity of wireless physical-layer identification," *IEEE Access*, vol. 5, pp. 3353–3368, 2017.

[229] A. Saeif, S. Savio, and O. Gabriele, "The day-after-tomorrow: On the performance of radio fingerprinting over time," in *Proc. Annual Comp Secur. Applications Conf.*, 2023, pp. 439–450.

[230] A. Elmaghbub and B. Hamdaoui, "No blind spots: On the resiliency of device fingerprints to hardware warm-up through sequential transfer learning," in *Proc. ACM Conf. Secur. Privacy Wireless Mobile Netw. (WiSec)*, 2024, pp. 134–144.

[231] D. Adesina, C.-C. Hsieh, Y. E. Sagduyu, and L. Qian, "Adversarial machine learning in wireless communications using RF data: A review," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 1, pp. 77–100, Jan. 2022.

[232] Z. Liu, C. Xu, Y. Xie, E. Sie, F. Yang, K. Karwaski, G. Singh, Z. L. Li, Y. Zhou, D. Vasisht *et al.*, "Exploring practical vulnerabilities of machine learning-based wireless systems," in *Proc. USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*, 2023, pp. 1801–1817.

[233] T. Zhao, X. Wang, J. Zhang, and S. Mao, "Explanation-guided backdoor attacks on model-agnostic RF fingerprinting," in *Proc. IEEE INFOCOM*, 2024, pp. 221–230.

[234] T. Zhao, N. Wang, Y. Wu, W. Zhang, and X. Wang, "Backdoor attacks against low-earth orbit satellite fingerprinting," in *Proc. IEEE INFOCOM Workshops*, 2024, pp. 01–06.

[235] Z. Bao, Y. Lin, S. Zhang, Z. Li, and S. Mao, "Threat of adversarial attacks on dl-based iot device identification," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 9012–9024, Sep. 2021.

[236] B. Liu, H. Zhang, Y. Wan, F. Zhou, Q. Wu, and D. W. K. Ng, "Robust adversarial attacks on deep learning-based rf fingerprint identification," *IEEE Wireless Commun. Lett.*, vol. 12, no. 6, pp. 1037–1041, Dec. 2023.

[237] J. Ma, J. Zhang, G. Shen, A. Marshall, and C.-H. Chang, "White-box adversarial attacks on deep learning-based radio frequency fingerprint identification," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2023, pp. 3714–3719.

[238] L. Papangelo, M. Pistilli, S. Sciancalepore, G. Oligeri, G. Piro, and G. Boggia, "Adversarial machine learning for image-based radio frequency fingerprinting: Attacks and defenses," *IEEE Commun. Mag.*, 2024.

[239] W. Li, S. Wang, Y. Zhang, L. Guo, Y. Liu, Y. Lin, and G. Gui, "Slpa: Single-line pixel attack on specific emitter identification using time-frequency spectrogram," *IEEE Trans. Veh. Technol.*, 2024.
[240] G. Yin, J. Zhang, X. Yi, and X. Wang, "Evasion attacks and countermeasures in deep learning-based Wi-Fi gesture recognition," *IEEE Trans. Mobile Comput.*, 2025.
[241] L. F. Abanto-Leon, A. Bäuml, G. H. Sim, M. Hollick, and A. Asadi, "Stay connected, leave no trace: Enhancing security and privacy in wifi via obfuscating radiometric fingerprints," *Proc. ACM on Meas. and Anal. of Comput. Syst.*, vol. 4, no. 3, pp. 1–31, 2020.
[242] T. M. Pham, L. Senigagliesi, M. Baldi, G. P. Fettweis, and A. Chorti, "Machine learning-based robust physical layer authentication using angle of arrival estimation," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2023, pp. 13–18.
[243] C. Zhao, H. Du, D. Niyato, J. Kang, Z. Xiong, D. I. Kim, X. Shen, and K. B. Letaief, "Generative ai for secure physical layer communications: A survey," *IEEE Trans. on Cogn. Commun. Netw.*, pp. 3 – 26, 2025.
[244] N. Wang, T. Zhao, S. Mao, and X. Wang, "AI generated wireless data for enhanced satellite device fingerprinting," in *Proc. IEEE Int. Conf. on Commun. Workshops*, 2024, pp. 88–93.
[245] H. Zhou, C. Hu, Y. Yuan, Y. Cui, Y. Jin, C. Chen, H. Wu, D. Yuan, L. Jiang, D. Wu, X. Liu, C. Zhang, X. Wang, and J. Liu, "Large language model (LLM) for telecommunications: A comprehensive survey on principles, key techniques, and opportunities," *IEEE Commun. Surveys Tuts.*, pp. 1–1, 2024.
[246] K. Merchant and B. Nousain, "Securing IoT RF fingerprinting systems with generative adversarial networks," in *Proc. of Military Commun. Conf. (MILCOM)*, 2019, pp. 584–589.
[247] X. Chen, Z. Liu, X. Zhang, Y. Wang, D. Shi, and X. Liu, "Optic fingerprint: Enhancing security in visible light communication networks," in *Proc. IEEE INFOCOM Workshops*, 2024, pp. 1–6.

Junqing Zhang received a Ph.D. degree in Electronics and Electrical Engineering from Queen's University Belfast, UK in 2016. From Feb. 2016 to Jan. 2018, he was a Postdoctoral Research Fellow at Queen's University Belfast. From Feb. 2018 to Oct. 2022, he was a Tenure Track Fellow and then a Lecturer (Assistant Professor) at the University of Liverpool, UK. Since Oct. 2022, he has been a Senior Lecturer (Associate Professor) at the University of Liverpool. His research interests include the Internet of Things, wireless security, physical layer security, key generation, radio frequency fingerprint identification, and wireless sensing. Dr. Zhang is a co-recipient of the IEEE WCNC 2025 Best Workshop Paper Award. He is a Senior Area Editor of IEEE Transactions on Information Forensics and Security and an Associate Editor of IEEE Transactions on Mobile Computing.

Francesco Ardizzon (Member, IEEE) received the B.Sc. degree in 2016, the M.Sc. degree in 2019, and the Ph.D. degree in Information Engineering in 2023 from the University of Padova, Italy. In 2022, he was a visiting scientist at the ESA European Space Research and Technology Centre. He is currently an Assistant Professor at the University of Padova. His current research interests include authentication for global navigation satellite systems, physical layer security, and underwater acoustic communications.

Mattia Piana (Student Member, IEEE) received a B.Sc. in Information Engineering and an M. Sc in Telecommunication Engineering from the University of Padova in 2021 and 2023, respectively. In 2023, he was at National Instruments (Dresden, Germany), where he collaborated on the development of novel techniques for mmWave antenna characterization. He is currently a PhD student at the University of Padova within the EU ROBUST-6G project, and his research interests include physical layer security and reflective intelligent surfaces.

Guanxiong Shen received the B.Eng degree from Xidian University, Xi'an, China, in 2019, and the Ph.D degree from the University of Liverpool, UK, in 2023. He is currently an Associate professor at Southeast University, Nanjing, China. His research interests include the Internet of Things, wireless security, physical layer security, and radio frequency fingerprint identification.

Stefano Tomasin received the Ph.D. degree from the University of Padova, Italy (2003), where he is now a Full Professor. During his career, he has visited IBM Research (Switzerland), Philips Research (Netherlands), Qualcomm (California), the Polytechnic University in Brooklyn (New York), and Huawei (France). His current research interests include physical layer security, security of global navigation satellite systems, signal processing for wireless communications, synchronization, and scheduling of communication resources. He is a senior member of IEEE and a member of EURASIP. He is or has been an Editor of the IEEE Transactions on Vehicular Technologies, the IEEE Transactions on Signal Processing (2017-2020), the EURASIP Journal of Wireless Communications and Networking, and the IEEE Transactions on Information Forensics and Security.