1-2-2023

Channel-Resilient Deep-Learning-Driven Device Fingerprinting Through Multiple Data Streams

Nora Basha
Oregon State University

Bechir Hamdaoui
Oregon State University

Kathiravetpillai Sivanesan
Intel Corporation

Mohsen Guizani
Mohamed Bin Zayed University of Artificial Intelligence

Recommended Citation

N. Basha, B. Hamdaoui, K. Sivanesan and M. Guizani, "Channel-Resilient Deep-Learning-Driven Device Fingerprinting Through Multiple Data Streams," in *IEEE Open Journal of the Communications Society,* vol. 4, pp. 118-133, 2023, doi: 10.1109/OJCOMS.2022.3233372.

Received 23 November 2022; accepted 21 December 2022. Date of publication 2 January 2023; date of current version 13 January 2023.

Digital Object Identifier 10.1109/OJCOMS.2022.3233372

Channel-Resilient Deep-Learning-Driven Device Fingerprinting Through Multiple Data Streams

NORA BASHA[1], BECHIR HAMDAOUI[1] (Senior Member, IEEE), KATHIRAVETPILLAI SIVANESAN[2], AND MOHSEN GUIZANI[3]

[1]EECS, Oregon State University, Corvallis, OR 97331, USA

[2]Intel Labs, Intel Corporation, Hillsboro, OR 97124, USA

[3]Machine Learning Department, Mohamed Bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE

CORRESPONDING AUTHOR: N. BASHA (e-mail: bashano@oregonstate.edu)

This work was supported in part by the U.S. National Science Foundation through NSF under Award 1923884 and through NSF/Intel under Award 2003273.

ABSTRACT Enabling accurate and automated identification of wireless devices is critical for allowing network access monitoring and ensuring data authentication for large-scale IoT networks. RF fingerprinting has emerged as a solution for device identification by leveraging the transmitters' inevitable hardware impairments that occur during manufacturing. Although deep learning is proven efficient in classifying devices based on hardware impairments, the performance of deep learning models suffers greatly from variations of the wireless channel conditions, across time and space. To the best of our knowledge, we are the first to propose leveraging MIMO capabilities to mitigate the channel effect and provide a channel-resilient device classification framework. We begin by showing that for AWGN channels, combining multiple received signals improves the testing accuracy by up to 30%. We then show that for more realistic Rayleigh channels, blind channel estimation enabled by MIMO increases the testing accuracy by up to 50% when the models are trained and tested over the same channel, and by up to 69% when the models are tested on a channel that is different from that used for training.

INDEX TERMS Automated network access, deep learning, IoT device fingerprinting, multiple-input multiple-out (MIMO).

I. INTRODUCTION

AS THE IoT paradigm pervasively expands into various sectors like healthcare, home automation, and power grids, massive numbers of wireless devices are being connected to the Internet, widening the attack surface of emerging IoT networks [1], [2], [3], [4], [5]. Therefore, enabling these networks with automated device authentication methods is becoming a necessity [6], [7], [8]. Deep Neural Network (DNN)-based device fingerprinting has emerged as a promising technique for fulfilling such a necessity [9], [10], which leverages transceiver hardware impairments to provide a fingerprint for each device that, unlike high-layer features such as IP or MAC addresses [11], is difficult to spoof or replicate. These hardware imperfections, which are inevitably inherited during device manufacturing, impair the transmitted waveform in a way that provides transmitters with fingerprints and signatures that can uniquely separate them from one another [12], [13], [14], [15].

DNN-based approaches have proven efficient in classifying wireless devices from captured RF signals [12], [13]. However, when the testing data are collected over a channel that is different from that used during training, their performance degrades significantly [16], [17], [18]. This is due to the severe impact of the wireless channel on the device fingerprints. That is, DNN models tend to extract their features from channel-distorted device impairments, thereby degrading their achieved accuracy substantially when training and testing are done on different channels. Our work leverages MIMO to mitigate such a channel effect by first exploiting the multiple received signal streams to restore less-distorted versions of the devices' originally transmitted signals, and then using them for classification.

If an RF signal is transmitted from a single antenna over an AWGN channel and received by multiple receiving antennas, the multiple received signals experience additive white Gaussian noise with zero mean and unit variance and a less noisy estimate of the originally transmitted signal could be estimated from the multiple received signals. In this scenario, multiple receiving antennas can be exploited to mitigate the effect of degraded SNR on the classification accuracy. Now, for flat fading channels, estimating the transmitted signal from multiple received signals is more challenging since the received signals over the fading channel are not independent and each received signal consists of multiple paths arriving at the receiver at different time instances. The problem of estimating the originally transmitted signal without any knowledge of the channel and by only observing multiple received signals boils down to a blind channel estimation problem. Space-Time Block Codes (STBCs) and MIMO can provide solutions to such a problem, by estimating a less distorted versions of the transmitted signals blindly from the received signals, which can then enable reliable RF devices classification without requiring channel state information (CSI). In this work, we propose to leverage these MIMO benefits to mitigate the impact of flat fading in Rayleigh channels and degraded SNRs in AWGN channels on DNN-based device fingerprinting.

A. RELATED WORKS

Radio Frequency (RF) Fingerprinting has emerged to enable secure, authenticated communication by identifying wireless devices based on unique fingerprints from the devices' signals [19]. The inefficiency of deep learning-based device RF fingerprinting under time- and location-varying channel conditions has been a well-recognized challenge within the device fingerprinting research community. For instance, experimental results on WiFi device fingerprinting [16] show that the wireless channel condition severely degrades the classification accuracy, dropping it from 85% to 9%. They also show that equalizing the IQ data can improve the accuracy by up to 23%. Similarly, recent experimental studies on LoRa device fingerprinting [17] also show accuracy degradation, due to channel condition variation, from about 65% to about 20% when considering raw IQ data and from about 75% to about 4% when considering FFT data as the input to the deep neural nets. Sankhe et al. [12] benefit from the adaptivity feature of software-defined radios and modify the transmitter chain of these radios such that their respective demodulated symbols acquire unique characteristics that make the CNN robust to channel changes (the signal unique characteristics dominate the channel changes). Restuccia et al. [20], on the other hand, show that a carefully-optimized digital finite impulse response filter (FIR) at the transmitter's side, applying tiny modifications to the waveform to strengthen its fingerprint based on current channel conditions, can improve the accuracy from about 40% to about 60% in case of training on 5 devices. However, these research attempts depend on modifying the transmitted signals by either adding artificial impairments that are immune to the channel variations, or by filtering to alter the transmitted signals to maximize the model accuracy, thereby resulting in a potential impact on the BER. In addition, these techniques require changes to be made at the transmitters' side. Exploiting data augmentation techniques [21] and exploring new CNN architectural designs [22] have also been tried to mitigate the channel effect on RF fingerprinting. For instance, in [21], Slotani et al. propose a data augmentation step within the training pipeline that exposes the DNN to many simulated channel and noise variations that are not present in the original dataset. Testing is later performed on the collected IQ traces. This data augmentation technique shows 75% improvement in accuracy. To bypass the complicated manual features extraction and to achieve resilient RF fingerprinting, Peng et al. in [23] propose using heat constellation trace figure (HCTF) and slice integration cooperation (SIC) to extract more features automatically from the RF signal. The authors show that the ensemble learning-inspired approach achieves an identification accuracy of 91% at degraded AWGN channels of 0 dB and 100% at SNR values greater than 5 dB. However, the authors did not investigate their proposed approach over realistic fading channels. Elmaghbub et al. [15], [17], [18] propose a new technique that increases the robustness of fingerprinting against the channel condition variations, without requiring changes to be made at the transmitters nor the receiver. Their idea basically leverages spectrum emissions that are caused by inherent transceiver hardware impairments and that occur in the band surrounding the signal's original band (referred to as out-of-band spectrum emissions) to improve the robustness and insensitivity of device fingerprinting to wireless channel variations. Their proposed technique is evaluated using LoRa device datasets collected using 25 IoT devices and a USRP B210 receiver [17], [24] while considering both indoor and outdoor environment settings. Meneghello et al. in [25] propose DeepCSI, a technique that investigates the use of MIMO beamforming for RF fingerprinting in multi-user MIMO (MU MIMO) settings. DeepCSI leverages the beamforming feedback matrix information, estimated by beamformees (i.e., WiFi stations) and sent back to the WiFi access points (APs), to capture and identify fingerprints of the transmitting WiFi APs. DeepCSI relies, however, on the fact that APs are resourceful and on the feedback sent from the receiver to the sender. The Experimental results indicate that the proposed approach correctly identifies the transmitter with an accuracy of up to 98% and that the classification is immune to inter-users and inter-streams interference in multi-user MIMO systems. Basha et al. [26], [27] propose a new framework that leverages MIMO systems hardware capabilities to mitigate the channel effect and showed that, for Rayleigh channels, MIMO enabled blind partial channel estimation increases the testing accuracy by up to 40% when the CNN models are trained and tested over the same channel, and by up to 60% when the models are

tested on a channel that is different from that used for training.

B. CONTRIBUTIONS

In this paper, we propose a new technique that mitigates the impact of channel variation on device fingerprinting through the exploitation of MIMO diversity capabilities. Unlike previous fingerprinting approaches, our technique does not rely on software-defined radios, nor on changing the hardware impairments and altering the transmitted signals, nor on CSI to be collected from the receiver. More specifically, the contributions of this work are as follows:

- We show that for AWGN channels, combining the multiple received noisy versions of the transmitted signal in a SIMO (transmitters each has a single antenna, receiver has multiple antennas) system improves the training accuracy substantially compared to the conventional SISO (all devices each has single antenna) system. We also show that when tested on a different AWGN channel with degraded SNR, the testing accuracy improves by up to 30% compared to the SISO system.
- We show that combining the multiple received signals enabled via SIMO allows the DNN model to be trained on channels with lower SNR values without compromising the obtained classification accuracy. But at low SNR values, the same DNN model fails to distinguish between different devices as the impairments are overshadowed by the noisy channel effect in the case of the conventional SISO approach. This improvement increases with the number of receiving antennas.
- For flat fading channels, we expand the work in [26], [27] by leveraging MIMO capabilities to mitigate the channel effect without the need for altering the transmitted signals, nor impacting the bit error rates when using MIMO-enabled full blind channel estimation.
- We show that full blind channel estimation performed by leveraging the combined capability of MIMO and STBC improves the training accuracy by 50% over Rayleigh flat fading channel when compared to SISO. We also show that when tested on a different fading channel, the accuracy of the DNN model is improved by up to 69% when compared to SISO.
- We study the impact of the number of devices and the intensity of hardware impairments on RF fingerprinting. In addition, we study the effect of channel variations on RF fingerprinting in dynamic scenarios where the transmitter and/or the receiver are in relative motion.

Note that it is more practical to rely on resourceful devices (like access points/base stations) to perform the device classification task and to consider classifying transmitters that are SISO-enabled only. One of the techniques proposed in this work focuses on and assumes, however, that transmitters are multi-antenna equipped, and hence works in this scenario only. That said, we would like to mention that MIMO technology is being more and more adopted by both TX and RX ends in emerging standards like IEEE 802.11ax, and hence we believe that investigating identification techniques for these systems is also important and will be useful in emerging next-generation wireless networks.

The rest of the paper is organized as follows. Section II provides a MIMO background. Section III explains the techniques used for mitigating the effect of AWGN and flat fading channel models. Section IV presents the simulation results and performance evaluation. Finally, Section V concludes the paper.

II. BACKGROUND ON MIMO AND SPATIAL DIVERSITY

Among other well-known benefits [28], MIMO links improve SNR of multipath fading channels through spatial diversity by means of combining the output signals received on multiple uncorrelated antenna elements in presence of fading caused by multipath propagation [29], [30]. The improvement in the SNR achieved through diversity is characterized by [30]:

- Array gain, which measures the increase in average output SNR relative to the single-branch average SNR.
- Diversity gain, which measures the increase in the error rate slope as a function of the SNR.

Two types of diversity could be realized via MIMO: receive diversity and transmit diversity.

A. RECEIVE DIVERSITY

This can be realized with SIMO (single-input, multiple-output) links, i.e., when the receiver is equipped with multiple antennas and the transmitter is equipped with a single antenna. Receive diversity could be obtained by two combining methods: (i) selection combining, where the receiving antenna element whose signal offers the highest SNR is chosen for detection, and (ii) gain combining, where all receiving antenna signals are optimally combined to increase the overall SNR [30]. Compared to conventional single-receiving antenna systems, receive diversity yields a higher average SNR value (array gain), and a lower symbol error rate. In our studied device identification problem, the SNR improvement due to spatial diversity is exploited for mitigating the degradation of the classification accuracy that occurs due to AWGN channel variations (i.e., the DNN model is tested on a channel whose conditions are different from those used for training). Therefore, MIMO is leveraged in our framework to devise classification approaches that are agnostic to channel condition variations.

B. BLIND CHANNEL ESTIMATION AND SPACE-TIME BLOCK CODING

STBC (Space Time Block Coding) is a coding technique that achieves transmit diversity by spreading information symbols in space using multiple transmitting antennas and in time using pre-coding [30], [31]. Spreading in space and time is achieved by an $M \times K \quad (K \leq M)$ code matrix $\mathcal{C}$, where K is the time diversity of the code and M is the number of

transmitting antennas. Each column i of $\mathcal{C}$ corresponds to the signals transmitted by all the transmitting antennas at time epoch i, $1 \leq i \leq K$. The Alamouti scheme [32] is an example of STBCs that uses two transmit antennas, with a code matrix

$$\mathcal{C}_{Alamouti} = \begin{bmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{bmatrix}$$

where c_1 and c_2 are the symbols transmitted (each on one of the two transmit antennas) at the first time epoch, and $-c_2^*$ and c_1^* are the symbols transmitted at the second time epoch. When sending an STBC data matrix $\mathbf{S}$ using an $M \times L$ MIMO over a flat fading channel, the received signal matrix after K time epoch is given by:

$$\mathbf{R} = (\mathbf{I}_K \otimes \mathbf{H})\mathbf{C}\mathbf{S} + \mathbf{N} \tag{1}$$

where $\mathbf{C} = \begin{bmatrix} \mathbf{C}_1^T & \mathbf{C}_2^T & \cdots & \mathbf{C}_K^T \end{bmatrix}^T$ is a block matrix whose elements are $\mathbf{C_k}$ for $k = 1, \ldots, K$, and $\mathbf{C_k}$ is a matrix calculated from the code matrix $\mathcal{C}$ for $k = 1, \ldots, K$, $\mathbf{H}$ is the channel matrix, $\mathbf{N}$ is the noise matrix, and $\otimes$ denotes the Kronecker product operation [33], [34]. STBCs enable the blind estimation of the channel by observing only the received signals [33], [34], [35]. Eq. (1) shows that, for a MIMO system transmitting an STBC signal over a flat fading channel, each receiving antenna receives a signal that is a mixture of the signals transmitted by all the transmitting antennas, and each of the transmitted signals contributes to the mixture with a weight dictated by the channel matrix. The problem of estimating the transmitted signals given only the received signals and the properties of the transmitted signals, referred to as the blind source separation/blind channel estimation problem, essentially boils down to finding the inverse of the channel matrix, which can then be used to recover the transmitted signals [33], [34], [35], [36]. The recovered transmitted signals are less affected by the channel and are expected to achieve higher classification accuracy compared to the unprocessed received signals when used for RF fingerprinting.

C. BLIND ESTIMATION ALGORITHMS

This section provides background about two blind algorithms we used in this framework to enable reliable device classification and study the impact of the wireless channel on RF fingerprinting. The two MIMO-enabled RF fingerprinting approaches we studied in this paper are referred to as **MIMO 1** and **MIMO 2**. The first considered blind estimation algorithm [35], which enables the studied **MIMO 1** approach, blindly finds a closed-form estimation for the channel matrix using the orthogonal space-time block codes properties (OSTBC) and the second order statistics of the received signals (the received signal covariance matrix). Unlike the blind estimation method in [33], this method fully estimates the channel matrix. The vector form of the estimated channel matrix is given by [35]

$$\hat{\mathbf{h}} = \frac{\sqrt{tr\left\{\mathbf{A}\left(\tilde{\mathbf{h}_{\mathbf{opt}}}\right)^T \hat{\mathbf{G}} \mathbf{A}\left(\tilde{\mathbf{h}_{\mathbf{opt}}}\right)\right\} - K\sigma^2}}{tr\{\Gamma_s\}} \tilde{\mathbf{h}_{\mathbf{opt}}}$$
$$\tilde{\mathbf{h}_{\mathbf{opt}}} = \mathcal{P}\left\{\Phi^T\left(\mathbf{I}_{2K} \otimes \hat{\mathbf{G}}\right)\Phi\right\} \tag{2}$$

where $\mathbf{G}$ is the received signal covariance matrix, A is a matrix derived from the space-time block code, and Φ is a 4KMT x 2MN matrix where the k^{th} column of Φ is: $[\Phi]_k = A(e_k)$ and e_k is canonical vector with 1 in the k^{th} entry and zeroes otherwise. Γ_s is the diagonal covariance matrix of the transmitted symbols $\mathbf{s}$ before space-time coding, and it is known at the receiver. The operators $\mathcal{P}\{.\}$ and $tr\{.\}$ denote the normalized principal eigenvector of a matrix, and the trace of a matrix, respectively. The second blind estimation algorithm [33], [34] used in the **MIMO 2** approach aims at determining the subspace of the channel matrix. First, the algorithm starts by finding $\mathbf{N}_L$, the left null space of $\mathbf{R}$. Eq. (1) yields the following blind equation [33], [34]

$$\mathbf{N}_L^H(\mathbf{I}_K \otimes \mathbf{H})\mathbf{C} = \mathbf{0} \tag{3}$$

which represents a homogeneous linear equation in the unknown $\mathbf{H}$, and the uniqueness of the solution depends on the matrix $\mathbf{C}$ of the STBC and the rank of the matrix $\mathbf{H}$ as explained in [33], [34]. Second, the blind algorithm decouples the channel matrix subspace from Eq. (3) to get:

$$\overbrace{\left(\sum_{k=1}^{K} \mathbf{C}_k^T \otimes \left(\mathbf{N}_L^H \mathbf{E}_k\right)\right)}^{\triangleq \Delta} \mathbf{h} = 0 \text{ with } \mathbf{E}_k \triangleq \begin{bmatrix} \mathbf{0}_{L(k-1)\times L} \\ \hline \mathbf{I}_L \\ \hline \mathbf{0}_{L(K-k)\times L} \end{bmatrix} \tag{4}$$

Eq. (4) shows that the channel subspace, $\mathbf{h}$, lies in the null space of Δ. Yet, this blind method estimates the channel partially up to some ambiguity, and the actual channel cannot be identified from its rotated versions due to the remaining complex ambiguity [34].

III. LEVERAGING MIMO FOR CHANNEL-AGNOSTIC WIRELESS DEVICE IDENTIFICATION

Prior deep learning-based fingerprinting approaches suffer from the impact of time- and location-varying channel conditions. In other words, although they show promising results when the learning models are trained and tested on data collected under the same channel conditions, these approaches perform poorly when training and testing are done on data collected under different channel conditions. In this section, we propose new fingerprinting approaches enabled by the MIMO systems that are resilient to channel condition changes. The novelty of our techniques lies in leveraging the capabilities offered by MIMO systems to mitigate the distortions in the received RF signals caused by the wireless channel, thereby making the learning models agnostic to the underlying channel. We consider both AWGN and flat

fading Rayleigh channel models in this framework, which we present next.

A. AWGN CHANNEL-AGNOSTIC DEVICE FINGERPRINTING

For AWGN channels, we leverage the capability of SIMO systems (the receiver is equipped with multiple antennas but transmitters are each single-antenna equipped) to combat the impact of channel variations. For a SIMO system, with L receiving antennas, sending a signal $s(t)$ over an AWGN channel with zero mean and σ^2 noise power, the received signal by the ith receiving antenna is $r_i(t) = s(t) + n_i(t)$, where $n_i(t)$ is the noise seen at antenna i. Our technique proposed for AWGN channels consists of averaging the received signals over all the receiving antennas to achieve SNR gains that mitigate the channel noise, and using this averaged received signal $r(t) = s(t) + \frac{1}{L}\sum_{i=1}^{L} n_i(t)$ for training the DNN models. This yields a classification accuracy that is less sensitive to noise, and more agnostic to AWGN channels. In addition, testing the models over an AWGN channel with a lower SNR using the averaged $r(t)$ is less affected by the noise, and the reduction in the testing accuracy due to the change of the channel is lower than what a single-antenna receiver achieves. Since the noise samples received by the antenna elements are i.i.d., the noise power decreases by a factor of L and $r(t)$ is considered an unbiased, consistent estimator of the originally transmitted waveform $s(t)$. Increasing the number of receiving antennas L makes $r(t)$ more immune to the channel noise when compared to a waveform received by a single antenna.

B. FLAT FADING CHANNEL-AGNOSTIC DEVICE FINGERPRINTING

For the Rayleigh flat fading channel, the channel matrix transforms the transmitted waveforms causing the accuracy degradation for pre-trained classification models. The channel matrix mixes the transmitted signals, and the received signal by each of the receiving antennas is a mixture of the signals transmitted by all the transmitting antennas. Estimating the transmitted signals given the received signals and the properties of the transmitted signals is a problem known as blind source separation/blind channel estimation. Blind source separation methods have shown efficiency in removing the channel effects on the transmitted signals without the need for training pilots. In this work, we investigate two blind estimation methods: The blind algorithm in [35], **MIMO 1**, fully estimates the channel matrix from the received signals. The estimated channel matrix is then used to estimate the originally transmitted signal used for classification. The blind algorithm in [33], [34], **MIMO 2**, partially estimates the channel matrix up to some complex ambiguities and finds the channel matrix subspace. If the estimated channel subspace is used to reconstruct the originally transmitted signal, the reconstructed transmitted signal is expected to be less affected by the channel, and the effect of the channel is less when the number of remaining ambiguities is minimized. To guarantee the minimum ambiguity in the estimated channel matrix, i.e., a single complex ambiguity in the estimated channel, we simulate a 3×3 MIMO system that transmits a QPSK signal using Tarokh STBC of rate $1/2$ (code length $= 8$) [34]. A single complex ambiguity is expected to have a minor effect on the classification accuracy. We exploit Convolution Neural Network (CNN) high dimensional feature mapping capabilities and high performance in classifying RF devices to achieve accurate classification from the reconstructed transmitted signals despite the remaining ambiguities.

At the transmitter side, every 80 symbols are modulated using QPSK, then the modulated symbols are encoded into blocks using Tarokh STBC of rate $1/2$ (code/block length $= 8$) such that each transmitted block encodes 4 QPSK modulated symbols. For 20 transmitted blocks, we construct the STBC symbol matrix $\mathbf{S}$ with the size of 24×20, where the columns represent the transmitted blocks from the 3 transmitting antennas. The signals transmitted by each of the transmitting antennas are obtained by reshaping $\mathbf{S}$ into a 3×160 matrix $\mathcal{S}$. Each row in $\mathcal{S}$ represents the signal transmitted by each of the 3 transmitting antennas. The transmitted signals are then impaired by the MIMO Rayleigh channel. The channel is fixed for 20 transmitted blocks. At the receiver side, we first apply each of the blind algorithms to determine the channel matrix subspace as explained in II. We first collect 20 received blocks from the 3 receiving antennas in a matrix $\mathcal{R}$ with the size of 3×160 to construct the matrix Δ. Then we construct the matrix $\mathbf{R}$ for the received STBC blocks. R is a 24×20 matrix, where the columns represent the received blocks from the 3 receiving antennas.

For the MIMO-enabled classification approach based on blind full channel estimation, **MIMO 1**, to find $\tilde{\mathbf{h}}_{\mathbf{opt}}$, we calculate the matrix $\mathbf{G}$ and Γ_s from $\mathbf{R}$ and $\mathbf{S}$, respectively. $\tilde{\mathbf{h}}_{\mathbf{opt}}$ is determined via SVD. Then, Eq. (2) is used to find the estimated channel matrix. For the MIMO-enabled classification approach based on blind partial channel estimation, **MIMO 2**, we compute matrix Δ in Eq. (4) from the received signal matrix $\mathbf{R}$ and matrix $\mathbf{C_k}$ determined by the STBC used from the transmission. The channel subspace $\mathbf{h}$ is found using SVD of matrix Δ. For both methods, **MIMO 1** and **MIMO 2**, we reconstruct the transmitted signals using the inversion of the estimated channel subspace:

$$\hat{\mathcal{S}} = \mathbf{H}^{-1}\mathcal{R}$$

The reconstructed transmitted signals $\hat{\mathcal{S}}$ are then sampled with a window size of 160 for each of the simulated devices to create the training, validation, and testing detests. The reconstructed transmitted signals $\hat{\mathcal{S}}$ are less affected by the fading channel and hence are expected to achieve higher accuracy compared to the raw IQ data when used to train and test the DNN models. The reconstructed signals are also expected to be more immune to the channel conditions variations when used for classifying devices using previously trained DNN models on varying channel conditions.

TABLE 1. Hardware impairments used to simulate 10 different devices [26], [27].

Device	Phase Noise	Frequency Offset	IQ Gain Imbalance	IQ Phase Imbalance	AMAM	AMPM	Real DC Offset	Imaginary DC Offset
DV1	-60	20	0.08	0.1	[2.1587,1.1517]	[4.0033,9.104]	0.1	0.15
DV2	-60.15	20.01	0.1	0.09	[2.1687,1.1617]	[4.1033,9.124]	0.11	0.14
DV3	-59.9	20.2	0.09	0.09	[2.1789,1.1317]	[4.0933,9.151]	0.1	0.11
DV4	-60.1	20	0.108	0.109	[2.1987,1.1217]	[4.1033,9.194]	0.1	0.1
DV5	-60	20.09	0.1	0	[2.1587,1.1717]	[4.093,9.094]	0.089	0.1008
DV6	-59.95	20.1	0.12	0.15	[2.1487,1.1117]	[4.1033,9.156]	0.1	0.098
DV7	-59.93	20.11	0.11	0.11	[2.1897,1.1237]	[4.1133,9.135]	0.111	0.1011
DV8	-60.13	20.099	0.101	0.14	[2.1387,1.1627]	[4.1533,9.096]	0.12	0.099
DV9	-59.89	19.9	0.099	0.08	[2.1548,1.1917]	[4.09833,9.10056]	0.09	0.0999
DV10	-59.91	19.98	0.111	0.105	[2.1777,1.09874]	[4.0987,9.123]	0.101	0.10015

TABLE 2. Hardware impairments sets.

	IQ Gain mean	IQ Gain Standard Deviation	IQ Phase Mean	IQ Phase Standard Deviation
Low Impairments Set	0.1	0.01	0.09	0.02
Moderate Impairments Set	0.1	0.055	0.09	0.11
High Impairments Set	0.1	0.1	0.09	0.2

IV. PERFORMANCE EVALUATION AND ANALYSIS

For all the techniques proposed in this work to mitigate the impact of AWGN and flat fading channels on classification accuracy, we use MATLAB R2021b to build our wireless communication system model and generate the IQ datasets.

A. SYSTEM SCENARIOS

We consider studying two system scenarios:

- *Scenario 1—SIMO Over AWGN Channels:* In this scenario, we leverage SIMO hardware capabilities to mitigate the effect of AWGN on RF fingerprinting accuracy. All 10 devices to be classified are equipped with a single transmitter antenna, and the authenticator (e.g., access point) is equipped with multiple receiving antennas. MATLAB WLAN toolbox waveform generator is used to simulate different RF devices' waveforms impaired with the impairments sets described in Table 1. The impairments are unique for each device, and the devices are slightly different to mimic the slight variations that inevitably occur during manufacturing and distinguish each device [15], [37]. For each device, the transmitted signal is impaired by the AWGN channel, and the receiver (the authenticator) receives the signal using multiple receiving antennas and averages all the received signals. We varied the number of receiving antennas and the channel SNR to study the impact of channel degradation on RF devices classification accuracy. For each device, we collected 5000 frames, with each frame having the size of 160. Then, we split the real and the imaginary parts of the signal and reshaped the frames as 2×160 vectors to be fed to the input layer of the CNN. The dataset was divided into 80% for training, 10% for validation, and 10% for testing.
- *Scenario 2—MIMO and STBC Over Rayleigh Flat Fading Channels:* In this scenario, we leverage MIMO and STBC-enabled blind channel estimation to mitigate the effect of Rayleigh flat fading on the fingerprinting accuracy. In order to ensure that the flat fading channels are blindly identifiable from the received signals for both studied blind estimation techniques, **MIMO 1** and **MIMO 2** explained in Section III, all devices to be classified are each equipped with 3 transmitting antennas, and the receiver/authenticator is equipped with 3 receiving antennas. MATLAB WLAN toolbox is used to add hardware impairments to the transmitted signals. The 3×3 MIMO system sends QPSK symbols encoded via Tarokh STBC $\mathcal{O}_3$ [31], and the authenticator performs blind channel estimation to recover the signals. The estimated signals are then used for training and testing the CNN model. In this scenario, we studied the impact with 3 different impairments sets that differ in intensity as shown in Table 2.

In the simulation, we assume that the devices are WiFi-enabled and that the hardware impairments for a specific device are fixed and do not change within a specific tolerance over time. Such an assumption is widely adopted in the context of RF fingerprinting. We also assume that the effect of the hardware impairments on the authenticator/receiver is negligible, which is a valid assumption since the receiver/authenticator is expected to be high-end/performing devices with minimal impairments. We also assume that, for each device, the effect of the impairment is the same for all the transmitting antennas. We set the values of the hardware impairments based on previous works on RF impairments modeling and previous works on RF fingerprinting. We followed the impairments modeling provided in [11], [38] and the parameter settings used in [10], [38]. Essentially, these parameters are chosen in such a way that the device distinguishability can be easily controlled.

B. CNN ARCHITECTURE

We used the CNN architecture in [12] as a benchmark to evaluate the proposed MIMO-enabled approaches and compare them to the SISO/conventional approach. The CNN

architecture consists of two convolution layers and two fully connected layers. The 2×160 input is fed into the first convolution layer that consists of 50 1×7 filters. This layer produces 50 feature maps from the entire input. The second convolution layer consists of 50 2×7 filters, where each filter is convoluted with the 50-D volumes obtained from the first layer. The second convolution layer learns variations over both I and Q dimensions. Each convolution layer is followed by a ReLU activation function to add non-linearity, and a 2-stride max pooling layer to prevent overfitting. The first fully connected layer consists of 256 nodes whose output is fed into the second fully connected layer of 80 nodes. Each fully connected layer has a ReLU activation. The last layer is a softmax classifier to generate the classification probabilities. At the classifier output, the cross-entropy loss is calculated and the back-propagation algorithm is used to find the network parameters that minimize the prediction error. The CNN uses Adam optimizer with a learning rate of 0.0001.

C. PERFORMANCE METRICS

We consider the following metrics in this evaluation.

- *Training Accuracy*, the percentage of the correctly classified samples to the total number of samples in the training datatset.
- *Testing Accuracy*, the percentage of the correctly classified samples to the total number of samples in the testing datatset.
- *Relative Different channel Testing Gap (RDTG)*, the percentage of reduction in the testing accuracy occurred when the testing channel is different from the training channel. Precisely, RDTG is defined as

$$\text{RDTG} = \frac{\text{same channel testing acc.} - \text{different channel testing acc.}}{\text{same channel testing acc.}} \%$$

  If the classification technique is perfectly channel-agnostic, then the RDTG value is zero, and the deviation from zero quantifies the effect of the channel variation on accuracy. RDTG values are only calculated when the training and the testing channels are different.

In the evaluation of the proposed fading channel-agnostic techniques, we also vary the following parameters:

- *Training APG*, the average path gain (APG) of the Rayleigh flat fading channel used for training. APG is varied from -20 dB to 20 dB.
- *Testing APG*, the average path gain (APG) of the Rayleigh flat fading channel used for testing. APG is varied from -20 dB to 20 dB.
- *Training MDS*, the maximum Doppler shift (MDS) of the Rayleigh flat fading channel used for training. MDS has two values: 0 Hz to 1 Hz.
- *Testing MDS*, the maximum Doppler shift (MDS) of the Rayleigh flat fading channel used for testing. MDS has two values: 0 Hz to 1 Hz.
- *Number of Devices*, the number of devices to be classified (number of classes). We simulated 10 devices and 20 devices to study the effect of the number of devices to be classified on the classification accuracy.
- *Impairments Intensity - The IQ Imbalance Standard Deviation* , the IQ- gain and -phase imbalances standard deviation values. A low standard deviation value is used to generate random impairments that mimic devices that are almost similar whereas high values correspond to devices that are more distinguishable. We used three standard deviation values, and hence 3 impairments sets. The impairments sets are shown in Table 2.

In this section, we study the resilience of the proposed MIMO-enabled classification approaches. We use MIMO to refer to the case when the receiver (the authenticator/AP) and all transmitters (the devices to be classified) are each equipped with multiple antennas, SIMO to refer to the case when the receiver (the authenticator/AP) is equipped with multiple antennas and the transmitters (the devices to be classified) are each equipped with a single antenna, and SISO to refer to the conventional case when the receiver (the authenticator/AP) and all transmitters (the devices to be classified) are each equipped with a single antenna. For scenario 2, we study the resilience of two MIMO-enabled approaches: **MIMO 1**, which is based on the blind full estimation approach proposed in [35], and **MIMO 2**, which is based on the blind partial estimation approach proposed in [33], [34] over flat fading channel. The proposed MIMO-enabled approaches are compared against the conventional SISO fingerprinting technique. It is worth noting here that prior literature on the use of MIMO to enable RF fingerprinting is limited and hence finding a prior MIMO-based framework to serve as a benchmark was not possible. To the best of our knowledge, the only one prior related work we found is DeepCSI [25], which investigates the use of MIMO beamforming in the context of multi-user MIMO to identify 802.11ac/ax access points (APs). Unfortunately, DeepCSI cannot be used as a benchmark for our work, because of its different scope and system settings, as DeepCSI is meant for multi-user systems with MIMO being leveraged for enabling multi-user streams via beamforming. Our proposed technique, on the other hand, is meant for single-user systems with MIMO being leveraged for spatial diversity via STBC. Besides, DeepCSI intends to classify WiFi APs as the transmitters via beamforming, and hence relies on the fact that APs are resourceful, while our work intends to classify WiFi stations as non-resourceful transmitters. In addition, DeepCSI relies on feedback to be sent by the receiver to the sender, whereas our approach does not require any feedback from the receiver.

D. SCENARIO 1: SIMO SYSTEM OVER AWGN CHANNEL RESULTS

In this section, we consider studying the benefit of exploiting the receiver's (the authenticator's) multiple antenna capability of the SIMO system in overcoming the impact of channel impairments in AWGN channel models. The

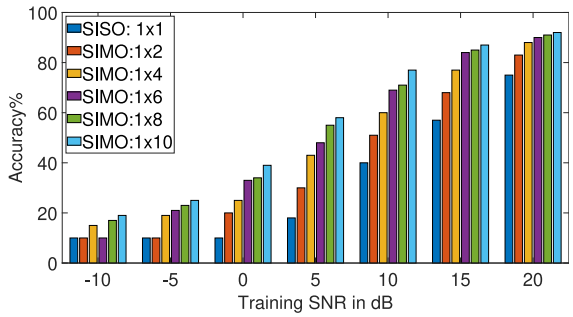


FIGURE 1. Training accuracy obtained when varying the number of receiving antennas.

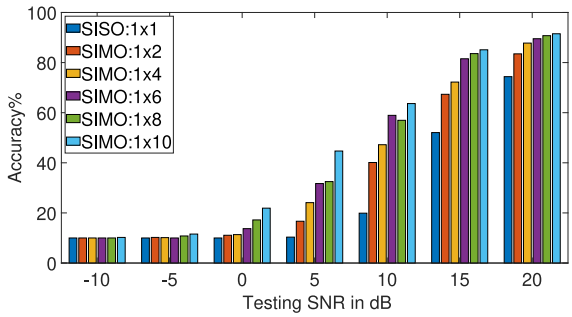


FIGURE 2. Testing accuracy obtained when the training channel SNR is fixed at 20 dB while the testing SNR is varied from −10 dB to 20 dB.

obtained results are collected under scenario 1 and are compared to conventional SISO systems.

For each of the 10 simulated devices, whose hardware impairments values are shown in Table 1, the transmitted IQ signals are impaired with AWGN channels of SNR values ranging from −20 dB to 20 dB and then received and sampled at the receiver by its multiple receiving antennas. Once received, the waveform is averaged as explained in Section III, and the averaged signal is used to create the dataset used for training, testing, and validation.

1) SAME TRAINING AND TESTING AWGN CHANNELS

Fig. 1 shows the training accuracy obtained under different SNR values for SIMO links with varying number of receiving antennas. First and as expected, the figure shows that the worse the AWGN channel (i.e., the smaller the SNR value), the less accurate the CNN training outcome. For instance, a SIMO system with 2 receiving antennas achieves a training accuracy of 84% at 20 dB, but only a 51% accuracy at 10 dB. Second, it is also observed that increasing the number of receiving antennas yields higher training accuracy when considering the same SNR. For example, training a 4-antenna SIMO system at 15 dB increases the training accuracy from 57% to 77% compared to the SISO system, whereas, a 10-antenna SIMO system increases the training accuracy from 57% to about 87%. Third, the figure shows that the improved training accuracy of SIMO systems over the SISO system is more significant at SNR values ranging from 15 dB to 0 dB. However, at severe noisy channels with SNR values less than 0 dB, increasing the number of receiving antennas from 1 to 10 does not achieve a training accuracy higher than 40% reflecting that the SNR gain acquired from combining the multiple received signals is no longer compensating the noisy channel effect and that the impairments are significantly overshadowed by the noise.

2) DIFFERENT TRAINING AND TESTING AWGN CHANNELS

In Fig. 2, we show the testing accuracy obtained under different SNR values also while considering a SIMO system with a varying number of receiving antennas. Here, CNN models are trained at SNR = 20 dB, but tested at different SNRs varying from −10 dB to 20 dB. We make three observations from this figure. First, we observe that a decrease in the SNR value of the testing channel results in a decrease in the classification accuracy, and such a decrease is more profound for lower SNR values. This is true regardless of the number of receiving antennas. This observation clearly shows again the challenging impact that channel quality variations have on the classification accuracy, where achieving high accuracy when the channel used for testing has the same quality as that used for training does not guarantee a similar accuracy performance when the channel used for testing is different from that used for training. However, such a discrepancy is less significant under the proposed SIMO approach, where leveraging multiple receiving antennas plays a key role in maintaining high accuracy even in the presence of channel quality degradation. For example, while testing over a 10 dB channel makes the SISO system accuracy drop from about 75% to about 20%, it makes the accuracy of a SIMO system with 6 receiving antennas drop from about 90% to about 59% only. This demonstrates the benefit of SIMO in overcoming the impact of channel quality variations, and in making our proposed technique more channel agnostic. Second, the figure also shows that the higher the number of receiving antennas is, the higher the achieved accuracy is, and this is regardless of the testing SNR. For example, at a tested channel of 15 dB, a SIMO system with 2, 4, 6, and 10 antennas achieve classification accuracy of about 67, 72, 82, and 85%, respectively. Third, we observe that for severely degraded channels (i.e., low SNRs), the CNN models perform poorly regardless of the number of receiving antennas, and the achieved classification accuracy is as good (or as bad) as what random classification would be. This can clearly be seen from Fig. 2, where the accuracy is similarly low for all studied systems.

The observed accuracy gain is achieved by the increased SNR value resulting from the averaging method used by the proposed SIMO approach. Such an SNR gain mitigates the noise effect allowing the CNN models to identify devices in the presence of impairments in the waveform. However, for severe noisy channels where the SNR gain obtained from increasing the number of receiving antennas is no longer enough to mitigate the increase in the noise power, the SIMO

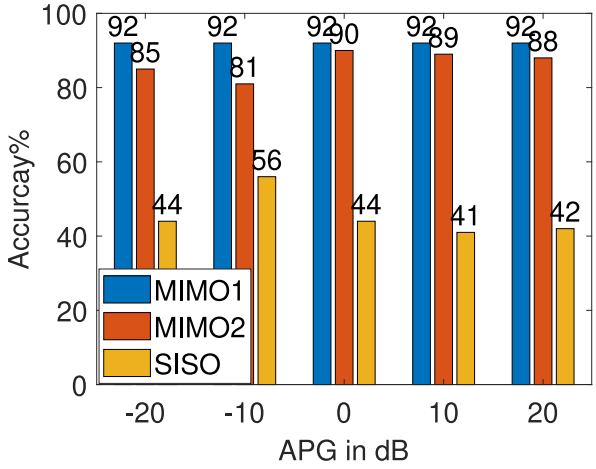


FIGURE 3. Testing accuracy when the same Rayleigh channel (same APG value) is used for training and testing. 10 devices are used for classification.

approach performance degrades and shows no improvement over the SISO system. The results for both training and testing are commensurate, as both indicate that the averaging method using SIMO systems increases the SNR of the averaged signal as previously explained. Thus the CNN is able to capture the impairments effects on the waveform more efficiently despite the noise.

E. SCENARIO 2: MIMO SYSTEM WITH STBC OVER RAYLEIGH FLAT FADING CHANNEL RESULTS

In this scenario, we consider studying the benefit of exploiting MIMO and STBC capabilities in overcoming the impact of channel impairments in Rayleigh flat fading channel models. The obtained results are collected under scenario 2 are compared to conventional SISO systems.

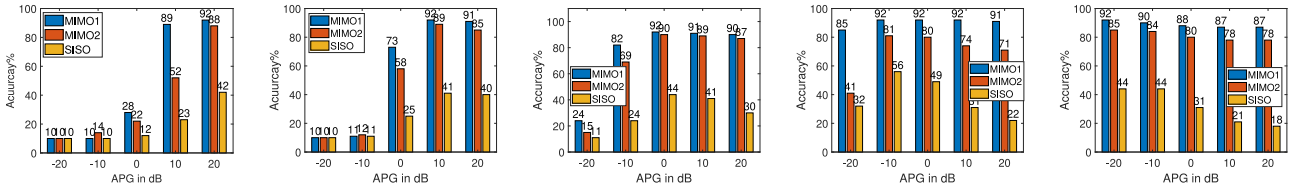
1) SAME TRAINING AND TESTING RAYLEIGH CHANNELS

In this section, we analyze the results obtained when training and testing are performed over the same Rayleigh fading channel for **MIMO 1**, **MIMO 2**, and SISO approaches. We used 10 devices impaired with the low impairments set for classification. Fig. 3 shows the obtained testing accuracy while varying the APG values of the channel. This figure clearly shows that the MIMO-enabled approaches **MIMO 1** and **MIMO 2** increases the testing accuracy significantly compared to the conventional/SISO approach. For instance, we observe that at training and testing APG of -20 dB, the MIMO-enabled approaches increase the testing accuracy from 44% to up to 92% for **MIMO 1** and 85% for **MIMO 2** compared to the SISO approach, thereby doubling the obtained accuracy. In addition, we observe that the accuracy improvement that MIMO-enabled approaches offer over the conventional approach is consistent across the entire APG value range, i.e., our proposed MIMO-enabled classification approaches double the accuracy regardless of the APG value.

2) DIFFERENT TRAINING AND TESTING RAYLEIGH CHANNELS

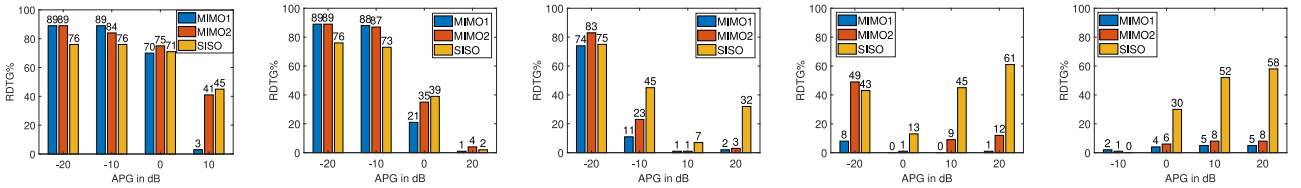
We now present, analyze and compare the results obtained when the channel used for training is different from that used for testing. This scenario mimics real-world settings where the classification models are trained under certain channel conditions but then used later for real-time device classification under possibly different channel conditions.

1) Impact of APG (Testing Accuracy): Fig. 4 shows the testing accuracy over Rayleigh channels where the models are trained and tested over channels with different APG values, ranging from 20 to -20 dB. For this figure, MDS is set to 0 Hz. This figure shows the effect of the wireless channel on the device classification accuracy. For instance, in Fig. 4(e), we observe that when the conventional/SISO method is used and the models are trained over a channel with training APG of -20 dB and tested over a channel with testing APG of 20 dB, the accuracy drops from 44% to 18%. Our observation indicating the seriousness of the channel impact on device classification accuracy is well aligned with previous work findings as discussed in Section I. The figure also shows that the proposed MIMO-enabled approaches, **MIMO 1** and **MIMO 2**, significantly overcome the impact of channel variations by improving the testing accuracy over the conventional/SISO approach when the CNN is trained and tested on different channels, and this is especially true when the training channels exhibit severe fading (small APG values). In addition, we observe that the MIMO-based approaches testing accuracy is more stable when the CNN is trained at severe fading channels (small APG values). The figure also shows that the **MIMO 1** approach that is based on full channel blind estimation outperforms the **MIMO 2** approach that is based on the partial blind estimation up to a complex ambiguity. Looking at Fig. 4(a), which depicts the testing accuracy under varied testing APG values of the testing channel while fixing the training APG to 20 dB, we observe that for testing APG greater than 0 dB, MIMO achieves improved performance over SISO. However, when the testing APG is less than 0 dB, the testing accuracy is unreliable, and both SISO and the MIMO approaches are equivalent. Now in Fig. 4(b), which depicts the testing accuracy under varied APG values of the testing channel while fixing the APG of the training channel to 10 dB, we observe that the MIMO-based approaches achieve significantly higher testing accuracy compared to the SISO approach when the testing APG is greater than 0 dB. For instance, when the training APG is 10 dB and the testing APG is 20 dB (channel with less severe fading), **MIMO 1** system achieves a testing accuracy of about 91%, whereas only about 40% is achieved under SISO. Moreover, when the testing APG $= 0$ dB (channel with more severe fading), the testing accuracy improves from 25% to about 73% when considering the **MIMO 1** approach versus the SISO approach. Fig. 4(c) and Fig. 4(d) capture the testing accuracy under varied APG values of the tested channel



(a) Training APG = 20dB (b) Training APG = 10dB (c) Training APG = 0dB (d) Training APG = -10dB (e) Training APG = -20dB

FIGURE 4. Impact of APG on accuracy. MDS is fixed at 0 Hz. Number of devices = 10 with the low impairments set.



(a) Training APG = 20 dB (b) Training APG = 10 dB (c) Training APG = 0 dB (d) Training APG = -10 dB (e) Training APG = -20 dB

FIGURE 5. Impact of APG on RDTG. MDS is fixed at 0 Hz. Number of devices = 10 with the low impairments set.

while fixing the training APG to 0 and 10 dB, respectively. From the two figures, we observe the same trends of Fig. 4(b), and the MIMO-based approaches achieve higher classification accuracy compared to the SISO/conventional approach when the testing APG is different from the training APG. Fig. 4(e) captures the testing accuracy when varying the testing APG values at a fixed training APG of -20 dB. From this figure, we first observe that when testing on different channels with testing APG values varying from -10 dB to 20 dB, the MIMO-based classification approaches achieve an improved and stable testing accuracy when compared to the SISO approach. For instance, when the training APG is -20 dB and the testing APG is 20 dB, **MIMO 1** approach achieves up to 70% increase in the testing accuracy over the SISO approach. Second, when the training APG is -20 dB (severe fading channel), the testing accuracy of the MIMO approaches at different channels with testing APG varying from -10 dB to 20 dB does not go below 78% compared to the SISO approach where the testing accuracy degrades to 18%.

2) Impact of APG (RDTG): We now assess the robustness of the proposed MIMO-enabled fingerprinting approaches against Rayleigh channel condition variations through the study of the RDTG performance metric. Fig. 5 shows the RDTG values when training and testing are done over Rayleigh channels with varying APGs. First, we observe that compared to the SISO approach, the MIMO-based approaches provide a much higher resilience to channel variations, i.e., yield smaller RDTG values, when the training channel exhibits severe fading conditions (training APG $= -10$ dB; Fig. 5(d), and training APG $= -20$ dB; Fig. 5(e)). In Fig. 5(d), and Fig. 5(e), the RDTG values can go as high as 61% under the SISO approach but do not exceed 12% under the proposed MIMO-enabled approaches. Second, for high training APG values (e.g., 20 dB in Fig. 5(a) and 10 dB in Fig. 5(b)), we observe that although the MIMO approaches still outperform the SISO approach in terms of testing accuracy, the RDTG gap for SISO is less compared to the MIMO-based approaches. For instance, in Fig. 5(a), when testing APG $= -20$ dB, **MIMO 1** achieves an RDTG value of 89% and the SISO approach achieves a value of 76%. We observe that the closer the testing APG values are to the training ones, the smaller the RDTG values achieved under MIMO. This observation is commensurate with the previous observations made in Fig. 4 about the MIMO-based approaches testing accuracy degradation when the models are tested on more severe fading channels. Fig. 5(e), depicting RDTG values when training APG $= -20$ dB, shows that when the fading conditions of the training channel become worse, the SISO approach continues to degrade significantly, but not so for the MIMO approaches. As an example, when testing APG $= 0$ dB, the SISO approach yields an RDTG value of 30% whereas **MIMO 1** and **MIMO 2** achieve RDTG values of only 4% and 6%, respectively. Moreover, as the testing APG continues to increase, while the MIMO-based approaches maintain stable RDTG values (about 8%), SISO testing accuracy continues to degrade significantly, reaching RDTG values of up to 58%. One explanation for the high RDTG values for the proposed MIMO-enabled approaches when the channel used for testing exhibits severe fading (high APG values) compared to the channel used for training is that despite the MIMO-enabled blind estimation, there remains some ambiguity (scalar ambiguity for **MIMO 1** and complex ambiguity for **MIMO 2**) in the estimated channel. The remained ambiguity affects the learning models in the training phase, and the models learn features that are extracted from both channel and device impairments. Therefore, training at less severe flat fading channels while testing at more severe fading channels yields a significant reduction in accuracy. However, the MIMO-based approaches still outperform the SISO approach

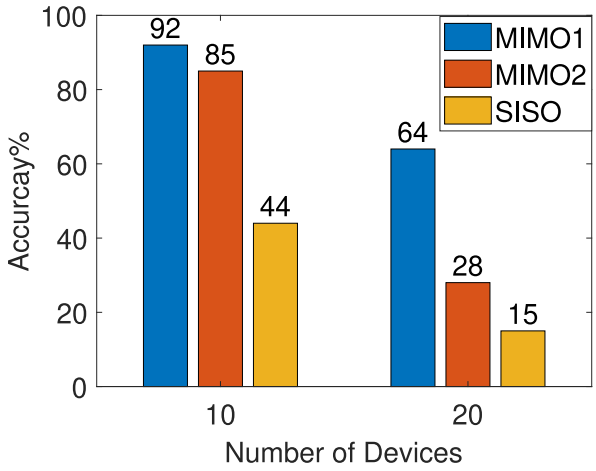


FIGURE 6. Impact of the number of devices on the testing accuracy when the same Rayleigh channel (APG = −20 dB) is used for training and testing.

in spite of these unsolved ambiguities in terms of testing accuracy.

F. IMPACT OF NUMBER OF DEVICES

In this section, we study the impact of the number of devices on the classification accuracy. We doubled the number of devices while fixing the impairments intensity/ IQ imbalance standard deviation; i.e., the devices are impaired using the low impairments sets. By doing this, we make the devices less distinguishable and the task of classification over fading channels more challenging.

1) SAME TRAINING AND TESTING RAYLEIGH CHANNEL

Fig. 6 shows the testing accuracy obtained when the training and testing APG = −20 dB, and the number of devices is varied. We observe that the classification accuracy decreases when the number of devices increases for the MIMO-enabled approaches as well as the SISO approach. This reduction in accuracy with the number of devices limits the RF/device fingerprinting scalability. However, the figure shows that while the conventional SISO approach accuracy decreases from 44% to 15% when the number of devices is doubled, the proposed **MIMO 1** approach accuracy decreases from 92% to 64%. This improvement in accuracy over SISO grants a more scalable RF fingerprinting system. The figure also shows that **MIMO 2** approach accuracy decreases severely from 85% to 28% when the number of devices is doubled. This observation is expected as the **MIMO 2** approach partially estimates the channel blindly up to a complex ambiguity before classification, whereas the **MIMO 1** approach completely estimates the channel before classification. This observation proves the severe effect of the channel on the classification accuracy. We also observe that the identification decline rate of the proposed MIMO approaches when the number of devices is doubled is greater compared to the SISO approach. However, this observation is expected due to the higher recognition rate of the proposed MIMO approaches as it is normal for a model with a high recognition rate to have a faster drop in accuracy compared to a model with a lower recognition rate at the beginning.

2) DIFFERENT TRAINING AND TESTING RAYLEIGH CHANNELS

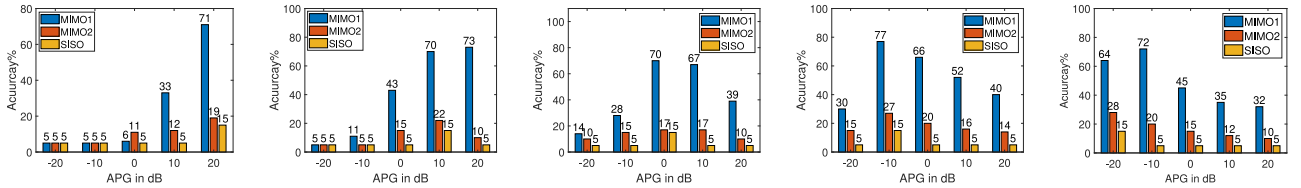
The limited scalability of the conventional SISO approach is even more demonstrable when considering the effect of channel variations. Fig. 7 shows the effect of channel variations on the testing accuracy when the number of devices is doubled. The figure shows that the **MIMO 1** approach outperforms **MIMO 2** and the conventional SISO approach. The figure also indicates that the full channel estimation leads to higher classification accuracy. For instance, in Fig. 7(e) when the training APG = −20 dB, **MIMO 1** approach where blind channel estimation is performed before the classification achieves up to 72% testing accuracy, whereas the partial blind channel estimation in **MIMO 2** achieves only up to 28% testing accuracy. The conventional SISO approach completely fails to mitigate the channel variation effects, and the testing accuracy is as low as random guessing. The previous observations show that full channel estimation is required for reliable RF fingerprinting. Fig. 8 shows the impact of channel variation on the RDTG value when classifying 20 devices impaired with the low impairments sets. The figures show that, when the channel used for training has a smaller APG value (severe fading) compared to the channel used for testing (Fig. 8(e)), **MIMO 1** approach is more resilient to channel variations, i.e., small RDTG values, compared to **MIMO 2** and the conventional SISO approach. We also observe that when the training APG is high compared to the testing APG (Fig. 8(a)), the MIMO-based approaches have higher RDTG values compared to SISO. However, this observation could be explained by the high same-channel testing accuracy obtained by the MIMO-based approaches compared to SISO. Also, note that the MIMO-based approaches still outperform the conventional SISO approach in terms of the classification accuracy.

G. IMPACT OF IMPAIRMENTS INTENSITY: IQ IMBALANCE STANDARD DEVIATION

In this section, we investigate the impact of the IQ imbalance standard deviation on the classification accuracy when classifying 20 devices. The IQ imbalance standard deviation values we used are shown in Table 2. A high IQ imbalance standard deviation value generates random impairments that are more distinguishable and vice versa. Therefore, the classification accuracy is expected to increase with the increase in the IQ imbalance standard deviation for the same number of devices.

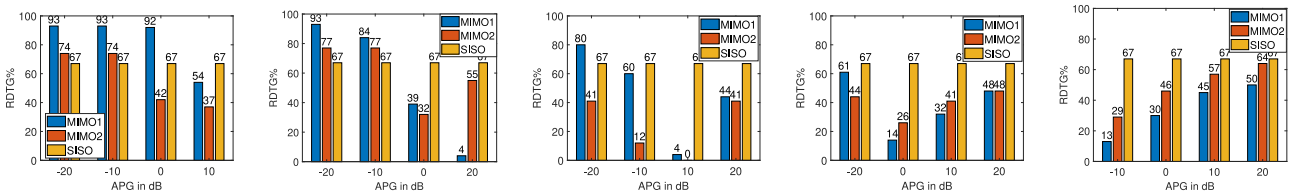
1) SAME TRAINING AND TESTING RAYLEIGH CHANNEL

Fig. 9 shows the impact of the IQ imbalance standard deviation on the testing accuracy for classifying 20 devices when



(a) Training APG = 20dB (b) Training APG = 10dB (c) Training APG = 0dB (d) Training APG = -10dB (e) Training APG = -20dB

FIGURE 7. Impact of APG on accuracy. MDS is fixed at 0 Hz. Number of devices = 20 with the low impairments set.



(a) Training APG = 20 dB (b) Training APG = 10 dB (c) Training APG = 0 dB (d) Training APG = -10 dB (e) Training APG = -20 dB

FIGURE 8. Impact of APG on RDTG. MDS is fixed at 0 Hz. Number of devices = 20 with the low impairments set.

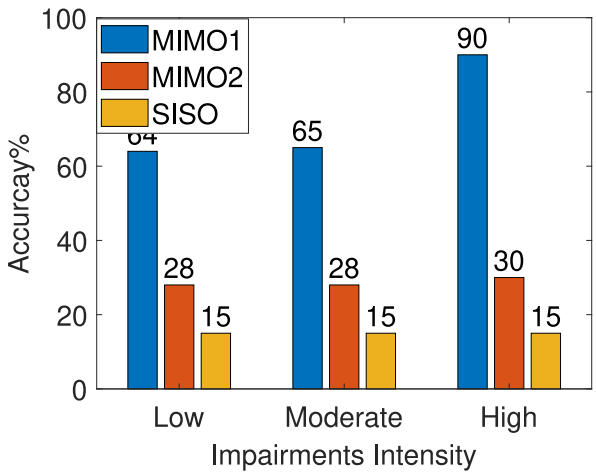
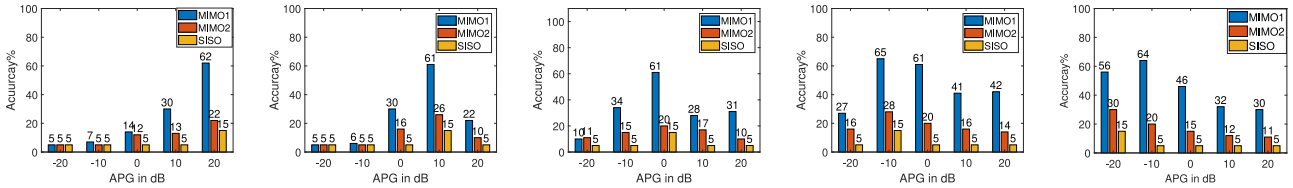


FIGURE 9. Impact of IQ Imbalance Standard Deviation on the testing accuracy when the same Rayleigh channel (APG = −20 dB) is used for training and testing. 20 devices are used for classification.

the CNN is trained and tested at APG = −20 dB. The figure shows that for the **MIMO 1** approach, increasing the IQ imbalance standard deviation increases the testing accuracy, while the testing accuracy does not improve with increasing the IQ imbalance standard deviation for **MIMO 2** and the conventional SISO approach. This observation shows that for the conventional SISO approach and for the partial blind estimation approach **MIMO 2**, multiplying the IQ imbalance standard deviation by a factor of 10 to simulate a more distinguishable set of devices is insufficient to restore the reduction in accuracy occurred when the number of devices is doubled. For instance, with low and moderate IQ imbalance standard deviations, **MIMO 2** and SISO achieve the same testing accuracy of 28% and 15%, while **MIMO 1** achieves 65%. This observation also reflects the scalability problem of RF fingerprinting over fading channels.

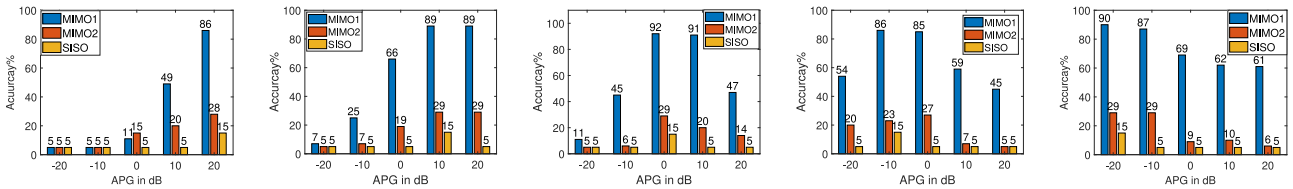
2) DIFFERENT TRAINING AND TESTING RAYLEIGH CHANNELS

Fig. 10 and Fig. 11 show the impact of channel variation on the testing accuracy when classifying 20 devices impaired with the moderate and the high impairments sets. The figures show that the high IQ imbalance standard deviation leads to higher classification accuracy values when the channel used for training is different from the channel used for testing. For instance, the **MIMO 1** approach achieves testing accuracy of up to 65% when the devices are impaired with the moderate impairments set and up to 92% when the devices are impaired with the high impairments set. The figures also show that the **MIMO 1** approach achieves higher classification accuracy compared to **MIMO 2** and the conventional SISO approach. For instance, in Fig. 11(e), which depicts the classification accuracy for the devices impaired with the high impairments set when the training APG = −20 dB and the testing APG = 20 dB, **MIMO 1** approach achieves up to 60% improvement in the testing accuracy over **MIMO 2** and the conventional SISO approach. We also observe that for the conventional SISO approach, which does not perform channel estimation before classification, the testing accuracy is 5%. These two observations show that full blind channel estimation enabled in **MIMO 1** outperforms the partial channel estimation in **MIMO 2**, and the conventional SISO where no estimation is performed. This improvement indicates the effect of the fading channel on the classification accuracy, and that removing the channel effect from the received signals is necessary for reliable RF fingerprinting. Fig. 12 and Fig. 13 show the impact of channel variation on the RDTG value when classifying 20 devices impaired with the moderate and the high impairments sets. The figures show that **MIMO 1** approach is more resilient to channel variations, i.e., small RDTG values, compared to **MIMO 2** and the conventional SISO



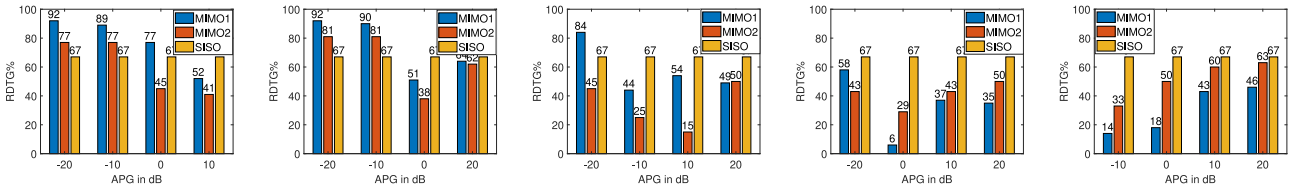
(a) Training APG = 20dB (b) Training APG = 10dB (c) Training APG = 0dB (d) Training APG = -10dB (e) Training APG = -20dB

FIGURE 10. Impact of APG on accuracy. MDS is fixed at 0 Hz. Number of devices = 20 with the moderate impairments set.



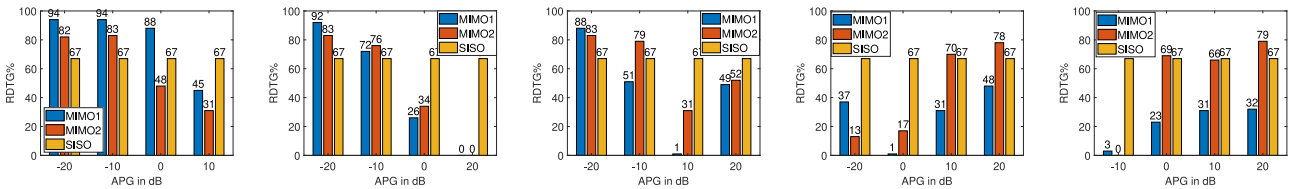
(a) Training APG = 20dB (b) Training APG = 10dB (c) Training APG = 0dB (d) Training APG = -10dB (e) Training APG = -20dB

FIGURE 11. Impact of APG on accuracy. MDS is fixed at 0 Hz. Number of devices = 20 with the high impairments set.



(a) Training APG = 20 dB (b) Training APG = 10 dB (c) Training APG = 0 dB (d) Training APG = -10 dB (e) Training APG = -20 dB

FIGURE 12. Impact of APG on RDTG. MDS is fixed at 0 Hz. Number of devices = 20 with the moderate impairments set.



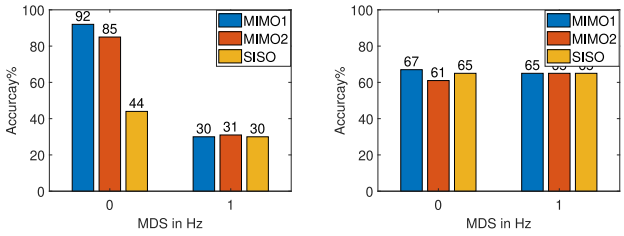
(a) Training APG = 20 dB (b) Training APG = 10 dB (c) Training APG = 0 dB (d) Training APG = -10 dB (e) Training APG = -20 dB

FIGURE 13. Impact of APG on RDTG. MDS is fixed at 0 Hz. Number of devices = 20 with the high impairments set.

approach, especially when the channel used for training has a smaller APG value compared to the channel used for testing. Note that when the training APG is high compared to the testing APG, the MIMO-based approaches have higher RDTG values and low channel resiliency compared to SISO. However, the MIMO-based approaches still outperform the conventional SISO approach in terms of the classification accuracy. The figures also show that the MIMO-based approaches are more resilient to channel variations when the devices are distinguishable, i.e., high impairments intensity leads to small RDTG values. For instance, in Fig. 12(e) and Fig. 13(e) which capture the impact of APG on the RDTG values when the training APG $= -20$ dB for 20 devices impaired with the moderate and the high impairments sets, respectively, we observe that for **MIMO 1**, when the testing APG $= 20$ dB, the RDTG value is 46% with the moderate impairments compared to 33% with the high impairments.

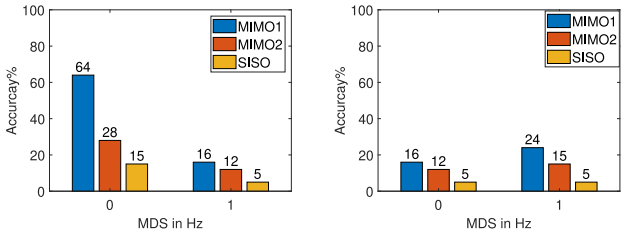
H. IMPACT OF THE TRANSMITTER-RECEIVER RELATIVE SPEED

In a dynamic scenario where the transmitter and/or the receiver are in motion, the relative speed between the transmitter and the receiver results in channel variations that are expected to impact the RF fingerprinting accuracy. In this section, we use the maximum Doppler shift (MDS) as a measure of the relative speed between the transmitter and the receiver in the channel. For instance, for crowded indoor environments, the MDS values could exceed 30 Hz at 3.6 GHz [39]. To study the impact of MDS on the classification accuracy, we slightly change the MDS value of the channel for the proposed MIMO-enabled approaches and the



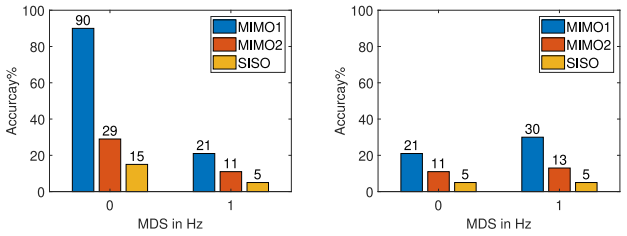
(a) Training MDS = 0 Hz (b) Training MDS = 1 Hz

FIGURE 14. Impact of MDS on accuracy. APG is fixed at −20 dB. 10 devices with low impairments set.



(a) Training MDS = 0 Hz (b) Training MDS = 1 Hz

FIGURE 15. Impact of MDS on accuracy. APG is fixed at −20 dB. 20 devices with low impairments set.



(a) Training MDS = 0 Hz (b) Training MDS = 1 Hz

FIGURE 16. Impact of MDS on accuracy. APG is fixed at −20 dB. 20 devices with high impairments set.

conventional SISO approach. We also studied the impact of the number of devices and the impairment intensity in the dynamic scenario.

1) IMPACT OF MDS: TESTING ACCURACY

Fig. 14 captures the MDS impact on the classification accuracy when classifying 10 devices impaired with the low impairment set when the training MDS = 0 Hz (Fig. 14(a)) and when the training MDS = 1 Hz (Fig. 14(b)). In Fig. 14(a), we first observe that when the training and testing MDS = 0 Hz (no channel variations), the MIMO-based approaches achieve higher classification accuracy (92%, 85%) compared to the conventional SISO approach (44%). However, a slight change of 1 Hz in the testing MDS value results in a severe degradation in the classification accuracy for all approaches, and the MIMO approaches are as accurate as the SISO approach with a 30% classification accuracy. This observation shows that the blind estimation methods fail to mitigate the channel effect caused by the relative speed between the transmitter and the receiver. In Fig. 14(b), which depicts the classification accuracy obtained when the training MDS = 1 Hz, we observe that when the testing MDS = 1 Hz (no channel variations), the MIMO-based approaches are as good as the conventional SISO approach with a testing accuracy of 65%. This observation is commensurate with the results in Fig. 14(a) and, indeed, blind estimation is not mitigating the channel effect and the MIMO approaches are as accurate as the conventional approach even when the training and testing MDS are the same. We also observe that when the training MDS = 1 Hz and the testing MDS = 0 Hz (a static channel), all approaches achieve the same accuracy of 65%.

2) IMPACT OF NUMBER OF DEVICES

The effect of the relative speed between the transmitter and the receiver appears when the number of devices increases as shown in Fig. 15. In Fig. 15(a), we observe that for 20 devices, when the training MDS = 0 Hz and the testing MDS = 1 Hz, the testing accuracy decreases to 16% for **MIMO 1** compared to 30% for **MIMO 1** when only 10 devices are classified. We also observe that the SISO approach is randomly classifying the devices. In Fig. 15(b), which shows the testing accuracy obtained when the training MDS = 1 Hz, we observe that when the testing MDS = 1 Hz (no channel variations) the testing accuracy for **MIMO 1** and **MIMO 2** is 24% and 15%, compared to 5% for the conventional SISO approach. The figure also shows that when the testing MDS = 0 Hz, the testing accuracy is 16%. These observations indicate the severe degradation in the testing accuracy resulting from the relative speed between the transmitter and the receiver when the number of devices increases.

3) IMPACT OF IMPAIRMENTS INTENSITY

In this section, we evaluate the MIMO-enabled approaches for classifying 20 devices impaired with the high impairments set, while considering the dynamic scenario. The results show that increasing the IQ gain imbalance standard deviation does not improve the classification accuracy when the transmitter and/or the receiver are in motion. For instance, from Fig. 16, which captures the testing accuracy obtained when classifying 20 devices impaired with the high impairments set, we observe that when the training and testing MDS values are 0 Hz, **MIMO 1** approach achieves a testing accuracy of 90%. However, when the training and testing MDS values are 1 Hz, **MIMO 1** approach achieves a testing accuracy of 30%. This observation indicates that the increase in the impairment intensity does not improve the classification accuracy when the MDS value is greater than zero, thereby confirming that the reduction in accuracy is caused by the relative speed between the transmitter and the receiver and that the blind estimation cannot mitigate the dynamic channel effect on the classification accuracy.

V. CONCLUSION

Although RF/device fingerprinting provides a lightweight technique proven resilient to spoofing that allows to identify illegitimate devices, its performance severely degrades by wireless channel condition variations, when the training channel differs from the testing channel. In this paper, we proposed a deep learning-based MIMO-enabled RF/device classification approach, and showed that the MIMO hardware capabilities can indeed mitigate the wireless channel effect and improve the RF fingerprinting accuracy for AWGN and flat fading channels. We showed that, for an AWGN channel, averaging multiple received signals at the receiving end of a SIMO system of L receiving antennas increases the SNR by a factor of L, leading to improved classification accuracy. For channels with low SNR values, we conclude that the SNR gain provided by the SIMO-based approach compensates for the SNR reduction. We also showed that for flat fading channels, MIMO-based approaches improve the classification accuracy by up to 69% compared to the conventional/SISO approach, and that the improvement the MIMO approaches achieve over the conventional approach is more significant and stable when the models are tested on less severe fading channels compared to the channel used for training. We showed that fading impacts the scalability of RF fingerprinting, and that doubling the number of devices reduces the testing accuracy to random guessing for the conventional SISO approach, while the MIMO-based approaches can classify 20 devices with an accuracy of up to 64% in the presence of fading. We conclude that the CNN models tend to learn and extract features from both channel and device impairments. The learned channel features are undesirable and result in a degradation in the classification accuracy. We also conclude that despite the significant improvement the MIMO approaches achieve in classification accuracy, the approaches are more suitable for slow flat fading channels, and mitigating the effect of the relative speed between the transmitter and the receiver is still an open challenge that requires further investigation.

REFERENCES

[1] S. Singh and N. Singh, "Internet of Things (IoT): Security challenges, business opportunities reference architecture for E-commerce," in *Proc. Int. Conf. Green Comput. Internet Things (ICGCIoT)*, Oct. 2015, pp. 1577–1581.

[2] A. S. Yeole and D. R. Kalbande, "Use of Internet of Things (IoT) in healthcare: A survey," in *Proc. ACM Symp. Women Res.*, New York, NY, USA, Mar. 2016, pp. 71–76. [Online]. Available: https://doi.org/10.1145/2909067.2909079

[3] B. Hamdaoui, M. Alkalbani, A. Rayes, and N. Zorba, "IoTShare: A blockchain-enabled IoT resource sharing on-demand protocol for smart city situation-awareness applications," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10548–10561, Oct. 2020.

[4] A. Rachedi, M. H. Rehmani, S. Cherkaoui, and J. J. P. C. Rodrigues, "IEEE access special section editorial: The plethora of research in Internet of Things (IoT)," *IEEE Access*, vol. 4, pp. 9575–9579, 2016.

[5] T. Qiu, N. Chen, K. Li, M. Atiquzzaman, and W. Zhao, "How can heterogeneous Internet of Things build our future: A survey," *IEEE Commun. Surveys Tut.*, vol. 20, no. 3, pp. 2011–2027, 2018.

[6] Z.-K. Zhang, M. C. Y. Cho, C.-W. Wang, C.-W. Hsu, C.-K. Chen, and S. Shieh, "IoT security: Ongoing challenges and research opportunities," in *Proc. IEEE 7th Int. Conf. Service-Orient. Comput. Appl.*, Nov. 2014, pp. 230–234.

[7] J. Bao, B. Hamdaoui, and W.-K. Wong, "IoT device type identification using hybrid deep learning approach for increased IoT security," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, 2020, pp. 565–570.

[8] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki, "Network intrusion detection for IoT security based on learning techniques," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2671–2701, 3rd Quart., 2019.

[9] T. Jian et al., "Deep learning for RF fingerprinting: A massive experimental study," *IEEE Internet Things Mag.*, vol. 3, no. 1, pp. 50–57, Mar. 2020.

[10] A. Elmaghbub and B. Hamdaoui, "Leveraging hardware-impaired out-of-band information through deep neural networks for robust wireless device classification," 2020, *arXiv:2004.11126*.

[11] L. Smaini, *RF Analog Impairments Modeling for Communication Systems Simulation: Application to OFDM-Based Transceivers*. New York, NY, USA: Wiley, 2012. [Online]. Available: http://ebookcentral.proquest.com/lib/osu/detail.action?docID=990619

[12] 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 IEEE Conf. Comput. Commun.*, Paris, France, Apr. 2019, pp. 370–378. [Online]. Available: https://ieeexplore.ieee.org/document/8737463/

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

[14] N. T. Nguyen, G. Zheng, Z. Han, and R. Zheng, "Device fingerprinting to enhance wireless security using nonparametric Bayesian method," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 1404–1412.

[15] A. Elmaghbub, B. Hamdaoui, and A. Natarajan, "WideScan: Exploiting out-of-band distortion for device classification using deep learning," in *Proc. IEEE Global Commun. Conf. GLOBECOM*, 2020, pp. 1–6.

[16] A. Al-Shawabka et al., "Exposing the fingerprint: Dissecting the impact of the wireless channel on radio fingerprinting," in *Proc. IEEE INFOCOM IEEE Conf. Comput. Commun.*, Jul. 2020, pp. 646–655.

[17] 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. 142893–142909, 2021.

[18] B. Hamdaoui, A. Elmaghbub, and S. Mejri, "Deep neural network feature designs for RF data-driven wireless device classification," *IEEE Netw.*, vol. 35, no. 3, pp. 191–197, May/Jun. 2021.

[19] N. Soltanieh, Y. Norouzi, Y. Yang, and N. C. Karmakar, "A review of radio frequency fingerprinting techniques," *IEEE J. Radio Freq. Identification*, vol. 4, no. 3, pp. 222–233, Sep. 2020.

[20] F. Restuccia et al., "DeepRadioID: Real-time channel-resilient optimization of deep learning-based radio fingerprinting algorithms," Apr. 2019, *arXiv:1904.07623*.

[21] 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.

[22] A. Ivanov, K. Tonchev, V. Poulkov, H. Al-Shatri, and A. Klein, "Hybrid noise-resilient deep learning architecture for modulation classification in cognitive radio networks," in *Future Access Enablers Ubiquitous Intell. Infrastructures* (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering), V. Poulkov, Ed. Cham, Switzerland: Springer, 2019, pp. 214–227.

[23] Y. Peng, P. Liu, Y. Wang, G. Gui, B. Adebisi, and H. Gacanin, "Radio frequency fingerprint identification based on slice integration cooperation and heat constellation trace figure," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 543–547, Mar. 2022.

[24] A. Elmaghbub and B. Hamdaoui, "Comprehensive RF dataset collection and release: A deep learning-based device fingerprinting use case," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2021, pp. 1–7.

[25] F. Meneghello, M. Rossi, and F. Restuccia, "DeepCSI: Rethinking Wi-Fi radio fingerprinting through MU-MIMO CSI feedback deep learning," Apr. 2022. [Online]. Available: https://arxiv.org/abs/2204.07614v1.

[26] N. Basha, B. Hamdaoui, and K. Sivanesan, "Leveraging MIMO transmit diversity for channel-agnostic device identification," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2022, pp. 2254–2259.
[27] N. Basha and B. Hamdaoui, "Leveraging multiple transmissions and receptions for channel-agnostic deep learning-based network device classification," Sep. 2021, *arXiv:2109.03799*.
[28] B. Hamdaoui and K. G. Shin, "Characterization and analysis of multi-hop wireless MIMO network throughput," in *Proc. 8th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2007, pp. 120–129.
[29] J. H. Winters, "The diversity gain of transmit diversity in wireless systems with rayleigh fading," *IEEE Trans. Veh. Technol.*, vol. 47, no. 1, pp. 119–123, Feb. 1998.
[30] B. Clerckx and C. Oestges, *MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems*. San Diego, CA, USA: Elsevier Sci. Technol., 2013. [Online]. Available: http://ebookcentral.proquest.com/lib/osu/detail.action?docID=1117221
[31] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999. [Online]. Available: http://ieeexplore.ieee.org/document/771146/
[32] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
[33] N. Ammar and Z. Ding, "On blind channel identifiability under space-time coded transmission," in *Proc. Conf. Rec. 36th Asilomar Conf. Signals Syst. Comput.*, vol. 1, Nov. 2002, pp. 664–668.
[34] N. Ammar and Z. Ding, "Blind channel identifiability for generic linear space-time block codes," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 202–217, Jan. 2007. [Online]. Available: http://ieeexplore.ieee.org/document/4034106/
[35] S. Shahbazpanahi, A. Gershman, and J. Manton, "Closed-form blind MIMO channel estimation for orthogonal space-time block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4506–4517, Dec. 2005.
[36] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York, NY, USA: Wiley, 2001. [Online]. Available: http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10272411
[37] "RF Analog Impairments Modeling for Communication Systems Simulation: Application to OFDM-Based Transceivers |wiley." Accessed: Oct. 26, 2022. [Online]. Available: https://www.wiley.com/en-us/RF+Analog+Impairments+Modeling+for+Communication+Systems+Simulati%3A+Application+to+OFDM+based+Transceivers-p-9781119999072
[38] A. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. 29, no. 11, pp. 1715–1720, Nov. 1981.
[39] B. Hanssens, E. Tanghe, L. Martens, C. Oestges, and W. Joseph, "Measurement-based analysis of delay-doppler characteristics in an indoor environment," *IEEE Trans. Antennas Propag.*, vol. 64, no. 1, pp. 370–374, Jan. 2016.

NORA BASHA received the M.S. degree in ECE from Oregon State University, USA, in 2021, where she is currently pursuing the Ph.D. degree. Her research interests are wireless communication networks, data-enabled intelligent network access, and IoT and autonomous systems security.

BECHIR HAMDAOUI (Senior Member, IEEE) received the M.S., C.S., and Ph.D. degrees in ECE from the University of Wisconsin-Madison in 2002, 2004, and 2005, respectively. He is a Professor of Electrical Engineering and Computer Science with Oregon State University. His research interests are in the general areas of intelligent networked systems, wireless and network security, and data communication networks. He won several awards, including the 2009 NSF CAREER Award, the 2016 EECS Outstanding Research Award, the ICC 2017 Best Paper Award, and the ISSIP 2020 Distinguished Recognition Award. He served as a Distinguished Lecturer for the IEEE Communication Society for 2016 and 2017, and currently serves as the Chair for the IEEE Communications Society's Wireless Technical Committee.

KATHIRAVETPILLAI (SIVA) SIVANESAN received the Ph.D. degree in wireless communications in 2004. He is a Senior Research Scientist with Intel Labs. He is driving connected vehicles and edge computing research for transportation and smart city applications. He has authored more than 30 publications. He holds more than 70 patents and has several patents pending.

MOHSEN GUIZANI received the B.S., M.S., and Ph.D. degrees in ECE from Syracuse University, USA, in 1985, 1987, and 1990, respectively. He is a Professor and an Associate Provost with the Mohamed Bin Zayed University of Artificial Intelligence, Abu Dhabi, UAE. His research interests include applied machine learning and artificial intelligence, Internet of Things, smart cities, and cybersecurity. He won several awards, including the Best ComSoc Journal Paper Award in 2021. He is currently the IEEE ComSoc Distinguished Lecturer.