

Machine Learning for the Detection and Identification of Internet of Things (IoT) Devices: A Survey

Yongxin Liu, *Graduate Student Member, IEEE*, Jian Wang, Jianqiang Li, Shuteng Niu, and Houbing Song, *Senior Member, IEEE*

Abstract—The Internet of Things (IoT) is becoming an indispensable part of everyday life, enabling a variety of emerging services and applications. However, the presence of rogue IoT devices has exposed the IoT to untold risks with severe consequences. The first step in securing the IoT is detecting rogue IoT devices and identifying legitimate ones. Conventional approaches use cryptographic mechanisms to authenticate and verify legitimate devices' identities. However, cryptographic protocols are not available in many systems. Meanwhile, these methods are less effective when legitimate devices can be exploited or encryption keys are disclosed. Therefore, non-cryptographic IoT device identification and rogue device detection become efficient solutions to secure existing systems and will provide additional protection to systems with cryptographic protocols. Non-cryptographic approaches require more effort and are not yet adequately investigated. In this paper, we provide a comprehensive survey on machine learning technologies for the identification of IoT devices along with the detection of compromised or falsified ones from the viewpoint of passive surveillance agents or network operators. We classify the IoT device identification and detection into four categories: device-specific pattern recognition, Deep Learning enabled device identification, unsupervised device identification, and abnormal device detection. Meanwhile, we discuss various ML-related enabling technologies for this purpose. These enabling technologies include learning algorithms, feature engineering on network traffic traces and wireless signals, incremental learning, and abnormality detection.

Index Terms—Internet of Things, Security, Physical-layer Security, Malicious Transmitter Identification, Radiometric signature, Non-cryptographic identification, Physical-layer identification.

I. INTRODUCTION

AS a rapidly evolving field, the Internet of Things (IoT), involves the interconnection and interaction of smart objects, i.e., IoT devices with embedded sensors, onboard data processing capabilities, and means of communication, to provide automated services that would otherwise not be possible [1]. Trillions of network-connected IoT devices are expected to emerge in the global network around 2020 [2]. The

Yongxin Liu was with the Security and Optimization for Networked Globe Laboratory (SONG Lab), Embry-Riddle Aeronautical University, Daytona Beach, FL 32114 USA, and is with the Department of Computer Science, Auburn University at Montgomery, Montgomery, AL 36117 USA.

Jian Wang, Shuteng Niu and Houbing Song are with the Security and Optimization for Networked Globe Laboratory (SONG Lab), Embry-Riddle Aeronautical University, Daytona Beach, FL 32114 USA

Jianqiang Li is with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, Guangdong 518060 China

Corresponding authors: Jianqiang Li, Houbing Song

Manuscript received October 18, 2020; revised XXX.

IoT is becoming an indispensable part of smart cities, enabling a variety of emerging services and applications in cities and communities [3], including in health [4], transportation [5]–[8], energy/utilities, and other areas. Furthermore, big data analytics enables the move from the IoT to real-time control [9]–[11].

However, the IoT is subject to threats stemming from increased connectivity [12], [13]. For example, rogue IoT devices, *defined as devices claiming a falsified identity or compromised legitimate devices*, can expose the IoT to untold risks with severe consequences. Rogue IoT devices could conduct various attacks: forging the identity of trusted entities to access sensitive resources, hijacking legitimate devices to participate in distributed denial of service (DDoS) attacks [13], and etc. The problem of rogue devices becomes even more hazardous in wirelessly connected IoT, as the network traffic is easier to be intercepted and falsified. Hence, from the perspective of network operators, the first step in securing the IoT against rogue devices is identifying known (or unknown) devices and detecting compromised ones. This survey defines the term *Device Detection and Identification to contain two perspectives: a) Identity verification of known devices. b) Detection of falsified or compromised devices*.

Conventional cryptographic mechanisms use message authentication code, digital signatures, challenge-response sessions, and etc. to authenticate legitimate peers or verify the identities of message senders [14]. These methods make it mathematically impossible for the malicious to forge the legitimate identities. Even though cryptographic mechanisms are effective as long as secret keys are securely protected, security requirements may not be fully satisfied in pervasively distributed IoT. Reports have shown that it is possible to use reverse engineering to access encryption keys or conduct further exploitations [15]–[19]. Moreover, it is impossible to install cryptographic protocols into the huge amount of insecure systems or devices in a short time. Some of those insecure systems have already become part of critical infrastructures [20]–[25]. Finally, cryptographic approaches become less effective in dealing with hijacked devices. Therefore, as a supplementary to existing cryptography mechanisms, non-cryptographic Device Identification with Rogue Device Detection are needed to secure the IoT ecosystem especially from the perspective of network operators and cybersecurity surveillance agents.

Non-cryptographic device identification and rogue device

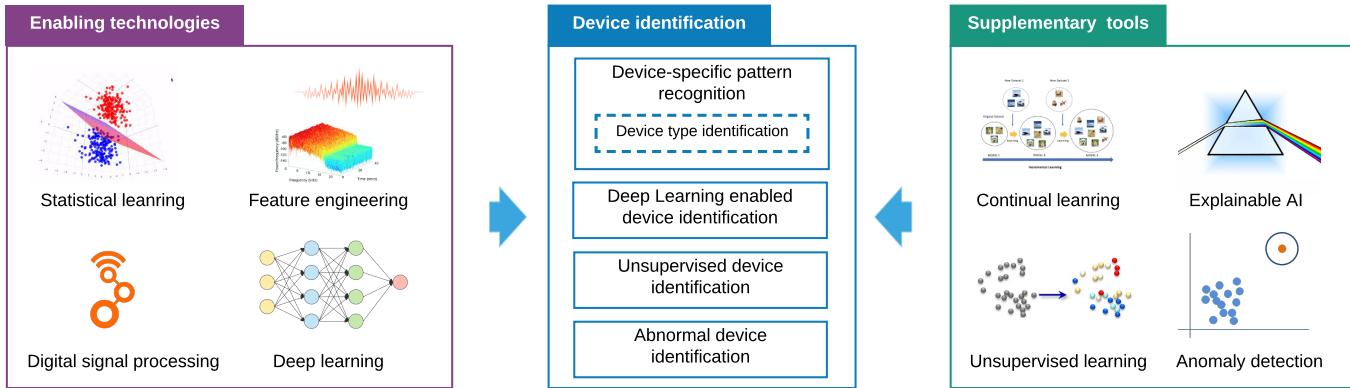


Fig. 1. Overview of ML for the Detection and Identification of Rogue IoT Devices

detection have emerged as essential requirements in safety-critical IoT [26]–[28] or physical layer authentication [29]. Compared with cryptographic approaches, non-cryptographic approaches aim to identify known devices and detect rogue devices by exploiting device-specific signal patterns or behavior characteristics [30]. More importantly, non-cryptographic approaches do not require modifications to the existing systems that can not be upgraded easily, e.g., ADS-B (Automatic Dependent Surveillance Broadcasting [31]), AIS (Automatic Identification System [32]) and etc.

Non-cryptographic device identification and detection are still challenging. Firstly, the flexible deployment scenarios and diverse specifications of devices make it challenging to provide a general solution to derive distinctive features from signals or network traffic. Moreover, even though machine learning (ML) and Deep Learning (DL) have the potential to automatically discover distinctive latent features for accurate device identification, state-of-art algorithms require intensive modifications to be utilized in IoT [33]. Therefore, these gaps motivate us to conduct a comprehensive survey as a summary of existing works and anticipate the future directions from the perspective of machine learning.

The scope of this paper and related surveys are compared in Table I. In general, existing surveys focus on presenting broad overviews of threats and countermeasures in IoT. In this paper, we focus on a more specific perspective by providing a comprehensive survey of machine learning for the detection and identification of devices in IoT using passively collected traffic traces and wireless signals, which are easily accessible to network operators and surveillance agents. Figure 1 presents an overview of ML for the detection and identification of IoT devices with relations between concepts in Figure 2. We divide the IoT device identification and detection into four categories: device-specific pattern recognition, Deep Learning enabled device identification, unsupervised device identification, and abnormal device detection. We identify various ML-related enabling technologies and tools for this purpose, including statistical learning, feature engineering, digital signal processing, and deep learning. These tools include incremental learning, unsupervised learning, and anomaly detection.

The remainder of this paper is structured as follows. Section

TABLE I
A COMPARISON WITH EXISTING SURVEYS

Surveys	Year	FD	DL	DT	UD	RD
[34]	2020	•	•			•
[35]	2019	•		•		•
[36]	2017	•	•	•		
[37]	2016	•		•	•	
[38]	2012	•				•
[39]	2010	•			•	•
This paper	2021	•	•	•	•	•

FD: Feature-based specific device identification; DL: Deep Learning enabled specific device identification; DT: Device type identification; UD: Unsupervised device identification; RD: Rogue device detection.

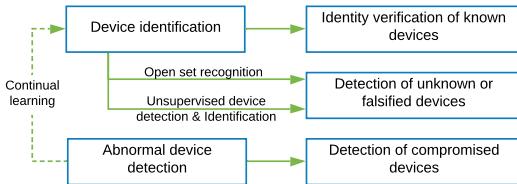


Fig. 2. Key concepts in this survey.

II presents a general threat model and attack chain of rogue devices in IoT. In Section III, we review device type identification (Section III-A) and statistical learning on device-specific feature identification (Section III-B), including conventional radiometric signature and statistical learning. In Section III-C we review state-of-the-art Deep Learning (DL) based methods for device identification with a focus on emerging issues such as incremental learning, abnormality detection, hyperparameter, and architecture search. A novel emerging approach, unsupervised device detection, is reviewed in Section III-D. In Section IV, we present methodologies to detect compromised wireless devices using anomaly detection algorithms, which is complementary to device-specific identification. Section V pinpoints the challenges and future research directions with discussions on enabling technologies. Section VI concludes this paper.

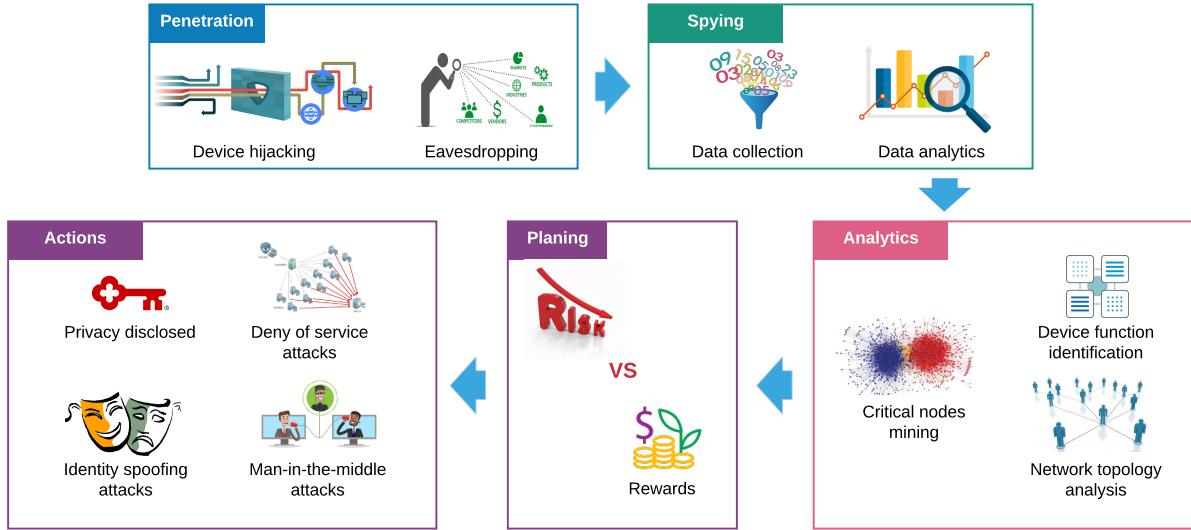


Fig. 3. Attack chain in the IoT.

II. THREAT MODE OF ROGUE DEVICES IN IoT

This section briefly reviews the threat modes of rogue devices along with countermeasures in IoT. We analyze the attack chain and identify the requirements of IoT device detection and identification: verifying legitimate devices' identity, detecting unknown or falsified devices, and detecting compromised (hijacked) devices with abnormal behaviors.

The cyberinfrastructure of IoT allows sharing information and collaborating among devices with different capacities and vulnerabilities. On the one hand, this scheme cultivates a large open system with low entry restrictions. On the other hand, adversaries can conduct rogue activities with great convenience [40]. Generally, the attack modes of adversaries in IoT are in two folds: passive attack and proactive attacks. In a passive attack, adversaries do not cause damage or performance degradation for a long time. Still, they passively analyze devices' communication and activity patterns, providing plans for attacks in the future. If we regard passive attackers as spies secretly and peacefully gathering intelligence, the proactive attackers do whatever possible to conduct malicious activities. In practical attacks, proactive and passive attacks are combined. A typical attack chain in IoT is shown in Figure 3 with a more specific demonstration of spoofing attacks depicted in

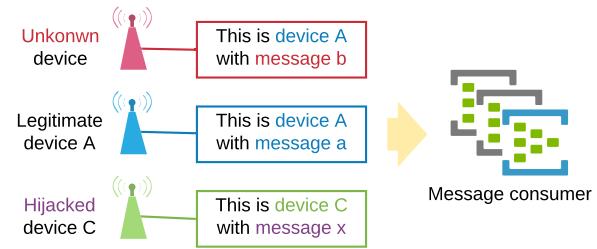


Fig. 4. Identity spoofing attacks.

Figure 4. We divided the whole attack chain into five stages as follows:

- 1) *Penetration*: In this stage, the rogue devices try to eavesdrop on communication channels or attain the control privileges of vulnerable devices for further actions. Research in [41] shows that using ARP (Address Resolution Protocol) spoofing, the malicious can easily observe ongoing traffic generated by connected IoT devices from more than 20 manufacturers. Nowadays, it is still challenging to develop software stacks with assured security [42].
- 2) *Spying*: In this stage, the malicious will observe the

TABLE II
COMPARE OF CRYPTOGRAPHIC AND NON-CRYPTOGRAPHIC COUNTERMEASURES

Methods	Principles	Advantages	Challenges
Cryptographic	Use shared secrecy to mathematically make the decryption of sensitive information and forge of identity computationally expensive.	<ul style="list-style-type: none"> • Device independent • Protects both confidentiality and can verify identity 	<ul style="list-style-type: none"> • Disclosure of secret keys. • Re-distribution of secret keys. • Needs special adaption to existing systems.
Non-cryptographic	Extract and verify device-specific features from received messages to assure that messages are from known sources.	<ul style="list-style-type: none"> • Device-specific. • Can identify Hijacked devices with abnormal behaviors. • compatible with existing IoT 	<ul style="list-style-type: none"> • Computationally expensive. • Identity disclosure.

ongoing activities by exploiting penetrated devices as its agents. As in [41], more than 50% of tested popular smart home IoT devices contain at least one vulnerable ports.

- 3) *Data analytics*: The malicious attackers analyse the behaviors and evaluate the vulnerabilities of the IoT from multiple perspectives. An example in [43] reveals that even if encryption mechanisms are employed, an attacker can still extract sensitive information, such as manufacture, device functionality, and etc.
- 4) *Planning*: In this stage, the adversaries perform strategic planning and wait for the best time to minimize their risk and maximize the rewards.
- 5) *Attack*: In this stage, prevalent attacks are in action.

From the perspective of network operators or cybersecurity surveillance agents, if we can prevent the adversaries from successfully impersonating legitimate devices in the first stage (penetration) or can identify hijacked devices in the second stage (spying). Network operators and surveillance agents can destroy the whole attack chain easily.

Various countermeasures can be applied in IoT device identification and detection. Both cryptographic and non-cryptographic methods can be applied. A brief comparison of them is presented in Table II. Cryptographic methods are widely used in computer networks and telecommunication systems. However, special modifications are needed to deploy cryptographic protocols to existing systems without cryptographic protocols such as ADS-B, AIS, and etc. Non-cryptographic methods require higher computational capacity to derive device-specific fingerprints, but they are transparently compatible with existing systems.

III. LEARNING-ENABLED DEVICE IDENTIFICATION IN IoT

This section reviews methods to recognize devices' identities and types in IoT. Most of them are based on network traffic and wireless signal pattern recognition. We first review device type identification methods, which are widely used in identifying commercial IoT devices. We then discuss and compare the corresponding signal feature based device recognition approaches. Especially, we discuss Deep Learning in device identification with emerging issues extensively. Finally, we review the unsupervised device identification and its open issues. A brief summary of open datasets for wireless device identification is provided in section V-A4.

A. Device type identification

Even though device types are not directly related to devices' identities, they still provide essential information for network management and risk control. A brief diagram of typical IoT devices considering their protocols is in Figure 5, and comparisons of their Physical Layer, Data Link Layer as well as aggregated data transmission characteristics are presented in [44], [45] and [46]. As in Figure 5, WiFi is pervasively utilized in smart home devices while smart city facilities prefer reliable cellular networks. Device type identifications are frequently performed on network, transportation, and application layers and have been implemented in Software Defined Network

(SDN) controllers or Software Routers [47]–[49]. Device types reveal functionality and activity profiles. A taxonomy of features for device type identification is presented in Figure 6.

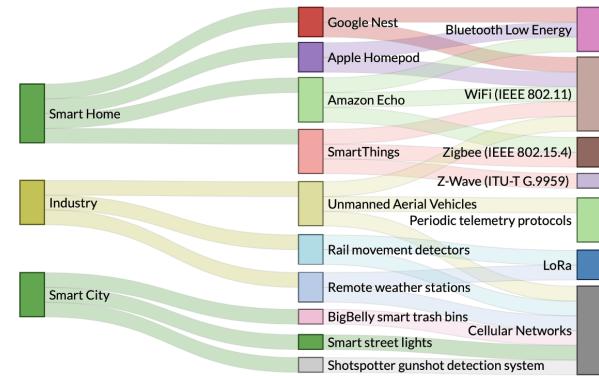


Fig. 5. Typical IoT devices and protocols.

As in Figure 6, remote service is a popular attack surface to disclose the device type or even identity. The reason is that the IoT devices communicate with remote service providers through the REST API [50]. Even though sensitive data are encrypted, some unique strings in their Web requests can still be exploited to infer device types. Authors in [51] present that using only port numbers, domain names, and cipher suite information, a Naive Bayesian classifier can reach high accuracy in classifying 28 commercial IoT devices.

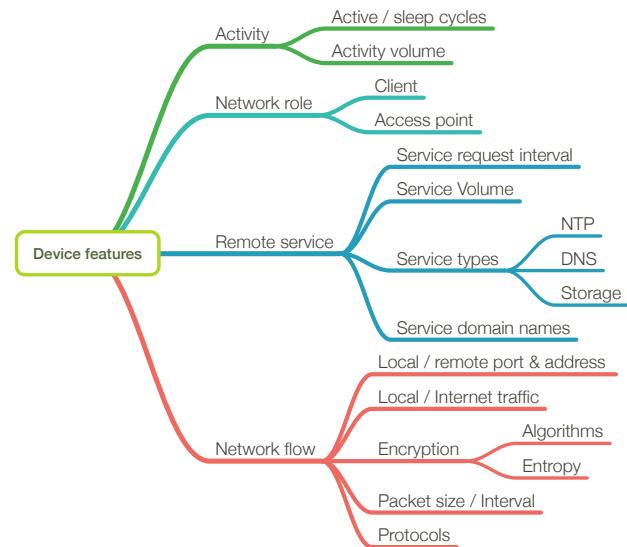


Fig. 6. Features for device type identification.

However, remote service information may not work if devices interact with anonymous service providers. For alleviation, device activity and data flow patterns can be utilized. Authors in [52] propose that their Random Forest classifier reaches a high accuracy of 95% in identifying 20 IoT devices when the features of activities, network data flows, and remote service requests are utilized simultaneously. In [53], devices' types are identified based on the periodicity of activities.

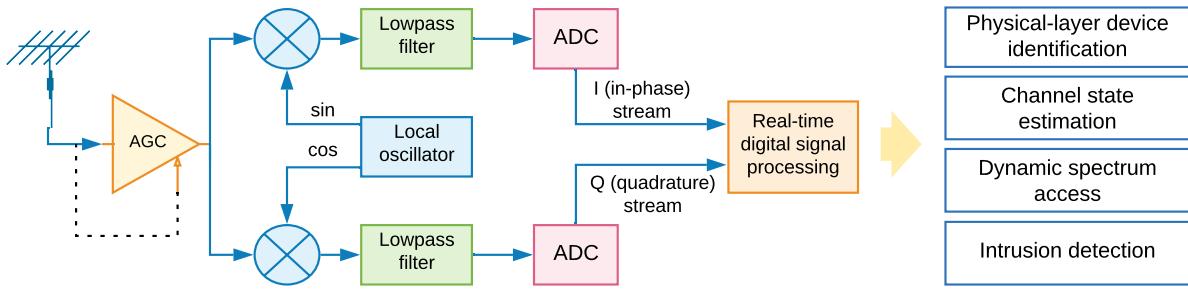


Fig. 7. General pipeline of Software-Defined wireless signal identification.

The authors first used the Discrete Fourier Transform (DFT) and discrete autocorrelation to find the dominant periods in protocol-specific activities. They then used statistical and stability metrics to model the devices' behavioral patterns. Finally, the Bayesian-optimized k-Nearest Neighbor algorithm was employed for classification. In [54] and [55], the authors extracted the protocols and network flow properties within a sliding window to generate fingerprints of devices. They used one-versus-rest classifiers to identify commercial devices. In [56], The authors first provided a Random Forest classifier using TCP/IP stream features. They incorporated confidence thresholds and averaged decisions within a sliding window to identify known or unknown device types. Similar research is presented in [56] and [57]. In [58], the authors also present that network traffic, device types, and their operation states (boot, active, and idle) can be inferred simultaneously.

An extra benefit of modeling device activity patterns is increasing the chances of identifying behavioral variations. Such benefit directly contributes to the detection of compromised devices or network attacks, which will be discussed in section IV.

Deriving devices' benign flow characteristics is nontrivial, therefore, the IETF standard Manufacturer Usage Description (MUD) profile [59] is proposed as an initial static profile to describe IoT device network behavior and support the making of security policies. A collection of MUD profiles from 30 commercial devices in [60]. The MUD profiles can be used to either verify device types or detect devices under attack or being compromised [61]. However, one issue of using the static profiles is that long observation time is needed to make decisions.

Device identifiers based on network flow and activity patterns may encounter emerging issues. First, IoT devices are becoming smart devices where new extensions can be installed, and firmware upgrades can happen periodically, thereby changing activity patterns or network flow statistics, as suggested in [62], [63] and [51]. Second, deriving relevant and distinctive features is not always easy and straightforward. To automate the processes of deriving useful features, in [57], the authors proposed a Genetic Algorithm (GA) enabled feature selector. Furthermore, a Deep Neural Network approach, which does not require complicated feature engineering, is presented in [64]. Third, device types do not necessarily correlate with their identities. Therefore, behavior-independent

specific device identification is of great significance..

B. Feature-based statistical learning for specific device identification

IoT device identification can be formalized as a classification problem. In this section, we first introduce the generic pipeline for signal reception and then focus on feature-based statistical learning approaches for specific device identification from raw signals and their open issues.

1) *Generic wireless signal reception pipeline for device identification:* Software-Defined Radios (SDR) are multipurpose front-ends to deal with various modulation and baseband encoding schemes in wireless device identification. Fundamental technologies in SDR are quadrature modulation and demodulation [65].

Generally, the wireless signals of IoT devices can be represented as: $S(t) = I(t) \cdot \cos[2\pi(f_c + f')t] + Q(t) \cdot \sin[2\pi(f_c + f')t]$, where $I(t)$ and $Q(t)$ are denoted as in-phase and quadrature components, respectively. The key idea is use $I(t)$ and $Q(t)$ to represent different modulation schemes.

A brief quadrature demodulation pipeline is given in Figure 7. We denote the reconstructed version of $I(t)$ and $Q(t)$ as $\hat{I}(t)$ and $\hat{Q}(t)$, respectively. We can derive the signals instantaneous amplitude, phase, and frequency by $\hat{m}(t) = \sqrt{\hat{I}^2(t) + \hat{Q}^2(t)}$, $\hat{\phi}(t) = \tan^{-1}(\hat{Q}(t)/\hat{I}(t))$ and $\hat{f}(t) = \partial\hat{\phi}(t)/\partial t$. Manufacturing imperfections and channel characteristics can cause $\hat{m}(t)$, $\hat{\phi}(t)$ and $\hat{f}(t)$ to deviate from its original form, providing side channels to identify wireless devices. A brief overview of features for IoT device identity verification using wireless signals in Physical Layer is given in Figure 8. The features for wireless device identification are also named Radiometric Fingerprints.

2) *Hardware imperfections:* Heterogeneous imperfections exist in IoT devices' wireless frontends. These imperfections do not necessarily degrade the communication performance but influence signal waveforms, thereby providing a side channel to identify different devices. Such features enclosed in transmitted signals are named Physical Unclonable Features (PUF) [66], [67] since regular users can not clone or forge the characteristics of these manufacturing imperfections.

a) *Error / noise patterns:* The errors between expected rational signals and actual received signals can disclose useful

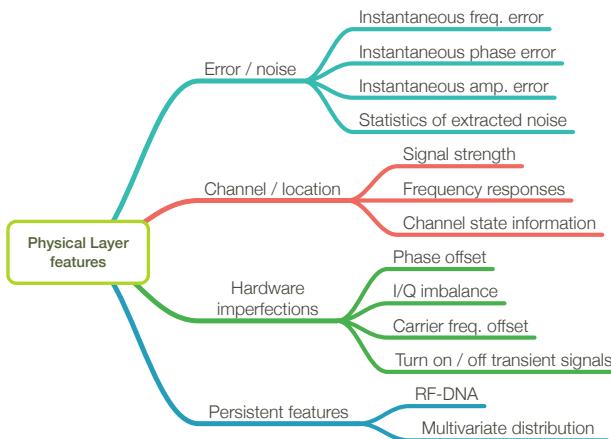


Fig. 8. Physical Layer device-specific features.

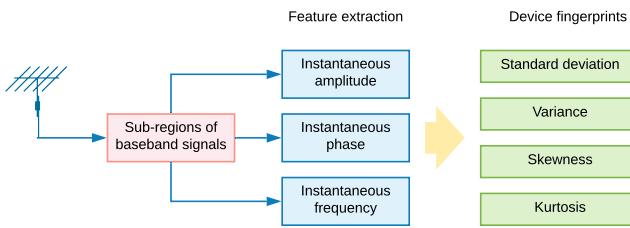


Fig. 9. A brief dataflow of RF-DNA.

device-specific information. In [68] and [69], the authors used the phase errors of Phase Lock Loop (PLL) in transmitters as a distinctive feature. Their simulations indicate promising results even with low SNR (Signal-to-Noise Ratio). In [70], the authors used the instantaneous differences between received I/Q signals and theoretically expected templates to construct error vectors. They combined error vectors' statistics and time-frequency domain statistics to synthesize the fingerprints of RF transmitters.

In [71]–[73], the authors used the differential constellation trace figure (DCTF), carrier frequency offset, phase offset, and I/Q offset to identify different Zigbee devices. They developed a low-overhead classifier, which learns how to adjust feature weights under different SNRs. The behaviors of their classifiers are similar to k-NN algorithms. Authors in [74] used the odd harmonics of center frequencies as fingerprints for RFID transmitters.

b) Persistent patterns: Persistent pattern recognition assumes that the statistics of consecutive subregions in received signals can disclose identity-related information. A typical method is named as RF-DNA (Distinctive Native Attributive [75], [76]. The basic idea is to use the statistical metrics of signals' consecutive subregions to form device fingerprints. A brief dataflow of RF-DNA is given in Figure 9. In [77]–[79], the authors captured the preamble of WPAN (Wireless Personal Area Network) signals and extract the variance, skewness, and kurtosis of signals' subregions (bins) as signatures. Research in [80] showed that RF-DNA can even be applied to

model the Fourier spectrum of devices.

From the perspective of Stochastic Process, a sequence of signal symbols can be regarded as a sample from a multivariate distribution. The parameters of such distribution represent the unique fingerprints of a wireless transmitter. With this idea, the authors in [81] used the Central Limit Theorem and proposed a repetitive stacking symbol-based algorithm. They modeled that the preamble of each packet as a sample from a specific multivariate distribution. They extracted statistics from the preambles of ZigBee devices and employed Mahalanobis Distance and nearest neighbor algorithm to identify 50 Zigbee devices.

Regional statistic vectors from complete messages can unintentionally embed protocol-dependent features and result in unreliable device identification models. Therefore, if we only extract persistent features from the protocol-agnostic part of the signals (e.g., preambles), the resulting device identification model will focus on signal features rather than communication protocols.

c) Transient patterns: Compared with persistent statistics of signals' subregions, transient patterns are more difficult to forge in terms of wireless channels [82]. An example of transient periods in wireless communication is given in Figure 10. Transient periods are commonly seen at the beginning and end of wireless packet transmission. In [83], the authors employed the nonlinear in-band distortion and spectral regrowth of the received signals (potentially caused by power amplifiers of transmitters) to distinguish the masquerading device. In [84], the authors derived the energy spectrum from transmitters' turn-on transient amplitude envelopes to classify eight different devices. The results show that frequency-domain features are more reliable than time-domain features. In [85] and [86], the time-domain statistical metrics and wavelet features of transmitters' turn-on transient signals were transformed into devices' RF fingerprints. Finally, it is notable that the authors in [87] captured the turn-on transient signal of Bluetooth devices and extracted 13 time-frequency domain features (via Hibert-Huang spectrum) to construct the devices' fingerprints. Their experiments have shown that well-designed fingerprints provide promising results even without using complicated machine learning models.

The merit of transient features is that an adversary could not forge such nonlinear features unless they can accurately forge the coupled characteristics of pair-wise wireless channels and RF front-ends between victims and surveillance agents. In other words, the transient features can be influenced by the locations of devices, as different locations can result in vari-

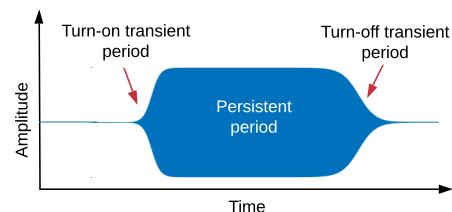


Fig. 10. Transient periods during wireless communication.

ation of RF channel characteristics, e.g., transient responses, machine learning algorithms can produce accurate but unreliable device identification results by exploiting RF channel characteristics rather than learning device-specific features.

3) *Channel state features*: From the perspective of signal propagation, the nonlinear characteristics of radio channels can cause recognizable distortions to the received signals. Those distortions can become unique profiles of transmitters. Therefore, the channel state recognition approach's basic idea is to: a) mathematically or statistically describe the nonlinear characteristics of the propagation channel within receivers and transmitters. b) Estimate whether a wireless device's signals' distortions comply with specific channel characteristics. A typical solution was provided in [94], the authors used a kernel regression method to model the nonlinear pattern of signals' propagation channels. Their basic idea is that the combination of frequency offsets and special channel characteristics may not be forged easily, and therefore, can be used as a profile for wireless devices.

Channel state features are commonly seen in Orthogonal Frequency-Division-Multiplexing OFDM modulated communication systems. In the OFDM and MIMO schemes of wireless communication, the channel state information (CSI) [95], [96] can provide rich information on the time-varying characteristics of radio channels. IEEE 802.11 receivers estimate CSI during the reception of each packet's preamble. For each packet, its CSI is expressed as a complex-valued T_n by R_m by K matrix \mathbf{H} along with a noise component $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{S})$, where T_n denotes the number of transmitter' antennas, R_m denotes the number of receivers' antennas, K denotes the number of sub-carriers and n denotes the complex-valued Gaussian random variable with mean zero and covariance matrix \mathbf{S} . Each complex-valued element in \mathbf{H} provides the instantaneous phase and amplitude response of antenna-wise channels at specific subcarriers.

Channel state information directly reveals the phase, frequency, and amplitude responses of radio channels and has been utilized to identify fixed-position wireless transmitters. Specifically, CSI is affected by propagation obstacles, signal reflections, and even baseband data patterns [96]. In [97], a CSI based device identification scheme was proposed. The authors used averaged CSI to construct an SVM based profile for each legitimate device to prevent and identify spoofing attacks. They compared CSI and RSS based approaches and demonstrated the superiority of CSI. Another merit of their solution is utilizing the two-cluster k-means algorithm to detect the presence of rogue IoT transmitters when constructing legitimate devices' profiles. Similar research was presented in [98], legitimate devices' CSI from multiple locations are collected to train a more robust device identification model. Comparably, in [99], the authors used the information from CSI to model the radiometric signatures of obstacles within the signals' propagation path. They provided an iterative differentiation approach to derive the weights and factor out the multipath components in the received signals.

Except for wireless channel characteristics, CSI can disclose RF transmitter-specific information for persistent feature-based device identification. Related researches are as follows:

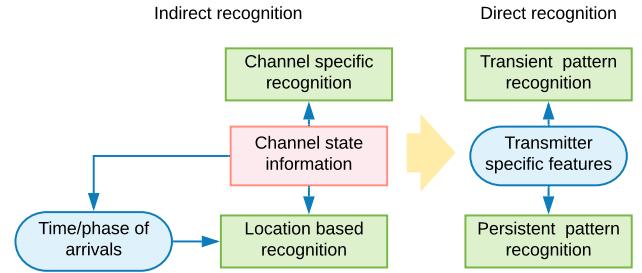


Fig. 11. A brief overview of channel state recognition and related approaches.

- *Carrier Frequency Offsets (CFOs)*: In [100], the authors propose to derive Carrier Frequency Offsets (CFOs) from CSI as devices' fingerprints. Their primitive hypothesis was that the constant CFOs can cause a linearly varying trend in instantaneous phases in received signals. Specifically, the authors first used phase measurements on specifically selected subcarriers to eliminate phase shifts at the receiver of the device identification oracle. They then used the differentiated phases from adjacent packets to eliminate the phase shifts introduced by the relative positions of transmitters. Finally, they derived the carriers' frequency offsets by the slope (relative to the time intervals of adjacent packets) of the purified instantaneous phase.
- *Phase errors*: Authors in [101] used the summation of selected subcarriers' instant phases to extract the rationale arrival phases of subcarriers. They then estimated and subtracted the rationale arrival phases and receivers' insertion phase lag to derive the phase error caused by the transmitters' internal imperfections. A drawback of their approach is they need to estimate the Time of Flight (ToF) of received packets.

A summary of device identification based on channel state features is in Figure 11. The drawbacks of channel state features are apparent. For one thing, researches show that channel state features can even be influenced by the motions of obstacles in subcarriers' propagation path [102]–[104]. On the other hand, the channel characteristics are environment-oriented. Consequently, using channel state features based device identifier in indoor or mobile environments with human activities is still challenging [105], [106].

It should be noted that a great majority of CSI enabled researches depend on limited categories of Network Interface Cards (NICs) for data collection, owing to the limitation of CSI Tools [95]. However, the authors in [107] provided a new way. They used generic SDR transceivers to extract the Long Training Sequences (LTS) in the preambles of IEEE 802.11n pilot carriers and successfully identified more than 50 Network Interface Cards. They showed that by exploiting the frequency offsets and comparing LTS frequency responses of adjacent pilot carriers, they even derived a location-agnostic device identification model.

4) *Cross domain features*: Many researchers convert signals to other domains that are more distinguishable. A straight-

TABLE III
INFLUENTIAL FACTORS FOR FEATURE-BASED SPECIFIC DEVICE IDENTIFICATION

Influential factors ¹	Persistent feature recognition	Transient feature recognition	Channel status recogniton	Cross-domain recognition	Hybrid approaches	Countermeasures	Reference
Stationary noise	Median (Exc. noise pattern)	Median	Low	Median	Low	<ul style="list-style-type: none"> • Denoise filtering. • Data argumentation 	[81], [88]
Rx imperfections	Median	Median	Median	Median	Median	<ul style="list-style-type: none"> • Adaptive filtering. • Calibrations 	[89], [90]
Co-channel devices	High	High	Low	High	High	<ul style="list-style-type: none"> • MIMO receivers. • Blind signal separation 	[91], [92]
Channel features	Median	Median	High	Low	Low	<ul style="list-style-type: none"> • Adaptive filtering 	[89]
Baseband patterns	Median (Exc. noise pattern)	Low	Median	Low	Low	<ul style="list-style-type: none"> • Message-independent features 	[93]

¹ High: solutions include hardware modifications; Median: solutions are software-based but require high capacity processors; Low: Software-based optimal solutions are available and compatible with regular processors;

forward way is to remap signals into the time-frequency domain [108]. In [109], the authors used the STFT (Short-Time Fourier Transform) with the SVM algorithm to identify four different transceivers. This research is comparable to [110], where Discrete Gabor Transform (Gaussian windowed STFT) was employed.

Other signal transformation methods can be utilized. In [111], [112], the authors utilized the wavelet transform as well as classifiers (SVM and Probabilistic Neural Network) to construct a device identifier, compared with [109], they also used the PCA algorithm to reduce the redundancy of the extracted data. In [113], the authors provided a normal frequency-based method along with PCA and SVM to distinguish devices in the GSM band. They compared their method with Hibert-Huang Transform based method in [114]. Similar work presented in [115], showed that Variation Mode Decomposition theoretically provides even better performance than the conventional EMD method for relaying scenarios. Please note that Bispectrum is also widely utilized. In [116], the energy entropy and color moments of the Bispectrum combined with Support Vector Machine (SVM) were employed to simulate the possibility of device identification. Their results indicated that higher-order statistics can theoretically improve the performance of identification under low SNR. However, other authors [117] claimed that compared to Bispectrum, the squared integral bispectra (SIB) is more robust to noise while providing the same amount of information as the Bispectrum. In [118], the authors employed singular values of the Axial Integrated Wigner bispectrum (AIWB) feature to identify spoofing signals from genuine signals in navigation satellite systems (GNSS).

5) *Hybrid methods*: A large number of device-specific features have been discovered along with different signal transform techniques. Hybrid methods aim to find the optimized combinations of features from different domains to derive robust identification models. In [119], the authors extracted the signals' energy distribution from wavelet coefficients [120] and used k-NN and SVM to identify eight devices. Their test showed that this k-NN requires higher SNR than SVM. In [121], the authors applied Intrinsic Time-Scale Decomposition (ITD) [122] to input signals. They extracted factual,

bispectrum, and energy features from all subchannels of ITD decomposition subsignals, their test on SVM shows that more features can significantly improve device identifiers' performance.

Although integrating signals' features from multiple domains can provide promising device identification results, the redundant information within the integrated features requires complicated models and considerable processing capacity. Therefore, automatic feature selection is introduced and becomes an indispensable part. Research in [77] demonstrated that properly selected features, particularly from the F-test and MLF methods, enable a significant (80%) reduction of redundancy. In [123], the authors captured the pilot tones of the OFDM signals and extract a series of features relative to the rational signal. They used an information-theoretic approach to select useful features for device identification. In [124], four types of features, scramble seed similarity, carrier frequency offset, sampling clock offset, and transient pattern, were used for the physical layer fingerprints of WiFi devices.

A comparison of device-specific feature-based approaches in Table III, hybrid approaches have superior performance under various influential factors, since the automatic feature selection methods can remove irrelevant information and provide an optimal combination of features. However, hybrid features could bring side effects, especially in statistical learning algorithms: a) The complicated combination of a large number

TABLE IV
A BRIEF COMPARE OF CLASSIFIERS IN DEPLOYABLE WIRELESS TRANSMITTER IDENTIFICATION SYSTEMS

Approach	Application overhead	Continual learning	Abnormality detection
k-NN	Depends on the size of fingerprint library.	Natively supported	Clustering or statistical models
SVM	Depends on the number of feature dimensions	Knowledge replay	One-class SVM
Random forest	Depends on the number of decision trees.	Knowledge replay	Isolation forest
Neural network	Depends on structural complexity	Section III-C2d	Section III-C2c

of features can result in a highly accurate identifier with its internal mechanism not interpretable. b) High dimension features can potentially result in complicated models that are computationally expensive to retrain for operational variations. We can make better use of hybrid features in Deep Neural Networks, which will be discussed in Section III-C.

6) *Open issues*: In general, the following issues need to be investigated in feature-based statistical learning for specific device identification:

- 1) These methods require efforts to manually extract features or high-order statistics, the quality of handcraft features dominates device identification performances. E.g., authors in [129] showed that the combination of permutation entropy [130] and K-NN even surpasses combination of bispectrum [131] and SVM in [116].
- 2) Experiments are conducted in rational environments with a limited number (less than 30) of IoT devices. Therefore, publicly available datasets containing signals from a larger number of IoT devices are needed to provide a reliable benchmark. Currently, publicly available datasets for IoT device identification from wireless signals are still limited. Some small datasets are available in [132], [133] and [134] while a larger dataset but with only ADS-B signals is available in [135].
- 3) There's no guarantee whether a specific type of feature is time-invariant. Therefore, this type of system should incorporate wireless channel estimation approaches to identify real device-specific patterns.
- 4) A brief comparison of the device-specific feature-based wireless device identification with influential factor is given in Table III, co-channel devices have the most significant impacts among all solutions. Unfortunately, there's limited research in dealing with it.
- 5) A deployable wireless device identification system should have the capacity to report unknown abnormalities and continually evolve and adapt to operational variations. A comparison of frequently employed statistical learning algorithms on incremental learning and abnormality detection is in Table IV. Among these algorithms, only k-NN provides intuitive and native supports for incremental learning and abnormality detection. However, k-NN is insufficient in handling complicated features. Though SVM or Random Forest could handle more complicated features, they lack the incremental learning and abnormality detection abilities and explainability.

C. Deep Learning enabled specific device identification

The feature-based statistical learning approaches require manual selection of useful transforms or features. In contrast, deep neural networks (DNN) can incorporate existing features or directly deal with raw inputs and derive latent distinctive features. Therefore, Deep Learning enabled device identification mechanisms are increasingly investigated. A brief comparison of device-specific feature-based statistical learning and deep learning based approaches are presented in Table VII. In this section, we discuss typical deep learning enabled wireless device identification solutions and then focus on open

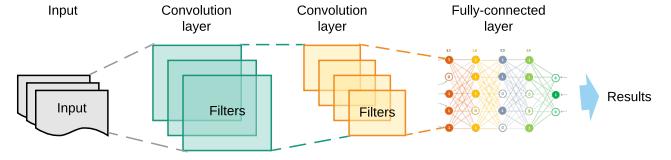


Fig. 12. Typical architecture of deep neural network classifiers

issues that impede the application of deep learning in IoT device identification.

1) *Case studies and comparisons*: A typical Deep Neural Network enabled classifier is depicted in Figure 12. Generally, it employs convolutional layers to extract latent features and uses fully connected dense layers to produce final results. Deep Neural Networks with convolutional layers are also referred as Covolutional Neural Networks (CNN).

Deep neural networks can be seamlessly integrated with existing feature engineering methods. In [127], the authors used the differential error between the reconstructed rational signals and received signals to train Deep Neural Networks to distinguish Zigbee transceivers. In [136], the authors compared the effects of short-time Fourier features and wavelet features for device identification, and their results show that wavelet features can outperform Fourier features. In [126], the authors extracted the 1-D Regions of Interest (ROIs) from 54 Zigbee devices' pREAMbles under different SNRs and then resampled the signals within ROIs into various substreams with different sample rates. Finally, the substreams were fed into a convolutional neural network for identification. Similar ideas are in [125], [137] and [138].

Compared with the conventional fully connected neural network, convolutional layers apply filters (a.k.a. kernels) with much fewer parameters to obtain distinctive information. In [88], the authors proposed a combined solution to denoise signals and identify devices simultaneously using an autoencoder and a CNN network. The authors used their encoder to automatically extract relevant features from the received signals and use the derived features to train another deep neural network for device identification. Similar methods are presented in [139]. In [128], the authors provided an optimized Deep Convolutional Neural Network approach to classify wireless devices in 2.4 GHz channels and compare the performance with SVM and Logistic Regression. Their results showed that, even by using raw I/Q digital baseband signals, CNN can achieve high accuracy and surpass the best performance of SVM and Logistic Regression. In [132], neural networks were trained on raw IQ samples using the open dataset¹ from CortexLab. Their results also showed that CNN can achieve promising results even on raw I/Q signals, but the movement of devices and the varying amplitudes can degrade CNN's performance.

An extensively discussed topic for Deep Learning based device identification is preventing the network from learning only trivial features, such as protocol identifiers, unique identifiers,

¹<https://wiki.cortexlab.fr/doku.php?id=tx-id>

etc. Generally, three types of countermeasures are applied, and their comparisons are provided as in Table V.

Compared with feature-based device identification approaches, Deep Learning methods usually require a much larger dataset to initialize the network. To know how large the training data is needed. In [140], CNN models were used to classify different devices' signals with controlled difficulty levels. The classification accuracy of a fixed CNN network with different dataset sizes was predicted using a power-law model and the Levenberg-Marquardt algorithm. Their results show that the dataset size should be at least 10,000 to 30,000 times the number of devices to be identified. However, this conclusion is only a rough estimation.

New architectures in Deep Learning are emerging and can significantly influence the performance of device identification systems. In [125], the authors used Convolutional Deep Complex-valued Neural Network (CDCN) and Recurrent Deep Complex-valued Neural Network [141] to address the device identification problem. Their networks utilized fragments of raw I/Q symbols as input, and their test was conducted on both WiFi and ADS-B datasets. Their experiments show that the Complex-valued neural networks surpass regular real-valued deep neural networks. In [142], [143], a zero-bias dense layer was proposed. The authors have shown that their solution enables deep neural networks' final decision stage to be transparent. Their zero-bias deep neural network maintains equivalent identification accuracy and outperforms regular DNN and one-class SVM in detecting unknown devices.

2) *Open issues in Deep Learning for IoT device identification:* Deep Learning is becoming a promising technology in this domain. However, as in other domains, Deep Learning encounters several challenges. Although researches in IoT device identification rarely cover the issues, we briefly discuss their current states and solutions.

a) *Hyperparameter searching:* One critical problem for using deep neural networks is hyperparameter tuning. Hyperparameters such as learning rate, mini-batch size, dropout rate, etc. are used to initialize the training process. Hyperparameters can significantly impact the performance of deep neural networks. For instance, in [151], the authors compared the performance of Deep Neural Networks, Convolutional Neural Network, and the LSTM (Long Short Term Memory) in device identification using the raw I/Q signals directly. Their results showed that CNN has the best performance, followed by DNN

and LSTM. They have pointed out that the hyper-parameters of Deep Learning, especially for network architectural parameters, significantly influence the upper bound of performance.

Obtaining optimized hyperparameters is computationally expensive. Several strategies are proposed for efficient hyperparameter searching, such as grid search, random search, prediction-based approaches, and evolutionary algorithms. Their characteristics are as follows:

- **Grid search:** Grid search divides the whole parameter space into identical intervals and performs brute-force trials to find optimal parameter combinations. However, this strategy is inefficient since useless combinations of parameters can not be pruned rapidly.
- **Random search:** In random search, sample points are distributed uniformly in the search space. This strategy increases the variation and outperforms the grid search when only a small number of parameters can impact the network performance.
- **Prediction-based:** In prediction-based approaches, the algorithms first perform random trials at the beginning to model the relation between the network performances with hyperparameters. Then the algorithms perform new trials based on parameters that are more probable to yield better results. Such trial-model-predict paradigm is conducted repeatedly [152]. A typical prediction strategy is the Bayesian optimization process [153], in which the algorithms model the target outcome space as Gaussian processes.
- **Evolution based:** In evolutionary algorithm based approaches, the heuristic searches are performed as in other nonlinear optimization problems. In [154], the authors used the Genetic Algorithm to find the optimal hyperparameters of a neural network. Compared with prediction-based approaches, evolutionary algorithms provide the best guess with bio-inspired strategies. However, there is no guarantee for the performance of evolutionary algorithms.

b) *Neural network Architecture search:* Network Architecture Search (NAS) is another challenging task in designing neural networks. Network architecture defines the flow of tensors and could significantly affect the complexity and performance of neural networks [155], [156]. At the current stage, most network architectures are specified manually or with trial-and-error.

TABLE V
COUNTERMEASURES TO PREVENT LEARNING FROM TRIVIAL FEATURES

Reference	Methodology	Description	Challenges
[125]	Fragmenting	The raw I/Q signals are split into small signal fragments or only use the preambles of packets..	Long range dependent features will be destroyed after fragmenting
[126]	Masking	One can directly mask or remove the trivial parts in raw signals.	The position and length of the masking bits or discontinuity can leak protocol information
[127], [128]	Randomization	One has to force transmitters to send random contents	One has to gain the access of large number of transmitters to train a reliable classifier.

Architecture searching algorithms are provided by several Automatic Machine Learning (AutoML) platforms. A brief comparison of their functionality and performance on different datasets is in [157]. A collection of recent literature and open-source tools are given in [158] and [159] respectively. These efforts can be classified into three categories: (i) network pruning [160], (ii) progressively growing [161], and (iii) heuristic network architecture search [162]. Their features are as follows:

- **Network pruning:** Network pruning algorithms use group sparsity regularizers [163] to remove unimportant connections from a regularly trained network. Then the pruned network will be retrained to fine-tune the weights of the remaining connections [164], [165]. A key benefit of network pruning is that it can greatly compress neural networks and make them suitable to deploy in low capacity IoT devices.
- **Progressively growing:** This strategy grows a neural network architecture during training. It is effective in simple networks with only one hidden layer [166], [167]. More recent advances employ growing strategies to progressively add nodes and layers to increase the network's approximation ability [168], [169].
- **Heuristic network search:** In heuristic network search, the architecture of the Deep Neural Network (which can either be block-wise [170] or element-wise [171]) can first be represented in a high dimension space with billions of parameters. Next, heuristic searching algorithms are applied to transverse this search space to find the optimal solutions. Examples are given in [162], [172] and [173]. The authors used the Genetic Algorithm to find the possible structure of neural networks. Notably, the Genetic Algorithm fits perfectly in NAS problems since it allows using length-varying variables (genes) to encode the candidate solutions. An empirical example is the NeuroEvolution of Augmenting Topologies (NEAT) algorithm [172].
- **Reinforcement Learning:** Reinforcement learning (RL) has become a popular strategy in NAS [174]–[176]. The basic idea is to let a deep learning-enabled agent explore network architectures' representative space and

use validation accuracy or other metrics as rewards to adjust the agents' solutions. Ideally, as an RL process moves on, an agent can find an optimal searching strategy and discover a novel architecture. Intuitively, evolution algorithms use a fixed strategy to discover the optimal architecture while RL agents learn their own strategies and have better capabilities in avoiding bad solutions.

- **Differentiable space search:** Aforementioned, NAS strategies use discrete space to encode the architecture of neural networks, which is not differentiable and lacks efficiency. Therefore, differentiable spaces to represent the Neural Networks' architectures are proposed, in which efficient off-the-shelf optimization algorithms can be used. Typical solutions are given in [177], [178]. The algorithm, DART (Differentiable Architecture search), is presented. The authors used the Softmax function to represent the discrete selections in a numerically continuous domain. They then used a gradient descent algorithm to explore the search space. Similar work with an enhanced stochastic adaptive searching strategy is in [179]. Block-wise representations of the neural network and differentiable searching space together are bringing NAS to practice.

Network architecture search has become an emerging topic for deep neural network research with publicly available benchmarking tools in [180] and [181], respectively.

c) *Openset recognition:* A critical problem for learning based device identification is that classifiers only recognize pretrained devices' signals but can not deal with novel ones that are not in the training dataset. In [150], the authors formulated it as a semi-supervised learning problem. They first trained a CNN model with the last layer as a Softmax output on a collection of known devices. They then removed the Softmax function and turn the neural network into a nonlinear feature extractor. Finally, they used the DBSCAN algorithm to perform cluster analysis on the remapped features of raw I/Q signals. Their results showed that such a semi-supervised learning method has the potential of detecting a limited number of untrained devices. Comparably, in [182], the authors used an incremental learning approach to train neural networks to classify newly registered devices.

From the perspective of Artificial Intelligence, this issue

TABLE VI
METHODS FOR UNKNOWN DEVICE RECOGNITION

Methods	Description	Complexity	Memory	Pros & Cons	Reference
GAN	Use the discriminator from GAN model as an outlier detector.	High ¹	Depends on final network	• Can catch deep latent features. • Hard to design and train.	[137], [144]
Autoencoder	Train a deep Autoencoder on known signals and use its reconstruction error to judge outliers.	High ¹	Depends on final network	• Can catch deep latent features. • Easier than GAN to design and train	[145], [146]
Statistic metrics	Measure the confidence of whether a signal or its fingerprint is generated by a given category.	Low	Low	• Provide explainable results. • Accuracy depends on the fingerprinting methods.	[143], [147]–[149]
Clustering	Perform clustering analysis on known signals' fingerprints to judge whether it is in an identical cluster where the known ones are in.	Median ²	Depends on the number of existing fingerprints.	• Provide explainable results • Accuracy depends on the fingerprinting methods.	[147], [150]

¹ Needs to specify both network architecture and hyperparameters.

² Needs to specify the clustering algorithms to use.

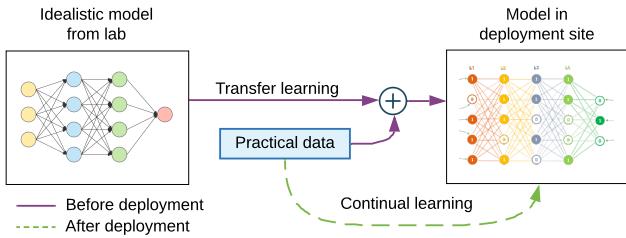


Fig. 13. Transfer learning and incremental learning.

is categorized to the Open Set Recognition [183], [184] and the Abnormality Detection problem. The taxonomy of existing approaches is given in table VI. In [137], the authors used the Generative Adversarial Network (GAN) to generate highly realistic fake signals. Then they extracted the discriminator network to distinguish whether an input is from an abnormal source. In [147], the authors provided two methods to deal with unknown devices: i) Reuse trained convolutional layers to transform signals to feature vectors, and then use Mahalanobis distance to judge the outliers. ii) Reuse pretrained convolutional layers to transform signals to feature vectors, and then perform k-means ($k = 2$) clustering to group outliers.

d) incremental learning: In practical scenarios, deep neural networks would have to evolve to adapt to operational variations continuously. Intuitively, a deep learning enabled IoT device identifier has to learn to distinguish new devices' characteristics during its life cycle. Therefore, such functionalities are defined as lifelong learning. Generally, there are two ways to achieve this goal: Transfer Learning (TL) and Incremental Learning (IL). In Transfer Learning, neural networks are pre-trained in the lab and then fine-tuned for deployment using practical data [186], [187]. In incremental learning, neural networks are trained incrementally as new data come in progressively [188]. Incremental learning does not allow neural networks to forget what they have learned in the early stages compared with transfer learning. The phenomenon in which a neural network forgets what it has previously learned after training on new data is named Catastrophic Forgetting. Therefore, transfer learning is useful when deploying new signal identification systems, and incremental learning is useful in regular maintenance, as depicted in Figure 13. The strategies to implement incremental learning for deep neural networks are as follows:

- **Knowledge replay:** An intuitive solution for incremental learning is to replay data from old tasks while training neural networks for new tasks. However, such a solution requires longer training time and larger memory consumption. Besides, one can not judge how many old samples are enough to catch sufficient variations. Therefore, some studies employ data generator networks to replay data from old tasks. For instance, in [189], Generative Adversarial Network (GAN) based scholar networks were proposed to generate old samples and mixed with the current task. In this way, the deep neural network can be trained on various data without using substantial memories to retain old training data.

- **Regularization:** Initially, regularization is employed to prevent models from overfitting by penalizing the magnitude of parameters [190]. In incremental learning, regularization is employed to prevent model parameters from changing dramatically. In this way, the knowledge (represented by weights) learned from the old tasks will be less likely to vanish when an old network is trained on new tasks. There are two types of regularization strategies: global regularization and local regularization. Global regularization penalizes the whole network's parameters from rapid change but impedes the network from learning new tasks. In local regularization strategies, such as Elastic Weight Consolidation (EWC) [191], the algorithms identify important connections and protect them from changing dramatically, in which the non-critical connections are used to learn new tasks.

- **Dynamic network expansion:** Network expansion strategies lock the weights of existing connections and supplement additional structures for new tasks. For instance, the Dynamic Expanding Network (DEN) [192] algorithm first trains an existing network on a new dataset with regularization. The algorithm compares the weights of each neuron to identify task-relevant units. Finally, critical neurons are duplicated and to allow network capacity expansion adaptively.

- **Orthogonal memory organization:** Some recent researches have shown that real biology brains organize memory representation vectors within a mutually orthogonal manner to minimize the interference and avoid catastrophic forgetting during incremental learning. Authors in [193], [194] discovered and proved that this phenomenon also exists in DNNs. Consequently, they invented a Channel Separation Incremental Learning framework based on orthogonal memory organization as in [195].

Incremental learning algorithms, as well as abnormality detection, together provide great potential for deploying neural networks in complex, uncertain scenarios.

e) Summary: A brief comparison of Deep Learning and other statistical learning methods is given in Table IV. Compared with statistical learning, Although Deep Learning is not yet an idealistic solution, its unified development pipeline, and the capability of dealing with complex features are making it easy to use. Furthermore, for practical issues such as incremental learning and abnormality detection, deep learning provides better performance than the majority of statistical learning algorithms. In one word, although deep learning is not theoretically novel, it gains its place by providing the most balanced merits.

D. Unsupervised device detection and identification

Feature-based statistical learning and deep learning are supervised learning schemes, where classifiers must learn the features of legitimate devices in advance. Unsupervised device detection and identification are required in scenarios where the identities of devices are not directly available [196]. Generally, the methods in this topic can be divided into two folds, device behavior modeling and signal propagation pattern modeling.

TABLE VII
BRIEF COMPARE OF IoT DEVICE IDENTIFICATION AND DETECTION METHODS

Device identification approaches	Technology branch	Feature requirement	Model explanability	Continuous learning	Anomaly detection	Challenges
Feature based device identification	Supervised learning	High ¹	Strong (k-NN) / median (SVM)	Easy (k-NN) / median (PCA-SVM)	Low (k-NN) Median (k-Means)	Can not discover latent feature.
Deep learning enabled device identification	Supervised learning	Low	Weak ²	Hard (EWC) ³	High (Autoencoder) / Median (clustering)	Learning from trivial features
Unsupervised device detection and identification	Unsupervised learning	High ¹	Strong	N/A	Low	Can not be applied to complex environment

¹ Requires an extra feature engineering process.

² Please refer to Explainable AI (XAI) in [185]

³ Please refer to section III-C2d

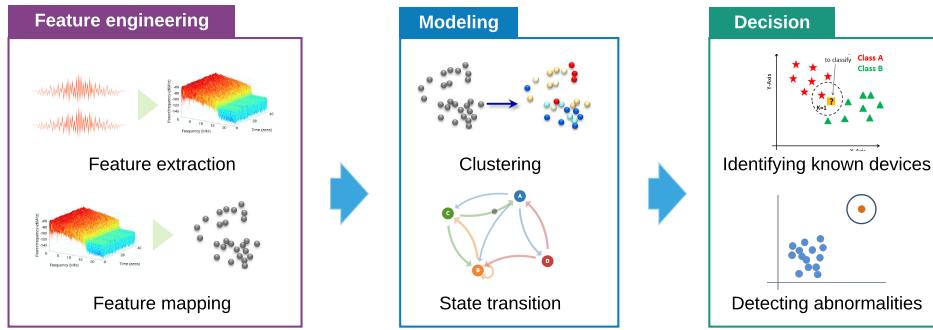


Fig. 14. Unsupervised device detection and identification

the essence of unsupervised device detection is to map devices' signals or activity profiles into latent representative spaces, where different devices are represented by separated clusters or probabilistic distributions. If behavior or signal propagation patterns are strictly correlated with specific devices, extracted behavior or signal features can be used to verify the identity of devices. Comparisons of the supervised and unsupervised learning based device identification are (also in Table VII):

- The training data does not directly indicate device specific information (device identifier, device type, and etc.).
- The number of devices may not be known in advance.

As depicted in Figure 14, the work flow of unsupervised learning enabled device detection and identification is made up of three steps: a) Feature engineering on IoT devices' signals or behavior profiles, including feature selection and mapping. b) Modeling the latent spaces, this step finds out cluster centers, probabilistic distributions, related decision boundaries, or state transition models. c) Matching input signal or behavior profiles to the most likely clusters or reporting abnormalities.

1) *Device behavior modeling:* Device behavior modeling extracts distinctive features from the input data and finds out the number of different devices using unsupervised learning algorithms. However, the physical layer does not provide much information for device behavior modeling. Therefore, the methods are more frequently employed in the upper layers with related techniques employed are unsupervised feature engineering, clustering, and Software-Defined Networking [49].

In [197] and [198], the data traffic attributes were obtained from flow-level network telemetry to recognize different IoT devices. The authors utilized Principle Component Analysis

along with an adaptive one-class clustering algorithm to find the optimal representative components and cluster centers for each device. They provided a conflict resolution mechanism to associate different types of devices to corresponding cluster centers in the representative space. A similar approach using Deep Learning is presented in [199]. The authors used TCP data traffics for each device to train an LSTM-enabled autoencoder to map inputs into a representative feature space. They then used a clustering algorithm to divide the training samples into their natural clusters. Finally, they used probabilistic modeling to associate new data with known clusters for device identification. Unfortunately, their results show that unsupervised behavior identification may not work once there are devices with an identical model.

2) *Signal propagation pattern modeling:* In the Physical Layer, signal propagation patterns provide information for device identification. On the one hand, if devices positions are unique and known in advance, we may directly use wireless localization algorithms to specify whether a received data packet is from its claimed identity. Corresponding surveys on wireless device localization are available in [204]–[206], and we provide a brief comparison of the widely employed methods in Table VIII.

On the other hand, signal propagation modeling derives the path loss or attenuation patterns of received signals to detect different devices using unsupervised learning algorithms [39]. In [207], the authors used the signals' propagation path effect, and they discovered that the received signal strength from transmitters in the same location would have very similar varying trends. They converted the signal strength metrics into time series and incorporated the Dynamic Time Warping algo-

TABLE VIII
COMPARISON OF DEVICE LOCALIZATION METHODS IN IoT

Methods	Requirements	Unit cost ¹	Precision	Weakness	References
Signal propagation modeling	Multiple collaborative transmitters to construct signal strength map.	Low	Depends on environmental features and refresh rate of respondent data.	<ul style="list-style-type: none"> Depends highly on signal propagation models of certain area. Results do not directly indicate certain device types or identities. 	[200]
Coherent TDoA	At least 4 coherent receivers and 5 receivers are recommended to linearize the computational process.	Median	Depends on the estimation of signals' Time of Arrival (ToA).	<ul style="list-style-type: none"> Receivers needs to be strictly synchronized. 	[201]
Sync-free TDoA	At least 4 receivers and receivers are able to communicate mutually.	Median	Same as coherent TDoA	<ul style="list-style-type: none"> Needs specific hardware with known response latency. 	[202], [203]

¹ Low: Does not require extra RF receivers; Median: Requiring commercially available RF receivers; High: Requiring dedicated hardware and specific processing stacks. ² Requiring multiple distributed receivers.

rithm to align and find differences between received signals. Finally, they applied a clustering algorithm to identify signals from active transmitters. In [208], the authors assumed that the received signals' Power Spectrum Density coefficients of each device, in a specific time window, form a mixture model of Gaussian distributions and propagation path related Relay distributions. In this way, they used the Expectation-Maximum algorithm to estimate the composition (different transmitters) of the received signals.

Signal propagation pattern modeling only provides an indirect evaluation on whether specific signals come from devices in close locations or with similar propagation paths. Although related methods are not widely utilized in commercial IoT devices owing to their complicated deployment environments, the methods provide a useful solution in preventing identity spoofing attacks in ADS-B systems [209], [210].

3) *Open issues:* Unsupervised device identification provides a novel solution when the identities of devices are not directly available. In essence, the unsupervised device identification and detection are similar to automatic knowledge discovery with the following issues to be addressed:

1) **Feature engineering:** Unsupervised device identification relies on feature engineering since representative vectors of devices are supposed to form distinctive clusters. Feature selection is still conducted manually, and there is no guarantee on whether the outputs of the mapped feature can form distinctive clusters.

2) **Clustering:** Clustering in the latent space can be challenging if the number of devices is unknown. Although one may use adaptive algorithm such as DBSCAN [211], Optics [212] or X-Means [213], the proper configurations of these algorithms is still difficult, similar obstacles are seen in setting hyperparameters in Deep Neural Networks (section III-C2a).

3) **Decision boundaries** Even if we know the number of devices, we can still get clusters with uncertain shapes or density, in which decision boundaries between different devices are difficult to define, as indicated in [197].

4) **Direct identity verification:** Researches on unsupervised device identification using behavior-independent and location-agnostic device specific features are still limited. Although unsupervised behavioral modeling has

shown promising results in identifying different types of devices, whether these methods are still effective in distinguishing devices from the same model needs further investigation.

Therefore, we believe learning-based unsupervised device detection is promising with great novelty, but the topic needs substantial investigation.

IV. LEARNING-ENABLED ABNORMAL DEVICE DETECTION

Previous sections have discussed methods to identify specific IoT devices. However, detection of compromised devices with abnormal behaviors is needed to alert ongoing attacks and discover system vulnerabilities.

In general, abnormal device detection algorithms are deployed in network and application layers. The detection algorithms first collect a certain amount of normal operation data from devices to create reference models (or signatures). Then IoT devices' operational data are collected and compared with reference models to judge whether significant deviations appear. Compared with device-specific identification schemes, the key methods are: abnormality detection, unsupervised learning [214], and supervised learning [215].

A. Statistical Modeling

Statistical modeling aims to judge whether devices are in abnormal situations. In [216], Markov models are utilized to judge whether IEEE 802.11 devices are compromised by calculating the probabilities of its sequential transitions of the protocol state machines. In [217], the authors modeled the Electronic Magnetic (EM) harmonics peaks of medical IoT devices as probabilistic distributions to assess whether a specific device is under attack. They assumed that when devices are operated under an abnormal scenario (with rogue shellcodes executing), its EM radiometric signals can deviate from known scenarios. However, statistical modeling requires manual selection of potentially informative features.

To reduce the cost of modeling IoT devices' normal behavior, Manufacturer Usage Description (MUD) profile [59] is proposed. A collection of MUD profiles for 30 commercial devices is provided in [60]. The MUD profiles enable operators to know devices' network flow patterns and dynamically

monitor their behavioral changes. Several open-source tools are provided to dynamically generate, validate, and compare IoT devices' MUD profiles in [61]. Besides, the authors presented that by comparing the deviation of devices' runtime MUD profiles with static ones, they can identify their behavioral deviations or even identify device types. In [218], the authors translated MUD profiles of devices into flowtable rules. The authors then used PCA to map each device's data traffic from side windows into its own representative one-class space, where X-Means [213] and Markov chains were used to partition the space and model the state transition in cluster centers. Finally, an exception will be triggered by a specific detector on either the mapped traffic pattern is out of boundaries or the state transitions do not comply with the reference model. Their results showed the accurate detection of several types of volumetric attacks.

B. Reconstruction Approaches

Reconstruction approaches aim to learn and reconstruct domain-specific patterns from devices' normal operation records. In other words, we need to develop a model to "memorize" the normal schemes of IoT devices by producing low reconstruction errors. Simultaneously, the model is supposed to produce high reconstruction errors for unknown scenarios or encounters behavioral deviations. This goal is generally achieved using deep autoencoders. Since an encoder removes a great amount of information, the corresponding decoder needs to reconstruct the lost information according to domain-specific memories. Consequently, once abnormal inputs are given to a well-trained autoencoder, its decoder would not be able to reconstruct such unknown inputs and yields a high abnormal score (reconstruction error). In [219]–[221], the authors utilized a deep autoencoder to detect abnormal activities by modeling the data traffic and content of IoT devices once abnormal activities are detected. In [222], the authors have shown that compared with other anomaly detection methods (one-class SVM [223], Isolation Forest [224] and Local Outlier Factor [225]), deep autoencoder yields the best result in terms of reliability and accuracy.

C. Prediction Approaches

Prediction approaches utilize temporal information in devices' operation records. Corresponding methods model each IoT device's operational data as multi-dimension time series. Then, device-specific prediction models are trained using time series from normal schemes. When devices are hijacked for rogue activities, they are not supposed to behave as predicted, causing the corresponding time series predictors to output high prediction errors.

In [226], the authors employed a CNN based predictor to analyze the abnormal behaviors in devices' network traffics. The results showed that predictors trained without abnormal data are sensitive (yield high prediction error) to anomalies. Similar work is shown in [227], and the authors used an autoregression model to capture the normal varying trend of devices' traffic volumes. However, modeling a single variable can not be sufficient in dealing with complicated scenarios.

Recent studies combine deep Autoencoder with Long Short Term Memory (LSTM) to derive abstracted representations of complex scenarios and make predictions. In [228] and [229], Deep Predictive Coding Neural Network [230] was used to predict consecutive frames of time-frequency video streams of wireless devices. They can even specify the type of attacks using the spatial distribution of error pixels in the reconstructed frames.

D. Open issue

Methods in this topic overlap with the methods of open set recognition in Deep Learning. We briefly list several open issues in this topic:

- **Selection of behavioral features:** Manual feature selection along with dimension reduction are applied extensively. A concern is that we can not guarantee the selected features are sensitive to unknown intrusions in the future.
- **Processing of abnormality metrics:** Generally, intrusion detection approaches provide metrics corresponding to the degree of deviation. However, the output error metrics require a posterior process, e.g., selecting appropriate decision thresholds or aggregation window length, which balances between the true positive, false negative, and response latency. One solution is to regard the corresponding parameters as hyperparameters and use cross-validation to tune them. The processing of error metrics remains a case-specific open issue.

V. CHALLENGES AND FUTURE RESEARCH DIRECTIONS

Our literature review has shown that device detection and identification provide another layer of security features to IoT. However, the existing solutions are still far from perfect. This section summarizes the existing challenges of IoT device identification and detection as well as future research directions.

A. Challenges in machine learning models

1) *Unknown device recognition:* Existing works focus on the accuracy they can obtain using a fixed dataset with all devices labeled, in which Black-Box models (e.g., Deep Learning and SVM) are commonly employed. In practical scenarios, these models can produce wrong answers when encountering novel devices. Additional mechanisms are needed to identify unknown signals. Although we can use the one-versus-rest technique to train a group of classifiers and avoid producing results on unknown devices. However, once we have new devices to register, all classifiers in the group are supposed to be retrained from scratch. Therefore, we need to provide a solution to verify the known devices. Meanwhile, we need to identify:

- Devices that are exactly not in the scope of the identification system.
- Unknown devices that are from identical manufacturers. Devices of the same model from an identical manufacturer can share similar behavior patterns, e.g., network flow characteristics. Such similarities can impede identity

TABLE IX
EXISTING DATASET FOR PHYSICAL LAYER IOT DEVICE IDENTIFICATION

Dataset	Number of transmitters	Sample rate	Frequency	Protocol / device	Fingerprint injection
[231]	6	50 MS/s	Multiple	GNSS	N
[232]	100+	10 MSa/s	1090 MHz	ADS-B	N
[135]	100+	8 MSa/s	1090 MHz	ADS-B	N
[133]	4	40 MSa/s	2.4 GHz	Drones	N
[233]	17	20 GSa/s	2.4 GHz	Drone remote controllers	N
[234]	27	5 to 20 GSa/s	2.4 GHz	Bluetooth	N
[235]	16	5 Ms/s	2.4 GHz	USRP X310	Y
[236]	4	5 Ms/s	900 MHz	USRP B200mini	Y
[237]	4	5 MSa/s to 7.69 MS/s	2.685 GHz	USRP X310	Y
[238]	7	10 MSa/s	2.4 GHz	DJI M100 drones	N

verification in the network, transportation, or application layers.

The latter is more challenging and requires extracting behavior-independent characteristics. We believe that without the capability of unknown device recognition, these types of systems are still far from practice.

2) *Incremental learning with new devices*: Incremental learning [188] in this domain emphasizes that an identification or detection model should be able to learn newly registered devices without retraining on a large dataset containing new and old devices. Because retaining the old dataset or deriving generators for knowledge replay is computationally expensive. This topic faces several challenges:

- Knowing the maximum number of devices a model can memorize, especially for the Deep Neural Networks.
- Expanding models dynamically as new devices are being added. Incremental learning is natively supported in Nearest Neighbor algorithms but is challenging to implement in Deep Neural Networks.

3) *Deployment of device identification models*: The deployment sites and model providers' lab can differ dramatically, in which identification accuracy can be impaired. This issue is more severe in device identification models using wireless signals due to the difference of wireless channel characteristics. For alleviation, extra works are needed:

- Deriving features that are independent of wireless channels or deployment sites. Authors in [235], [239] suggested that neural networks can only learn about channel-specific features rather than device-specific features.
- Occasional finetunes are needed with the help of incremental or transfer learning to adapt to variations.
- Model providers need to use data augmentation methods to simulate operational variations during lab training, as suggested in [240].
- Model providers can use multi-domain training to derive multi-purpose feature extractors, which will be utilized as building blocks for domain-specific device identification models. Diverse training from different domains could provide more robust feature extractors.

4) *Reliable benchmark datasets*: The IoT device identification is a pattern recognition problem on signals or communication patterns. A common benchmark dataset is critical for comparing various methods in device identification and rogue

device detection reliably. However, by the end of this survey, we only find a limited number of datasets providing devices' raw signals or network traffic traces in diverse scenarios. Some datasets are available in [132], [133] and [134], respectively. For physical layer device identification, some popular datasets are summarized in Table IX. As summarized, many of them only contain a limited number of devices, while a larger dataset containing only ADS-B raw signals from more than 100 airborne transponders are provided in [135]. Another dataset containing more than 30 IoT devices' traffic traces under volumetric attack and benign scenarios are in [60]. These dataset are important because they provide fair comparisons between algorithms. Additionally, models trained on large datasets can be efficiently transferred to more specific applications [241], [242].

B. Challenges in feature engineering

1) *The robustness of features*: Although many existing works claim the effectiveness of their discovered features, only very few evaluate the features' robustness under various scenarios in terms of device mobility pattern, temperature, obstacles, etc. Feature robustness has a limited influence on device type identification in the network or higher layers. However, in the Physical Layer identification of wireless devices, the robustness of features would severely impair the final model. Currently, a popular way to enforce robust feature discovery is through data augmentation to simulate various scenarios. Besides, in neural networks, regularization and dropout methods can encourage models to make full use of input data and discover robust latent features, but their effectiveness needs further study.

2) *Making use of time-varying features*: Some device detection and identification models use protocol-agnostic and behavior-independent features from physical layer wireless signals. However, in mobile environments, devices' movements can result in time-varying channel conditions, in which device identification methods based on static channel characteristics can be impaired. On the other hand, varying patterns of channels, signal strength, etc. also encode valuable features, e.g., location, distance, noise pattern, and etc., to help distinguish IoT devices [243], [244]. Therefore, both discovering time-invariant features and making use of time-

varying features are still an open issue in device identification and detection.

3) *Challenges from deep generative attackers*: The utilization of GAN brings challenges to device identification, especially in the Physical Layer. Using GAN models, an attacker can train highly realistic signal or data packet generators to mimic its victims' signal characteristics. Research in [245] showed that GAN can increase the success rate of spoofing attacks from less than 10% to approximately 80%. Fortunately, a simple remedy is to use MIMO receivers and wireless localization methods to estimate whether a transmitter is at an expected position. Additionally, controlled imperfections can be dynamically imprinted into the devices' signals or data flows in a Pseudorandom Noise Code driven time-varying manner [239], which is cryptographically impossible to predict.

C. Future research trends

1) *Deep identification models with explainable behaviors and assured performances*: The conveniences of Deep Neural Network make it a versatile tool to implement IoT device identification and rogue device detection systems, but more efforts have to be made, especially for model explainability and performance assurability. On the one hand, we have limited knowledge of the decision process, especially on how a deep neural network makes its final decisions and corresponding decision boundaries. Without knowing the decision process and decision boundaries, there is no way to assure its performance. On the other hand, researches on the explainability of Deep Neural Networks focus on explaining models' behaviors but do not provide guidelines on deriving assurable performance. Without explainability, we can not assure the performance of models.

2) *Unsupervised and incremental deep learning for device identification*: With a large number of devices being connected to IoT, device identification and detection models need to incrementally adapt to operational variations in real-time. A solution can be the seamless integration of the feature abstraction capability of deep neural networks, incremental learning, and unsupervised learning. The knowledge of using deep neural networks to perform unsupervised learning for IoT device identification and detection is currently limited. Meanwhile, incremental learning in deep models for device identification and detection is also rarely investigated.

3) *Controlled imprinting of verifiable patterns*: Compared with passive non-cryptographic device identification and detection methods in this survey, a proactive way is imprinting verifiable patterns into devices' transmitted signals or activity patterns. As suggested in [236], controlled imperfections are utilized as verifiable patterns. Embedded these patterns in signals could significantly enhance the performance of device identification. However, a critical concern is how to prevent the adversaries from collecting and learning about the imprinted identity verification information. As suggested in [235], a possible solution is to dynamically change the identity verification patterns according to a pair of synchronized pseudorandom code generators, where the initialization keys are only shared

among the device and corresponding device identifiers. Methods are still limited in imprinting verifiable patterns that are difficult to learn.

VI. CONCLUSION

This survey aims to provide a comprehensive on the existing technologies on IoT device detection and identification from passively collected network traffic traces and wireless signal patterns. We discuss existing non-cryptographic IoT device identification mechanisms from the perspective of machine learning and pinpoint several key developing trends such as incremental learning, abnormality detection, and deep unsupervised learning with explainability. We found that a multi-perspective IoT wireless device detection and identification framework is needed. Future research for rogue IoT device identification and detection needs to cope with challenges beyond signal processing and borrow ideas from advanced topics in Artificial Intelligence and Knowledge Discovery.

ACKNOWLEDGMENT

This research was partially supported through Embry-Riddle Aeronautical University's Faculty Innovative Research in Science and Technology (FIRST) Program and the National Science Foundation under Grant No. 1956193.

REFERENCES

- [1] S. Jeschke, C. Brecher, H. Song, and D. Rawat, *Industrial Internet of Things: Cybermanufacturing Systems*. Cham, Switzerland: Springer, 2017.
- [2] Y. Sun, H. Song, A. J. Jara, and R. Bie, "Internet of things and big data analytics for smart and connected communities," *IEEE Access*, vol. 4, pp. 766–773, 2016.
- [3] H. Song, R. Srinivasan, T. Sookoor, and S. Jeschke, *Smart Cities: Foundations, Principles and Applications*. Hoboken, NJ: Wiley, 2017.
- [4] Y. Zhang, L. Sun, H. Song, and X. Cao, "Ubiquitous wsn for healthcare: Recent advances and future prospects," *IEEE Internet of Things Journal*, vol. 1, no. 4, pp. 311–318, Aug 2014.
- [5] Y. Sun and H. Song, *Secure and Trustworthy Transportation Cyber-Physical Systems*. Singapore: Springer, 2017.
- [6] K. Zhang, Y. Jiang, D. Liu, and H. Song, "Spatio-temporal data mining for aviation delay prediction," in *2020 IEEE 39th International Performance Computing and Communications Conference (IPCCC)*. IEEE, 2020, pp. 1–7.
- [7] Y. Liu, J. Li, Z. Ming, H. Song, X. Weng, and J. Wang, "Domain-specific data mining for residents' transit pattern retrieval from incomplete information," *Journal of Network and Computer Applications*, vol. 134, pp. 62–71, 2019.
- [8] Y. Jiang, S. Niu, K. Zhang, B. Chen, C. Xu, D. Liu, and H. Song, "Spatial-temporal graph data mining for iot-enabled air mobility prediction," *IEEE Internet of Things Journal*, pp. 1–1, 2021.
- [9] G. Dartmann, H. Song, and A. Schmeink, *Big Data Analytics for Cyber-Physical Systems: Machine Learning for the Internet of Things*. Elsevier, 2019.
- [10] Y. Jiang, Y. Liu, D. Liu, and H. Song, "Applying machine learning to aviation big data for flight delay prediction," in *2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*. IEEE, 2020, pp. 665–672.
- [11] K. Zhang, Y. Liu, J. Wang, H. Song, and D. Liu, "Tree-based airspace capacity estimation," in *2020 Integrated Communications Navigation and Surveillance Conference (ICNS)*. IEEE, 2020, pp. 5C1–1.
- [12] H. Song, G. A. Fink, and S. Jeschke, *Security and Privacy in Cyber-Physical Systems: Foundations, Principles and Applications*. Chichester, UK: Wiley-IEEE Press, 2017.

- [13] I. Butun, P. Österberg, and H. Song, "Security of the internet of things: Vulnerabilities, attacks and countermeasures," *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2019.
- [14] J. Wang, Y. Liu, S. Niu, H. Song, W. Jing, and J. Yuan, "Blockchain enabled verification for cellular-connected unmanned aircraft system networking," *Future Generation Computer Systems*, vol. 123, pp. 233–244, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167739X21001461>
- [15] J. Wurm, K. Hoang, O. Arias, A.-R. Sadeghi, and Y. Jin, "Security analysis on consumer and industrial iot devices," in *2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC)*. IEEE, 2016, pp. 519–524.
- [16] O. Schwartz, Y. Mathov, M. Bohadana, Y. Elovici, and Y. Oren, "Reverse engineering iot devices: Effective techniques and methods," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4965–4976, 2018.
- [17] A. Gupta, *The IoT Hacker's Handbook*. Springer, 2019.
- [18] K. Liu, M. Yang, Z. Ling, H. Yan, Y. Zhang, X. Fu, and W. Zhao, "On manually reverse engineering communication protocols of linux based iot systems," *arXiv preprint arXiv:2007.11981*, 2020.
- [19] L. Costa, J. P. Barros, and M. Tavares, "Vulnerabilities in iot devices for smart home environment," in *Proceedings of the 5th International Conference on Information Systems Security e Privacy, ICISSP 2019*, vol. 1. SCITEPRESS, 2019, pp. 615–622.
- [20] J. Pollack and P. Ranganathan, "Aviation navigation systems security: Ads-b, gps, iff," in *Proceedings of the International Conference on Security and Management (SAM)*. The Steering Committee of The World Congress in Computer Science, Computer ..., 2018, pp. 129–135.
- [21] M. R. Manesh and N. Kaabouch, "Analysis of vulnerabilities, attacks, countermeasures and overall risk of the automatic dependent surveillance-broadcast (ads-b) system," *International Journal of Critical Infrastructure Protection*, vol. 19, pp. 16–31, 2017.
- [22] D. Dimitrios, M. Baldauf, and M. Kitada, "Vulnerabilities of the automatic identification system in the era of maritime autonomous surface ships," 06 2018.
- [23] C. Ray, R. Gallen, C. Iphar, A. Napoli, and A. Bouju, "Deais project: detection of ais spoofing and resulting risks," in *OCEANS 2015-Genova*. IEEE, 2015, pp. 1–6.
- [24] F. Lakew Yihunie, A. K. Singh, and S. Bhatia, "Assessing and exploiting security vulnerabilities of unmanned aerial vehicles," in *Smart Systems and IoT: Innovations in Computing*, A. K. Soman, R. S. Shekharwati, A. Mundra, S. Srivastava, and V. K. Verma, Eds. Singapore: Springer Singapore, 2020, pp. 701–710.
- [25] A. Koubaa, A. Allouch, M. Alajlan, Y. Javed, A. Belghith, and M. Khalgui, "Micro air vehicle link (mavlink) in a nutshell: A survey," *IEEE Access*, vol. 7, pp. 87658–87680, 2019.
- [26] L. Y. Paul, J. S. Baras, and B. M. Sadler, "Physical-layer authentication," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 38–51, 2008.
- [27] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, "Wireless physical-layer identification: Modeling and validation," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 2091–2106, 2016.
- [28] H. A. Khattak, M. A. Shah, S. Khan, I. Ali, and M. Imran, "Perception layer security in internet of things," *Future Generation Computer Systems*, vol. 100, pp. 144–164, 2019.
- [29] L. Bai, L. Zhu, J. Liu, J. Choi, and W. Zhang, "Physical layer authentication in wireless communication networks: A survey," *Journal of Communications and Information Networks*, vol. 5, no. 3, pp. 237–264, 2020.
- [30] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless device identification with radiometric signatures," in *Proceedings of the 14th ACM international conference on Mobile computing and networking*. ACM, 2008, pp. 116–127.
- [31] J. Sun, "An open-access book about decoding mode-s and ads-b data," <https://mode-s.org/>, May 2017.
- [32] B. J. Tetreault, "Use of the automatic identification system (ais) for maritime domain awareness (mda)," in *Proceedings of Oceans 2005 Mts/Ieee*. IEEE, 2005, pp. 1590–1594.
- [33] M. A. Amanullah, R. A. A. Habeeb, F. H. Nasaruddin, A. Gani, E. Ahmed, A. S. M. Nainar, N. M. Akim, and M. Imran, "Deep learning and big data technologies for iot security," *Computer Communications*, vol. 151, pp. 495–517, 2020.
- [34] M. A. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for internet of things (iot) security," *IEEE Communications Surveys & Tutorials*, 2020.
- [35] N. Wang, P. Wang, A. Alipour-Fanid, L. Jiao, and K. Zeng, "Physical-layer security of 5g wireless networks for iot: Challenges and opportunities," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8169–8181, 2019.
- [36] G. Baldini and G. Steri, "A survey of techniques for the identification of mobile phones using the physical fingerprints of the built-in components," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1761–1789, 2017.
- [37] Q. Xu, R. Zheng, W. Saad, and Z. Han, "Device fingerprinting in wireless networks: Challenges and opportunities," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 94–104, 2016.
- [38] B. Danev and S. Capkun, "Physical-layer identification of wireless sensor nodes," *Technical report/ETH Zürich, Department of Computer Science*, vol. 604, 2012.
- [39] K. Zeng, K. Govindan, and P. Mohapatra, "Non-cryptographic authentication and identification in wireless networks," *network security*, vol. 1, p. 3, 2010.
- [40] K.-H. Wang, C.-M. Chen, W. Fang, and T.-Y. Wu, "On the security of a new ultra-lightweight authentication protocol in iot environment for rfid tags," *The Journal of Supercomputing*, vol. 74, no. 1, pp. 65–70, 2018.
- [41] F. Loi, A. Sivanathan, H. H. Gharakheili, A. Radford, and V. Sivaraman, "Systematically evaluating security and privacy for consumer iot devices," in *Proceedings of the 2017 Workshop on Internet of Things Security and Privacy*, 2017, pp. 1–6.
- [42] D. S. Guamán, J. M. Del Alamo, and J. C. Caiza, "A systematic mapping study on software quality control techniques for assessing privacy in information systems," *IEEE Access*, vol. 8, pp. 74 808–74 833, 2020.
- [43] J. Ren, D. J. Dubois, D. Choffnes, A. M. Mandalari, R. Kolcun, and H. Haddadi, "Information exposure from consumer iot devices: A multidimensional, network-informed measurement approach," in *Proceedings of the Internet Measurement Conference*, 2019, pp. 267–279.
- [44] S. Al-Sarawi, M. Anbar, K. Alieyan, and M. Alzubaidi, "Internet of things (iot) communication protocols," in *2017 8th International conference on information technology (ICIT)*. IEEE, 2017, pp. 685–690.
- [45] I. Jawhar, N. Mohamed, and J. Al-Jaroodi, "Networking architectures and protocols for smart city systems," *Journal of Internet Services and Applications*, vol. 9, no. 1, p. 26, 2018.
- [46] F. Metzger, T. Hoßfeld, A. Bauer, S. Kounev, and P. E. Heegaard, "Modeling of aggregated iot traffic and its application to an iot cloud," *Proceedings of the IEEE*, vol. 107, no. 4, pp. 679–694, 2019.
- [47] S. Han, K. Jang, K. Park, and S. Moon, "Packetshader: a gpu-accelerated software router," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 4, pp. 195–206, 2010.
- [48] F. Hu, Q. Hao, and K. Bao, "A survey on software-defined network and openflow: From concept to implementation," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2181–2206, 2014.
- [49] W. Rafique, L. Qi, I. Yaqoob, M. Imran, R. ur Rasool, and W. Dou, "Complementing iot services through software defined networking and edge computing: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, 2020.
- [50] L. Gao, C. Zhang, and L. Sun, "Restful web of things api in sharing sensor data," in *2011 International Conference on Internet Technology and Applications*. IEEE, 2011, pp. 1–4.
- [51] A. Sivanathan, "Iot behavioral monitoring via network traffic analysis," *arXiv preprint arXiv:2001.10632*, 2020.
- [52] A. Sivanathan, D. Sherratt, H. H. Gharakheili, A. Radford, C. Wijenayake, A. Vishwanath, and V. Sivaraman, "Characterizing and classifying iot traffic in smart cities and campuses," in *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2017, pp. 559–564.
- [53] S. Marchal, M. Miettinen, T. D. Nguyen, A.-R. Sadeghi, and N. Asokan, "Audi: Toward autonomous iot device-type identification using periodic communication," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1402–1412, 2019.
- [54] M. Miettinen, S. Marchal, I. Hafeez, N. Asokan, A.-R. Sadeghi, and S. Tarkoma, "Iot sentinel: Automated device-type identification for security enforcement in iot," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2017, pp. 2177–2184.
- [55] Y. Meidan, M. Bohadana, A. Shabtai, J. D. Guarino, M. Ochoa, N. O. Tippenhauer, and Y. Elovici, "Profiliot: a machine learning approach for iot device identification based on network traffic analysis," in

Proceedings of the symposium on applied computing, 2017, pp. 506–509.

[56] Y. Meidan, M. Bohadana, A. Shabtai, M. Ochoa, N. O. Tippenhauer, J. D. Guarnizo, and Y. Elovici, “Detection of unauthorized iot devices using machine learning techniques,” *arXiv preprint arXiv:1709.04647*, 2017.

[57] A. Aksoy and M. H. Gunes, “Automated iot device identification using network traffic,” in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 2019, pp. 1–7.

[58] A. Sivanathan, H. H. Gharakheili, and V. Sivaraman, “Managing iot cyber-security using programmable telemetry and machine learning,” *IEEE Transactions on Network and Service Management*, vol. 17, no. 1, pp. 60–74, 2020.

[59] E. Lear, R. Droms, and D. Romascanu, “Rfc 8520: Manufacturer usage description specification,” *Internet Engineering Task Force*, 2019.

[60] H. Hassan, S. Vijay, H. Ayyoob, S. Arunan, and P. Arman, “IoT device profiles, attack traces and traffic traces,” <https://iotanalytics.unsw.edu.au/index>, Accessed on Oct. 2020.

[61] A. Hamza, D. Ranathunga, H. H. Gharakheili, T. A. Benson, M. Roughan, and V. Sivaraman, “Verifying and monitoring iots network behavior using mud profiles,” *IEEE Transactions on Dependable and Secure Computing*, 2020.

[62] Y. Ashibani and Q. H. Mahmoud, “A user authentication model for iot networks based on app traffic patterns,” in *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, 2018, pp. 632–638.

[63] ———, “Design and evaluation of a user authentication model for iot networks based on app event patterns,” *Cluster Computing*, pp. 1–14, 2020.

[64] J. Kotak and Y. Elovici, “Iot device identification using deep learning,” *arXiv preprint arXiv:2002.11686*, 2020.

[65] Z. Zvonar and J. Mitola, *Software radio technologies: selected readings*. Wiley-IEEE Press, 2001.

[66] B. Chatterjee, D. Das, S. Maity, and S. Sen, “Rf-puf: Enhancing iot security through authentication of wireless nodes using in-situ machine learning,” *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 388–398, 2018.

[67] C. Herder, M.-D. Yu, F. Koushanfar, and S. Devadas, “Physical unclonable functions and applications: A tutorial,” *Proceedings of the IEEE*, vol. 102, no. 8, pp. 1126–1141, 2014.

[68] A. C. Polak and D. L. Goeckel, “Wireless device identification based on rf oscillator imperfections,” *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2492–2501, 2015.

[69] M. Azarmehr, A. Mehta, and R. Rashidzadeh, “Wireless device identification using oscillator control voltage as rf fingerprint,” in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE, 2017, pp. 1–4.

[70] 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 *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*. ACM, 2018, pp. 261–272.

[71] L. Peng, A. Hu, Y. Jiang, Y. Yan, and C. Zhu, “A differential constellation trace figure based device identification method for zigbee nodes,” in *2016 8th International Conference on Wireless Communications & Signal Processing (WCSP)*. IEEE, 2016, pp. 1–6.

[72] 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 of Things Journal*, vol. 6, no. 1, pp. 349–360, 2018.

[73] L. Peng, J. Zhang, M. Liu, and A. Hu, “Deep learning based rf fingerprint identification using differential constellation trace figure,” *IEEE Transactions on Vehicular Technology*, 2019.

[74] G. Zhang, L. Xia, S. Jia, and Y. Ji, “Identification of cloned hf rfid proximity cards based on rf fingerprinting,” in *2016 IEEE Trustcom/BigDataSE/ISPA*. IEEE, 2016, pp. 292–300.

[75] W. E. Cobb, E. W. Garcia, M. A. Temple, R. O. Baldwin, and Y. C. Kim, “Physical layer identification of embedded devices using rf-dna fingerprinting,” in *2010-Milcom 2010 Military Communications Conference*. IEEE, 2010, pp. 2168–2173.

[76] C. Wang, Y. Lin, and Z. Zhang, “Research on physical layer security of cognitive radio network based on rf-dna,” in *2017 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*. IEEE, 2017, pp. 252–255.

[77] 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 Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1862–1874, 2016.

[78] B. W. Ramsey, T. D. Stubbs, B. E. Mullins, M. A. Temple, and M. A. Buckner, “Wireless infrastructure protection using low-cost radio frequency fingerprinting receivers,” *international journal of critical infrastructure protection*, vol. 8, pp. 27–39, 2015.

[79] C. K. Dubendorfer, B. W. Ramsey, and M. A. Temple, “An rf-dna verification process for zigbee networks,” in *MILCOM 2012-2012 IEEE Military Communications Conference*. IEEE, 2012, pp. 1–6.

[80] P. K. Harmer, M. A. Temple, M. A. Buckner, and E. Farquhar, “Using differential evolution to optimize ‘learning from signals’ and enhance network security,” in *Proceedings of the 13th annual conference on Genetic and evolutionary computation*. ACM, 2011, pp. 1811–1818.

[81] X. Zhou, A. Hu, G. Li, L. Peng, Y. Xing, and J. Yu, “Design of a robust rf fingerprint generation and classification scheme for practical device identification,” in *2019 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2019, pp. 196–204.

[82] B. Danev, H. Luecken, S. Capkun, and K. El Defrawy, “Attacks on physical-layer identification,” in *Proceedings of the third ACM conference on Wireless network security*. ACM, 2010, pp. 89–98.

[83] A. C. Polak and D. L. Goeckel, “Identification of wireless devices of users who actively fake their rf fingerprints with artificial data distortion,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 5889–5899, 2015.

[84] M. Köse, S. Taşçıoğlu, and Z. Telatar, “Rf fingerprinting of iot devices based on transient energy spectrum,” *IEEE Access*, vol. 7, pp. 18 715–18 726, 2019.

[85] Z. Zhang, Y. Li, C. Wang, M. Wang, Y. Tu, and J. Wang, “An ensemble learning method for wireless multimedia device identification,” *Security and Communication Networks*, vol. 2018, 2018.

[86] Y. Tu, Z. Zhang, Y. Li, C. Wang, and Y. Xiao, “Research on the internet of things device recognition based on rf-fingerprinting,” *IEEE Access*, vol. 7, pp. 37 426–37 431, 2019.

[87] A. M. Ali, E. Uzundurukan, and A. Kara, “Assessment of features and classifiers for bluetooth rf fingerprinting,” *IEEE Access*, vol. 7, pp. 50 524–50 535, 2019.

[88] J. Yu, A. Hu, F. Zhou, Y. Xing, Y. Yu, G. Li, and L. Peng, “Radio frequency fingerprint identification based on denoising autoencoders,” *arXiv preprint arXiv:1907.08809*, 2019.

[89] J. M. T. Romano, R. Attux, C. C. Cavalcante, and R. Suyama, *Unsupervised signal processing: channel equalization and source separation*. CRC Press, 2018.

[90] F. Restuccia, S. D’Oro, A. Al-Shawabka, M. Belgiovine, L. Angioloni, S. Ioannidis, K. Chowdhury, and T. Melodia, “Deepradioid: Real-time channel-resilient optimization of deep learning-based radio fingerprinting algorithms,” *arXiv preprint arXiv:1904.07623*, 2019.

[91] C. Xu, G. Huang, and X. Hou, “Research on communication fm signal blind separation under jamming environments,” in *2nd International Forum on Management, Education and Information Technology Application (IFMEITA 2017)*. Atlantis Press, 2018.

[92] S. N. Baliarsingh, A. Senapati, A. Deb, and J. S. Roy, “Adaptive beam formation for smart antenna for mobile communication network using new hybrid algorithms,” in *2016 International Conference on Communication and Signal Processing (ICCP)*. IEEE, 2016, pp. 2146–2151.

[93] S. S. Hanna and D. Cabric, “Deep learning based transmitter identification using power amplifier nonlinearity,” in *2019 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2019, pp. 674–680.

[94] T. Zheng, Z. Sun, and K. Ren, “Fid: Function modeling-based data-independent and channel-robust physical-layer identification,” in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 199–207.

[95] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, “Tool release: Gathering 802.11 n traces with channel state information,” *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, pp. 53–53, 2011.

[96] Y. Ma, G. Zhou, and S. Wang, “Wifi sensing with channel state information: A survey,” *ACM Computing Surveys (CSUR)*, vol. 52, no. 3, p. 46, 2019.

[97] H. Liu, Y. Wang, J. Liu, J. Yang, and Y. Chen, “Practical user authentication leveraging channel state information (csi),” in *Proceedings of the 9th ACM symposium on Information, computer and communications security*. ACM, 2014, pp. 389–400.

[98] S. Zaman, C. Chakraborty, N. Mehajabin, M. Mamun-Or-Rashid, and M. A. Razzaque, “A deep learning based device authentication scheme using channel state information,” in *2018 International Conference on Innovation in Engineering and Technology (ICIET)*. IEEE, 2018, pp. 1–5.

[99] Y. Zou, Y. Wang, S. Ye, K. Wu, and L. M. Ni, "Tagfree: Passive object differentiation via physical layer radiometric signatures," in *2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2017, pp. 237–246.

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

[101] 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 *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 190–198.

[102] Z. Wang, B. Guo, Z. Yu, and X. Zhou, "Wi-fi csi-based behavior recognition: From signals and actions to activities," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 109–115, 2018.

[103] F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, "Wfid: Passive device-free human identification using wifi signal," in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*. ACM, 2016, pp. 47–56.

[104] Y. Zeng, P. H. Pathak, C. Xu, and P. Mohapatra, "Your ap knows how you move: fine-grained device motion recognition through wifi," in *Proceedings of the 1st ACM workshop on Hot topics in wireless*. ACM, 2014, pp. 49–54.

[105] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures," in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 617–628.

[106] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2016.

[107] G. Li, J. Yu, Y. Xing, and A. Hu, "Location-invariant physical layer identification approach for wifi devices," *IEEE Access*, vol. 7, pp. 106 974–106 986, 2019.

[108] C. Xu, B. Chen, Y. Liu, F. He, and H. Song, "Rf fingerprint measurement for detecting multiple amateur drones based on stft and feature reduction," in *2020 Integrated Communications Navigation and Surveillance Conference (ICNS)*. IEEE, 2020, pp. 4G1–1.

[109] S. Chen, F. Xie, Y. Chen, H. Song, and H. Wen, "Identification of wireless transceiver devices using radio frequency (rf) fingerprinting based on stft analysis to enhance authentication security," in *2017 IEEE 5th International Symposium on Electromagnetic Compatibility (EMC-Beijing)*. IEEE, 2017, pp. 1–5.

[110] D. R. Reising, M. A. Temple, and J. A. Jackson, "Authorized and rogue device discrimination using dimensionally reduced rf-dna fingerprints," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 6, pp. 1180–1192, 2015.

[111] Y. Li, L. Chen, J. Chen, F. Xie, S. Chen, and H. Wen, "A low complexity feature extraction for the rf fingerprinting process," in *2018 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2018, pp. 1–2.

[112] X. Wu, Y. Shi, W. Meng, X. Ma, and N. Fang, "Specific emitter identification for satellite communication using probabilistic neural networks," *International Journal of Satellite Communications and Networking*, vol. 37, no. 3, pp. 283–291, 2019.

[113] D. Sun, Y. Li, and Y. Xu, "Specific emitter identification based on normalized frequency spectrum," in *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*. IEEE, 2016, pp. 1875–1879.

[114] Y. Yuan, Z. Huang, H. Wu, and X. Wang, "Specific emitter identification based on hilbert–huang transform-based time–frequency–energy distribution features," *IET Communications*, vol. 8, no. 13, pp. 2404–2412, 2014.

[115] U. Satija, N. Trivedi, G. Biswal, and B. Ramkumar, "Specific emitter identification based on variational mode decomposition and spectral features in single hop and relaying scenarios," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 581–591, 2018.

[116] X. Wang, J. Duan, C. Wang, G. Cui, and W. Wang, "A radio frequency fingerprinting identification method based on energy entropy and color moments of the bispectrum," in *2017 9th International Conference on Advanced Infocomm Technology (ICAIT)*. IEEE, 2017, pp. 150–154.

[117] Y.-K. Lei, "Individual communication transmitter identification using correntropy-based collaborative representation," in *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 2016, pp. 1194–1200.

[118] M. Sun, L. Zhang, J. Bao, and Y. Yan, "Rf fingerprint extraction for gnss anti-spoofing using axial integrated wigner bispectrum," *Journal of Information Security and Applications*, vol. 35, pp. 51–54, 2017.

[119] Z. Zhang, Y. Li, and C. Wang, "Research on individual identification of wireless devices based on signal's energy distribution," in *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*. IEEE, 2018, pp. 1–5.

[120] N. E. Huang, *Hilbert-Huang transform and its applications*. World Scientific, 2014, vol. 16.

[121] W. Wang, H. Liu, J. Yang, and H. Yin, "Specific emitter identification using decomposed hierarchical feature extraction methods," in *2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*. IEEE, 2017, pp. 1639–1643.

[122] M. G. Frei and I. Osorio, "Intrinsic time-scale decomposition: time–frequency–energy analysis and real-time filtering of non-stationary signals," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 463, no. 2078, pp. 321–342, 2006.

[123] Y. Shi and M. A. Jensen, "Improved radiometric identification of wireless devices using mimo transmission," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 4, pp. 1346–1354, 2011.

[124] T. D. Vo-Huu, T. D. Vo-Huu, and G. Noubir, "Fingerprinting wi-fi devices using software defined radios," in *Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks*. ACM, 2016, pp. 3–14.

[125] I. Agadakos, N. Agadakos, J. Polakis, and M. R. Amer, "Deep complex networks for protocol-agnostic radio frequency device fingerprinting in the wild," *arXiv preprint arXiv:1909.08703*, 2019.

[126] J. Yu, A. Hu, G. Li, and L. Peng, "A robust rf fingerprinting approach using multi-sampling convolutional neural network," *IEEE Internet of Things Journal*, 2019.

[127] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for rf device fingerprinting in cognitive communication networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 160–167, 2018.

[128] S. Riyaz, K. Sankhe, S. Ioannidis, and K. Chowdhury, "Deep learning convolutional neural networks for radio identification," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 146–152, 2018.

[129] G. Huang, Y. Yuan, X. Wang, and Z. Huang, "Specific emitter identification for communications transmitter using multi-measurements," *Wireless Personal Communications*, vol. 94, no. 3, pp. 1523–1542, 2017.

[130] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical review letters*, vol. 88, no. 17, p. 174102, 2002.

[131] C. L. Nikias and M. R. Raghubeer, "Bispectrum estimation: A digital signal processing framework," *Proceedings of the IEEE*, vol. 75, no. 7, pp. 869–891, 1987.

[132] C. Morin, L. Cardoso, J. Hoydis, J.-M. Gorce, and T. Vial, "Transmitter classification with supervised deep learning," *arXiv preprint arXiv:1905.07923*, 2019.

[133] M. S. Allahham, M. F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, and A. Erbad, "Dronerf dataset: A dataset of drones for rf-based detection, classification and identification," *Data in brief*, vol. 26, p. 104313, 2019.

[134] C. Kaushik, "Genesys lab mldatasets," <http://genesys-lab.org/mldatasets.au/index>, Accessed on Oct. 2020.

[135] Y. Liu, J. Wang, H. Song, S. Niu, and Y. Thomas, "A 24-hour signal recording dataset with labels for cybersecurity and IoT," 2020. [Online]. Available: <http://dx.doi.org/10.21227/gt9v-kz32>

[136] G. Baldini, C. Gentile, R. Giuliani, and G. Steri, "Comparison of techniques for radiometric identification based on deep convolutional neural networks," *Electronics Letters*, vol. 55, no. 2, pp. 90–92, 2018.

[137] D. Roy, T. Mukherjee, M. Chatterjee, E. Blasch, and E. Pasiliao, "Rfaf: Adversarial learning for rf transmitter identification and classification," *IEEE Transactions on Cognitive Communications and Networking*, 2019.

[138] S. Gopalakrishnan, M. Cekic, and U. Madhow, "Robust wireless fingerprinting via complex-valued neural networks," *arXiv preprint arXiv:1905.09388*, 2019.

[139] J. Huang, Y. Lei, and X. Liao, "Communication transmitter individual feature extraction method based on stacked denoising autoencoders under small sample prerequisite," in *2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC)*. IEEE, 2017, pp. 132–135.

[140] T. Oyedare and J.-M. J. Park, "Estimating the required training dataset size for transmitter classification using deep learning."

[141] J. S. Dramsch, M. Lüthje, and A. N. Christensen, "Complex-valued neural networks for machine learning on non-stationary physical data," *CoRR*, vol. abs/1905.12321, 2019. [Online]. Available: <http://arxiv.org/abs/1905.12321>

[142] Y. Liu, J. Wang, J. Li, H. Song, T. Yang, S. Niu, and Z. Ming, "Zero-bias deep learning for accurate identification of internet of things (iot) devices," *IEEE Internet of Things Journal*, 2020.

[143] Y. Liu, J. Wang, S. Niu, and H. Song, "Deep learning enabled reliable identity verification and spoofing detection," in *International Conference on Wireless Algorithms, Systems, and Applications*. Springer, 2020, pp. 333–345.

[144] M. Li, O. Li, G. Liu, and C. Zhang, "Generative adversarial networks-based semi-supervised automatic modulation recognition for cognitive radio networks," *Sensors*, vol. 18, no. 11, p. 3913, 2018.

[145] E. Marchi, F. Vesperini, F. Eyben, S. Squartini, and B. Schuller, "A novel approach for automatic acoustic novelty detection using a denoising autoencoder with bidirectional lstm neural networks," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 1996–2000.

[146] S. S. Khan and B. Taati, "Detecting unseen falls from wearable devices using channel-wise ensemble of autoencoders," *Expert Systems with Applications*, vol. 87, pp. 280–290, 2017.

[147] Y. Shi, K. Davaslioglu, Y. E. Sagduyu, W. C. Headley, M. Fowler, and G. Green, "Deep learning for rf signal classification in unknown and dynamic spectrum environments," in *2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE, 2019, pp. 1–10.

[148] 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 *2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE, 2019, pp. 1–10.

[149] Y. Kim, S. An, and J. So, "Identifying signal source using channel state information in wireless lans," in *2018 International Conference on Information Networking (ICOIN)*. IEEE, 2018, pp. 616–621.

[150] L. J. Wong, W. C. Headley, S. Andrews, R. M. Gerdes, and A. J. Michaels, "Clustering learned cnn features from raw iq data for emitter identification," in *MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM)*. IEEE, 2018, pp. 26–33.

[151] H. Jafari, O. Omotere, D. Adesina, H.-H. Wu, and L. Qian, "Iot devices fingerprinting using deep learning," in *MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM)*. IEEE, 2018, pp. 1–9.

[152] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Advances in neural information processing systems*, 2011, pp. 2546–2554.

[153] M. Pelikan, D. E. Goldberg, E. Cantú-Paz *et al.*, "Boa: The bayesian optimization algorithm," in *Proceedings of the genetic and evolutionary computation conference GECCO-99*, vol. 1, 1999, pp. 525–532.

[154] S. R. Young, D. C. Rose, T. P. Karnowski, S.-H. Lim, and R. M. Patton, "Optimizing deep learning hyper-parameters through an evolutionary algorithm," in *Proceedings of the Workshop on Machine Learning in High-Performance Computing Environments*, 2015, pp. 1–5.

[155] T. Elsken, J. H. Metzen, and F. Hutter, "Neural architecture search: A survey," *arXiv preprint arXiv:1808.05377*, 2018.

[156] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, 2017.

[157] A. Truong, A. Walters, J. Goodfellow, K. Hines, B. Bruss, and R. Farivar, "Towards automated machine learning: Evaluation and comparison of automl approaches and tools," *arXiv preprint arXiv:1908.05557*, 2019.

[158] M. Lindauer, "Literature on neural architecture search," <https://www.automl.org/automl/literature-on-neural-architecture-search/>, Feb. 2020.

[159] X. Dong, "Awesome-nas: A curated list of neural architecture search (nas) resources," <https://github.com/D-X-Y/Awesome-NAS>, Jan. 2020.

[160] C. Zanchettin and T. B. Ludermir, "A methodology to train and improve artificial neural networks' weights and connections," in *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. IEEE, 2006, pp. 5267–5274.

[161] M. M. Islam, M. A. Sattar, M. F. Amin, X. Yao, and K. Murase, "A new adaptive merging and growing algorithm for designing artificial neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 3, pp. 705–722, 2009.

[162] H. Lam, S. Ling, F. H. Leung, and P. K.-S. Tam, "Tuning of the structure and parameters of neural network using an improved genetic algorithm," in *IECON'01. 27th Annual Conference of the IEEE Industrial Electronics Society (Cat. No. 37243)*, vol. 1. IEEE, 2001, pp. 25–30.

[163] J. M. Alvarez and M. Salzmann, "Learning the number of neurons in deep networks," in *Advances in Neural Information Processing Systems*, 2016, pp. 2270–2278.

[164] Y. Guo, A. Yao, and Y. Chen, "Dynamic network surgery for efficient dnns," in *Advances in neural information processing systems*, 2016, pp. 1379–1387.

[165] S. Han, J. Pool, J. Tran, and W. Dally, "Learning both weights and connections for efficient neural network," in *Advances in neural information processing systems*, 2015, pp. 1135–1143.

[166] M. M. Islam, X. Yao, and K. Murase, "A constructive algorithm for training cooperative neural network ensembles," *IEEE Transactions on neural networks*, vol. 14, no. 4, pp. 820–834, 2003.

[167] P. L. Narasimha, W. H. Delashmit, M. T. Manry, J. Li, and F. Maldonado, "An integrated growing-pruning method for feedforward network training," *Neurocomputing*, vol. 71, no. 13–15, pp. 2831–2847, 2008.

[168] L. Ma and K. Khorasani, "A new strategy for adaptively constructing multilayer feedforward neural networks," *Neurocomputing*, vol. 51, pp. 361–385, 2003.

[169] C. Cortes, X. Gonzalvo, V. Kuznetsov, M. Mohri, and S. Yang, "Adanet: Adaptive structural learning of artificial neural networks," in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR.org, 2017, pp. 874–883.

[170] C. Liu, B. Zoph, M. Neumann, J. Shlens, W. Hua, L.-J. Li, L. Fei-Fei, A. Yuille, J. Huang, and K. Murphy, "Progressive neural architecture search," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 19–34.

[171] K. O. Stanley and R. Miikkulainen, "Efficient evolution of neural network topologies," in *Proceedings of the 2002 Congress on Evolutionary Computation CEC'02 (Cat. No. 02TH8600)*, vol. 2. IEEE, 2002, pp. 1757–1762.

[172] ———, "Evolving neural networks through augmenting topologies," *Evolutionary computation*, vol. 10, no. 2, pp. 99–127, 2002.

[173] J. Liang, E. Meyerson, B. Hodjat, D. Fink, K. Mutch, and R. Miikkulainen, "Evolutionary neural automl for deep learning," in *Proceedings of the Genetic and Evolutionary Computation Conference*, 2019, pp. 401–409.

[174] B. Baker, O. Gupta, N. Naik, and R. Raskar, "Designing neural network architectures using reinforcement learning," *arXiv preprint arXiv:1611.02167*, 2016.

[175] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, "Mnasnet: Platform-aware neural architecture search for mobile," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2820–2828.

[176] C.-H. Hsu, S.-H. Chang, J.-H. Liang, H.-P. Chou, C.-H. Liu, S.-C. Chang, J.-Y. Pan, Y.-T. Chen, W. Wei, and D.-C. Juan, "Monas: Multi-objective neural architecture search using reinforcement learning," *arXiv preprint arXiv:1806.10332*, 2018.

[177] H. Liu, K. Simonyan, and Y. Yang, "Darts: Differentiable architecture search," *arXiv preprint arXiv:1806.09055*, 2018.

[178] X. Chen, L. Xie, J. Wu, and Q. Tian, "Progressive differentiable architecture search: Bridging the depth gap between search and evaluation," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 1294–1303.

[179] X. Dong and Y. Yang, "Searching for a robust neural architecture in four gpu hours," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 1761–1770.

[180] C. Ying, A. Klein, E. Real, E. Christiansen, K. Murphy, and F. Hutter, "Nas-bench-101: Towards reproducible neural architecture search," *arXiv preprint arXiv:1902.09635*, 2019.

[181] X. Dong and Y. Yang, "Nas-bench-102: Extending the scope of reproducible neural architecture search," *arXiv preprint arXiv:2001.00326*, 2020.

[182] K. Youssef, L. Bouchard, K. Haigh, J. Silovsky, B. Thapa, and C. Vander Valk, "Machine learning approach to rf transmitter identification," *IEEE Journal of Radio Frequency Identification*, vol. 2, no. 4, pp. 197–205, 2018.

[183] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, "Toward open set recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 7, pp. 1757–1772, 2012.

[184] A. Bendale and T. E. Boult, "Towards open set deep networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1563–1572.

[185] D. Gunning, "Explainable artificial intelligence (xai)," *Defense Advanced Research Projects Agency (DARPA), nd Web*, vol. 2, 2017.

- [186] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," in *International conference on artificial neural networks*. Springer, 2018, pp. 270–279.
- [187] S. Niu, Y. Liu, J. Wang, and H. Song, "A decade survey of transfer learning (2010–2020)," *IEEE Transactions on Artificial Intelligence*, vol. 1, no. 2, pp. 151–166, 2020.
- [188] G. I. Parisi, R. Kemker, J. L. Part, C. Kanan, and S. Wermter, "Continual lifelong learning with neural networks: A review," *Neural Networks*, 2019.
- [189] H. Shin, J. K. Lee, J. Kim, and J. Kim, "Continual learning with deep generative replay," in *Advances in Neural Information Processing Systems*, 2017, pp. 2990–2999.
- [190] F. Girosi, M. Jones, and T. Poggio, "Regularization theory and neural networks architectures," *Neural computation*, vol. 7, no. 2, pp. 219–269, 1995.
- [191] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska *et al.*, "Overcoming catastrophic forgetting in neural networks," *Proceedings of the national academy of sciences*, vol. 114, no. 13, pp. 3521–3526, 2017.
- [192] J. Yoon, E. Yang, J. Lee, and S. J. Hwang, "Lifelong learning with dynamically expandable networks," *arXiv preprint arXiv:1708.01547*, 2017.
- [193] Y. Liu, J. Wang, J. Li, S. Niu, and H. Song, "Formal Proofs of Orthogonality for Class-Incremental Learning for Wireless Device Identification in IoT," 5 2021. [Online]. Available: https://www.techrxiv.org/articles/preprint/Formal_Proofs_of_Orthogonality_for_Class-Incremental_Learning_for_Wireless_Device_Identification_in_IoT/14559765
- [194] ———, "Formal Proofs of Orthogonality for Class-Incremental Learning for Wireless Device Identification in IoT," 5 2021. [Online]. Available: https://www.techrxiv.org/articles/preprint/Formal_Proofs_of_Orthogonality_for_Class-Incremental_Learning_for_Wireless_Device_Identification_in_IoT/14559765
- [195] ———, "Class-incremental learning for wireless device identification in iot," *IEEE Internet of Things Journal*, pp. 1–1, 2021.
- [196] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE signal processing magazine*, vol. 29, no. 3, pp. 101–116, 2012.
- [197] A. Sivanathan, H. H. Gharakheili, and V. Sivaraman, "Detecting behavioral change of iot devices using clustering-based network traffic modeling," *IEEE Internet of Things Journal*, 2020.
- [198] ———, "Inferring iot device types from network behavior using unsupervised clustering," in *2019 IEEE 44th Conference on Local Computer Networks (LCN)*. IEEE, 2019, pp. 230–233.
- [199] J. Ortiz, C. Crawford, and F. Le, "Devicemien: network device behavior modeling for identifying unknown iot devices," in *Proceedings of the International Conference on Internet of Things Design and Implementation*, 2019, pp. 106–117.
- [200] G. Liu, R. Zhang, C. Wang, and L. Liu, "Synchronization-free gps spoofing detection with crowdsourced air traffic control data," in *2019 20th IEEE International Conference on Mobile Data Management (MDM)*. IEEE, 2019, pp. 260–268.
- [201] M. Monteiro, A. Barreto, T. Kacem, D. Wijesekera, and P. Costa, "Detecting malicious ads-b transmitters using a low-bandwidth sensor network," in *2015 18th International Conference on Information Fusion (Fusion)*. IEEE, 2015, pp. 1696–1701.
- [202] B. Xu, G. Sun, R. Yu, and Z. Yang, "High-accuracy tdoa-based localization without time synchronization," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 8, pp. 1567–1576, 2012.
- [203] M. Schäfer, M. Strohmeier, V. Lenders, I. Martinovic, and M. Wilhelm, "Bringing up opensky: A large-scale ads-b sensor network for research," in *Proceedings of the 13th international symposium on Information processing in sensor networks*. IEEE Press, 2014, pp. 83–94.
- [204] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, 2007.
- [205] I. Guvenc and C.-C. Chong, "A survey on toa based wireless localization and nlos mitigation techniques," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 3, pp. 107–124, 2009.
- [206] G. Han, H. Xu, T. Q. Duong, J. Jiang, and T. Hara, "Localization algorithms of wireless sensor networks: a survey," *Telecommunication Systems*, vol. 52, no. 4, pp. 2419–2436, 2013.
- [207] Q. Li, H. Fan, W. Sun, J. Li, L. Chen, and Z. Liu, "Fingerprints in the air: unique identification of wireless devices using rf rss fingerprints," *IEEE Sensors Journal*, vol. 17, no. 11, pp. 3568–3579, 2017.
- [208] M. Zheleva, R. Chandra, A. Chowdhery, A. Kapoor, and P. Garnett, "Txminer: Identifying transmitters in real-world spectrum measurements," in *2015 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE, 2015, pp. 94–105.
- [209] Y. A. Nijssure, G. Kaddoum, G. Gagnon, F. Gagnon, C. Yuen, and R. Mahapatra, "Adaptive air-to-ground secure communication system based on ads-b and wide-area multilateration," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3150–3165, 2015.
- [210] M. Monteiro, A. Barreto, T. Kacem, J. Carvalho, D. Wijesekera, and P. Costa, "Detecting malicious ads-b broadcasts using wide area multilateration," in *2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC)*. IEEE, 2015, pp. 4A3–1.
- [211] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "Dbscan revisited, revisited: why and how you should (still) use dbscan," *ACM Transactions on Database Systems (TODS)*, vol. 42, no. 3, pp. 1–21, 2017.
- [212] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "Optics: ordering points to identify the clustering structure," *ACM Sigmod record*, vol. 28, no. 2, pp. 49–60, 1999.
- [213] D. Pelleg, A. W. Moore *et al.*, "X-means: Extending k-means with efficient estimation of the number of clusters," in *Icm*, vol. 1, 2000, pp. 727–734.
- [214] R.-H. Hwang, M.-C. Peng, C.-W. Huang, P.-C. Lin, and V.-L. Nguyen, "An unsupervised deep learning model for early network traffic anomaly detection," *IEEE Access*, vol. 8, pp. 30387–30399, 2020.
- [215] R.-H. Hwang, M.-C. Peng, and C.-W. Huang, "Detecting iot malicious traffic based on autoencoder and convolutional neural network," in *2019 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2019, pp. 1–6.
- [216] H. Alipour, Y. B. Al-Nashif, P. Satam, and S. Hariri, "Wireless anomaly detection based on ieee 802.11 behavior analysis," *IEEE transactions on information forensics and security*, vol. 10, no. 10, pp. 2158–2170, 2015.
- [217] N. Sehatbakhsh, M. Alam, A. Nazari, A. Zajic, and M. Prvulovic, "Syndrome: Spectral analysis for anomaly detection on medical iot and embedded devices," in *2018 IEEE international symposium on hardware oriented security and trust (HOST)*. IEEE, 2018, pp. 1–8.
- [218] A. Hamza, H. H. Gharakheili, T. A. Benson, and V. Sivaraman, "Detecting volumetric attacks on iot devices via sdn-based monitoring of mud activity," in *Proceedings of the 2019 ACM Symposium on SDN Research*, 2019, pp. 36–48.
- [219] V. L. Thing, "Ieee 802.11 network anomaly detection and attack classification: A deep learning approach," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2017, pp. 1–6.
- [220] R. Sridharan, R. R. Maiti, and N. O. Tippenhauer, "Wadac: Privacy-preserving anomaly detection and attack classification on wireless traffic," in *Proceedings of the 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, 2018, pp. 51–62.
- [221] T. Luo and S. G. Nagarajan, "Distributed anomaly detection using autoencoder neural networks in wsn for iot," in *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, pp. 1–6.
- [222] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, A. Shabtai, D. Breitenbacher, and Y. Elovici, "N-baito—network-based detection of iot botnet attacks using deep autoencoders," *IEEE Pervasive Computing*, vol. 17, no. 3, pp. 12–22, 2018.
- [223] Y. Chen, X. S. Zhou, and T. S. Huang, "One-class svm for learning in image retrieval," in *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, vol. 1. IEEE, 2001, pp. 34–37.
- [224] Z. Ding and M. Fei, "An anomaly detection approach based on isolation forest algorithm for streaming data using sliding window," *IFAC Proceedings Volumes*, vol. 46, no. 20, pp. 12–17, 2013.
- [225] H.-P. Kriegel, P. Kröger, E. Schubert, and A. Zimek, "Loop: local outlier probabilities," in *Proceedings of the 18th ACM conference on Information and knowledge management*, 2009, pp. 1649–1652.
- [226] M. S. Parvez, D. B. Rawat, and M. Garuba, "Big data analytics for user-activity analysis and user-anomaly detection in mobile wireless network," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 2058–2065, 2017.
- [227] M. Wei and K. Kim, "Intrusion detection scheme using traffic prediction for wireless industrial networks," *Journal of communications and networks*, vol. 14, no. 3, pp. 310–318, 2012.
- [228] N. Tandiya, A. Jauhar, V. Marojevic, and J. H. Reed, "Deep predictive coding neural network for rf anomaly detection in wireless networks,"

in 2018 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, 2018, pp. 1–6.

[229] S. Rajendran, W. Meert, V. Lenders, and S. Pollin, “Saife: Unsupervised wireless spectrum anomaly detection with interpretable features,” in 2018 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN). IEEE, 2018, pp. 1–9.

[230] W. Lotter, G. Kreiman, and D. Cox, “Deep predictive coding networks for video prediction and unsupervised learning,” *arXiv preprint arXiv:1605.08104*, 2016.

[231] R. Morales-Ferre, W. Wang, A. Sanz-Abia, and E.-S. Lohan, “Identifying gnss signals based on their radio frequency (rf) features? a dataset with gnss raw signals based on roof antennas and spectracom generator,” *Data*, vol. 5, no. 1, 2020. [Online]. Available: <https://www.mdpi.com/2306-5729/5/1/18>

[232] Y. Liu, J. Wang, S. Niu, and H. Song, “Ads-b signals records for non-cryptographic identification and incremental learning.” 2021. [Online]. Available: <https://dx.doi.org/10.21227/1bxc-ke87>

[233] M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc, “Drone remote controller rf signal dataset,” 2020. [Online]. Available: <https://dx.doi.org/10.21227/ss99-8d56>

[234] E. Uzundurukan, Y. Dalveren, and A. Kara, “A database for the radio frequency fingerprinting of bluetooth devices,” *Data*, vol. 5, no. 2, p. 55, 2020.

[235] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, “Oracle: Optimized radio classification through convolutional neural networks,” in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 370–378.

[236] S. Mohanti, N. Soltani, K. Sankhe, D. Jaisinghani, M. Di Felice, and K. Chowdhury, “Airid: Injecting a custom rf fingerprint for enhanced uav identification using deep learning,” in *IEEE GLOBECOM 2020-IEEE Global Communications Conference*. IEEE, 2020, pp. 370–378.

[237] G. Reus-Muns, D. Jaisinghani, K. Sankhe, and K. Chowdhury, “Trust in 5g open rans through machine learning: Rf fingerprinting on the powder pawr platform,” in *IEEE Globecom 2020-IEEE Global Communications Conference*. IEEE, 2020.

[238] N. Soltani, G. Reus-Muns, B. Salehi, J. Dy, S. Ioannidis, and K. Chowdhury, “RF Fingerprinting Unmanned Aerial Vehicles With Non-Standard Transmitter Waveforms,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 15 518–15 531, 2020.

[239] K. Sankhe, M. Belgiovine, F. Zhou, L. Angioloni, F. Restuccia, S. D’Oro, T. Melodia, S. Ioannidis, and K. Chowdhury, “No radio left behind: Radio fingerprinting through deep learning of physical-layer hardware impairments,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 165–178, 2019.

[240] N. Soltani, K. Sankhe, J. Dy, S. Ioannidis, and K. Chowdhury, “More is Better: Data Augmentation for Channel-Resilient RF Fingerprinting,” *IEEE Communications Magazine*, vol. 58, no. 10, pp. 66–72, 2020.

[241] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[242] S. Niu, J. Wang, Y. Liu, and H. Song, “Transfer learning based data-efficient machine learning enabled classification,” in *2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*. IEEE, 2020, pp. 620–626.

[243] P. Nguyen, H. Truong, M. Ravindranathan, A. Nguyen, R. Han, and T. Vu, “Matthan: Drone presence detection by identifying physical signatures in the drone’s rf communication,” in *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2017, pp. 211–224.

[244] J. Wang, Y. Liu, and H. Song, “Counter-unmanned aircraft system(s) (c-uas): State of the art, challenges, and future trends,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 36, no. 3, pp. 4–29, 2021.

[245] Y. Shi, K. Davaslioglu, and Y. E. Sagduyu, “Generative adversarial network for wireless signal spoofing,” in *Proceedings of the ACM Workshop on Wireless Security and Machine Learning*. ACM, 2019, pp. 55–60.



YongXin Liu (LIU11@my.erau.edu) is currently an assistant professor at department of Computer Science, Auburn University at Montgomery. He received his second Ph.D. from the Department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL in June 2021. He received the first Ph.D. from South China University of Technology, 2018. His major research interests include machine learning, data mining, wireless networks, the Internet of Things, and Unmanned Systems.



Jian Wang (wangj14@my.erau.edu) is a Ph.D. student in the Department of Electrical, Computer, Software, and Systems Engineering (ECSSE), Embry-Riddle Aeronautical University (ERAU), Daytona Beach, Florida, and a graduate research assistant in the Security and Optimization for Networked Globe Laboratory (SONG Lab, www.SONGLab.us). He received his M.S. from South China Agricultural University (SCAU) in 2017 and B.S. from Nanyang Normal University in 2014. His major research interests include wireless networks, unmanned aerial systems, and machine learning.



Jianqiang Li (lijq@szu.edu.cn) received his B.S. and Ph.D. degrees from the South China University of Technology in 2003 and 2008, respectively. He is a Professor with the College of Computer and Software Engineering, Shenzhen University, Shenzhen, China. He is leading two projects funded by the National Natural Science Foundation of China and two projects funded by the Natural Science Foundation of Guangdong, China. His major research interests include Internet of Things, robotic, hybrid systems, and embedded systems.



Shuteng Niu (shutengn@my.erau.edu) is a Ph.D. student in the Department of Electrical, Computer, Software, and Systems Engineering (ECSSE), Embry-Riddle Aeronautical University (ERAU), Daytona Beach, Florida, and a graduate research assistant in the Security and Optimization for Networked Globe Laboratory (SONG Lab, www.SONGLab.us). He received his M.S. from Embry-Riddle Aeronautical University (ERAU) in 2018 and B.S. from Civil Aviation University of China (CAUC) in 2015. His major research interests include machine learning, data mining, and signal processing.



Houbing Song (M'12-SM'14) received the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, VA, in August 2012.

In August 2017, he joined the Department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL, where he is currently an Assistant Professor and the Director of the Security and Optimization for Networked Globe Laboratory (SONG Lab, www.SONGLab.us). He has served as an Associate Technical Editor for IEEE Communications Magazine (2017-present), an Associate Editor for IEEE Internet of Things Journal (2020-present) and IEEE Journal on Miniaturization for Air and Space Systems (J-MASS) (2020-present), and a Guest Editor for IEEE Journal on Selected Areas in Communications (J-SAC), IEEE Internet of Things Journal, IEEE Transactions on Industrial Informatics, IEEE Sensors Journal, IEEE Transactions on Intelligent Transportation Systems, and IEEE Network. He is the editor of six books, including Big Data Analytics for Cyber-Physical Systems: Machine Learning for the Internet of Things, Elsevier, 2019, Smart Cities: Foundations, Principles and Applications, Hoboken, NJ: Wiley, 2017, Security and Privacy in Cyber-Physical Systems: Foundations, Principles and Applications, Chichester, UK: Wiley-IEEE Press, 2017, Cyber-Physical Systems: Foundations, Principles and Applications, Boston, MA: Academic Press, 2016, and Industrial Internet of Things: Cybermanufacturing Systems, Cham, Switzerland: Springer, 2016. He is the author of more than 100 articles. His research interests include cyber-physical systems, cybersecurity and privacy, internet of things, edge computing, AI/machine learning, big data analytics, unmanned aircraft systems, connected vehicle, smart and connected health, and wireless communications and networking. His research has been featured by popular news media outlets, including IEEE GlobalSpec's Engineering360, USA Today, U.S. News & World Report, Fox News, Association for Unmanned Vehicle Systems International (AUVS), Forbes, WFTV, and New Atlas.

Dr. Song is a senior member of ACM and an ACM Distinguished Speaker. Dr. Song was a recipient of the Best Paper Award from the 12th IEEE International Conference on Cyber, Physical and Social Computing (CPSCom-2019), the Best Paper Award from the 2nd IEEE International Conference on Industrial Internet (ICII 2019), the Best Paper Award from the 19th Integrated Communication, Navigation and Surveillance technologies (ICNS 2019) Conference, the Best Paper Award from the 6th IEEE International Conference on Cloud and Big Data Computing (CBDC 2020), and the Best Paper Award from the 15th International Conference on Wireless Algorithms, Systems, and Applications (WASA 2020).