Device Fingerprinting for Cyber-Physical Systems: A Survey

VIJAY KUMAR and KOLIN PAUL, IIT Delhi

The continued growth of the cyber-physical system (CPS) and Internet of Things technologies raises device security and monitoring concerns. For device identification, authentication, conditioning, and security, device fingerprint/fingerprinting (DFP) is increasingly used. However, finding the correct DFP features and sources to establish a unique and stable fingerprint is challenging. We present a state-of-the-art survey of DFP techniques for CPS device applications. We investigate the numerous DFP features, their origins, characteristics, and applications. Additionally, we discuss the DFP characteristics and their sources in detail, taking into account the physical contexts of various entities (i.e., machines, sensors, networks, and computational devices), as well as their software and applications for the CPS. We believe that this article will provide researchers and developers with insights into the DFP and its applications, sources, aggregation methods, and factors affecting its use in CPS domains.

CCS Concepts: • **Computer systems organization** → **Sensor networks**; **Embedded systems**; **Maintainability and maintenance**; • **Hardware** → Sensor applications and deployments; Networking hardware; Wireless integrated network sensors; • **Security and privacy** → **Distributed systems security**; **Tamper-proof and tamper-resistant designs**; *Authorization*; • **Networks** → *Sensor networks*; • **Applied computing** → Surveillance mechanisms;

Additional Key Words and Phrases: Device Fingerprint/Fingerprinting (DFP), device characteristic, identification and authentication, Internet of Things (IoT), security and privacy, distributed systems, condition monitoring, Cyber Physical System (CPS)

ACM Reference format:
Vijay Kumar and Kolin Paul. 2023. Device Fingerprinting for Cyber-Physical Systems: A Survey. *ACM Comput. Surv.* 55, 14s, Article 302 (July 2023), 41 pages.
https://doi.org/10.1145/3584944

1 INTRODUCTION

A fingerprint is a trace of the information left by someone. It is a well-known term in various biometric applications such as personal identification, access control, and security. When computers, networks, and security are discussed, it is referred to as a digital fingerprint, browser fingerprint, **Device Fingerprint/Fingerprinting (DFP)**, or machine fingerprint [54, 56, 81, 104, 188, 206]. Digital fingerprints are digital traces of information carried by computer network devices that can be used to uniquely identify devices based on unique patterns of features generated by remote devices.

Authors' address: V. Kumar, Amar Nath and Shashi Khosla School of Information Technology, Indian Institute of Technology Delhi, New Delhi,110016, India; email: vijay.kumar@sit.iitd.ac.in; K. Paul, Department of Computer Science and Engineering, Indian Institute of Technology Delhi, New Delhi, 110016, India; email: kolin@cse.iitd.ac.in.

0360-0300/2023/07-ART302 $15.00
https://doi.org/10.1145/3584944

With the advancement of technologies in embedded systems, communication protocols, storage, data analysis, sensors, and actuators, the physical processes/machines can interact through cyber systems. The new form of such systems is called the **Internet of Things (IoT)**. Today, IoT has several key applications, such as in automobiles and aviation, manufacturing, healthcare, transportation, energy, civil infrastructure, and consumer electronics [46, 81, 109]. IoT is often referred to as the **Cyber-Physical System (CPS),** as it integrates the cyber system with physical devices and systems [59]. In the CPS, control and monitoring of physical entities (system, process, or machine) is done via the cyber system using sensors and actuators [35, 125, 130, 204]. With the advancement of CPSs, more and more devices have been connected day by day in the past few years. As a result, the global market for IoT end user solutions was valued at $100 billion in 2017 and is expected to grow to $1.6 trillion by 2025 [82]. Hence, the importance of the DFP has increased many fold and can be used in a variety of CPS applications, including device identification [83, 189, 199], forensic applications [13, 99, 104], condition monitoring [20, 93, 123, 177], and security [14, 74, 89, 115, 159, 164], as well as fault analysis and diagnosis [107, 127, 142, 192]. Therefore, CPS applications require a detailed structured study and evaluation of different DFP techniques.

1.1 DFP and Identification

Significant increase in the frequent attacks on the CPS and IoT system have been noted, and recent examples are Turkish oil pipeline attacks (in 2008) [23], Ukraine power supply utility attacks (in 2015) [60], and the Mirai botnet attack (in 2016) [92]. The most significant concern with these attacks is the threat to the CPS, which can directly or indirectly impact the physical world. Additionally, the device is installed in a remote location with minimal physical security, poor computation resources, and restricted power. Attacks on these interconnected network nodes can have a cascading effect on the entire system. For example, an attacker may interrupt communication between the sensor and the control unit, impairing the physical system's performance. In 2010, attackers in Iran exploited the memory and computational limitations of the programmable logic controller to perform attacks and destroy the whole system [159]. Detecting such attacks using conventional networking techniques is not a simple task.

Therefore, it requires a security solution to ensure security at every step of the CPS network's system. One approach to providing such a security solution is described in the work of Yadav et al. [207], which provides an end-to-end automatic penetration testing framework for CPS networks to identify all possible attack vectors using target graphs. Unfortunately, providing complete security assurance in a real-world system is difficult, particularly in the CPS, which consist of numerous distributed computational and physical systems interconnected by a computer network. Detecting and monitoring threats and vulnerabilities in this vast and diverse environment are complex tasks. The existing cryptographic-based technique is also inefficient for the CPS due to the high computational demands and large amount of memory required. These methods are insufficient to secure the system from real-world exploits such as denial-of-service (DoS), man-in-the-middle, node tampering, node jamming, malicious system injection, node cloning, and security breaches [46, 81]. A CPS is a distributed network that supports various communication protocols, such as LAN, WAN, CAN, WiFi, BT, 3G/4G/5G, GSM, LoRA, and ZigBee, with multiple network architectures for wired or wireless networks, such as mesh networks, cognitive radio networks, wireless sensor networks, and small cell networks [48, 206]. With this diversity and complexity, the CPS network becomes highly vulnerable to fake and insider intruders, as malicious devices or users may get security credentials from legitimate users.

One of the most effective solutions to deal with the security threat mentioned previously is DFP. In **Information and Communication Technology (ICT)** applications like social networking

and personalized marketing, DFPs are often used to track and recognize users and their gadgets. Specially, in ICT, DFPs can represent unique trace of specific device and network attributes such as the operating system, firmware, network configuration, time zone, type and version of applications, and information from various remote system devices or applications. Development of the CPS has expanded the scope of DFP such that it now includes physical devices and networking equipment, making use of the more comprehensive CPS features. Consequently, in a CPS, DFP is a unique information pattern generated by the features of both cyber and physical system entities [54, 56, 81, 206]. The heterogeneity and complexity of the CPS make DFP-based identification challenging. In recent years, several authors have used DFP techniques that leverage the CPS's unique physics, electrical, operational, and behavioral characteristics [54, 56, 81, 188].

1.2 DFP Applications

DFP-based systems can find their application in various types of ICT and CPSs to enhance device control management and security. In this section, we specifically discuss DFP applications for CPSs.

Condition Monitoring. Condition monitoring is a technique that continuously monitors the health of machinery, equipment, and applications to predict critical events such as device malfunctions, failures, or glitches. The process of condition monitoring is closely related to the DFP methods because they utilize the unique characteristics of the device, machine, or system for any given application. This similarity between condition monitoring and DFP enables direct inference of the device's condition from fingerprint data. In recent years, the advancements in sensing and machine monitoring technology have also promoted the development of condition monitoring technologies such as vibration analysis and diagnostics, lubricant analysis, acoustic emission, Infrared (IR) imaging, ultrasound, motor position monitoring and motor current signature analysis (MCSA), and the model-based voltage and current system (MBVI system) [145, 174, 195]. Condition monitoring is currently one of the essential components of the **Industrial IoT (Industry 4.0)**, as it enables people to keep an eye on their equipment and machines [20, 93, 123, 177]. Condition monitoring is often used in industrial settings for machinery, auxiliary systems, and other equipment like electric motors, pumps, air compressors, and furnaces, where it is important to check for possible faults or breakdowns ahead of time [142]. Applications for monitoring systems are not limited to the industrial sector; they can also be used to monitor the health, behavior, and performance of household appliances, infrastructure, heavy machinery, sophisticated equipment, and software [107, 147].

Several authors [93, 145, 174] proposed DFP in industrial activities like predictive maintenance of pumps and motors on the factory floor using DFP. Supervisory systems on the factory floor use DFP to monitor the health of machines and predict their failure. This allows supervisory systems to build machine profiles and provide personalized and targeted services to machines prior to a breakdown. Likewise, DFP can be used to diagnose and predict the health conditions and security breaches of home appliances in the CPS.

Tracing and Forensic Applications. In forensic applications, it is standard practice to apply DFP techniques. To conduct a logical investigation of a crime scene, security organizations and forensic specialists use fingerprinting procedures. The use of fingerprinting analysis processes is totally dependent on the nature of the crime and the available evidence. An investigation of a crime focuses mostly on the persons present at the crime site. Electronic devices such as mobile phones, tablets, laptops, smart speakers, TVs with built-in Internet access, smart locks, and smart lighting are omnipresent. Therefore, one may utilize the temporary location information given by a GPS-enabled device to locate the crime scene [104]. Similarly, remote human-machine interfacedevices in an industrial setting may be examined, managed, and identified for tracing and forensic applications

using cookies and canvas data in a web application. Kohno et al. [99] monitored remote devices at specific public access points using minor hardware variations, such as clock skew. CPSs use sensors like a camera, microphone, GPS, temperature, humidity, and so on to look at a crime scene and figure out who did it.

Security. Additionally, fingerprinting technology may categorize devices, networks, and platforms according to their security or vulnerability [13]. Detecting and generating DFPs is crucial for identifying vulnerabilities, enhancing cyber security, identifying malicious behavior, detecting counterfeit components, and determining physical tampering. It is also possible to classify devices based on their potential security threats. Some important uses of DFP security are listed next:

- *Device identification*: Users, suppliers, and vendors are continuously at risk from counterfeit devices and software. Users will always be concerned about its dependability, security, and trustworthiness due to the fact that it impacts the business and reputation of both the supplier and the manufacturer. Therefore, there is a need for an identification system capable of accurately and unambiguously identifying any given device. However, hard-coded identifiers (i.e., unique identification numbers (UIDs) and serial numbers) are insecure because they are simple to clone and alter. Some authors [83, 189, 199] use DFP for device identification, which is unique, strong, and hard to change. Gu et al. [60] used device physics based features such as programmable logic controller response time, device dynamics, and open or closed latching relays of different operation types to classify the device type.
- *Physical unclonable function*: A **Physical Unclonable Function (PUF)** is a cryptographic entity that authenticates devices using a challenge-response pair [199]. PUF methods exploit the device's stable and unique feature set for a secret key generation as a DFP. For example, in BoardPUF, the author used the **Printed Circuit Board (PCB)** fabrication variance effect to make volatile secret keys [199]. Additionally, clock skew, delay, the received signal strength indicator, web plugin information, system cookies, canvas data, and device physics have been used in the PUF system by several authors [12].
- *Identifying threats*: Devices with low computation, memory, and power capacities are more vulnerable to compromise. Additionally, the CPS is a significant target for malicious activities because of its various legacies and scattered devices. Administrators of **Industrial Control Systems (ICSs)** or security patches must be able to identify vulnerable devices in advance of an attack (cyber or physical) so that they may be isolated from the network. Behavioral fingerprints and other DFP-based methodologies may be used to anticipate malicious activity in compromised networks and devices [115, 164]. It generates anomalies and detects intrusions by use of fingerprinting and data-generated methods [14, 74, 89]. This includes information on network traffic [60, 198, 205], system logs [208], event logs [30], and device physics [60, 202] that enables the development of autonomous threat detection and prediction of potential attacks on physical components of a CPS [4, 175].

1.3 Problem Definition

DFP-based applications have grown rapidly in CPS domains. Some of these applications are related to forensics, health conditioning, and security. Specifically, in Industry 4.0, the CPS is an essential element of the cyber and physical worlds for monitoring and managing the physical environment in the industry. The CPS is used to achieve optimal operational efficiency, productivity, and enhanced industrial asset and equipment maintenance management via the application of DFP methods. Additionally, DFP can also be used to monitor and control the physical environment in areas such as water supply, smart grid, healthcare, intelligent transportation, process control monitoring, aviation, and defense. As a result, research communities have also given special attention

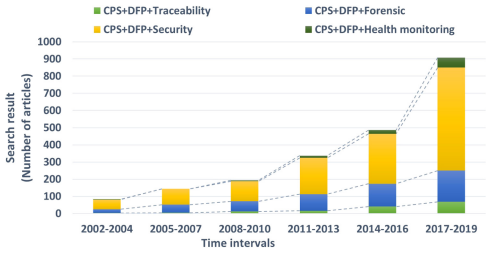

Fig. 1. Research trends on DFP.

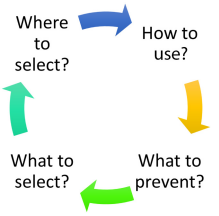

Fig. 2. 4-Q procedure for DFP implementation.

in this area to ensure DFP-based development to meet the requirements. Figure 1 shows the total number of results provided by Google Scholar for the search queries "CPS + fingerprinting + traceability," "CPS+ fingerprinting + forensic," "CPS+ fingerprinting + security," and "CPS+ fingerprinting + health monitoring" during the past two decades. Between 2002 and 2004, fewer than 100 articles on DFP applications such as traceability, forensics, security, and health monitoring were published. The statistics from the subsequent years, however, show that interest in this field of study has exponentially increased.

The DFP technique generates a unique fingerprint by analyzing features of cyber and physical system components. Table 1 illustrates the specific characteristics that are considered, including information related to communication networks, physical devices, storage, embedded systems, software, and applications. However, determination of the appropriate devices' characteristics and their sources is a time-consuming job, particularly in the complex scenario of the CPS. Therefore, we need a state-of-the-art study on DFP to provide insights to researchers and developers to understand CPS devices' features, extracting procedures, and potential applications. This survey mainly aims to answer four fundamental queries related to the design and development of DFP applications in the CPS depicted in Figure 2. The first query is "What to select?" and it is addressed by explaining the unique characteristics of the device in detail. The second query, "Where to select?", is answered by describing the device feature sources and methods for collection. The third query, "How to use?", provides a complete view of DFP applications and explains the interconnection between DFP features, their sources, and uses. Finally, the fourth query, "What to prevent?", considers the difficulties in the implementation of DFPs in a CPS.

1.4 Related Surveys

Table 2 summarizes some of the most cited and recent surveys, which also briefs our contribution toward this survey and shows how this work differs from the existing survey work. The referred surveys in this table include only the cyber system's entities (i.e., the application and web browser, computer network, and human-machine interfaces or sensors) in their consideration. However, one of the crucial entities of the CPS—the physical system—is not included. In recent years, device physics based fingerprinting techniques have grown their importance significantly. For example, in smart manufacturing, machines are being monitored by their operational data as well as their physics-based information. As a result, whenever any spike, abnormalities, or malicious pattern is visible in their physical features, the device can be categorized as a malicious or fake or failed. Specially, for the CPS, the physics-based features are required so that the physical device can communicate with the cyber system.

1.5 Outline of the Survey

This survey provides a near-complete and up-to-date assessment of DFP and related technologies for CPSs. As shown later in Figure 4, the colored portions of the DFP-generating system comprise

Table 1. DFP Techniques in CPS

Method	Features	Source	Gathering Techniques	Applications
Web browser	Operating system, hardware configuration, active plugins, screen resolution, font, user agent, language, time zone, geolocation	Desktop/laptop, smart phone, tablets, or any end device (human interface)	HTTP header, JavaScript, Flash, APIs	Fingerprinting, identification, tracking, hardware configuration
Communication standard (network)	Network packet size, packet direction, packet length frequency, packet ordering, inter-packet interval time, packet count, burst size, traffic data rates, burst size-surge time	Network traffic	Network scanning monitoring, communication protocols, APIs	Localization, identification, tracking, security, forensics, network management
Communication standard (wireless)	Carrier frequency offset, wireless channel information, radio signal strength, carrier frequency difference, phase shift difference	Communication channel and protocols' frame information	Wireless network scanning, communication protocols, APIs	Localization, identification, tracking, security, forensics, network management, detection of radio transmitter
Device physics or local attributes	Refer to Table 4	Device physical (i.e., electrical, thermal, optical, magnetic and physical) characteristics	Machine, computing and network device, sensors	Localization, identification, tracking, condition monitoring
Behavior	Human activity, network activity, network traffic pattern, malicious activity, fault/failure activity, machine/device operational pattern	System logs and operational historical datasets	Device historical data collected by application	Identification, predictive maintenance and monitoring, security and threat prediction

the device's unique features, collection techniques, and applications of the CPS, all of which are discussed in detail in this survey. To the best of our knowledge, for the first time, this survey looks at the physical setting of the CPS to better explain DFP.

We have gathered and accessed several research articles related to CPS devices, their hardware, software, applications, and communication and networking protocols linked with fingerprinting through various resources. These include free online search engines such as Google, Google Scholar, and Research Gate, as well as paid publishers such as IEEE, Elsevier, Science Direct, and ACM for academic and scientific conferences, journals, and research articles. They are classified according to the DFP sources, acquisition techniques, and applications. We have reviewed these articles by their title, abstract, and full text, in case the work provides a novel idea. Then, we correlate various work done in a different dimension of CPS DFP, resulting in corresponding taxonomy. The functions related to each dimension are discussed in the respective sections.

Table 2. Comparison of Surveys Similar to This Work

Survey	Domain			DFP Features		Applications				Survey Focus
	ICT	IoT	CPS	Cyber	Physical	COND	TRACE	ID	Security	
[143]	✓	X	X	NW	X	X	✓	X	X	Network security characteristics
[131]	✓	X	X	WEB	X	X	X	✓	X	Browser extension and HTTP proxies
[188]	✓	X	X	WEB	X	X	X	✓	X	Web-based client hardware information
[206]	✓	X	X	WiFi	X	X	X	✓	X	Consider wireless network features
[78]	✓	X	X	WiFi	X	X	X	✓	X	Indoor positioning using WiFi
[47]	✓	X	X	NW	X	X	X	✓	✓	Operating system fingerprinting using comm. protocols
[211]	✓	X	X	WiFi	X	X	✓	X	X	Human activity recognition
[104]	✓	✓	X	WEB	X	X	X	✓	✓	Browser fingerprinting
[117]	✓	✓	✓	NW	X	X	X	X	X	IoT device security in a huge network
[164]	✓	✓	✓	NW, RF	X	✓	✓	✓	✓	Device behavioral fingerprinting
[86]	✓	✓	✓	RF	X	X	X	✓	✓	Survey on RF fingerprinting and traditional approaches, deep learning, and open challenges
[160]	✓	✓	✓	NW	X	X	X	✓	✓	IoT profiling, fingerprinting, and identification
[33]	✓	✓	X	NW	X	X	X	✓	✓	DFP approaches for resource-constrained IoT devices
This work	✓	✓	✓	NW, WEB, RF	PHY, OPR	✓	✓	✓	✓	General survey on DFP, DFP features, acquisition techniques, applications, and constraints

TRACE, traceability or forensic; COND, condition monitoring; ID, device identification; NW, network; WEB, web browser; RF, radio frequency; PHY, device physics; OPR, device operation.

1.6 Contribution

The main contributions of this study are as follows:

- We explore DFP concepts in the CPS, their properties, quality evaluation, and matching metrics to differentiate between distinct DFPs.
- We provide a comprehensive analysis of DFP for the CPS to understand better the various issues related to appropriate devices' feature selection, acquisition, and usage.
- We study how the external and internal properties of distinct CPS entities are exploited for DFP generation while they are linked with different CPS applications. Furthermore, we investigate their origins and aggregation techniques and their merits and weaknesses.
- We also explore the potential application areas of DFP for CPSs. Furthermore, we study how the different features can be applied to several CPS applications.

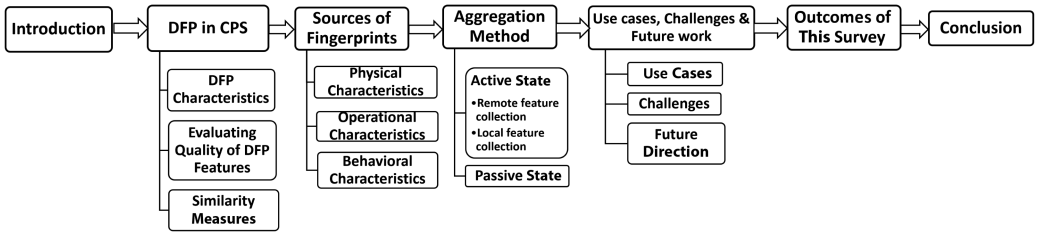

Fig. 3. Taxonomy of survey.

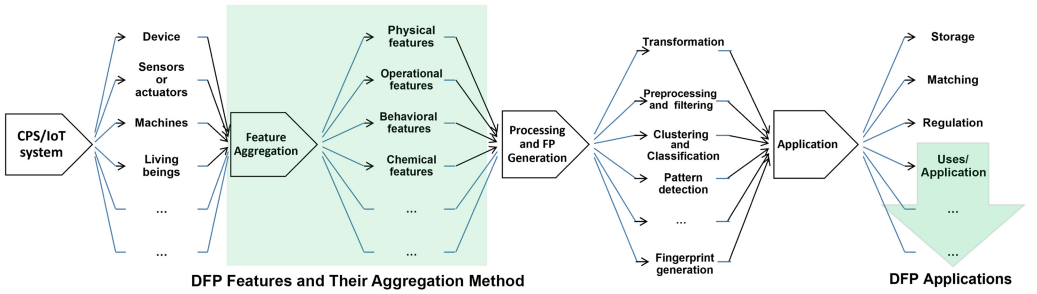

Fig. 4. Block diagram of a typical fingerprint generator.

- Finally, this article discusses various challenges and open issues caused due to the deployment and evolution of the DFP system. We also discuss potential solutions for addressing challenges and problems that have yet to be resolved. Moreover, we also propose a DFP feature selection and tabular collection strategy that would be useful while designing a CPS application.

1.7 Taxonomy

We propose a taxonomy for DFP in the CPS, shown in Figure 3. The CPS-specific DFP generation system consists of the CPS, feature aggregation, processing, DFP generation, and application (see Figure 4). Then, we discuss the CPS-specific DFP characteristics and matching metrics, as explained in Section 2. Analysis of DFP sources is necessary to understand the use of a DFP system in CPS applications. We describe different sources of device features required for the DFP and study DFP origins based on the three-axis taxonomy (see Figure 6). The DFP features are classified based on device physics, operation, and behavior. Furthermore, we analyze the device characteristics used directly or indirectly for the DFP. Section 3 illustrates several DFP methods and features based on taxonomy.

Device feature aggregations are used to extract device features from the CPS to generate and apply fingerprints. These attributes are classified according to the mechanism by which the systems acquire them, including local and remote monitoring. We further explain the acquisition form based on the active and passive interaction with the CPS. This study helps to associate the device features, their sources, and collection methods. A detailed study of device feature aggregations is done in Section 4. Additionally, we highlight various factors influencing the growth of DFP-based applications in the CPS in Section 5. In Section 6, we present a summary of the findings of the survey as well as a feature selection chart for implementing DFPs in CPS environments. Finally, we conclude the survey in Section 7.

2 DFP IN THE CPS

Fingerprinting is one of the most significant biometrics for user identification, mainly due to its persistence and uniqueness. The U.S. government extended the same concept to device identification in the 1960s to identify and track unique mobile transmitters [40]. Moreover, this approach is popular in forensic applications and device debugging. The DFP unique pattern used for device identification is collected and created from their unique features or signatures with the objective of device identification [54, 81, 206]. For example, we consider a DFP-based health monitoring system that is fitted on a water-lifting unit in public water distribution systems. For this, the DFP system utilizes operational, physical, and surrounding information of installed equipment and machinery (the water pump, control unit, and water valve) at the water-lifting unit. The DFP system performs different data transformations, modifying the format, structure, or values of raw data to provide a format that is clean, verified, and usable [98]. In the preprocessing and filtering stage, the DFP system processes the data to remove redundant or noisy information to optimize computational and storage resources [98, 191]. The DFP system utilizes various algorithms for information processing, clustering [191, 196], pattern recognition [181], and classification techniques [213] to identify a consistent and unique pattern in these information for a specific device, machine, or equipment [86, 106, 191, 196, 213]. This unique pattern is also called the *fingerprint* of a device [41, 85, 95, 104, 164, 206]. The generated DFP can have variety of applications in security, condition monitoring, and device identification.

Figure 4 briefly illustrates the DFP in the CPS with four essential steps: source identification, aggregation, processing, and application of equipment, machinery, or equipment-specific device characteristics [22, 85, 104, 117, 164, 191, 206]. The source identification step identifies the relevant device feature sources that may be utilized to generate a stable and unique DFP. In the feature aggregation phase, the device's features are gathered from various CPS entities. However, due to the CPS's complexity, heterogeneity, and diversity, it might be challenging to ascertain which device characteristic and aggregation methods are the most effective. In Sections 3 and 4, we explain how to choose device features and aggregate them. The aggregated features are then utilized to generate unique DFPs, which follow preprocessing, transformation, clustering, classification, and pattern recognition of features. The application phase includes different application areas of the DFP, as discussed in Section 1.2. The processing step performs preprocessing and fingerprint generation on the device features to generate a stable and unique fingerprint. For this purpose, rule-based [9, 209], statistical [51, 116, 182, 212], knowledge-based [17, 77, 162, 201], data-driven **Machine Learning (ML)** and **Deep Learning (DL)** [33, 85, 156, 164, 197], and time series approaches [33, 85, 136, 156, 164, 197] have been proposed. However, selection of the processing approach for DFP generation depends on the amount of data, the available resources, and the complexity of feature correlations, and consequently some algorithms are better than others in CPS scenarios [164]. For example, a simple CPS application, like a household water distribution system (i.e., a water pump, water storage, distribution line, water pressure or level meter, and taps or consumer points) with a limited and known set of actions, can be modeled with a rule-based approach, leveraging the limited resources in the components. In contrast, a city water distribution system is complex and hard to model using rules, and an ML/DL-based approach exploiting the correlation in the available data sources would be more successful.

2.1 DFP Characteristics

There is a need for standardized criteria for determining the characteristics of devices in the CPS since devices have several characteristics related to various entities, including network and communication characteristics, applications, and services. Some works [15, 25, 52, 97] outline the

application-specific selection criteria and quality evaluation techniques for CPS devices, which are as follows:

(1) *Diversity*: This metric assesses how well a device can be distinguished and identified from a group of similar devices in the CPS [96, 97].
(2) *Stability*: This determines the rate at which a fingerprint's values change over time and in different working environments. For fingerprinting, there are four factors proposed by Kobusinska et al. [97]. This report includes the total number of alterations, the average time between alterations, the number of devices for which at least one change is identified, and the average percentage ratio, representing the number of samples modified for the devices.
(3) *Efficiency*: This evaluates the efficiency with which various resources, such as fingerprint collecting, preprocessing, detection algorithms, processing time, and memory, are used. More specifically, Kobusinska et al. [97] identified three characteristics for determining fingerprint efficiency: length of execution code, execution time, and fingerprint length.
(4) *Supportable*: This characteristic implies that the algorithm for DFP should be simple to develop, deploy, and update so that the system administrator may quickly configure or modify DFP applications to match their requirements and use cases.

2.2 Evaluating Quality of DFP Features

The quality of the data acquired and utilized for DFP generation is a common factor across industries that rely on a CPS DFP for different applications. In this section, we refer to data as device features. In the fast-growing fields of **Artificial Intelligence (AI)** and Industry 4.0, quality assessment is crucial for identifying a device feature's relevance for DFP-driven applications, including condition monitoring, security, device tracking, and forensics. The device features in the CPS face multiple challenges caused by diversity in the feature sources; large dataset; missing, noisy, or biased features due to sensor anomaly; communication failure; inconsistent data format; or incorrect data profiling [24, 38].

Numerous metrics developed and published in academic literature may be used to monitor and address CPS data quality issues. These issues are due to a wrong schema, duplicate data, lagged data, anomalous data, and so forth [24, 38, 42, 132]. In Table 3, the CPS data quality measures are listed under five major classes: availability, usability, reliability, relevance, and presentation. For ensuring these data quality measures, the rule-based technique is typically used [42, 132]. However, due to the complexity of the CPS with its high cardinality and multidimensional data, it require many rules. The data quality framework needs custom implementation and human intervention for every new defect or anomaly. ML-based algorithms have been used to evaluate data quality instead of rule-based approaches that require human intervention. ML algorithms can learn from vast amounts of data and find hidden patterns without human intervention. Some ML methods used for data quality evaluations include dimensionality reduction, clustering, anomaly detection, and association rule mining [33, 85, 156, 164, 164]. Researchers have utilized these algorithms for years to evaluate time series, textual, and audio/video data for text creation, sentiment classification, invalid text identification, video/image quality evaluation, and many others [85].

2.3 Fingerprint Similarity Measures

Without an effective matching mechanism, a system cannot distinguish between counterfeit and genuine devices using fingerprints. Therefore, the system requires an appropriate fingerprint matching method to identify, classify, and match the fingerprint of the unknown device with the fingerprint of the reference device. Several similarity assessment metrics are utilized in most fingerprinting applications [56, 100, 121, 193, 196]. These include the Hamming distance, Euclidean

Table 3. Data Quality Metrics of Device Features

Dimension	Metric	Definition/Indicators
Availability	Accessibility	How easily data is accessible for the DFP system.
	Timeliness	Within a given time interval, whether the data arrived and data is regularly updated or not.
		Whether data collection, processing, and release meet deadlines.
Usability	Credibility	How trustworthy your data is.
		Existence of data within the range of known or acceptable values.
		Specialists periodically audit and verify data content.
Reliability	Accuracy	How close a measurement is to its true value.
	Consistency	After processing, data concepts, value domains, and formats are unchanged.
	Integrity	The data format is clear and fits the requirements.
		Data is consistent with structural and content integrity.
	Completeness	Refers to how complete the information is.
		For multi-component data, how much of an effect would the absence of a certain component have on its usefulness, data accuracy. and integrity?
	Orderliness	Data randomness and entropy.
	Uniqueness	A metric for determining how much duplicate information exists for a given field, record, or dataset.
	Confidentiality	Prevents the unauthorized exposure of sensitive information or data.
	Anonymity	Measures the amount of privacy of digital data (device feature) acquired from a device in the CPS.
	Normality	Used to determine whether sample data has been drawn from a normally distributed population (within some tolerance).
	Validity	How accurate a measure is (whether the results really measure what they are supposed to measure).
	Efficiency	Coherence to achieve accuracy and completeness with existing resources.
Relevance	Fitness	The data gathered does not fit the theme perfectly, but it does provide more information about one part of the theme.
	Effectiveness	Ability to generate defined outcomes with precision and completeness to reach project objectives.
	Satisfaction	Extent to which the user is content with the aims, confident, and comfortable with the system.
	Context coverage	Extent to which a system can be used efficiently, effectively, satisfactorily, and without risk within a set of constraints.
	Freedom from risk	Degree to which a system mitigates the expected risk to the CPS in specific scenarios, equipment, environment, and physical assets.
Presentation	Readability	Clear and understandable data (in terms of content, format, etc.).
		Information provided is adequate, and this is easily determined.
		Data description, classification, and code content are easy to understand and meet requirements.
	Interpretability	How well data uses proper languages, symbols, units, and definitions.

distance [196], overlap coefficient, Jaccard distance, Lorentzian distance, correlation, Matthews correlation coefficient [73], Bhattacharyya distance, and cosine similarity [206], among others. These are derived from the probabilistic and spatial distribution of device features.

3 SOURCES OF FINGERPRINT

This section discusses the internal and external features of the device and their sources used for DFP. In the CPS, these feature sources are derived from different cyber and physical system attributes. As shown in Figure 5, the bottom of the pyramid consists of attributes connected to the cyber system's components acquired from hardware, software, applications, or networks. This includes the cyber system setup, used protocols, and available resources. The majority of the cyber

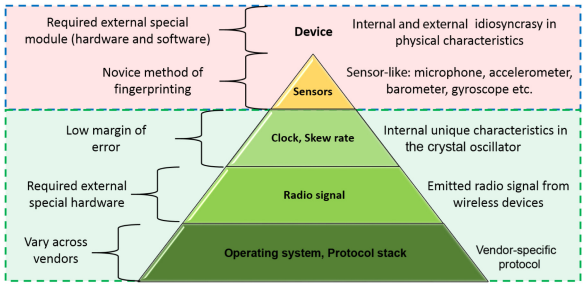

Fig. 5. Device features that can be exploited in CPS applications.

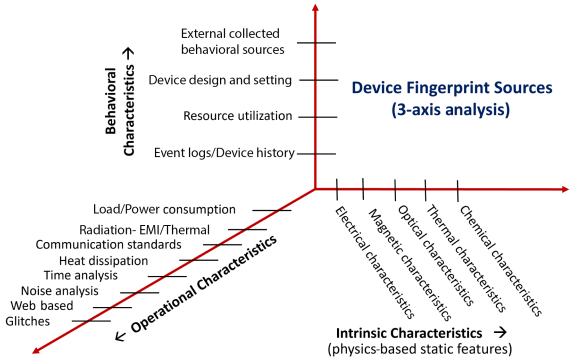

Fig. 6. Three-axis taxonomy of DFP sources in CPS.

system features sources used in DFP are supplied by the device's manufacturer and may be obtained from the device's datasheet. However, the radio signal and clock-related features of the device are based on internal characteristics and vary from device to device. The features shown at the top of the pyramid are based on the device physics or operational characteristics. These features require specialized hardware and software modules to extract, store, and analyze.

Figure 6 shows a three-axis classification approach to study the device features in which the features are classified along the three axes: intrinsic, operational, and behavioral characteristics. These characteristics are based on the source of fingerprint generation.

3.1 Intrinsic Characteristics

During the device fabrication, subtle flaws in its electrical, electronic, or mechanical modules and sensors introduce reasonable imperfections in the hardware of devices. As a result, no two modules performing the same job may ever be considered identical. Over the past few years, several researchers have exploited this imperfection for device identification and classification. We have analyzed imperfection-based aspects under five classes, described in Table 4, depending on the device's physics-based intrinsic features. Additionally, Table 5 provides a comparative overview of various DFP techniques, taking into consideration both cyber and physical features.

3.1.1 Electrical Properties. A PCB is the primary component of a device that integrates all of the electrical and electronic system's elements. Therefore, the electrical property is the main source of the unique signature of a device [76, 194, 199]. In a typical PCB layout, copper traces are routed across several layers in a conventional PCB layout, separated by dielectrics, such as epoxy glass. Copper or the conducting materials used in the PCB fabrication are not a perfect conductor

Table 4. Device Physics and Intrinsic Characteristics Features Used in Fingerprinting

Physical Context	Feature	Applications	Link
Electrical	Conductivity, inductance, impedance, permeability or capacitance	Device identification, device conditioning, fault analysis	[68, 69]
Thermal	Thermal conductivity (and resistance), heat capacity, density, thermal resistance, specific heat, emissivity and absorptivity	Device identification, fault analysis, status prediction, localization	[62, 67, 122]
Magnetic	Electromagnetic interference (EMI) shield, materials magnetic properties, PCB layout, material permittivity	Device identification, security, condition monitoring	[63, 72, 94, 173, 183, 217]
Optical	Radiance, absorption, reflection coefficient, refractive index, luminescence, intensity, irradiance, color	Device identification, conditioning monitoring, security, copyright compliance, licensing	[62, 70, 71, 197]
Chemical	Toxicity, reactivity, flammability, heat of combustion, acidity, radio activity, chemical stability, spectral signature	Machine fault analysis, condition monitoring, safety, localization	[53, 111, 146, 155]

and therefore obey all aspects of electrical conductivity and resistivity [79, 94, 173, 183, 184]. The fabrication and processing variation appears as electrical characteristics such as resistance, inductance, and capacitance [217].

3.1.2 Magnetic Properties. Thin copper traces in a PCB are similar to current-carrying conducting wires and generate a magnetic field proportional to the amount of current [94, 217]. Magnetic field based systems are often used to interact with the remote device or monitor device status. Table 6 summarizes different DFP approaches based on electromagnetic emissions that are used for condition monitoring or fault investigation [173, 183], security [72], and identification [58, 63]. The magnetic emissions from hard disk drives detected by a mobile phone's sensor can be used for precise channel monitoring in various applications [63]. Guri et al. [72] exploit the low-frequency magnetic field generated by the CPU for the malware attack, namely "ODINI." The magnetic field is used for PCB board fault localization and detection [183]. A magnetic induction diagnosis method is proposed by Sheen et al. [173], who use the concept of magnetic induction to energize antennas' conductive loop. This depends on the induced voltage that can be used to identify the separate components of the PCB board. Similar work has been reported by Goulette et al. [58], who use the PCB's electromagnetic emission through the array of electromagnetic measuring probes to classify the board based on their emission level.

With the development of magnetic imaging techniques, the fault diagnosis of complex and large electronic devices is becoming easier. Yao and Pan [210] use a non-contact method of the magnetic image for PCB fault diagnosis and its analysis. The overall characteristics of a device that is controlled by its size, shape, and length of a PCB can also be affected by the skin effect [217]. The skin effect can be used for device identification because of the variation in length, width, and height of the PCB due to randomness of the manufacturing process and the material used. Additionally, the magnetic characteristics of the materials used to make CPS devices influence their magnetic properties. Based on the magnetic behavior shown by the material, they are classified as diamagnetic, paramagnetic, ferromagnetic, diamagnetic, or super magnetism.

3.1.3 Thermal Properties. The conversion of electrical energy into thermal energy due to the current flowing in electrical conductors (PCB copper traces) and other electrical elements is responsible for heat generation. Perfect thermal design of a PCB is a challenge for any heat transfer designer, particularly for CPS devices. Heat dissipation in different parts of the PCB board (conducting layer traces, electronic elements, etc.) influences thermal performance and signature linked

Table 5. DFP Techniques and Their Comparison Based on Cyber and Physical Features

Date, Work	Feature	Gathering Method	Device	Uses	NET	Web Application	Physical Context
2005 [99]	Clock skews	TCP and ICMP timestamp	ICT (Web clients)	Forensic, TRK, SEC	Yes	Yes	No
2014 [151]	Device configuration	JavaScript and APIs	Web client	IDN	No	Yes	No
2015 [119]	Power usages	Battery status API	Mobile, IoT	TRK	No	No	Yes
2015 [199]	Manufacturing variation	Electrical properties	PCB (device)	PUF, SEC	No	No	Yes
2015 [214]	Manufacturing variation	Intrinsic trace impedance variations	PCB (device)	AUTH	No	No	Yes
2016 [102]	PDA configurations	Third-party apps	IoT (mobile)	IDN	Yes	Yes	Yes
2016 [29]	Acoustics	Temporal and frequency response	IoT (mobile), PCs	AUTH, SEC	No	No	Yes
2016 [140]	Electrical characteristics	Manufacturing variability PCB	IoT, CPS (PCB)	IDN, SEC	No	No	Yes
2016 [54]	Physical process	Data processing time	IoT, CPS (ICS n/w)	SEC	Yes	No	Yes
2016 [32]	ECU	Time analysis of in-vehicle messages	CPS (vehicular n/w)	SEC	Yes	No	No
2017 [76]	Manufacturing variation	Electrical characteristics	Device (PCB)	IDN	No	No	Yes
2017 [4]	Noise pattern	Sensor measurements	CPS (ICS)	IDN	No	No	Yes
2017 [13]	Intrinsic and physical	On-board sensor	IoT (mobile)	IDN	No	No	Yes
2017 [170]	Communication environment effects	Communication and processing parameters	IoT, CPS (R-pi)	IDN, SEC	No	No	Yes
2017 [84]	Physical characteristics	Sensor CRT, physical properties	IoT, CPS	SEC	No	No	Yes
2017 [4]	Noise pattern	Sensor measurement	IDN	CPS	Yes	Yes	Yes
2017 [83]	Manufacturing imperfection	Visually imperceptible random pattern	ICT (PCB)	AUTH	No	No	Yes
2018 [60]	Device physics context	Processes physical context	CPS (ICS)	SEC	No	No	Yes
2018 [190]	Controller and gateways coordinate	Behaviors of traffic, servers, and patches	IoT (nodes)	IDN	Yes	Yes	NA
2018 [85]	Radiometric extracted	RF capturing	IoT	IDN, SEC	Yes	Yes	No
2018 [165]	Time analysis	HTML5 cryptography and timing API	IoT devices	IDN, SEC	No	Yes	No
2019 [215]	Calibration fingerprinting	Sensor reading	IoT (mobile, R-pi)	IDN, SEC	NA	NA	Yes
2019 [52]	Overview of DFP	CPS	All	IDN, SEC, COND	Yes	Yes	Yes
2019 [21]	Cryptographic protocols	Network traffic	IoT, CPS	IDN, SEC	Yes	No	No
2019 [31]	Magnetic inductions	DeMiCPU sensor	IoT, CPS	AUTH	No	No	Yes
2019 [215]	Sensor calibration error	On-board sensor reading	IoT (mobile)	IDN			Yes

PDA, personal digital assistant; ECU, electronic control unit; CRT, cross-layer response time; SEC, security; IDN, identification; TRK, tracking; AUTH, authentication; R-pi, Raspberry Pi board; NET, network; COND, conditioning.

Table 6. DFP Using Magnetic Field and Electromagnetic Emission

Method	Generated by	ApplicationType	Work
ODINI	CPU low-frequency magnetic field	Air gap covert channel	[63]
Myhayun	Hard disk drive	Air gap covert channel	[72]
Spence	Magnetic field	Printed board diagnosis	[183]
Sheen	Magnetic induction	PCB tester	[173]
Goulette	Electromagnetic emission	PCB monitoring	[58]
AirHopper	Mobile phone radio frequency	Bridging of two devices	[65]
GSMem	GSM frequency of a cellular phone	Malware that can exfiltrate data through an air gap over cellular frequencies	[64]
USBee	Electromagnetic emission from USB	Exfiltrates data from a secure and air-gapped computer	[66]
Funtenna	GNU radio	Device identification	[37]

with layer-wise thermal conductivity, thermal mass distribution/nature of drill hole, soldering, and electrical heating. Temperature distribution may also be attributed to the randomness introduced during the manufacturing process of device PCBs [45, 184, 194]. In some works [8, 108, 128], the authors exploit the device PCB's temperature uncertainty to develop a unique signature for device identification and monitoring. Aside from electrical current in electrical elements, heat is also produced by mechanical friction, fuel combustion, and thermal heating of the electric coil in mechanical devices [11, 163, 186].

Additionally, in the open air, all things above absolute zero temperature produce radiation in the form of light, known as thermal and heat radiation. All objects generating thermal radiation follow an ideal blackbody radiation curve. At the same time, the surface material composition influences the thermal wave properties due to its interaction with the radiated energy. The total radiant heat Q incident on an object surface is divided into three components, namely radiant reflected heat (Q_r), radiant transmitted heat (Q_t), and radiant absorbed heat (Q_a). In either case, the internal heat generation or variations in thermal absorption by the assembled electronic components produce the thermal pattern/signature. These thermal patterns are frequently used to locate surface mount components, determine the quality of solder joints, help in identification of the board/device, and generate the signatures [8, 45, 108, 128, 184, 194].

In general, the thermal patterns of electrical and electronic equipment can be acquired through contact or non-contact methods [45]. The contact-based method captures the thermal signature of the equipment via embedded MEMS (or micro-electromechanical systems), sensors, or external sensor probes. This approach is complicated and time consuming with the higher possibility of device damage during measurement, especially in the external sensor probing. However, in the non-contact method, the thermal signature of the device can be acquired using thermal or visual imaging without direct contact with the board. Thermal imaging technology utilizes an infrared (IR) camera to capture the thermophysical patterns of electrical, electronic boards, or machinery through various modes of heat transfer, including radiation, conduction, and convection [45, 127, 163, 185].

3.1.4 Optical Properties. There are many ways to identify and recognize machines, tools, and other things, but one of the oldest is to use visible light (0.4–0.7 μm). The human eye captures the visible light reflected from the object surface, and the visual perception is based on the object's color, brightness, and intensity. Visible light categorizes and identifies fake and legitimate devices in different critical infrastructures, including manufacturing, energy, utilities, defense, healthcare, water supply, wastewater management, transportation, surveillance, and safety. With the advancement of image acquisition and processing technologies such as wide-spectrum (IR and visible) cameras and image analysis algorithms, imaging-based device recognition, classification, and condition monitoring have improved significantly [113, 127, 163, 186]. IR thermography is one such method. It is a thermal imaging camera that is used to record the thermal energy released by a device to do thermal profiling. It is determined by the amount of energy picked up by the IR sensor, which is proportional to the object's surface emissivity [113, 135, 166].

Over the years, optical fingerprints have evolved to be an effective technique for object detection, identification, and condition monitoring of devices in non-contact and non-invasive ways. The accuracy of the optical fingerprint based application depends on the resolution of the camera (thermal and standard cameras) used for imaging [135, 167]. The operating band of the electromagnetic wavelength of the light (IR and visible) spectrum have limited penetration capabilities [135]. Therefore, it is suitable for an object whose surface optical or thermal characteristics are unique. For example, the condition of a critical infrastructure's elements such as railway infrastructure elements [186], wind turbine blades [163], nuclear power plant cables [127], and induction motor [113] are monitored by their visual image or through thermal imaging techniques.

Table 7. Operational DFP Features, Sources, and Applications

Feature	Source	Use	Work
Radiation-EMI	Electrical coil, PCB, electric machine, power generation station, power transmission	Predictive maintenance and health condition, fault detection and classification, process management	[11, 31, 122, 177]
Thermal radiation or Heat dissipation	Electric and electronic devices (ohmic effect of current carrying element) and mechanical machines (mechanical friction)	Predictive maintenance and health condition, fault detection and classification	[11, 11, 20, 93, 113, 127, 163]
Power consumption	Electrical and electronic devices (uses electric power) and mechanical machinery (combustion of fuel)	Predictive maintenance and health condition, performance, malicious activity monitoring	[11, 69]
Sound and vibration	Generated due to the mechanical and kinetic motion of the device's components	Predictive maintenance and health condition, fault detection and classification, performance, behavioral analysis	[18, 148, 216]
Noise (mechanical)	Mechanical sound, surrounding noise, mechanical fault or breakdown	Predictive maintenance, fault detection, performance, behavioral analysis, localization	[5, 18, 39, 126, 126, 134, 134, 192, 216]
Noise (electronic or digital)	Information conversion (analog to digital), quantization error, information compression, signal processing, sensor calibration error, channel or processor noise	Predictive maintenance and health conditioning, fault detection and classification, performance, behavioral analysis, IDS, security and privacy, localization	[3, 5, 6, 134, 144, 148, 148]
Glitches	Caused by sudden and temporary malfunction in a system or equipment	Predictive maintenance and health condition of devices, fault detection, performance, behavioral analysis, security	[79, 123, 142, 158, 192]

3.1.5 Chemical Properties. Fingerprinting techniques based on chemical properties are crucial for the safety of high-risk critical infrastructure setups. Factors such as fire, explosions, high-risk radiation, and leakage or discharge of toxic or hazardous chemicals are the possible risks that might endanger human health and the environment [53]. Several sensors and actuators have been based on the chemical properties of materials. Examples include air quality sensors (which detect CO and CO_2 concentrations), non-dispersive IR sensors, electroacoustic sensors, and electrochemical sensors [146]. An automated monitoring system monitors the distribution pattern of different chemical elements to evaluate the plant floor or machine condition in industries such as petrochemicals, pharmaceuticals, and nuclear power plants. Chemical property based features in Industrial IoT applications may offer information on the health of equipment or devices on the factory floor. Chromatographic and X-ray analyses, as well as spectroscopic techniques (mass spectrometry, gas spectrometry, atomic emission spectrometry, and atomic absorption spectrometry), are used to detect the chemical composition of traces of chemical leaks in equipment and industries [111, 155].

3.2 Operational Characteristics

This section discusses the operational properties of CPS components, including networks, devices, sensors, actuators, and storage. Over the past several decades, researchers have developed many operational characteristics that may be used in this context, including power consumption, resource utilization, communication protocols, device noise, electromagnetic interference and

Table 8. Time Analysis Based DFP

Time Feature	Feature Sources	Application	Work
Clock skew	System/physical clock, network time	Device IDEN, fault analysis, clock synchronization	[34, 54, 114, 150, 212]
Network delay	Computer network: ICMP/TCP timestamp	Network characterization, device IDN, throughput calculation, network topology	[54, 60]
Processes or CPU time	Hardware, web script or APIs (bounded execution of instruction sets)	CPU fingerprinting, device performance analysis, CPU IDN, CPU behavior analysis	[54, 165]
Clock drift	Network data packet header, web app, cookies, canvas data, hardware data	Device IDN, clock drift analysis, reliability and performance evaluation, design improvement	[54, 165]

IDN, Identification; network latency/delay, time required for the bit data to travel in the network from sender to receiver and again from the receiver to process that request; clock skew, clock inaccuracies across the different devices/nodes; clock drift, amount of physical or system clock that differs from the real-time or reference clock; process time, time required by a device to process the prescribed task.

thermal radiation, and time analysis, among others. Table 7 summarizes the operational features and applications of CPS.

3.2.1 Radiation-Electromagnetic interference (EMI)/Thermal. The CPS is composed of electrical and electronic components that emit electromagnetic radiation whenever the current is flowing. However, the magnitude and effect of radiation depend upon the application and operating frequency. Usually, for electric machines, their electric coils and electronic control board emit a unique electromagnetic radiation pattern of low power and low frequency [11]. As a result, electromagnetic radiation has been recognized as a separate operational characteristic that is based on physics of the process. Device authentication, interference monitoring, condition monitoring, fault localization, security, and surveillance are all possible with electromagnetic radiation. CPS applications like manufacturing, transportation, smart grid, water treatment, medical devices and ICSs, and electromagnetic radiation patterns are often used for monitoring of the electrical machine condition and localization of fault [11]. Cheng et al. [31] proposed an electromagnetic-based DFP method that uses the electromagnetic radiation generated by the device electronic board CPU during the process to identify and classify device computation overhead and types. Aside from device identification applications, electromagnetic field or flux signatures can be utilized to monitor the health of induction motors and improve their overall operational efficiency [177].

3.2.2 Timing Analysis. In CPSs, oscillators (or timing circuits) synchronize all devices, including routers, modems, switches, sensors, and actuators. Moreover, the CPSs are decentralized because they are composed of devices that are physically isolated from one another and connected through a variety of network topologies and communication protocols. In this manner, the system synchronizes the components of the CPS, which are carried out using a variety of separate clocks [123, 124, 187]. Timing analysis of devices using DFP techniques involves microscopic changes in time across the hardware module, device, or system. Sánchez et al. [187] review different timing analysis techniques for fingerprinting computer networks, equipment, and computers, including clock skew, latency, packet or processing time, and clock drift. Table 8 summarizes the time- or clock-based DFP characteristics employed for automatic device classification, identification, management, and control applications.

3.2.3 Power Consumption. The term *power consumption* refers to the amount of energy required per unit of time. In a CPS, power consumption reflects the actual condition of the cyber system

(i.e., computing, storage, and communication and network devices) and the physical system (i.e., electrical, electronic, and mechanical devices). In a cyber system, the computing activities govern the device's or chip's power consumption—that is, the amount of data and the processing schedule. However, in physical devices, the power consumption pattern governs device operation activities such as operating time, breakdown/failure, downtime, performance, scheduling of tasks, kind of machine/jobs, efficiency, and type of power source (i.e., electrical, fossil fuel, nuclear fuel, or any other chemical composition). Hence, power analysis or differential power analysis of the CPS's elements is used to evaluate the behavioral patterns of the equipment, machinery, and operator. Guri et al. [69] identified the behaviors of household appliances by analyzing the power pattern of the home. Power signatures can also be employed by device architects, developers, or engineers to find and locate the fault of a device. Despite this, many devices in the CPS continue to function normally without electricity. In most cases, fossil fuels like natural gas or petroleum are used as a backup power source for this kind of machinery. Thus, the fuel consumption pattern is used to determine the power signature, which is then used to identify the machine and check its condition.

3.2.4 Heat Dissipation. We discussed heat dissipation and radiation earlier in this section. System characteristics such as hardware design, PCB layout, and component arrangement determine the total power dissipated. In addition to the electrical system, the mechanical system dissipates heat due to internal friction between several mechanical components. Therefore, the device's electrical, electronic, and mechanical system characteristics determine the heat dissipation pattern. This comprises the distribution, location, and amount of heat generated due to the ohmic effect of an electrical and electronic system, as well as the level of mechanical friction between various parts. The heat dissipation patterns of a system can predict the condition of CPS devices, machinery, and equipment [123].

3.2.5 Glitches. A glitch is a short-term, temporary malfunction that resolves itself on its own. It is an abrupt and short-lived breakdown in a system or piece of hardware. Electronic and computer systems are particularly vulnerable to disruptions. Disruptions can happen at any point in the CPS or IoT architecture, from physical devices, sensors, or actuators to data collection, processing, and application. Based on the IoT architecture, we have classified it into four primary groups, as shown in Table 9. These four types include hardware, software, network, and physical equipment faults that can arise due to spontaneous, unwanted, and unscheduled interruptions in the functioning parts of any CPS.

3.2.6 Noise Analysis. Noises such as mechanical sound, sensors, communication channels, white noise, and process noise are frequent in real-world CPS applications [5, 126, 134, 192, 216]. During operation, the electrical, electronic, and physical components of the CPS generate these noises. Depending on the signal and mechanism of generation, these noises are classified into four classes: ambient noise (or sound), sensor noise, processing noise, and channel or transmission noise. The noise analysis techniques used in CPS applications are summarized in Table 10. In some works [18, 39], the authors investigate ambient sounds for device localization on the factory floor and machine sounds for forecasting device health. Ambient noise is produced when a change in pressure is felt on the factory floor due to the impact of machines, surface vibration, or dynamics [134]. As a result, the operator can use the sound of the machine to estimate the machine's health on the factory floor. Additionally, characteristics like loudness, speed, and frequency are controlled by the device's condition and surrounding environmental aspects such as temperature, humidity, noise, and acoustic insulation [126]. As well, the process and electrical noise (A/D conversion, filtering, accuracy, and precision) are employed for device identification, conditioning, and navigation [144, 148, 148].

Table 9. Common Glitches That Occur in the CPS

Glitch Type	Description	Source
Hardware	Malfunctioning of physical components of the sensor, actuators, or ICS	Power, clock/timer, register overflow, spike, overheating, unusual noises, power interrupt
Software	Malfunctioning of software in the computational system	Software bugs, operators error, undetected invalid input or communications error, computer virus, Trojan attack and adversarial exploiting, stack overflow
Network	Malfunctioning of the networking system	Packet drop, bit-flip, interference/noise, channel noise, adversarial attacks, power interrupt, network overload/bandwidth congestion
Physical devices	Malfunctioning or breakdown of the physical device/machines	Mechanical breakdown, power interrupt, resource (supply chain disruption, labor shortage), constraints, accidents (fire, leakage of hazardous materials, wear and tear on machinery), tampering

Table 10. Noise-Based DFP

Work	Noise	Sensor	Feature	Application
SurroundSense [18]	Light and sound	Light sensor, microphone	Sound: frequency spectrum; light: intensity variance	Localization
[144]	Sound	Microphone and speaker	Signal processing related features	Identification of activities of daily living
SoundUAV [148]	Electromagnetic and mechanical	Microphone	Noise variance	Motor (or device) identification or fingerprinting
NoisePrint [5]	Sensor and process	Intrinsic error in sensors	Time domain and frequency domain	Sensor identification
NoiSense [3]	Sensor	Ultrasonic, flow meter, RADAR, level sensor	Time domain and frequency domain	Sensor identification
Noise matters [6]	Manufacturing imperfections	Sensor (e.g., ultrasonic, flow meter, RADAR)	Time domain and frequency domain	Sensor identification

3.2.7 Web-Based Fingerprinting. In a typical IoT or CPS, devices are linked through a complex and heterogeneous cyber system with several communication protocols, network topologies, and services [10, 139, 152, 161]. Due to the diversity of communication protocols, network topologies, physical processes, data types, and application scenarios, communication between devices in an inter-IoT/CPS environment is complicated. This problem is solved in the CPS by a homogeneous layer that uses web technologies [35, 137]. In web technologies, a web service running on back-end infrastructure provides various services through URLs or APIs that facilitate the configuration and maintenance of system architecture and components. Web-based applications are often used on human-machine interface devices, particularly handheld devices, to monitor operations and take action. This is also true in real-world CPS examples such as water treatment, transportation, manufacturing, power grid, manufacturing, healthcare, and smart city [96, 109, 125, 142, 178].

Over the past few decades, the web-based (or browser) fingerprint has been a popular technique for identification and authentication [104]. Browser fingerprinting is a technique that uses a web browser to collect information about a device (e.g., its operating system, hardware configuration, active plugins, screen resolution, font, user agent, language, time zone, geolocation, and other current settings) to create a fingerprint of that device [15, 25, 104, 188]. It is a powerful method frequently used in ICT and CPS applications for device and resource management.

3.2.8 Communication Standards. CPS devices use a variety of communication standards and technologies to send and receive information with other CPS or ICT devices or interfaces that are within range. There are different network communication protocols, standards, topologies, and ways for the CPS network to interact with other cyber or physical devices [161]. However, selecting the appropriate networking arrangement and protocol is challenging in every environment. Additionally, many IoT devices have limited resources (i.e., power supply, computation, memory, and storage) and obsolete networking and communication protocols. As a result, it may not be compatible with all CPS devices using older communication standards. In the past few years, numerous new communication protocols and topologies have been developed for devices that will support a wider variety of communication standards. For example, MQTT, CoAP, AMQP (Advanced Message Queuing Protocol), and API standards for application communication between devices or systems are used in CPSs on the manufacturing floor. Specifically, healthcare system devices employ short-range communication protocols like Zigbee or BLE to support the wireless body area network [75, 171]. As a result, fingerprinting techniques use the characteristics of communication systems to identify, monitor, and manage interconnected devices in a specific application setting [7, 91]. This includes making changes to communication standards, routing protocols, and the network and session levels of the communication protocol stack [118, 143, 206]. Multiple fingerprinting techniques have been developed over time based on the characteristics of CPS communication and networking components. Wireless and network fingerprinting, for example, use the features of a wireless device (the transmitter) and a network, respectively, to generate a unique DFP [60, 91, 154, 178, 179, 198, 205].

- *Wireless fingerprinting*: Wireless communication uses electromagnetic waves such as **Radio Frequency (RF)**, IR, visible light communication, and satellite to transmit data wirelessly. Due to the rapid development of the CPS and IoT, wireless technologies have emerged as essential components of today's computing platforms and embedded systems. These technologies provide low-cost personalized connectivity everywhere, at any time. Examples are Bluetooth, Zigbee, WiFi, LoRA, RFID, GPS, and WirelessHART, among others [36, 48, 73, 75, 110, 115, 157, 171, 191]. Currently, CPSs use a large number of wireless devices; thus, their security is essential. The use of wireless fingerprinting as a solution has the potential to be an effective method to address this issue. The wireless DFP procedure includes the identification of radio transmitters using device-specific artifacts of their transmitted signals. These artifacts are caused by natural variations in the hardware attributes of the transmitters. It is also known as radio fingerprinting or RF fingerprinting [73], which uses wireless device features through the communication channel and protocols' frame information.

  Many authors have reported wireless fingerprint work for transmitter node identification and classification in the past few years. For this, it exploits features of wireless communication protocols like carrier frequency offset, wireless channel information (i.e., channel state information, channel impulse response, channel frequency response, radio signal strength, carrier frequency difference, and phase shift difference [26, 73, 80]. Furthermore, to analyze these features, a device uses traditional and ML-based supervised or unsupervised learning techniques [26, 26, 57, 80, 206].
- *Network fingerprinting*: Network traffic fingerprinting, often known as NTF, allows users to remotely and automatically collect information about the hosts that are part of a network [49, 60]. NTF can assist with network administration, as well as the identification and isolation of vulnerable hosts.

  However, a traffic analysis attack involves monitoring network traffic to identify usage patterns that can be used for protection and surveillance of security threats (e.g., botnet

traffic), forensics (to identify a suspect or if a crime occurred), or malicious purposes like personal information breaches [14, 74, 205]. One of the methods of traffic analysis is traffic fingerprinting, which is sometimes characterized as a passive privacy-related side-channel attack [154]. It refers to comparing and matching characteristics derived from network traces (e.g., packet sizes, distribution, and other statistical metrics) with the existing models (fingerprints). Fingerprinting approaches often use an ML-based approach–classifiers–to improve the accuracy of traffic type recognition [198]. NTF uses network traffic characteristics, including packet size, packet direction, packet length frequency, packet ordering, inter-packet arrival time, packet count, burst size, traffic data rates, and burst size-surge period, among others, for analysis of networking techniques and connected devices for configuration, management, and security [105, 160, 198, 207].

3.3 Behavioral Characteristics

Real-time data analytics approaches and applications benefit from the development of data handling and processing methods (e.g., big data, AI, ML, DL, and computer vision), computing hardware (e.g., high-performance computing, GPU computing, and cloud computing), and algorithms. Behavior analysis is an example of a data analytics application that uses sensory information to predict the present state of the nodes [21, 164]. Real-time device conditioning and monitoring are carried out using behavioral analytic methodologies [142]. Consequently, several researchers have explored DFP by analyzing the device's behavioral properties (Table 11). Dogan and Incel [43] summarize DFP algorithms based on their data sources, application scenarios, processing, and assessment methodologies for IoT, mobile communication, computer network, and ICS use cases. Table 12 summarizes a few more fingerprint applications, where the devices' utilization patterns of various modules are listed. However, behavioral analysis requires historical information related to large amounts of device operational activities [147, 189]. Typically, it uses four types of behavioral feature sources: event logs, resource utilization, device design and its configuration, and externally collected features. These features are easily observed and collected in CPS applications using dedicated sensory systems or manually by operators.

3.3.1 Events Logs. Data generated by the CPS components consists of cyber-physical activities related to accessing the connected devices, establishing new connections, verifying device identity, tracking faults and error messages, and processing/operating pre-scheduled tasks. These logs represent a particular device's activities. Data analytics approaches are used to categorize and identify a variety of problems, such as monitoring device performance, classification of device type, and its condition using these logs [21, 164]. Additionally, system and network logs are used to create real-time security solutions (e.g., threat prediction, detection, and tracking) for complex networks like the Industrial IoT and CPS [208].

3.3.2 Resource Utilization. The key performance metric that is important describes how resources are utilized during the execution of an activity or operation. In the context of the CPS, computation, communication, memory, storage, energy, sensor, and actuation resources are required for their smooth operation. Moreover, these resources are limited and/or shared by multiple devices or machines. Therefore, the resource utilization feature can be used to forecast and track health conditions, system behaviors, and malicious activities, among others [119, 142, 164]. For example, uses of household electrical machines were approximated by assessing household power usage [44]. Monitoring the power consumption behavior of machines on the factory floor (electric motors), as detailed elsewhere [89, 130, 164], may be used to make predictions about the state of health of a mechanical device, including the amount of wear and tear that has occurred. Similarly, for non-electrical or electronic equipment, the ML model may forecast the system's behavior. For

Table 11. Behavioral Features Used in DFP

Behavior Source	Feature	Applications/Work
Event logs	1. Cyber system: system, security, application, directory service, network service, status, error, etc.	DI [28, 44, 85, 104, 117], MD [28, 81, 103, 141], FT [85, 104, 117, 141]
	2. Physical/Mechanical system: breakdown/failure, process logs, power supply, fault and SCADA data	DI [117], CM [38, 106, 208], DM [53, 158]
Resources utilization	1. Cyber system: CPU uses, CPU activity, System storage usage, file system, Scheduler, System memory usage, Virtual memory, I/O throughput per network interface, network, execution time	DI [104, 117, 118], MD [27, 81, 104, 118, 129]
	2. Physical/Mechanical system: power supply voltage, current, power factor, fuel consumption, material flow, overall equipment effectiveness	DI [19, 27, 117, 117], CM [38, 106, 185], MD [41, 53, 117]
Device design and its configuration	1. Cyber system: system architecture/knowledge, sensors, actuators, computation type, networking and storage device type, software (type/version), applications, sensor/device hardware specification, sensor measurements values and data type, software signature, device driver/firmware, process properties, code execution time, CPU performance, RTC drift, system calls and logs, process/task scheduler, protocols, physical-layer attributes, security and privacy setting, manufacturing variability, process model	DI [16, 47, 117, 118, 120], CM, MD [47, 81, 95, 120, 129, 203], FT [129, 143]
	2. Physical/Mechanical system: hardware specification, firmware type, sensor configuration, machine type, machine utilization frequency, process schedule/scheduling of the manufacturing system, surrounding conditions, operator/expert characteristics, manufacturing variability, contextual, essential processes, networking with other system/devices	DI [19, 38, 87, 117], MD [19, 27, 38, 53]
Externally collected behavioral reatures	1. Cyber: data collected through the sensor and actuators related to radiated emission, EMI, heat/thermal, sound, vibration, camera visible/IR image/video, packet header statistics, network flow statistics, signal frequency, packet payload data and statistics, network signal quality	DI [16, 44, 56, 85, 117, 118, 120, 169, 206], CM [106, 185], MD [16, 44, 81, 118, 120, 206], FT [143]
	2. Physical/Mechanical system: data collected through the sensor and actuators related to radiated emission, EMI, heat/thermal, sound, vibration, camera visible/IR image, mechanical noise, leakage of petrochemical or any other chemical materials, radiation, poisonous/hazardous materials, pressure, temperature, product quality	DI [117], CM [20, 93, 106, 113, 120, 127, 145, 163, 174, 185, 186, 195], MD [27, 53, 120, 122]

DI, device identification; CM, condition monitoring; MD, miss-behavior detection; RTC, real-time clock; EMI, electromagnetic induction; CPU, central processing unit, FT (forensic or tracking).

instance, a diesel generator transfers the chemical energy contained in the fuel to electrical energy through a diesel engine and generator (i.e., alternator). Therefore, in electrical generators, IoT-based systems use real-time resource consumption information (i.e., fuel, lubricant, and coolant) to condition their status (health) or behavioral fingerprinting [142]. However, acquiring device or machine data from legacy systems or networks requires appropriate sensing methods (hardware and application) that can monitor and communicate resource usage statistics.

3.3.3 Device Design and Its Configuration. Information collected by a CPS is highly dynamic and depends on the design and architecture of the various cyber and physical entities [102, 165]. In this context, the device designs rely on various factors, including its application and environmental

Table 12. Features Related to Various Deployment of CPS/IoT Applications Used for DFP

Type of Applications					Application domain	Physical Properties					Operational								Behavioural						Limiting Factors					
Identification	Security	Condition monitoring	Forensic	Device management control		Electrical	Magnetic	Optical	Thermal	Chemical	Load/power	Radiation-EMI	Radiation-thermal	Communication Std./protocols	Heat/thermal	Time analysis	Noise	Glitches/faults	System Logs/Device history	Resources utilization	Security and privacy	Network features	Software Applications	Hardware setting	DFP manufacturing standards	Overhead cost of gathering	Traceability and privacy	Security and privacy	Cyber act and regulation	Harsh operating conditions
✓	!	✓	!	!	Intelligent vehicles	✓	✓	!	!	!	✓	!	!	!	✓	✓	✓	✓	✓	!	✓	✓	✓	✓	X	✓	✓	!	X	X
✓	!	✓	!	✓	Manufacturing	✓	✓	✓	✓	✓	✓	✓	✓	!	✓	✓	✓	✓	✓	!	!	!	!	✓	X	X	!	!	!	X
!	X	!	X	!	Agriculture	!	!	!	!	!	!	!	!	!	X	X	!	!	!	!	X	!	!	!	!	!	!	!	!	!
✓	✓	✓	✓	✓	Smart city	!	!	!	!	!	!	!	!	!	!	!	✓	✓	!	!	✓	!	!	!	X	✓	✓	!	X	X
!	!	X	X	!	Education	X	X	!	X	X	!	X	X	X	X	!	!	!	!	!	!	!	!	!	!	!	!	!	!	✓
✓	!	✓	!	!	Healthcare	!	!	!	!	!	!	!	!	!	!	!	!	!	!	!	!	!	!	!	X	X	X	X	X	!
✓	!	✓	!	!	Remote sensing	!	!	!	✓	!	!	!	!	!	!	✓	!	✓	!	!	!	!	!	!	!	!	!	!	!	X
✓	✓	✓	!	✓	Aviation and space	!	✓	✓	✓	✓	✓	✓	✓	!	!	✓	✓	✓	!	✓	✓	!	!	✓	X	!	!	X	!	X
✓	✓	✓	✓	!	Defence	✓	✓	✓	✓	✓	✓	✓	✓	!	✓	✓	✓	✓	✓	✓	✓	!	!	✓	!	!	!	!	!	X

Color code symbols	X	X	X	!	!	!	✓	✓	✓
Level of favourable–>	Minimum								Maximum

aspects [54, 206]. Inaccuracies in the manufacturing process may have a negative effect not only on the performance and efficiency of different physical devices but also on the behavior and functions of such devices [61]. For example, an observed behavior is drift in a system, which is caused by the defect in the timing circuit of the hardware introduced during device manufacturing. Consequently, the time required to execute a particular code or function differs between devices and can be used to forecast the system's behavior [34, 47, 165]. Among other valuable features, the software and processes running on a device or system are used as valuable sources for device identification and classification [25]. Additionally, the CPS is integrated with several sensors and actuators for controlling and monitoring machines and physical systems. Therefore, each machine, system, or device in a CPS has a unique set of sensors and actuators optimized for its specific purpose and operating environment [106]. Thus, in such a setting, it is crucial to comprehensively understand all relevant environmental factors to evaluate any device or system effectively. Therefore, whenever it comes to identifying and classifying CPS devices, it is done by detecting abnormalities in the readings produced by the sensors and actuators [4, 5, 61, 106].

3.3.4 Externally Collected Behavioral Features. In this category, an external device monitors the behavior of the device. Therefore, network communications, emitted electromagnetic signals, radiated thermal radiation, and so on are the primary externally aggregated sources used to model device behavior. In the case of network-based data, the data is usually collected by a proxy or gateway, whereas a sensor collects electromagnetic signal based data through an antenna. At the same time, the radiated thermal radiation is collected by the onboard or external temperature sensor. In Section 3.2, we covered the operational characteristics of CPS devices or machines, and how external sources of equipment characteristics can be used to identify their behavioral characteristics. This information can be valuable in identifying, classifying, and detecting threats related to the device.

To summarize, Table 12 compares the main features of data directly collected from the modeled device. Hardware performance counters, CPU percentages, system calls, software signatures, and sensor readings are used for device identification and misbehavior detection. Additionally, low-level information about system CPUs and sensor hardware is solely used for device identification. As well, the expanded feature sources are based on their generality, including network communications, hardware events, resource usages, and software and processes. Finally, resource use and process attribute characteristics are only used in anomaly identification.

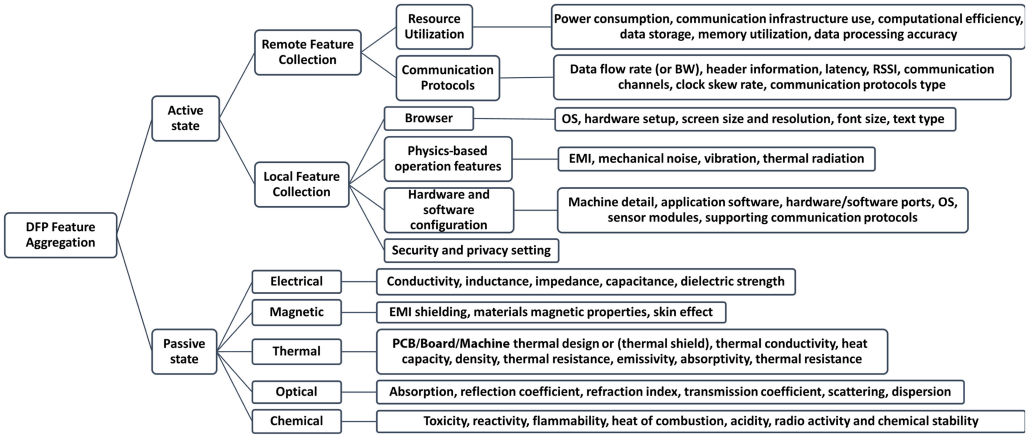

Fig. 7. Taxonomy of the DFP's features collection techniques or sources.

4 DFP FEATURE AGGREGATION

This section explores the different device feature gathering methods for DFPs, highlighting their advantages and disadvantages. A detailed classification of device feature aggregation is shown in Figure 7. Depending on the state of the fingerprinting device during fingerprinting or data acquisition, DFP feature extraction is performed in one of two different modes: active (the device is operational) or passive (the device is not operational) mode [56]. Consequently, in an active state, a cyber or physical device is operational (powered) and communicates with other network entities. However, these devices do not interact with other network-connected modules in the passive state (not powered). This happens when a device (hardware or software) is malfunctioning, non-functional, or in standby mode.

4.1 Active State

The active state of the device includes a CPS node that actively participates in network activities such as data transmission (receiving and transmitting) between multiple nodes. Therefore, the node must be powered by a battery or other energy source. Active mode is also known as operational mode because the device performs specific tasks, such as sensor data acquisition, processing, and transmission. The physical objects, including electrical and mechanical machinery, have not communicated directly with the cyber system even when active or operational. As a result, the sensory system continuously monitors the operational characteristics of these devices (e.g., power consumption, electromagnetic interference, heat radiation, vibration, and noises) and transmits this information to the network. Active state devices can be further subdivided into two categories based on how active fingerprinting technology is implemented. These categories are local and remote sensing facilities.

4.1.1 Remote Feature Collection. Remote feature collection approaches are widely used in modern communication networks to collect the device's unique features for their fingerprint. For example, web-based DFP techniques utilize web application APIs or scripts to collect remote device details such as the operating system, hardware, and operational characteristics [104]. We further classified DFP capturing methods based on resource usage, communication standards/protocols, and other operational by-products (i.e., radiation-EMI or thermal, heat dissipation, sound, vibration, and noises) based on local device feature generation techniques:

- *Resource utilization*: Section 3.3.2 demonstrates how the operating environment influences the device's resource use. Consequently, fingerprinting will consider the resource utilization characteristics of the device in the network, such as power consumption, communication bandwidth (analog or digital), communication infrastructure requirements, memory utilization, computational efficiency, and data processing accuracy. By analyzing household power consumption patterns, Michalevsky et al. [119] get the profile of home appliance usage. Similarly, the bandwidth use and data rate are analyzed to identify and monitor a device on a wireless network [206]. Additionally, transitory information about wireless networks may be utilized to determine their operation and behavior.
- *Communication protocols*: Over time, to address the complexity and needs of the CPS, researchers have developed different communication standards (or protocols) for CPSs to ensure the smooth operation of network activities [1, 176, 200]. These communication protocols have addressed device discovery, data processing, device management, semantics, and many other network-related tasks. Fingerprinting has eventually made use of this operating protocol's features. For example, the clock skew rate, data flow rate (or bandwidth), packet analysis-header information, session cookies information, latency, the received signal strength indicator, and web user-agent attributes are all used to generate its unique characteristics [44, 49, 73, 86, 104, 154, 191, 198].

4.1.2 Local Sensing Features. Local sensing technology gathers device functional attributes and configuration details in the local system using scripts, software modules, applications, or specific APIs. Traditional antivirus operations, for instance, depend on known malware signatures because antivirus software examines the disk for the digital signatures of software and computing equipment for dangerous intent and compares them to a database of known malware signatures [133]. Similarly, browser fingerprints are used to control the content of online applications by executing scripts on the local system to obtain information about the operating system, hardware setup, screen resolution, screen size, security and privacy setup, and many other factors [25, 104, 112, 188]. Additionally, connected sensor networks are used for physics-based operational feature collection. The sensor network is made up of many different sensors that measure things like thermal radiation, electromagnetic interference, noise, vibration, sound, efficiency, and energy/fuel consumption, and they are listed in Table 7. In local sensing, we collect device features or information in an aggressive or non-invasive manner. In invasive technology, the sensor acts as an integral part of the same host system or circuit board to obtain complete information about the operational physics-based characteristics of the device. For example, an integrated temperature sensor is connected to the host circuitry to collect the device's operating temperature. However, in non-invasive approaches, the sensing unit is isolated from the host circuitry (or device) and maintains an air gap. For example, thermal cameras are installed on factory floors to record the thermal signature of machinery.

Moreover, in a CPS, sensing techniques are developed for particular applications; as a result, the data collected by sensors is quite diverse, inconsistent, and application dependent in terms of data structures. For instance, the precision level of temperature monitoring is lower for a weather station than for an industrial process since it depends on the application's severity and relevance.

4.2 Passive State

DFPs are formed in a passive state by their physics-based attributes, such as their electrical, magnetic, and optical features, as we covered in Section 3. For their collection, specific types of sensors or devices are required, such as multimeters (i.e., voltmeter + ohmmeter + ammeter), electromagnetic field sensors or magnetometers (measure the magnetism of a material), and the photodetector or photosensor (light and radiation).

5 USE CASES, CHALLENGES, AND FUTURE RESEARCH DIRECTION

5.1 Use Cases

With IoTiaztion, IoT/CPS-based applications are growing in healthcare, manufacturing, transportation, security, smart grid, building management, defense, trade, supply chain, and more, from home appliances to industrial automation [10, 109]. The growing demand for such applications is due to the fact that it improves system performance. Therefore, to provide a better understanding of DFP applications, this section presents a concise explanation of CPSs' five popular use cases. This includes smart homes, smart power grids, medical and hospital facilities, smart factories (or Industry 4.0), and intelligent transportation systems.

5.1.1 Smart Home. A *smart home* is a modern home that includes appliances, lighting, infotainment systems, security systems, and other electronic devices that can be remotely controlled by the owner through a mobile application [35, 44]. In a smart home, different devices can communicate through the cyber layer by utilizing the embedded computing and communication unit. The intelligent security system in a smart home consists of digital locks, cameras, mobile applications, cloud, and communication systems distributed throughout the home. As a result, it requires a robust security solution to protect the distributed unit from physical and cyber threats. The system based on fingerprints could be used for device authentication and status monitoring to protect the system from possible threats [88, 149]. Furthermore, condition monitoring can be used to predict the health of home appliances.

5.1.2 Smart Grid. A *smart grid* is an autonomous power distribution system that uses automation, communication, and information technology capabilities to communicate, monitor, and analyze the entire process starting from power generation to its consumption [103]. A power grid system is a massive network that consists of transmission lines, microgrids, isolators, transformers, breakers, and other electrical components, hence monitoring and identifying any fault or malicious activity in the associated components is a challenging task. In such cases, fingerprinting-based methods can be utilized as a solution. Soltan et al. [180] discuss a smart grid attack scenario, where the adversary hacks the power-consuming appliances (e.g., motors, air conditioners, water heaters, washing machines, and refrigerators) to generate large amounts of electricity that cause fake loads. This results in the disruption of the grid's service due to sudden increase of such generated power demand. In these situations, fingerprinting-based approaches may be used as a probable solution. DFP-based identification and monitoring algorithms may predict, detect, and prevent malfunctions and malicious activity in the interconnected elements of the power distribution networks and energy-hungry equipment [90, 103, 127, 153, 177]. For example, DFP-based applications utilize the transformer's characteristics, such as load dynamics (peak load, average load, and percentage of time overload), location-related information, and instrument rating, to predict the transformer's health condition [90].

5.1.3 Smart Factory. A *smart factory* in the context of the Industrial IoT is a concept that views a production environment as a fully automated and intelligent technology network. This technology network allows the manufacturing facility's infrastructure, equipment, and supply chains to be handled automatically [10, 204]. This next era of industrial digital technology comprises CPSs (e.g., sensors, actuators, and computer networks), cloud computing, ML and DL, cognitive computing, and real-time data. In a smart factory, DFP-based approaches may be used to identify devices, monitor their conditions for predictive maintenance, and detect and track the malicious activity. For example, predictive maintenance applications gather information about a machine, such as its temperature, vibration, mechanical sound, motion, power consumption, electromagnetic radiation, and environmental characteristics, to forecast, detect, and monitor the health of the associated

devices [113, 134, 145, 174, 177]. Furthermore, DFP may be utilized to find malicious and fake network devices, which would stop data breaches and protect critical infrastructure [5, 192].

5.1.4 Smart Hospital. With the IoTization of the healthcare system, healthcare providers use a variety of medical solutions, including wearables, smart beds, biosensors, smart pills, blood pressure monitoring devices, glucose measurement devices, real-time health systems, equipment monitoring devices, remote monitoring devices, and many more. The development of IoT-enabled healthcare devices will also introduce a different type of security and privacy risk [2, 33, 50, 55]. Device tampering and counterfeiting are two such security concerns [55]. With the advancement of counterfeiting techniques, traditional device identification techniques cannot distinguish between original and fake devices [101]. Therefore, more sophisticated techniques for device identification are required. As such, the fingerprinting process applied to various devices may be seen as a means of identifying these devices. Implantable medical devices such as cardiac pacemakers, cardiac defibrillators, and insulin pumps pose complex challenges in device health monitoring and patient safety. Early detection of such equipment failure can save lives. In these cases, DFP techniques can be used to estimate the health status of implementable devices, allowing for predictive maintenance. Device predictive maintenance also applies to the equipment used in the healthcare system for screening, monitoring, or diagnosing.

5.1.5 Intelligent Transportation. An *intelligent transportation system* is equipped with the sophisticated IoT applications. Modern ICT is used in transportation and traffic management systems to make traffic safer and more efficient to reduce traffic congestion [172]. Recent advances in intelligent transportation systems, vehicle identification, monitoring, and vehicle conditioning present significant challenges. Timely prediction of vehicle system failure or fault can sometimes prevent unwanted traffic congestion. Information about the vehicle's status and surroundings can also be used to find its location, track it, map it, and predict traffic and route quality. With the advancement of 5G and IoT systems, the vehicle can communicate with any other vehicle or infrastructure. Therefore, we need a robust security platform that can quickly and easily check the authenticity of vehicles and allow them to talk. The current method of detection is always under the attacker's radar [54, 81, 138]. It will be harder to do because of recent improvements in cloning devices that make it easier to spot fakes [101]. In that case, the device may use internal DFP-based technologies for identification, authentication, and condition monitoring.

5.2 Challenges

In this section, we review the challenges that limit DFP use in CPSs.

5.2.1 Lack of a Standard for Feature Selection and Matching. In the CPS, feature extraction, classification, quality assessment, and matching are complex tasks due to the system's complexity, diversity, and massive size. As a result, CPSs do not have a single standard technique that can be used for feature extraction, categorization, and quality evaluation in DFP. A few evaluation methods for measuring stability were reviewed in Section 2.2. However, these methods are insufficient to provide a suitable platform for all DFP systems in CPSs. Without the correct matching technique, a system cannot differentiate between fake and real devices based on their DFPs [56, 100, 121, 193, 196].

5.2.2 Overhead Cost of Features Acquisition Technology or Its Module. The use of DFP in the IoT and CPS covers a wide range of tasks from feature collection to processing to fingerprint generation. Therefore, it require a suitable size of data storage and a computing facility to collect and process the aggregated data to be used for fingerprint applications such as device identification and conditioning monitoring. It also requires additional software patches and a scripting language

that supports feature generation and collection for DFP. Additionally, fingerprinting devices need a reliable communication system to interface with remote client systems to communicate in any situation. For example, it requires a high bandwidth connection so that the data collected in bulk can be transferred from one node to another node within a suitable time interval. Therefore, the fingerprint generation technique requires additional costs for hardware, software, and fingerprint generation applications [152].

5.2.3 Security and Privacy. During DFP generation, data privacy and security are of the most significant concerns in CPSs as clients share their personal information, such as hardware, software, and operational and security information with third-party remote devices or servers [168, 181]. The server utilizes shared features to generate client DFPs for their traceability, localization, or security applications. Therefore, security and privacy depend entirely on the third-party device. Once that device is malicious or compromised, the client's details are no longer secure [92, 115, 126, 181]. Additionally, as data analytics methods like ML and DL continue to expand, it becomes simpler for attackers to get confidential information through behavioral analysis of the functional and networking features of CPS devices [14, 74, 75, 115, 157, 164, 181, 205, 206]. As a result, there is a greater need for reliable methods of protecting users' personal information and data on these devices.

5.2.4 Cyber Acts and Regulation. Stable rules and regulations foster any technologies' growth because they create strong trust for that device or technologies. Therefore, the lack of regulation for DFP generation and uses is a hurdle in the growth of many applications based on DFP. The Cookie law was recently implemented in the European Union, mandating that all websites clearly disclose to visitors whether they need to use this tracking technology [97]. However, this solely addresses how browser apps handle cookie data. Therefore, there must be stringent laws and regulations for data management, privacy, and security of CPSs, which all stakeholders will enforce (i.e., users, manufacturers/vendors, and regulators) in the event of any malpractice or failure [168].

5.2.5 Harsh Operating Conditions and Aging. CPS are commonly subjected to extreme operating conditions that can significantly impact their performance and reliability. These conditions include exposure to a range of environmental factors, such as temperature, humidity, pressure, and vibration, as well as various types of hazardous radiation, including X-rays, radioactive waves, and infrared/thermal radiation. Additionally, CPS may also encounter electrical, magnetic, or electromagnetic fields, as well as acoustic and electromagnetic noise [60, 101, 199]. DFP stability is impacted by the device or equipment used, its working circumstances, and its applications. CPS applications, such as smart cities, intelligent transportation systems, intelligent manufacturing, intelligent power grids, intelligent defense, and medical equipment, and networks are dispersed and operate in harsh environments [54, 104, 143, 188, 206]. Over time, the physics-based characteristics of the CPS devices will deteriorate due to prolonged exposure to harsh working conditions. Some authors [101, 199] reported a bit-flip error caused by aging effects. Additionally, as distributed systems, attackers utilize tampering to expose a system, including a computer device, network, database, sensor, actuator, and server, to steal an asset or obtain unauthorized access [101]. As well, CPSs are vulnerable to various physical threats, including device tampering, RF interference, node jamming, malicious node jamming, and physical damage, among others [46, 96, 159].

5.3 Future Direction

5.3.1 Development of Efficient Techniques for Device Feature Extraction. Most of the electrical, electronic, and mechanical equipment in use today is outdated and cannot keep up with the

demands of the 21st century regarding processing power, network speeds, and data storage capacity. Researchers face a significant challenge in creating hardware and software modules that are compatible with one another to integrate legacy systems with cutting-edge CPS architectures. Developing a separate interface for each system is both resource intensive and inconvenient. Therefore, researchers working in CPSs need to focus on developing devices that can work with both modern and legacy systems simultaneously.

5.3.2 DFP Feature Analysis Techniques. There are many sources of noise and artifacts in extracted device features in their raw format. These include sensor abnormalities, power fluctuations, sudden changes in operation conditions, miscalibrated or biased sensors, and glitches, among others. Furthermore, the feature data structures that are collected will differ from one device to another depending on the type of sensor, the working principle of the sensor, and the vendor's specification. As a result, the collected device features must undergo feature processing to reduce the impact of noise and artifacts before being used in the DFP application. Typically, feature analysis employs rule-based digital signal processing and data analysis methods. However, the relationship between device features and DFP is sometimes unknown. AI, ML, and DL are all data-driven methods that can be used to determine the connection. Furthermore, data-driven approaches may be used for a wide range of applications, including data processing, feature engineering, classification, pattern recognition, abnormality, and failure prediction, among others.

5.3.3 Smart Sensor Nodes. A device or sensor node capable of independently gathering and processing data is referred to as an intelligent device or sensor node. This alleviates pressure on remote data centers by decreasing the amount of data that needs to be stored, processed, and transmitted over long distances. The ability of a system to make decisions independently on a local level can increase its autonomy, security, stability, and performance. Recent developments in distributed systems, such as mobile computing, fog computing, and edge computing, have increased the possibility that this dream will become a reality. Building an AI-enabled embedded system takes time and effort due to the constraints imposed by factors such as computation facilities, memory, storage, protection, and battery life.

5.3.4 Standards for Device Feature Selection and Matching. CPS refers to a network of electronic devices, machinery, and physical things linked together via a cyber system such as a cloud, server, computer network, and communication medium. DFP is responsible for collecting and processing data obtained from a wide variety of CPS entities. Therefore, finding suitable device features for consistent and novel DFP creation is difficult, as it depends on feature sources, selection, and processing methods. Currently, no method is generally accepted for determining where the feature originated. As a result, establishing a standard for feature selection in CPS contexts required the participation of all relevant stakeholders, including clients, researchers, policymakers, and developers.

5.3.5 Applying DFP Techniques to Uncover the Relationship Between Device Features with Device Conditions. Recent advances in sensing technology have made it possible to collect detailed data on the properties of a CPS device. In most instances, the relationship between device characteristics and the device's physical properties, operating conditions, and behavioral use is known. Unfortunately, this is not known in the majority of cases. Therefore, it is crucial to comprehend the connection between the device's characteristics and its current condition. Device attribute based algorithms offer a tremendous amount of potential for establishing a relationship between device attribute information and the device's condition, even if a direct connection between the two is not currently known to exist. Therefore, developing AI/ML-powered algorithms that can be

integrated into CPS devices with limited hardware and computational resources is an exciting area for research.

5.3.6 Favorable Ecosystem. In CPSs, deployment of DFP needs favorable infrastructure with the following capabilities:

- *Data storage and computing facilities*: Such facilities require a system that can store and process a huge amount of sensor data (i.e., DFP features) efficiently.
- *Cyber acts and regulation*: A persistent cyber act and a globally uniform, fair, and equal law are required for a given application. In addition to this, it outlines the responsibilities that are held by the various stakeholders in the process of developing, installing, and employing DFP systems in the face of potential risks to the scientific community, the industrial sector, and the legal system for violating a law, regulation, or contract.
- *Encouragement of interdisciplinary research activities*: DFP is generated using features aggregated from CPS elements (cyber and physical systems). A specialized sensing system and an interface are required to connect DFP modules to cyber (sensors, actuators, electronic and networking devices) and physical systems (mechanical machinery and equipment, living and non-living objects). Designing an effective DFP system will require a multidisciplinary capability that is able to interface with various components of the CPS. As a result, there will be a need for innovative sensing systems and interfaces (hardware and software) in the future to accelerate the development of DFP in CPSs.
- *Physical security*: CPSs are examples of distributed systems in which cyber and physical objects exchange data over the Internet. As a result, it is simple to execute both physical attacks and manipulation of devices. To overcome this challenge, a safe and secure working environment will be required to ensure the physical security of all equipment, networks, and objects.

5.3.7 Developing Interfaces (Hardware or Software) to Use Legacy Devices. Equipment based on obsolete or legacy technologies has limited facilities that do not support cutting-edge technologies. Therefore, it is important to develop appropriate hardware, software, and physical systems to facilitate a flawless interface between the two generations of users and devices.

5.3.8 Development of a New Computation or Interfacing System for Inter- or Intra-CPSs. Network topology, storage, communication systems, communication media, dataset types, software languages, security, privacy, and authentication procedures and methods vary between two different CPSs. In this case, an interfacing platform would be required to transfer and execute any DFP application from one CPS to another or within the same CPS.

6 OUTCOMES OF THIS SURVEY

This review focuses on four fundamental questions that need to be answered before using DFP in a CPS, as described in Section 1. The aforementioned circular approach, 4-Q, is depicted in Figure 2. In the first step of this approach, the application scenario of DFP in CPS deployment will be identified. After that, it determines various constraints associated with DFP implementation in the associated context, then it chooses an appropriate feature and its acquisition techniques. Table 12 presents a comprehensive picture of the DFP system for CPS applications, and it summarizes various characteristics, applications, deployment scenarios, and limiting factors of DFPs to provide quick answers to the questions in 4-Q. This will assist researchers and developers in efficiently answering the 4-Q questions. The features in DFP that are appropriate for a given field of application are represented by different colors and symbols.

The use of DFP in manufacturing is somewhat different from its applications in other fields, such as smart cities, intelligent transportation, and healthcare. The working environment setup (hardware and software), legal limits, and security processes in the manufacturing sector are different compared to those for other applications. In such conditions, DFPs might be used for identification, security, status monitoring, and control and administration of manufacturing floor machinery. In general, forensic applications are rarely used in manufacturing because manufacturing premises are protected that are not open to public activities. Therefore, in DFP feature selection in Table 12, these applications are highlighted in light red to suggest that they are the least desirable in the manufacturing sector. Likewise, crosses (X), exclamation marks (!), and checkmarks (✓) are used to show the level of recommendation for a DFP feature, such as "not recommended," "considerable," and "recommended."

When DFPs are generated in manufacturing, factors of the machine's physics are considered. Operational characteristics like bandwidth are not recommended parameters in DFP applications because the network quality in manufacturing operations is usually very good and meets industry standards. However, behavioral characteristics such as security and privacy are also limited influence since industrial activities include equipment and machinery with enough physical protection on the factory floor. In this setting, the machinery or equipment may be operated and supervised only by a person who has received appropriate training and authorization. Therefore, the DFP application may be deployed in an industrial environment with the least effort. Additionally, DFP applications in manufacturing operations are affected by variables such as a lack of proper regulation, device deployment costs, and severe working conditions.

As well, there are some other DFP use cases in which these feature selection criteria may not be strictly enforced. In these instances, when developing DFP applications, it is feasible to reconsider the device's characteristics and generate a suitable DFP selection matrix by examining the device's attributes.

7 CONCLUSION

This survey summarized existing hardware or software internal features based on security and its applications for CPS that may be used for DFP. Additionally, the study presented a novel classification of DFPs, their feature sources, and feature aggregation methods. Under DFP, we briefly described the different DFP characteristics, their measuring scale, and the necessary similarity metrics for fingerprint matching. The study consisted of a novel classification that describes specific device feature sources along a three-dimensional axis. The classification is based on the device's physics, operation, and behavioral features. It is more suitable for CPS than other techniques currently in use because it explains several characteristics of the CPS related to cyber, physical, and operational domains. To further describe the aggregation approach of DFP characteristics, we chose a second categorization scheme that divides CPS nodes into active (powered) and passive (un-powered) states depending on their involvement in network activities. At the end of the survey, a DFP feature selection chart was introduced, which will help research and development professionals by providing answers to four key questions concerning DFP deployment and its applications.

However, technical and regulatory constraints impede the development of DFP systems. For example, technical constraints include a complex heterogeneous network of CPSs with no common standard for quality assessment, legacy system modernization overhead cost, security and privacy, multi-dimensional and multi-variable data, highly dynamic and heterogeneous scenarios, and so on, whereas regulatory constraints include cyber acts and regulations for data management, privacy, and security, and an ecosystem that provides sufficient physical security to access the CPS remotely. Hence, this study emphasized the different constraints that affect the development of

DFP systems. Additionally, we outlined future research topics on the CPS that are essential for developing the DFP solution. This will be useful for researchers and developers in shaping the direction of their research effort swiftly to overcome the issues linked to CPSs that impede the expansion of DFP.

REFERENCES

[1] Postscapes. 2019. IoT Standards and Protocols (Creative Commons License (CC BY-NC-SA 4.0). Retrieved September 19, 2019 from https://www.postscapes.com/internet-of-things-protocols/.

[2] Robab Abdolkhani, Kathleen Gray, Ann Borda, and Ruth DeSouza. 2019. Patient-generated health data management and quality challenges in remote patient monitoring. *JAMIA Open* 2, 4 (2019), 471–478.

[3] Chuadhry Mujeeb Ahmed, Aditya Mathur, and Martin Ochoa. 2017. NoiSense: Detecting data integrity attacks on sensor measurements using hardware based fingerprints. *arXiv:1712.01598* [cs.CR] (2017).

[4] Chuadhry Mujeeb Ahmed and Aditya P. Mathur. 2017. Hardware identification via sensor fingerprinting in a cyber physical system. In *Proceedings of the 2017 IEEE International Conference on Software Quality, Reliability, and Security Companion (QRS-C'17)*. IEEE, Los Alamitos, CA, 517–524.

[5] Chuadhry Mujeeb Ahmed, Martin Ochoa, Jianying Zhou, Aditya P. Mathur, Rizwan Qadeer, Carlos Murguia, and Justin Ruths. 2018. *NoisePrint*: Attack detection using sensor and process noise fingerprint in cyber physical systems. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*. ACM, New York, NY, 483–497.

[6] Chuadhry Mujeeb Ahmed, Jianying Zhou, and Aditya P. Mathur. 2018. Noise matters: Using sensor and process noise fingerprint to detect stealthy cyber attacks and authenticate sensors in CPS. In *Proceedings of the 34th Annual Computer Security Applications Conference*. ACM, New York, NY, 566–581.

[7] A. T. Al Ghazo and R. Kumar. 2019. ICS/SCADA device recognition: A hybrid communication-patterns and passive-fingerprinting approach. In *Proceedings of the 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM'19)*. 19–24.

[8] Nabil El Belghiti Alaoui, Patrick Tounsi, Alexandre Boyer, and Amaud Viard. 2019. Detecting PCB assembly defects using infrared thermal signatures. In *Proceedings of the 26th International Conference on Mixed Design of Integrated Circuits and Systems (MIXDES'19)*. IEEE, Los Alamitos, CA, 345–349.

[9] Ramón Alcarria, Borja Bordel, Diego Martín, and Diego Sánchez De Rivera. 2017. Rule-based monitoring and coordination of resource consumption in smart communities. *IEEE Transactions on Consumer Electronics* 63, 2 (2017), 191–199.

[10] V. Alcácer and V. Cruz-Machado. 2019. Scanning the Industry 4.0: A literature review on technologies for manufacturing systems. *Engineering Science and Technology* 22, 3 (2019), 899–919. https://doi.org/10.1016/j.jestch.2019.01.006

[11] A. A. Alkahtani, S. M. Norzeli, and F. H. Nordin. 2019. Condition monitoring through temperature vibration and radio frequency emission. *Test Engineering and Management* 800 (2019), 5621–5636.

[12] Muhammad N. Aman, Kee Chaing Chua, and Biplab Sikdar. 2016. Position paper: Physical unclonable functions for IoT security. In *Proceedings of the 2nd ACM International Workshop on IoT Privacy, Trust, and Security*. ACM, New York, NY, 10–13.

[13] Irene Amerini, Rudy Becarelli, Roberto Caldelli, Alessio Melani, and Moreno Niccolai. 2017. Smartphone fingerprinting combining features of on-board sensors. *IEEE Transactions on Information Forensics and Security* 12, 10 (2017), 2457–2466.

[14] Vafa Andalibi. 2022. *Leveraging Machine Learning for End-User Security and Privacy Protection*. Ph. D. Dissertation. Indiana University.

[15] Elbren Antonio, Arnel Fajardo, and Ruji Medina. 2020. Tracking browser fingerprint using rule based algorithm. In *Proceedings of the 2020 16th IEEE International Colloquium on Signal Processing and Its Applications (CSPA'20)*. IEEE, Los Alamitos, CA, 225–229.

[16] Alberico Aramini, Marco Arazzi, Tullio Facchinetti, Laurence S. Q. N. Ngankem, and Antonino Nocera. 2022. An enhanced behavioral fingerprinting approach for the Internet of Things. In *Proceedings of the 2022 IEEE 18th International Conference on Factory Communication Systems (WFCS'22)*. IEEE, Los Alamitos, CA, 1–8.

[17] Maroua Ben Attia, Chamseddine Talhi, Abdelwahab Hamou-Lhadj, Babak Khosravifar, Vincent Turpaud, and Mario Couture. 2015. On-device anomaly detection for resource-limited systems. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing*. 548–554.

[18] Martin Azizyan and Romit Roy Choudhury. 2009. SurroundSense: Mobile phone localization using ambient sound and light. *ACM SIGMOBILE Mobile Computing and Communications Review* 13, 1 (2009), 69–72.

[19] Leonardo Babun, Hidayet Aksu, and A. Selcuk Uluagac. 2021. CPS device-class identification via behavioral fingerprinting: From theory to practice. *IEEE Transactions on Information Forensics and Security* 16 (2021), 2413–2428.

[20] Subramaniam Bagavathiappan, B. B. Lahiri, T. Saravanan, John Philip, and T. Jayakumar. 2013. Infrared thermography for condition monitoring—A review. *Infrared Physics & Technology* 60 (2013), 35–55.
[21] Bruhadeshwar Bezawada, Indrakshi Ray, and Indrajit Ray. 2021. Behavioral fingerprinting of Internet-of-Things devices. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 11, 1 (2021), e1337.
[22] Laksh Bhatia, Michael Breza, Ramona Marfievici, and Julie A. McCann. 2020. LoED: The LoRaWAN at the edge dataset: Dataset. In *Proceedings of the 3rd Workshop on Data: Acquisition to Analysis*. 7–8.
[23] Ariel Bogle. 2014. A Cyber Attack May Have Caused a Turkish Oil Pipeline to Catch Fire in 2008. Retrieved January 30, 2021 from https://slate.com/technology/2014/12/bloomberg-reports-a-cyber-attack-may-have-made-a-turkish-oil-pipeline-catch-fire.html.
[24] Li Cai and Yangyong Zhu. 2015. The challenges of data quality and data quality assessment in the big data era. *Data Science Journal* 14 (2015), 2.
[25] Yinzhi Cao, Song Li, and Erik Wijmans. 2017. (Cross-)browser fingerprinting via OS and hardware level features. In *Proceedings of the 2017 Network and Distributed System Security Symposium (NDSS'17)*. 1–15.
[26] Metehan Cekic, Soorya Gopalakrishnan, and Upamanyu Madhow. 2020. Wireless fingerprinting via deep learning: The impact of confounding factors. *arXiv preprint arXiv:2002.10791* (2020).
[27] Alberto Huertas Celdrán, Pedro Miguel Sánchez Sánchez, Gérôme Bovet, Gregorio Martínez Pérez, and Burkhard Stiller. 2022. CyberSpec: Intelligent behavioral fingerprinting to detect attacks on crowdsensing spectrum sensors. *arXiv preprint arXiv:2201.05410* (2022).
[28] Alberto Huertas Celdrán, Pedro Miguel Sánchez Sánchez, Miguel Azorín Castillo, Gérôme Bovet, Gregorio Martínez Pérez, and Burkhard Stiller. 2022. Intelligent and behavioral-based detection of malware in IoT spectrum sensors. *International Journal of Information Security* 2022 (2022), 1–21.
[29] Dajiang Chen, Ning Zhang, Zhen Qin, Xufei Mao, Zhiguang Qin, Xuemin Shen, and Xiang-Yang Li. 2016. S2M: A lightweight acoustic fingerprints-based wireless device authentication protocol. *IEEE Internet of Things Journal* 4, 1 (2016), 88–100.
[30] Long Cheng, Ke Tian, Danfeng Daphne Yao, Lui Sha, and Raheem A. Beyah. 2019. Checking is believing: Event-aware program anomaly detection in cyber-physical systems. *IEEE Transactions on Dependable and Secure Computing* 18, 2 (2019), 825–842.
[31] Yushi Cheng, Xiaoyu Ji, Juchuan Zhang, Wenyuan Xu, and Yi-Chao Chen. 2019. DeMiCPU: Device fingerprinting with magnetic signals radiated by CPU. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS'19)*. 1149–1170.
[32] Kyong-Tak Cho and Kang G Shin. 2016. Fingerprinting electronic control units for vehicle intrusion detection. In *Proceedings of the 25th USENIX Security Symposium (USENIX Security'16)*. 911–927.
[33] Rajarshi Roy Chowdhury and Pg Emeroylariffion Abas. 2022. A survey on device fingerprinting approach for resource-constraint IoT devices: Comparative study and research challenges. *Internet of Things* 20 (2022), 100632.
[34] W. E. Cobb, E. W. Garcia, M. A. Temple, R. O. Baldwin, and Y. C. Kim. 2010. Physical layer identification of embedded devices using RF-DNA fingerprinting. In *Proceedings of the 2010 Military Communications Conference (MILCOM'20)*. 2168–2173. https://doi.org/10.1109/MILCOM.2010.5680487
[35] Javier Criado, José Andrés Asensio, Nicolás Padilla, and Luis Iribarne. 2018. Integrating cyber-physical systems in a component-based approach for smart homes. *Sensors* 18, 7 (2018), 2156.
[36] Frankie A. Cruz. 2019. *Near Real-Time RF-DNA Fingerprinting for ZigBee Devices Using Software Defined Radios*. Technical Report. Air Force Institute of Technology, Wright-Patterson AFB, OH.
[37] Ang Cui. 2015. funtenna. Retrieved February 28, 2023 from https://github.com/funtenna.
[38] Yesheng Cui, Sami Kara, and Ka C. Chan. 2020. Manufacturing big data ecosystem: A systematic literature review. *Robotics and Computer-Integrated Manufacturing* 62 (2020), 101861.
[39] Anupam Das, Nikita Borisov, and Matthew Caesar. 2014. Do you hear what I hear? Fingerprinting smart devices through embedded acoustic components. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM, New York, NY, 441–452.
[40] Anupam Das, Nikita Borisov, and Matthew Caesar. 2016. Tracking mobile web users through motion sensors: Attacks and defenses. In *Proceedings of the 2016 Network and Distributed System Security Symposium (NDSS'16)*.
[41] José Paulo G. de Oliveira, Carmelo J. A. Bastos-Filho, and Sergio Campello Oliveira. 2022. Non-invasive embedded system hardware/firmware anomaly detection based on the electric current signature. *Advanced Engineering Informatics* 51 (2022), 101519.
[42] Jyoti Dhiman. 2021. Is Machine Learning the Future of Data Quality? Retrieved November 9, 2022 from https://towardsdatascience.com/have-you-started-using-machine-learning-for-data-quality-yet-c0136e0957ac.
[43] Kadriye Dogan and Ozlem Durmaz Incel. 2019. Mobile device identification via user behavior analysis. In *Proceedings of the International Conference on Big Data Innovations and Applications*. 32–46.

[44] Shuaike Dong, Zhou Li, Di Tang, Jiongyi Chen, Menghan Sun, and Kehuan Zhang. 2019. Your smart home can't keep a secret: Towards automated fingerprinting of IoT traffic with neural networks. *arXiv preprint arXiv:1909.00104* (2019).

[45] Dwight T. Dumpert. 1994. Infrared techniques for printed circuit board (PCB) evaluation. In *Infrared Methodology and Technology*. Gordon and Breach Science Publishers, Great Britain, 253–264.

[46] Mohammed El-Hajj, Ahmad Fadlallah, Maroun Chamoun, and Ahmed Serhrouchni. 2019. A survey of Internet of Things (IoT) authentication schemes. *Sensors* 19, 5 (2019), 1141.

[47] Omar E. Elejla, Bahari Belaton, Mohammed Anbar, and Basem O. Alijla. 2017. IPv6 OS fingerprinting methods: Review. In *Advances in Visual Informatics*, Halimah Badioze Zaman, Peter Robinson, Alan F. Smeaton, Timothy K. Shih, Sergio Velastin, Tada Terutoshi, Azizah Jaafar, and Nazlena Mohamad Ali (Eds.). Springer International Publishing, Cham, Switzerland, 661–668.

[48] Abdurrahman Elmaghbub and Bechir Hamdaoui. 2021. LoRa device fingerprinting in the wild: Disclosing RF data-driven fingerprint sensitivity to deployment variability. *IEEE Access* 9 (2021), 142893–142909.

[49] Xinlei Fan, Gaopeng Gou, Cuicui Kang, Junzheng Shi, and Gang Xiong. 2019. Identify OS from encrypted traffic with TCP/IP stack fingerprinting. In *Proceedings of the 2019 IEEE 38th International Performance Computing and Communications Conference (IPCCC'19)*. IEEE, Los Alamitos, CA, 1–7.

[50] Xiaotao Feng, Xiaogang Zhu, Qing-Long Han, Wei Zhou, Sheng Wen, and Yang Xiang. 2023. Detecting vulnerability on IoT device firmware: A survey. *IEEE/CAA Journal of Automatica Sinica* 10, 1 (2023), 25–41.

[51] Roman Ferrando and Paul Stacey. 2017. Classification of device behaviour in Internet of Things infrastructures: Towards distinguishing the abnormal from security threats. In *Proceedings of the 1st International Conference on Internet of Things and Machine Learning*. 1–7.

[52] Micro Focus. 2019. Device fingerprinting for low friction authentication. *White Paper, NETIQ*. Retrieved February 28, 2023 from https://www.netiq.com/docrep/documents/x6pyuysuqu/device_fingerprinting_for_low_friction_authentication_wp.pdf.

[53] National Security Staff Interagency Policy Coordination Subcommittee for Preparedness, Response to Radiological, and Nuclear Threats. 2021. *Planning Guidance for Response to a Nuclear Detonation* (3rd ed.). Draft.

[54] David Formby, Preethi Srinivasan, Andrew Leonard, Jonathan Rogers, and Raheem A. Beyah. 2016. Who's in control of your control system? Device fingerprinting for cyber-physical systems. In *Proceedings of the 2016 Network and Distributed System Security Symposium (NDSS'16)*.

[55] Michele Forzley. 2003. *Counterfeit Goods and the Public's Health and Safety*. Technical Report. International Intellectual Property Institute.

[56] Ke Gao, Cherita Corbett, and Raheem Beyah. 2010. A passive approach to wireless device fingerprinting. In *Proceedings of the 2010 IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'10)*. IEEE, Los Alamitos, CA, 383–392.

[57] Soorya Gopalakrishnan, Metehan Cekic, and Upamanyu Madhow. 2019. Robust wireless fingerprinting via complex-valued neural networks. In *Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM'19)*. IEEE, Los Alamitos, CA, 1–6.

[58] Richard R. Goulette, Stanilus K. Xavier, and Raymond L. Greenfield. 1991. Method and apparatus for monitoring electromagnetic emission levels. US Patent 5,006,788.

[59] Christopher Greer, Martin Burns, David Wollman, and Edward Griffor. 2019. *Cyber-Physical Systems and Internet of Things*. NIST.

[60] Qinchen Gu, David Formby, Shouling Ji, Hasan Cam, and Raheem Beyah. 2018. Fingerprinting for cyber-physical system security: Device physics matters too. *IEEE Security & Privacy* 16, 5 (2018), 49–59.

[61] Anton Gulenko, Marcel Wallschläger, Florian Schmidt, Odej Kao, and Feng Liu. 2016. A system architecture for real-time anomaly detection in large-scale NFV systems. *Procedia Computer Science* 94 (2016), 491–496.

[62] Mordechai Guri, Dima Bykhovsky, and Yuval Elovici. 2017. aIR-Jumper: Covert air-gap exfiltration/infiltration via security cameras & infrared (IR). *arXiv:1709.05742* [cs.CR] (2017).

[63] Mordechai Guri, Andrey Daidakulov, and Yuval Elovici. 2018. MAGNETO: Covert channel between air-gapped systems and nearby smartphones via CPU-generated magnetic fields. *arXiv:1802.02317* [cs.CR] (2018).

[64] Mordechai Guri, Assaf Kachlon, Ofer Hasson, Gabi Kedma, Yisroel Mirsky, and Yuval Elovici. 2015. GSMem: Data exfiltration from air-gapped computers over GSM frequencies. In *Proceedings of the 24th USENIX Security Symposium (USENIX Security'15)*. 849–864.

[65] Mordechai Guri, Gabi Kedma, Assaf Kachlon, and Yuval Elovici. 2014. AirHopper: Bridging the air-gap between isolated networks and mobile phones using radio frequencies. *arXiv:1411.0237* [cs.CR] (2014).

[66] Mordechai Guri, Matan Monitz, and Yuval Elovici. 2016. USBee: Air-gap covert-channel via electromagnetic emission from USB. *arXiv:1608.08397* [cs.CR] (2016).

[67] Mordechai Guri, Matan Monitz, Yisroel Mirski, and Yuval Elovici. 2015. BitWhisper: Covert signaling channel between air-gapped computers using thermal manipulations. *arXiv:1503.07919* [cs.CR] (2015).

[68] Mordechai Guri, Yosef Solewicz, Andrey Daidakulov, and Yuval Elovici. 2016. Fansmitter: Acoustic data exfiltration from (speakerless) air-gapped computers. *arXiv:1606.05915* [cs.CR] (2016).
[69] Mordechai Guri, Boris Zadov, Dima Bykhovsky, and Yuval Elovici. 2018. PowerHammer: Exfiltrating data from air-gapped computers through power lines. *arXiv:1804.04014* [cs.CR] (2018).
[70] Mordechai Guri, Boris Zadov, Dima Bykhovsky, and Yuval Elovici. 2019. CTRL-ALT-LED: Leaking data from air-gapped computers via keyboard LEDs. In *Proceedings of the 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC'19)*, Vol. 1. IEEE, Los Alamitos, CA, 801–810.
[71] Mordechai Guri, Boris Zadov, Andrey Daidakulov, and Yuval Elovici. 2017. xLED: Covert data exfiltration from air-gapped networks via router LEDs. *arXiv:1706.01140* [cs.CR] (2017).
[72] Mordechai Guri, Boris Zadov, Andrey Daidakulov, and Yuval Elovici. 2018. ODINI : Escaping sensitive data from Faraday-caged, air-gapped computers via magnetic fields. *arXiv:1802.02700* [cs.CR] (2018).
[73] Jose A. Gutierrez del Arroyo, Brett J. Borghetti, and Michael A. Temple. 2022. Considerations for radio frequency fingerprinting across multiple frequency channels. *Sensors* 22, 6 (2022), 2111.
[74] Adam Haavik. 2021. *Deep Learning-Based Traffic Classification for Network Penetration Testing*. Master's Thesis. Karlstads University.
[75] Muhammad Shadi Hajar, M. Omar Al-Kadri, and Harsha Kumara Kalutarage. 2021. A survey on wireless body area networks: Architecture, security challenges and research opportunities. *Computers & Security* 104 (2021), 102211.
[76] Jason R. Hamlet, Mitchell T. Martin, and Nathan J. Edwards. 2017. Unique signatures from printed circuit board design patterns and surface mount passives. In *Proceedings of the 2017 International Carnahan Conference on Security Technology (ICCST'17)*. IEEE, Los Alamitos, CA, 1–6.
[77] Ayyoob Hamza, Hassan Habibi Gharakheili, Theophilus A. Benson, and Vijay Sivaraman. 2019. Detecting volumetric attacks on lot devices via SDN-based monitoring of mud activity. In *Proceedings of the 2019 ACM Symposium on SDN Research*. 36–48.
[78] S. He and S. G. Chan. 2016. Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons. *IEEE Communications Surveys Tutorials* 18, 1 (Firstquarter 2016), 466–490. https://doi.org/10.1109/COMST.2015.2464084
[79] Sheng-Jen Hsieh. 2001. Thermal signature for printed circuit board stress-failure diagnosis. In *Thermosense XXIII*, Vol. 4360. International Society for Optics and Photonics, 60–70.
[80] Jingyu Hua, Hongyi Sun, Zhenyu Shen, Zhiyun Qian, and Sheng Zhong. 2018. Accurate and efficient wireless device fingerprinting using channel state information. In *Proceedings of the 2018 IEEE INFOCOM Conference on Computer Communications (INFOCOM'18)*. IEEE, Los Alamitos, CA, 1700–1708.
[81] Abdulmalik Humayed, Jingqiang Lin, Fengjun Li, and Bo Luo. 2017. Cyber-physical systems security—A survey. *IEEE Internet of Things Journal* 4, 6 (2017), 1802–1831.
[82] Mark Hung. 2017. *Leading the IoT, Gartner Insights on How to Lead in a Connected World*. Gartner Research.
[83] Taswar Iqbal and Kai-Dietrich Wolf. 2017. PCB surface fingerprints based counterfeit detection of electronic devices. *Electronic Imaging* 2017, 7 (2017), 144–149.
[84] Fehmi Jaafar. 2017. An integrated architecture for IoT fingerprinting. In *Proceedings of the 2017 IEEE International Conference on Software Quality, Reliability, and Security Companion (QRS-C'17)*. IEEE, Los Alamitos, CA, 601–602.
[85] Hossein Jafari, Oluwaseyi Omotere, Damilola Adesina, Hsiang-Huang Wu, and Lijun Qian. 2018. IoT devices fingerprinting using deep learning. In *Proceedings of the 2018 IEEE Military Communications Conference (MILCOM'18)*. IEEE, Los Alamitos, CA, 1–9.
[86] Anu Jagannath, Jithin Jagannath, and Prem Sagar Pattanshetty Vasanth Kumar. 2022. A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges. *Computer Networks* 219 (2022), 109455.
[87] Zhongfeng Jin, Nan Li, Chao Liu, Meimei Li, Shaohua An, and Weiqing Huang. 2021. A contextual and content features-based device behavioral fingerprinting method in smart grid. In *Proceedings of the 2021 IEEE 23rd International Conference on High Performance Computing and Communications, the 7th International Conference on Data Science and Systems, the 19th International Conference on Smart City, and the 7th International Conference on Dependability in Sensor, Cloud, and Big Data Systems and Application (HPCC/DSS/SmartCity/DependSys'21)*. IEEE, Los Alamitos, CA, 415–422.
[88] Arun Cyril Jose, Reza Malekian, and Ning Ye. 2016. Improving home automation security; integrating device fingerprinting into smart home. *IEEE Access* 4 (2016), 5776–5787.
[89] Khurum Nazir Junejo and David Yau. 2016. Data driven physical modelling for intrusion detection in cyber physical systems. In *Proceedings of the Singapore Cyber-Security Conference (SG-CRC'16)*. 43–57.
[90] Farzana Kabir, Brandon Foggo, and Nanpeng Yu. 2018. Data driven predictive maintenance of distribution transformers. In *Proceedings of the 2018 China International Conference on Electricity Distribution (CICED'18)*. IEEE, Los Alamitos, CA, 312–316.

[91] Anastasis Keliris and Michail Maniatakos. 2016. Remote field device fingerprinting using device-specific modbus information. In *Proceedings of the 2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS'16)*. IEEE, Los Alamitos, CA, 1–4.

[92] Sye Loong Keoh. 2016. Cyber-physical systems are at risk. *Infosecurity Magazine*. Retrieved July 25, 2020 from https://www.infosecurity-magazine.com/next-gen-infosec/cyberphysical-systems-risk-1/.

[93] Qasim Khan, Asfar A. Khan, and Furkan Ahmad. 2016. Condition monitoring tool for electrical equipment–Thermography. In *Proceedings of the 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT'16)*. IEEE, Los Alamitos, CA, 2802–2806.

[94] Raghbir Singh Khandpur. 2005. *Printed Circuit Boards: Design, Fabrication, Assembly and Testing*. McGraw-Hill Education.

[95] Sangjun Kim, Kyung-Joon Park, and Chenyang Lu. 2022. A survey on network security for cyber-physical systems: From threats to resilient design. *IEEE Communications Surveys & Tutorials* 24, 3 (2022), 1534–1573.

[96] Eric D. Knapp and Joel Thomas Langill. 2014. *Industrial Network Security: Securing Critical Infrastructure Networks for Smart Grid, SCADA, and Other Industrial Control Systems*. Syngress.

[97] Anna Kobusinska, Jerzy Brzezinski, and Kamil Pawulczuk. 2017. Device fingerprinting: Analysis of chosen fingerprinting methods. In *Proceedings of the 2nd International Conference on Internet of Things, Big Data, and Security (IoTBDS'17)*. 167–177.

[98] Andrzej Kochański. 2010. Data preparation. *Computer Methods in Materials Science* 10, 1 (2010), 25–29.

[99] Tadayoshi Kohno, Andre Broido, and Kimberly C. Claffy. 2005. Remote physical device fingerprinting. *IEEE Transactions on Dependable and Secure Computing* 2, 2 (2005), 93–108.

[100] Sven Kosub. 2019. A note on the triangle inequality for the Jaccard distance. *Pattern Recognition Letters* 120 (2019), 36–38.

[101] Vijav Kumar and Kolin Paul. 2021. DevFing: Robust LCR based device fingerprinting. In *Proceedings of the 2021 10th Mediterranean Conference on Embedded Computing (MECO'21)*. IEEE, Los Alamitos, CA, 1–6.

[102] Andreas Kurtz, Hugo Gascon, Tobias Becker, Konrad Rieck, and Felix Freiling. 2016. Fingerprinting mobile devices using personalized configurations. *Proceedings on Privacy Enhancing Technologies* 2016, 1 (2016), 4–19.

[103] Abraham Peedikayil Kuruvila, Ioannis Zografopoulos, Kanad Basu, and Charalambos Konstantinou. 2021. Hardware-assisted detection of firmware attacks in inverter-based cyberphysical microgrids. *International Journal of Electrical Power & Energy Systems* 132 (2021), 107150.

[104] Pierre Laperdrix, Nataliia Bielova, Benoit Baudry, and Gildas Avoine. 2019. Browser fingerprinting: A survey. *arXiv preprint arXiv:1905.01051* (2019).

[105] Sam Leroux, Steven Bohez, Pieter-Jan Maenhaut, Nathan Meheus, Pieter Simoens, and Bart Dhoedt. 2018. Fingerprinting encrypted network traffic types using machine learning. In *Proceedings of the 2018 IEEE/IFIP Network Operations and Management Symposium (NOMS'18)*. IEEE, Los Alamitos, CA, 1–5.

[106] Daoliang Li, Ying Wang, Jinxing Wang, Cong Wang, and Yanqing Duan. 2020. Recent advances in sensor fault diagnosis: A review. *Sensors and Actuators A: Physical* 309 (2020), 111990.

[107] Fudong Li, Nathan Clarke, Maria Papadaki, and Paul Dowland. 2014. Active authentication for mobile devices utilising behaviour profiling. *International Journal of Information Security* 13, 3 (2014), 229–244.

[108] Chia-Te Liao, Wen-Hao Lee, and Shang-Hong Lai. 2012. A flexible PCB inspection system based on statistical learning. *Journal of Signal Processing Systems* 67, 3 (2012), 279–290.

[109] Libelium. 2019. 50 Sensor Applications for a Smarter World. Retrieved September 25, 2020 from http://www.libelium.com/resources/top_50_iot_sensor_applications_ranking/.

[110] Weisong Liu, Xueqiong Li, Zhitao Huang, and Xiang Wang. 2021. Transmitter fingerprinting for VLC systems via deep feature separation network. *IEEE Photonics Journal* 13, 6 (2021), 1–7.

[111] Xindan Liu, Ying Zhang, Menghua Wu, Zhiguo Ma, and Hui Cao. 2020. Color discrimination and gas chromatography-mass spectrometry fingerprint based on chemometrics analysis for the quality evaluation of Schizonepetae Spica. *PLOS One* 15, 1 (2020), e0227235.

[112] Zengrui Liu, Prakash Shrestha, and Nitesh Saxena. 2021. Gummy browsers: Targeted browser spoofing against state-of-the-art fingerprinting techniques. *arXiv preprint arXiv:2110.10129* (2021).

[113] Amine Mahami, Chemseddine Rahmoune, Toufik Bettahar, and Djamel Benazzouz. 2021. Induction motor condition monitoring using infrared thermography imaging and ensemble learning techniques. *Advances in Mechanical Engineering* 13, 11 (2021), 16878140211060956.

[114] C. Mak, M. A. Hon, W. Lau, and W. Cheung. 2014. Refined Wi-Fi fingerprinting with tag-less proximity-based positioning technique. In *Proceedings of the 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN'14)*. 758–761. https://doi.org/10.1109/IPIN.2014.7275560

[115] Georgios Michail Makrakis, Constantinos Kolias, Georgios Kambourakis, Craig Rieger, and Jacob Benjamin. 2021. Industrial and critical infrastructure security: Technical analysis of real-life security incidents. *IEEE Access* 9 (2021), 165295–165325.

[116] Giuseppe Manco, Ettore Ritacco, Pasquale Rullo, Lorenzo Gallucci, Will Astill, Dianne Kimber, and Marco Antonelli. 2017. Fault detection and explanation through big data analysis on sensor streams. *Expert Systems with Applications* 87 (2017), 141–156.

[117] Noman Mazhar, Rosli Salleh, Muhammad Zeeshan, and M. Muzaffar Hameed. 2021. Role of device identification and manufacturer usage description in IoT security: A survey. *IEEE Access* 9 (2021), 41757–41786.

[118] K. Merchant, S. Revay, G. Stantchev, and B. Nousain. 2018. Deep learning for RF device fingerprinting in cognitive communication networks. *IEEE Journal of Selected Topics in Signal Processing* 12, 1 (Feb. 2018), 160–167. https://doi.org/10.1109/JSTSP.2018.2796446

[119] Yan Michalevsky, Aaron Schulman, Gunaa Arumugam Veerapandian, Dan Boneh, and Gabi Nakibly. 2015. Powerspy: Location tracking using mobile device power analysis. In *Proceedings of the 24th USENIX Security Symposium (USENIX Security'15)*. 785–800.

[120] M. Miettinen, S. Marchal, I. Hafeez, N. Asokan, A. Sadeghi, and S. Tarkoma. 2017. IoT SENTINEL: Automated device-type identification for security enforcement in IoT. In *Proceedings of the 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS'17)*. 2177–2184. https://doi.org/10.1109/ICDCS.2017.283

[121] Georgy Minaev, Ari Visa, and Robert Piché. 2017. Comprehensive survey of similarity measures for ranked based location fingerprinting algorithm. In *Proceedings of the 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN'17)*. IEEE, Los Alamitos, CA, 1–4.

[122] Yisroel Mirsky, Mordechai Guri, and Yuval Elovici. 2017. HVACKer: Bridging the air-gap by attacking the air conditioning system. *arXiv:1703.10454* [cs.CR] (2017).

[123] R. Keith Mobley. 2002. *An Introduction to Predictive Maintenance*. Elsevier.

[124] John V. Monaco. 2022. Device fingerprinting with peripheral timestamps. In *Proceedings of the 2022 IEEE Symposium on Security and Privacy (SP'22)*. IEEE, Los Alamitos, CA, 1018–1033.

[125] László Monostori, Botond Kádár, Thomas Bauernhansl, Shinsuke Kondoh, Soundar Kumara, Gunther Reinhart, Olaf Sauer, Gunther Schuh, Wilfried Sihn, and Kenichi Ueda. 2016. Cyber-physical systems in manufacturing. *CIRP Annals* 65, 2 (2016), 621–641.

[126] Chuadhry Mujeeb Ahmed and Jianying Zhou. 2020. Challenges and opportunities in CPS security: A physics-based perspective. *arXiv e-prints arXiv:2004.03178* (2020).

[127] Ehtasham Mustafa, Ramy S. A. Afia, and Zoltán Ádám Tamus. 2020. Condition monitoring uncertainties and thermal-radiation multistress accelerated aging tests for nuclear power plant cables: A review. *Periodica Polytechnica Electrical Engineering and Computer Science* 64, 1 (2020), 20–32.

[128] Jithendra P. R. Nayak, K. Anitha, and P. Rashmi. 2017. PCB fault detection using image processing. In *IOP Conference Series: Materials Science and Engineering*, Vol. 225. IOP Publishing, 012244.

[129] James Newsome, Elaine Shi, Dawn Song, and Adrian Perrig. 2004. The Sybil attack in sensor networks: Analysis & defenses. In *Proceedings of the 2004 3rd International Symposium on Information Processing in Sensor Networks (IPSN'04)*. IEEE, Los Alamitos, CA, 259–268.

[130] Oliver Niggemann, Gautam Biswas, John S. Kinnebrew, Hamed Khorasgani, Sören Volgmann, and Andreas Bunte. 2015. Data-driven monitoring of cyber-physical systems leveraging on big data and the Internet-of-Things for diagnosis and control. In *Proceedings of the International Workshop on the Principles of Diagnosis (DX@ Safeprocess'15)*. 185–192.

[131] Nick Nikiforakis, Alexandros Kapravelos, Wouter Joosen, Christopher Kruegel, Frank Piessens, and Giovanni Vigna. 2013. Cookieless monster: Exploring the ecosystem of web-based device fingerprinting. In *Proceedings of the 2013 IEEE Symposium on Security and Privacy (SP'13)*. IEEE, Los Alamitos, CA, 541–555.

[132] Nitesh Nitesh Varma Rudraraju and Varun Varun Boyanapally. 2019. *Data Quality Model for Machine Learning*. Master's Thesis. Blekinge Institute of Technology.

[133] NortonLifeLock. 2021. What is antivirus software? *Norton*. Retrieved August 9, 2021 from https://us.norton.com/internetsecurity-malware-what-is-antivirus.html.

[134] Elena Nyemkova. 2017. Invariants of noise in cyber-physical systems components. *Advances in Cyber-Physical Systems* 2, 2 (2017), 63–70.

[135] Roque Alfredo Osornio-Rios, Jose Alfonso Antonino-Daviu, and Rene de Jesus Romero-Troncoso. 2018. Recent industrial applications of infrared thermography: A review. *IEEE Transactions on Industrial Informatics* 15, 2 (2018), 615–625.

[136] Karl Ott and Rabi Mahapatra. 2019. Continuous authentication of embedded software. In *Proceedings of the 2019 18th IEEE International Conference on Trust, Security, and Privacy in Computing and Communications and the 13th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE'19)*. IEEE, Los Alamitos, CA, 128–135.

[137] Volker Paelke and Carsten Röcker. 2015. User interfaces for cyber-physical systems: Challenges and possible approaches. In *Design, User Experience, and Usability: Design Discourse*. Springer, 75–85.

[138] P. Paganini. 2016. 150,000 IoT Devices Behind the 1Tbps DDoS Attack on OVH. Retrieved February 28, 2023 from http://securityaffairs.co/wordpress/51726/cyber-crime/ovh-hit-botnet-iot.html.
[139] Arpan Pal, Hemant Kumar Rath, Samar Shailendra, and Abhijan Bhattacharyya. 2018. IoT standardization: The road ahead. In *Internet of Things: Technology, Applications and Standardization*, Jaydip Sen (Ed.). IntechOpen, 53.
[140] Steven Paley, Tamzidul Hoque, and Swarup Bhunia. 2016. Active protection against PCB physical tampering. In *Proceedings of the 2016 17th International Symposium on Quality Electronic Design (ISQED'16)*. IEEE, Los Alamitos, CA, 356–361.
[141] Jonathan Pan. 2021. IoT network behavioral fingerprint inference with limited network traces for cyber investigation. In *Proceedings of the 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC'21)*. IEEE, Los Alamitos, CA, 263–268.
[142] Michael G. Pecht and Myeongsu Kang. 2019. Predictive maintenance in the IoT era. In *Proceedings of the Cortana Analytics Workshop*.
[143] Emmanuel S. Pilli, R. C. Joshi, and Rajdeep Niyogi. 2010. Network forensic frameworks: Survey and research challenges. *Digital Investigation* 7, 1 (2010), 14–27. https://doi.org/10.1016/j.diin.2010.02.003
[144] Ivan Miguel Pires, Rui Santos, Nuno Pombo, Nuno M. Garcia, Francisco Flórez-Revuelta, Susanna Spinsante, Rossitza Goleva, and Eftim Zdravevski. 2018. Recognition of activities of daily living based on environmental analyses using audio fingerprinting techniques: A systematic review. *Sensors* 18, 1 (2018), 160.
[145] Peter Popaleny and Nicolas Péton. 2019. Cryogenic pumps monitoring, diagnostics and expert systems using motor current signature analyses and vibration analyses. In *Fluids Engineering Division Summer Meeting*, Vol. 59056. American Society of Mechanical Engineers.
[146] PROTRONIX. 2016. Operating Principles of Air Quality Sensors. Retrieved February 28, 2023 from https://www.careforair.eu/en/on-what-principles-do-air-quality-sensors-work/.
[147] Quan Qian, Jing Cai, Mengbo Xie, and Rui Zhang. 2016. Malicious behavior analysis for Android applications. *International Journal of Network Security* 18, 1 (2016), 182–192.
[148] Soundarya Ramesh, Thomas Pathier, and Jun Han. 2019. SoundUAV: Towards delivery drone authentication via acoustic noise fingerprinting. In *Proceedings of the 5th Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*. 27–32.
[149] V. P. Rasheeda. 2020. Improving home automation security integrating device fingerprinting into smart home. *International Journal of Research and Analytical Reviews* 7, 1 (2020), 858–863.
[150] M. Raspopoulos, C. Laoudias, L. Kanaris, A. Kokkinis, C. G. Panayiotou, and S. Stavrou. 2012. 3D ray tracing for device-independent fingerprint-based positioning in WLANs. In *Proceedings of the 2012 9th Workshop on Positioning, Navigation, and Communication*. 109–113. https://doi.org/10.1109/WPNC.2012.6268748
[151] Michael Rausch, Andrew Bakke, Suzanne Patt, Beth Wegner, and David Scott. 2014. Demonstrating a simple device fingerprinting system. In *Proceedings of the Midwest Instruction and Computing Symposium*.
[152] Partha Pratim Ray. 2018. A survey on Internet of Things architectures. *Journal of King Saud University—Computer and Information Sciences* 30, 3 (2018), 291–319.
[153] Jeffrey H. Reed and Carlos R. Aguayo Gonzalez. 2012. Enhancing smart grid cyber security using power fingerprinting: Integrity assessment and intrusion detection. In *Proceedings of the 2012 Future of Instrumentation International Workshop (FIIW'12)*. IEEE, Los Alamitos, CA, 1–3.
[154] Fatemeh Rezaei. 2021. *Design and Implementation of Algorithms for Traffic Classification*. Ph.D. Dissertation. University of Massachusetts Amherst.
[155] M. Rezaei and M. Gholami. 2014. The recognition chemicals in fingerprints by gas chromatography/mass spectrometry. *Scientific Quarterly of Research on Addiction* 8, 30 (2014), 69–80.
[156] Shamnaz Riyaz, Kunal Sankhe, Stratis Ioannidis, and Kaushik Chowdhury. 2018. Deep learning convolutional neural networks for radio identification. *IEEE Communications Magazine* 56, 9 (2018), 146–152.
[157] Florentin Rochet and Tariq Elahi. 2022. Towards flexible anonymous networks. *arXiv preprint arXiv:2203.03764* (2022).
[158] Miguel A. Rodríguez-López, Luis M. López-González, Luis M. López-Ochoa, and Jesús Las-Heras-Casas. 2018. Methodology for detecting malfunctions and evaluating the maintenance effectiveness in wind turbine generator bearings using generic versus specific models from SCADA data. *Energies* 11, 4 (2018), 746.
[159] Srikanth R. P. 2018. How device fingerprinting can help industrial control systems ward off cyber attacks from hackers. *Express Computer*. Retrieved February 28, 2023 from https://www.expresscomputer.in/news/how-device-fingerprinting-can-help-industrial-control-systems-ward-off-cyber-attacks-from-hackers/16237/.
[160] Miraqa Safi, Sajjad Dadkhah, Farzaneh Shoeleh, Hassan Mahdikhani, Heather Molyneaux, and Ali A. Ghorbani. 2022. A survey on IoT profiling, fingerprinting, and identification. *ACM Transactions on Internet of Things* 3, 4 (2022), Article 26, 39 pages.
[161] Tara Salman and Raj Jain. 2015. Networking protocols and standards for Internet of Things. In *Internet of Things and Data Analytics Handbook*. Wiley Telecom, 215–238.

[162] Areeg Samir and Claus Pahl. 2020. Detecting and localizing anomalies in container clusters using Markov models. *Electronics* 9, 1 (2020), 64.

[163] Hadi Sanati, David Wood, and Qiao Sun. 2018. Condition monitoring of wind turbine blades using active and passive thermography. *Applied Sciences* 8, 10 (2018), 2004.

[164] Pedro Miguel Sánchez Sánchez, Jose María Jorquera Valero, Alberto Huertas Celdrán, Gérôme Bovet, Manuel Gil Pérez, and Gregorio Martínez Pérez. 2021. A survey on device behavior fingerprinting: Data sources, techniques, application scenarios, and datasets. *IEEE Communications Surveys & Tutorials* 23, 2 (2021), 1048–1077.

[165] Iskander Sanchez-Rola, Igor Santos, and Davide Balzarotti. 2018. Clock around the clock: Time-based device fingerprinting. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. ACM, New York, NY, 1502–1514.

[166] Akshay A. Sarawade and Nadir N. Charniya. 2018. Infrared thermography and its applications: A review. In *Proceedings of the 2018 3rd International Conference on Communication and Electronics Systems (ICCES'18)*. IEEE, Los Alamitos, CA, 280–285.

[167] Akshay A. Sarawade and Nadir N. Charniya. 2019. Detection of faulty integrated circuits in PCB with thermal image processing. In *Proceedings of the 2019 International Conference on Nascent Technologies in Engineering (ICNTE'19)*. IEEE, Los Alamitos, CA, 1–6.

[168] Mohamed Seliem, Khalid Elgazzar, and Kasem Khalil. 2018. Towards privacy preserving IoT environments: A survey. *Wireless Communications and Mobile Computing* 2018, 2 (2018), 15.

[169] Yaman Sharaf-Dabbagh and Walid Saad. 2016. On the authentication of devices in the Internet of Things. In *Proceedings of the 2016 IEEE 17th International Symposium on a World of Wireless, Mobile, and Multimedia Networks (WoWMoM'16)*. IEEE, Los Alamitos, CA, 1–3.

[170] Yaman Sharaf-Dabbagh and Walid Saad. 2017. Demo abstract: Cyber-physical fingerprinting for Internet of Things authentication. In *Proceedings of the 2017 IEEE/ACM 2nd International Conference on Internet-of-Things Design and Implementation (IoTDI'17)*. IEEE, Los Alamitos, CA, 301–302.

[171] Pawan Kumar Sharma, Jaspreet Singh, Yogita, and Vipin Pal. 2021. Low power communication in wireless sensor networks and IoT. In *Smart Sensor Networks Using AI for Industry 4.0*, Soumya Ranjan Nayak, Biswa Mohan Sahoo, Muthukumaran Malarvel, and Jibitesh Mishra. CRC Press, Boca Raton, FL, 221–233.

[172] Shivani Sharma and Sateesh Kumar Awasthi. 2022. Introduction to intelligent transportation system: Overview, classification based on physical architecture, and challenges. *International Journal of Sensor Networks* 38, 4 (2022), 215–240.

[173] Timothy W. Sheen, Jiann-Neng Chen, Stephen A. Cohen, Michael A. Baglino, and Joseph F. Wrinn. 1997. Printed circuit board tester using magnetic induction. US Patent 5,631,572.

[174] Tanvir Alam Shifat and Jang Wook Hur. 2020. An effective stator fault diagnosis framework of BLDC motor based on vibration and current signals. *IEEE Access* 8 (2020), 106968–106981.

[175] Hyeok-Ki Shin, Woomyo Lee, Jeong-Han Yun, and Hyoung-Chun Kim. 2019. Implementation of programmable CPS testbed for anomaly detection. In *Proceedings of the 12th USENIX Workshop on Cyber Security Experimentation and Test (CSET'19)*.

[176] Bhagya Nathali Silva, Murad Khan, and Kijun Han. 2018. Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities. *Sustainable Cities and Society* 38 (2018), 697–713.

[177] M. L. Sin, W. L. Soong, and Nesimi Ertugrul. 2003. Induction machine on-line condition monitoring and fault diagnosis—A survey. In *Proceedings of the Australasian Universities Power Engineering Conference*, Vol. 28. 1–6.

[178] Ishwar Singh, Dan Centea, and Mo Elbestawi. 2019. IoT, IIoT and cyber-physical systems integration in the SEPT learning factory. *Procedia Manufacturing* 31 (2019), 116–122.

[179] Matthew Smart, G. Robert Malan, and Farnam Jahanian. 2000. Defeating TCP/IP stack fingerprinting. In *Proceedings of the 9th USENIX Security Symposium (USENIX Security'00)*.

[180] Saleh Soltan, Prateek Mittal, and H. Vincent Poor. 2018. BlackIoT: IoT botnet of high wattage devices can disrupt the power grid. In *Proceedings of the 27th USENIX Security Symposium (USENIX Security'18)*. 15–32.

[181] Pietro Spadaccino, Domenico Garlisi, Francesca Cuomo, Giorgio Pillon, and Patrizio Pisani. 2022. Discovery privacy threats via device de-anonymization in LoRaWAN. *Computer Communications* 189 (2022), 1–10.

[182] Georgios Spanos, Konstantinos M. Giannoutakis, Konstantinos Votis, and Dimitrios Tzovaras. 2019. Combining statistical and machine learning techniques in IoT anomaly detection for smart homes. In *Proceedings of the 2019 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD'19)*. IEEE, Los Alamitos, CA, 1–6.

[183] Hugh F. Spence. 1994. Printed circuit board diagnosis using artificial neural networks and circuit magnetic fields. *IEEE Aerospace and Electronic Systems Magazine* 9, 2 (1994), 20–24.

[184] Hugh F. Spence, Daniel P. Burris, and Robert A. Houston. 1995. Fault detection and diagnosis for printed circuit boards. US Patent 5,440,566.

[185] Anna Vladova Stoynova and Borislav Borisov Bonev. 2017. Infrared survey in electrical preventive maintenance. In *Proceedings of the 2017 XXVI International Scientific Conference Electronics (ET'17)*. IEEE, Los Alamitos, CA, 1–4.

[186] Krzysztof Stypułkowski, Paweł Gołda, Konrad Lewczuk, and Justyna Tomaszewska. 2021. Monitoring system for railway infrastructure elements based on thermal imaging analysis. *Sensors* 21, 11 (2021), 3819.

[187] Pedro Miguel Sánchez Sánchez, Jose María Jorquera Valero, Alberto Huertas Celdrán, Gérôme Bovet, Manuel Gil Pérez, and Gregorio Martínez Pérez. 2020. A survey on device behavior fingerprinting: Data sources, techniques, application scenarios, and datasets. *arXiv:2008.03343* [cs.CR] (2020).

[188] Ko Takasu, Takamichi Saito, Tomotaka Yamada, and Takayuki Ishikawa. 2015. A survey of hardware features in modern browsers: 2015 edition. In *Proceedings of the 2015 9th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*. IEEE, Los Alamitos, CA, 520–524.

[189] Vincent F. Taylor, Riccardo Spolaor, Mauro Conti, and Ivan Martinovic. 2017. Robust smartphone app identification via encrypted network traffic analysis. *IEEE Transactions on Information Forensics and Security* 13, 1 (2017), 63–78.

[190] Vijayanand Thangavelu, Dinil Mon Divakaran, Rishi Sairam, Suman Sankar Bhunia, and Mohan Gurusamy. 2018. DEFT: A distributed IoT fingerprinting technique. *IEEE Internet of Things Journal* 6, 1 (2018), 940–952.

[191] Joaquín Torres-Sospedra, Michiel Aernouts, Adriano Moreira, and Rafael Berkvens. 2022. LoRaWAN fingerprinting with *k*-means: The relevance of clusters visual inspection. In *Proceedings of the 2022 International Conference on Localization and GNSS (ICL-GNSS'22)*.

[192] Igor Ushakov, Alexey Bogomolov, Sergey Dragan, and Sergey Soldatov. 2021. Technology for predictive monitoring of the performance of cyber-physical system operators under noise conditions. In *Society 5.0: Cyberspace for Advanced Human-Centered Society*. Springer, 269–280.

[193] M. K. Vijaymeena and K. Kavitha. 2016. A survey on similarity measures in text mining. *Machine Learning and Applications* 3, 2 (2016), 19–28.

[194] Chaitali R. Wagh and Vijay B. Baru. 2013. Detection of faulty region on printed circuit board with IR thermography. *International Journal of Scientific & Engineering Research* 4, 11 (2013), 1–4.

[195] Geoff Walker. 2022. Benefits of Model-Based Voltage and Current (MBVI) Systems. Accessed May 16, 2022 from https://www.maintenanceandengineering.com/2022/02/24/benefits-of-model-based-voltage-and-current-mbvi-systems/.

[196] Boyuan Wang, Xuelin Liu, Baoguo Yu, Ruicai Jia, and Xingli Gan. 2019. An improved WiFi positioning method based on fingerprint clustering and signal weighted Euclidean distance. *Sensors* 19, 10 (2019), 2300.

[197] Xuyu Wang, Zhitao Yu, and Shiwen Mao. 2018. DeepML: Deep LSTM for indoor localization with smartphone magnetic and light sensors. In *Proceedings of the 2018 IEEE International Conference on Communications (ICC'18)*. IEEE, Los Alamitos, CA, 1–6.

[198] Zihao Wang, Kar Wai Fok, and Vrizlynn L. L. Thing. 2022. Machine learning for encrypted malicious traffic detection: Approaches, datasets and comparative study. *Computers & Security* 113 (2022), 102542.

[199] Lingxiao Wei, Chaosheng Song, Yannan Liu, Jie Zhang, Feng Yuan, and Qiang Xu. 2015. BoardPUF: Physical unclonable functions for printed circuit board authentication. In *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*. IEEE, Los Alamitos, CA, 152–158.

[200] Kazimierz Wilk, Michał Błaszczak, and IBCH PSNC. 2020. CPS communication. *DiH4CPS*. Retrieved May 15, 2022 from https://dih4cps.eu/2020/11/30/cps-communication/.

[201] Christian Wressnegger, Guido Schwenk, Daniel Arp, and Konrad Rieck. 2013. A close look on *n*-grams in intrusion detection: Anomaly detection vs. classification. In *Proceedings of the 2013 ACM Workshop on Artificial Intelligence and Security*. 67–76.

[202] Lars Wüstrich, Lukas Schröder, and Marc-Oliver Pahl. 2021. Cyber-physical anomaly detection for ICS. In *Proceedings of the 2021 IFIP/IEEE International Symposium on Integrated Network Management (IM'21)*. IEEE, Los Alamitos, CA, 950–955.

[203] Wei Xie, Yikun Jiang, Yong Tang, Ning Ding, and Yuanming Gao. 2017. Vulnerability detection in IoT firmware: A survey. In *Proceedings of the 2017 IEEE 23rd International Conference on Parallel and Distributed Systems (ICPADS'17)*. IEEE, Los Alamitos, CA, 769–772.

[204] Hansong Xu, Wei Yu, David Griffith, and Nada Golmie. 2018. A survey on Industrial Internet of Things: A cyber-physical systems perspective. *IEEE Access* 6 (2018), 78238–78259.

[205] Kuai Xu. 2022. Research frontiers of network behavior analysis. In *Network Behavior Analysis*. Springer, 119–163.

[206] Q. Xu, R. Zheng, W. Saad, and Z. Han. 2016. Device fingerprinting in wireless networks: Challenges and opportunities. *IEEE Communications Surveys Tutorials* 18, 1 (Firstquarter 2016), 94–104. https://doi.org/10.1109/COMST.2015.2476338

[207] Geeta Yadav, Alaa Allakany, Vijay Kumar, Kolin Paul, and Koji Okamura. 2019. Penetration testing framework for IoT. In *Proceedings of the 2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI'19)*. 477–482. https://doi.org/10.1109/IIAI-AAI.2019.00104

[208] Geeta Yadav and Kolin Paul. 2021. Global monitor using spatiotemporally correlated local monitors. In *Proceedings of the 2021 IEEE 20th International Symposium on Network Computing and Applications (NCA'21)*. IEEE, Los Alamitos, CA, 1–10.

[209] Poonam Yadav, Angelo Feraudo, Budi Arief, Siamak F. Shahandashti, and Vassilios G. Vassilakis. 2020. Position paper: A systematic framework for categorising IoT device fingerprinting mechanisms. In *Proceedings of the 2nd International Workshop on Challenges in Artificial Intelligence and Machine Learning for Internet of Things*. 62–68.

[210] Zhuting Yao and Hongxia Pan. 2012. Fault diagnosis using magnetic image of PCB. In *Proceedings of the 2012 UKACC International Conference on Control*. IEEE, Los Alamitos, CA, 702–707.

[211] Siamak Yousefi, Hirokazu Narui, Sankalp Dayal, Stefano Ermon, and Shahrokh Valaee. 2017. A survey of human activity recognition using WiFi CSI. *arXiv:1708.07129* [cs.AI] (2017).

[212] A. Yves and P. Hao. 2015. RSSI-based indoor localization using RSSI-with-angle-based localization estimation algorithm. *Sensor Networks and Data Communications* 4, 122 (2015), 2.

[213] Chenbin Zhang, Ningning Qin, Yanbo Xue, and Le Yang. 2020. Received signal strength-based indoor localization using hierarchical classification. *Sensors* 20, 4 (2020), 1067.

[214] Fengchao Zhang, Andrew Hennessy, and Swarup Bhunia. 2015. Robust counterfeit PCB detection exploiting intrinsic trace impedance variations. In *Proceedings of the 2015 IEEE 33rd VLSI Test Symposium (VTS'15)*. IEEE, Los Alamitos, CA, 1–6.

[215] Jiexin Zhang, Alastair R. Beresford, and Ian Sheret. 2019. SENSORID: Sensor calibration fingerprinting for smartphones. In *Proceedings of the 40th IEEE Symposium on Security and Privacy (SP'19)*. IEEE, Los Alamitos, CA.

[216] Zhe Zhou, Wenrui Diao, Xiangyu Liu, and Kehuan Zhang. 2014. Acoustic fingerprinting revisited: Generate stable device ID stealthily with inaudible sound. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM, New York, NY, 429–440.

[217] Hank Zumbahlen (Ed.). 2011. *Linear Circuit Design Handbook*. Newnes.

Received 27 September 2021; revised 31 January 2023; accepted 14 February 2023