UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

Identical IoT device identification via hardware performance fingerprinting and Machine Learning

Identificación de dispositivos IoT idénticos mediante fingerprinting del rendimiento del hardware y Machine Learning

D. Pedro Miguel Sánchez Sánchez

2024

UNIVERSIDAD DE MURCIA
ESCUELA INTERNACIONAL DE DOCTORADO

TESIS DOCTORAL

Identical IoT device identification via hardware performance fingerprinting and Machine Learning

Identificación de dispositivos IoT idénticos mediante fingerprinting del rendimiento del hardware y Machine Learning

Autor:
Pedro Miguel Sánchez Sánchez

Directores:
Dr. **Alberto Huertas Celdrán**, *Ph.D.*
Dr. **Gregorio Martínez Pérez**, *Ph.D.*

Murcia, 2024

UNIVERSIDAD DE MURCIA

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA EN MODALIDAD DE COMPENDIO O ARTÍCULOS PARA OBTENER EL TÍTULO DE DOCTOR

Aprobado por la Comisión General de Doctorado el 19-10-2022

D./Dña. Pedro Miguel Sánchez Sánchez

doctorando del Programa de Doctorado en

Informática

de la Escuela Internacional de Doctorado de la Universidad Murcia, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada:

Identical IoT device identification via hardware performance fingerprinting and Machine Learning / Identificación de dispositivos IoT idénticos mediante fingerprinting del rendimiento del hardware y Machine Learning

y dirigida por,

D./Dña. Alberto Huertas Celdrán

D./Dña. Gregorio Martínez Pérez

D./Dña.

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita, cuando se han utilizado sus resultados o publicaciones.

Además, al haber sido autorizada como compendio de publicaciones o, tal y como prevé el artículo 29.8 del reglamento, cuenta con:

- *La aceptación por escrito de los coautores de las publicaciones de que el doctorando las presente como parte de la tesis.*
- *En su caso, la renuncia por escrito de los coautores no doctores de dichos trabajos a presentarlos como parte de otras tesis doctorales en la Universidad de Murcia o en cualquier otra universidad.*

Del mismo modo, asumo ante la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada, en caso de plagio, de conformidad con el ordenamiento jurídico vigente.

En Murcia, a 21 de Diciembre de 2023

Fdo.: Pedro Miguel Sánchez Sánchez

Esta DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD debe ser insertada en la primera página de la tesis presentada para la obtención del título de Doctor.

Información básica sobre protección de sus datos personales aportados	
Responsable:	Universidad de Murcia. Avenida teniente Flomesta, 5. Edificio de la Convalecencia. 30003; Murcia. Delegado de Protección de Datos: dpd@um.es
Legitimación:	La Universidad de Murcia se encuentra legitimada para el tratamiento de sus datos por ser necesario para el cumplimiento de una obligación legal aplicable al responsable del tratamiento. art. 6.1.c) del Reglamento General de Protección de Datos
Finalidad:	Gestionar su declaración de autoría y originalidad
Destinatarios:	No se prevén comunicaciones de datos
Derechos:	Los interesados pueden ejercer sus derechos de acceso, rectificación, cancelación, oposición, limitación del tratamiento, olvido y portabilidad a través del procedimiento establecido a tal efecto en el Registro Electrónico o mediante la presentación de la correspondiente solicitud en las Oficinas de Asistencia en Materia de Registro de la Universidad de Murcia

Para ti mamá,
te quiero

The following PhD Thesis is a compilation of the next published articles, being the PhD student the main author in all of them:

- Pedro Miguel Sánchez Sánchez, José María Jorquera Valero, Alberto Huertas Celdrán, Gérôme Bovet, Manuel Gil Pérez, Gregorio Martínez Pérez. "**A Survey on Device Behavior Fingerprinting: Data Sources, Techniques, Application Scenarios, and Datasets.**", *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1048-1077, 2021.
  DOI: 10.1109/COMST.2021.3064259
  JIF 2021: 33.84 (D1)

- Pedro Miguel Sánchez Sánchez, José María Jorquera Valero, Alberto Huertas Celdrán, Gérôme Bovet, Manuel Gil Pérez, Gregorio Martínez Pérez. "**A methodology to identify identical single-board computers based on hardware behavior fingerprinting.**", *Journal of Network and Computer Applications*, vol. 212, pp. 103579, 2023.
  DOI: 10.1016/j.jnca.2022.103579
  JIF 2022: 8.7 (D1)

- Pedro Miguel Sánchez Sánchez, José María Jorquera Valero, Alberto Huertas Celdrán, Gérôme Bovet, Manuel Gil Pérez, Gregorio Martínez Pérez. "**LwHBench: A low-level hardware component benchmark and dataset for Single Board Computers.**", *Internet of Things*, vol. 22, pp. 100764, 2023.
  DOI: 10.1016/j.iot.2023.100764
  JIF 2022: 5.9 (Q1)

- Pedro Miguel Sánchez Sánchez, Alberto Huertas Celdrán, Gérôme Bovet, Gregorio Martínez Pérez, Burkhard Stiller. "**SpecForce: A Framework to Secure IoT Spectrum Sensors in the Internet of Battlefield Things.**", *IEEE Communications Magazine*, vol. 61, no. 5, pp. 174-180, 2023.
  DOI: 10.1109/MCOM.001.2200349
  JIF 2022: 11.2 (D1)

- Pedro Miguel Sánchez Sánchez, Alberto Huertas Celdrán, Gérôme Bovet, Gregorio Martínez Pérez. "**Adversarial attacks and defenses on ML- and hardware-based IoT device fingerprinting and identification.**", *Future Generation Computer Systems*, vol. 152, pp. 30-42, 2024.
  DOI: 10.1016/j.future.2023.10.011
  JIF 2022: 7.5 (D1)

- Pedro Miguel Sánchez Sánchez, Alberto Huertas Celdrán, Gérôme Bovet, Gregorio Martínez Pérez. "**Single-board Device Individual Authentication based on Hardware Performance and Autoencoder Transformer Models.**", *Computers & Security*, vol. 137, 103596, 2024.
  DOI: 10.1016/j.cose.2023.103596
  JIF 2022: 5.6 (Q2)

Contents

Acknowledgements

It has been six years since Chema and I, while looking for a Bachelor's Thesis on cybersecurity, stumbled upon Gregorio's office by chance. Unbeknownst to me, my life was about to change with that small decision. From the Bachelor's Thesis, I moved on to the Master's Thesis, and after giving it some thought, I decided to continue with the PhD thesis. I believe that has been one of the most accurate decisions of my life, both for the "academic" growth and for the personal growth I take away from these years. Therefore, I can only start this doctoral thesis by thanking my advisors, Alberto and Gregorio, for all the support, help, and work they have dedicated to me during these years. There have been many hours spent working on articles and figuring out how to guide this research, especially at the beginning of the thesis when not everything was going as we expected. But there have been even more good hours during these years, both on the trips we have been able to go on and in the small details of the day-to-day work.

I extend these thanks to the rest of the members of the research group that has been formed over these years, the CyberDataLab (CDL). I am proud to see how the group has evolved over the years and how it will continue to advance in the following years, thanks to the effort of those of us who form it. I would especially like to highlight Manuel because, although he has not been a tutor, he has always been there throughout this process, and Chema, for all the days of work since the early years of the Bachelor's degree.

I also want to thank Gérôme, and the rest of the team at the CyberDefence Campus in armasuisse S&T, for the opportunity to carry out this thesis in collaboration with their team. As well as for the good treatment during the two stays I have had the opportunity to do in Thun.

On a personal level, I want to first thank my father, Ginés, for his unconditional love, support, and sacrifices. Also to my mother, Salvadora, who left us just when this academic journey began, but has been present every day. I know you would be as proud of me as I am to have been your son.

Finally, I want to thank all my friends for your friendship, for the moments of disconnection, and for all the times I have been able to talk your ears off about my thesis during these years. You are an important pillar in my life.

Agradecimientos

Hace ya seis años desde que, buscando un TFG sobre ciberseguridad, Chema y yo dimos de casualidad con el despacho de Gregorio. Sin saberlo, mi vida estaba a punto de cambiar con esa pequeña decisión. Del TFG pasé al TFM y después de darle unas cuantas vueltas, decidí seguir con la tesis doctoral. Creo que esa ha sido una de las decisiones más acertadas de mi vida, tanto por el crecimiento "académico" como por el crecimiento personal que me llevo de estos años. Por lo tanto, solo puedo empezar esta tesis doctoral agradeciendo a mis directores, Alberto y Gregorio, por todo el apoyo, ayuda y trabajo que han dedicado a mí durante estos años. Han sido muchas las horas trabajando en los artículos y viendo cómo encaminar esta investigación, sobre todo al principio de la tesis cuando no todo salía como esperábamos. Pero más han sido los buenos momentos durante estos años, tanto en los viajes a los que hemos podido ir como en los pequeños detalles del trabajo día a día.

Extender estos agradecimientos al resto de miembros del grupo de investigación que se ha creado durante estos años, el CyberDataLab (CDL). Es un orgullo ver como el grupo ha evolucionado estos años, y lo que va a seguir avanzando en los siguientes, gracias al esfuerzo de quienes lo formamos. En especial me gustaría remarcar a Manuel porque, aunque no ha sido director, siempre ha estado ahí en todo este proceso, y a Chema, por todos los días de trabajo desde los primeros años del grado.

También agradecer a Gérôme, y al resto del equipo del CyberDefence Campus en armasuisse S&T, por la oportunidad de realizar esta tesis en colaboración con su equipo. Así como por el buen trato durante las dos estancias que he tenido oportunidad de hacer en Thun.

A nivel personal, quiero dar las gracias en primer lugar a mi padre, Ginés, por su amor incondicional, apoyo y sacrificios. También a mi madre, Salvadora, que aunque nos dejó justo cuando empezaba todo este trayecto académico, siempre ha estado presente cada día. Sé que estarías tan orgullosa de mí como yo lo estoy de haber sido tu hijo.

Por último, agradecer a todos mis amigos por vuestra amistad, por los momentos de desconexión y por todas las veces que os he podido calentar la cabeza con mi tesis durante estos años. Sois un pilar importante en mi vida.

Abstract

I Introduction and motivation

By 2025, nearly 64 billion Internet-of-Things (IoT) devices are expected to be connected across various cutting-edge environments, such as Smart Cities, Industry 4.0, and crowdsensing [1]. The growth of IoT promises not only to enhance service quality and accessibility but also to revolutionize the user's experience, as it fosters intelligent ecosystems that significantly reduce human intervention, thus streamlining processes, cutting costs, and mitigating errors. However, the unique objectives of these environments also increase the complexity of optimizing device and service performance.

From a scenario perspective, the vast array of IoT devices employed nowadays includes Single-Board Computer (SBC) devices like Raspberry Pi (RPi), which have gained popularity due to their flexibility, affordability, extensive support, and available peripherals [2]. However, the connectivity and resource constraints of SBCs, and IoT devices, create numerous cybersecurity concerns for diverse platforms [3]. A significant problem is the presence of unauthorized devices with identical hardware and software configurations as authorized nodes, launching attacks impacting application areas such as Industry 4.0 [4], smartphones [5], or Internet of Battlefield Things (IoBT) [6]. These malicious devices can be present as a consequence of various cybersecurity threats [7], including i) device spoofing, where an attacker replaces a legitimate device with a malicious one using the same identity; ii) unauthorized device deployment, involving the installation of a new device with an unregistered identity; and iii) Sybil attack, where a malicious device uses multiple identities to mimic numerous devices. Consequently, other threats like sensitive information leakage, data poisoning, and privilege escalation and lateral movements may emerge from spoofed devices.

To solve these arising cybersecurity problems, behavioral data science has expanded from studying human behaviors [8] to modeling device behaviors [9], with a focus on creating device behavior patterns (*fingerprints*) to optimize performance and detect potential issues early [10]. Two primary IoT application scenarios for constructing device fingerprints are i) device identification and authentication at different granularity levels [11], and ii) detecting cyberattacks, malfunction, or misbehavior [12, 13]. Various studies have applied device fingerprinting to both scenarios [14, 15]. In the identification scenario, behavioral data science has significantly enhanced device identification capabilities, transcending the constraints associated with conventional methodologies that predominantly rely on the utilization of names, identifiers, labels, or tags for device recognition [16]. A critical short-

coming of traditional strategies is their susceptibility to alterations and duplications. These conventional techniques often lack the dynamism to adapt to the rapidly evolving landscape of IoT scenarios, where many devices are being deployed with high mobility. These issues are especially relevant in scenarios with an exponential increase in the quantity of devices, such as smart industries or agriculture. Therefore, device identification based on behavior fingerprinting has significantly improved upon traditional solutions by focusing on type [11], model [17], and individual [18] granularity levels. On the other hand, detecting misbehavior or malfunction due to cybersecurity issues has seen the rise of device fingerprinting as a promising solution. Numerous works create "normal" behavioral fingerprints to detect changes caused by issues such as cyberattacks, malware execution, or device malfunctioning [12, 19].

At the individual identification granularity level, hardware behavior fingerprinting stands out as a promising avenue for uniquely identifying identical devices, a necessity in today's interconnected technological landscape. Despite its potential, this domain is still burgeoning, characterized by open challenges and a lack of dedicated research specifically targeting the identification of identical single-board devices [20]. The complexity of this task is amplified by the inherent similarities between such devices, necessitating innovative and precise identification methodologies. In the broader context, for devices that are not constrained by components and resources, the existing literature advocates for the adoption of *hardware behavioral fingerprinting* [21]. This technique is instrumental in discerning minor performance variances that are a byproduct of manufacturing imperfections [22]. By meticulously analyzing these subtle differences, hardware behavioral fingerprinting provides a granular level of device characterization, paving the way for accurate device identification. However, it is important to distinguish between identification and authentication in this context. While identification involves recognizing a device from a set based on its unique characteristics, authentication goes a step further by verifying the legitimacy of the device. Authentication addresses the question of whether a device is who it claims to be, offering a layer of security that mere identification does not.

The hardware behavior analysis for device identification has been partially performed in the literature but without following a clear methodology on the steps to be done in order to achieve a successful solution [20]. Therefore, there exists a necessity for this specific methodology, rooted in the critical need to bolster cybersecurity measures in network infrastructures that incorporate identical single-board computer devices, such as Raspberry Pi. The set of these devices working in a coordinated manner is what allows to offer IoT services based on crowdsourcing or joint data collection/processing. These devices, often deployed extensively and in settings where resources are scarce, are vulnerable to a myriad of security threats [23, 24] that affect the trustworthiness of the offered services. These range from attempts at malicious device impersonation to the introduction of unauthorized devices into a network. Ensuring the ability to identify and authenticate devices precisely becomes a non-negotiable requirement, pivotal for maintaining the integrity and resilience of the service. Therefore, a consistent security level is essential in order to have trustworthy services. Implementing hardware behavior fingerprinting in these contexts demands a level of precision and innovation that transcends conventional identification methods. The challenges stem from the subtle differences between identical devices and the constraints imposed by their limited resources [25]. Addressing these challenges necessitates a concerted research effort, aimed at unraveling the complexities of hardware behavior fingerprinting, selecting the most promising hardware performance characteristics, developing robust identification algorithms, and establishing best practices for practical application.

Moreover, having a complete identification solution based on hardware is critical for

IoT security. However, in real-world scenarios, individual identification solutions do not work isolated from other cybersecurity applications. They should be integrated with other network and behavior solutions that aim to cover heterogeneous cybersecurity issues such as malware. In this context, the integration of hardware behavior fingerprinting for single-board devices becomes a strategic component of a larger security framework. Uniquely identifying and verifying each device enables a solid foundation for secure communication and data integrity. However, other tools are still required to achieve a more complete security level. This level of security is not just about protecting information; it is about ensuring the operational effectiveness and success of IoT endeavors.

To bolster the effectiveness of hardware behavior fingerprinting for identical single-board devices, it is imperative to identify the most relevant data sources to solve this issue. Then, it is required to have a comprehensive benchmark and a well-curated dataset in place [26]. These resources serve as critical tools in the verification and validation of the methodology, ensuring its accuracy, reliability, and applicability in real-world scenarios. A precise benchmark is necessary to systematically evaluate the performance of the devices, providing a standardized measure that can be used to discern the subtle variances in hardware behavior [27]. This becomes particularly crucial when dealing with identical devices, where the differences are minuscule yet significant for accurate identification. However, there are no benchmarking applications available in the literature for low-level hardware behavior fingerprinting [26]. In tandem with a robust benchmark, a rich dataset plays a vital role in the process. It acts as a repository of hardware device behavior profiles, capturing the intricacies and unique characteristics of each device. This dataset becomes a reference point, aiding in the training of algorithms and serving as a benchmark for validation. The combination of a comprehensive benchmark and a detailed dataset ensures a holistic approach to device identification, fortifying the hardware behavior fingerprinting methodology. The motivation for establishing such rigorous verification tools stems from the evolving needs of the IoT and Edge computing paradigms, where devices are increasingly interconnected, and environment integrity is paramount. In these scenarios, the ability to precisely identify and authenticate devices is not just a security measure; it is a necessity for ensuring the seamless operation of the network and the trustworthiness of the data being exchanged. By investing in the development of precise benchmarks and comprehensive datasets, the hardware behavior fingerprinting methodology can be empowered, enhancing its precision and reliability. This, in turn, paves the way for a future where identical single-board devices can be seamlessly integrated into complex networks, operating securely and efficiently, and contributing to the robustness of the digital infrastructure.

Upon acquiring the fingerprints, a riveting domain of research emerges, centered on employing optimal techniques for processing and evaluating them. While statistical methods have held a dominant position in this field for decades, the advent of Artificial Intelligence (AI), specifically Machine Learning (ML) and Deep Learning (DL), has precipitated a paradigm shift, gaining the most prominence in the solutions being developed and deployed today [28]. The complexity and diversity of IoT device behaviors necessitate sophisticated DL models capable of capturing intricate patterns and variances. Techniques like Convolutional Neural Networks (CNNs) are being adapted to discern spatial relationships in hardware fingerprints [29], while Recurrent Neural Networks (RNNs), including their more sophisticated variant, Long Short-Term Memory (LSTM) networks, are being deployed for temporal data analysis, crucial for understanding device behavior over time [17]. Autoencoders have also carved a niche in this sector, enabling dimensionality reduction for high-volume fingerprint data and anomaly detection by reconstructing normal device behavior and highlighting deviations [30]. Additionally, recent advances in Graph Neural

Networks (GNNs) allow for the incorporation of topological data, facilitating the modeling of complex relationships within networks of interconnected IoT devices. Finally, attention-based transformer models are achieving extraordinary success in language-focused tools, such as ChatGPT [31], and can also be applied to other time series data. However, the applicability of all these modern techniques is conditioned by the lack of available datasets, as they require the existence of exhaustive datasets capable of training the required models [32]. Therefore, after having enough data, the next challenge is to explore the available ML/DL methods to find which one could provide the best performance in individual device identification. Here, new network architectures can be created to better model device behavior and enhance identification and authentication capabilities.

Besides, a fully functional and trustworthy solution should consider the possible security issues intrinsic to the hardware-based individual device identification [33]. Adversarial attacks arise as one of the main threats targeting identification solutions in IoT. These sophisticated attacks present a formidable challenge, as they are specifically crafted to manipulate or evade the security mechanisms in place, potentially leading to unauthorized access, data breaches, and compromised network integrity [34, 35]. Adversarial attacks in the context of single-board device identification exploit the mechanisms designed to ensure device authenticity and network security. By subtly altering device behavior or mimicking the characteristics of legitimate devices, adversaries can deceive identification systems, resulting in misclassification or false acceptance of rogue devices. This not only undermines the trustworthiness of the network but also opens the door to further malicious activities, jeopardizing the confidentiality, integrity, and availability of data and services. Delving deeper into the nature of these adversarial attacks, there is a rising trend in context-based and ML/DL-based evasion tactics. Context-based attacks cleverly manipulate the environmental conditions or operational context of devices, aiming to distort the data used for identification and thereby mislead the system [36]. These attacks are particularly insidious as they exploit the natural variability in device behavior due to changes in external factors, making them harder to detect and counteract. On the other hand, ML/DL-based evasion attacks represent a sophisticated and calculated assault on single-board device identification systems. Adversaries employing these techniques utilize advanced ML models to learn the patterns and characteristics of legitimate devices. Armed with this knowledge, they craft adversarial samples that closely mimic authentic devices, effectively blurring the lines between legitimate and rogue devices [37]. These samples are then used to probe and deceive the identification system, leading to erroneous classifications and potentially granting unauthorized access to the network. These attacks exploit the inherent complexities and variabilities in device behavior and the advanced capabilities of ML/DL models to deceive and compromise network security. Addressing these challenges requires a comprehensive and adaptive security strategy capable of anticipating, detecting, and mitigating the myriad of threats posed by adversarial attacks.

Apart from all the work to be done in the individual identification of IoT devices context, the authentication problem is an open challenge on its own due to its more complex requirements [38]. While hardware behavioral fingerprinting offers a robust foundation for identification, extending this approach to include authentication mechanisms is vital for ensuring a higher degree of security in interconnected systems. However, the devices from the same model can exhibit very similar or even overlapping behavioral characteristics due to standardized production processes. This similarity in hardware behavior makes it challenging to distinguish between legitimate and unauthorized devices based solely on hardware behavioral data, as data distributions merge together. Therefore, authentication requires a much finer granularity of data analysis. Here, traditional processing

approaches and ML/DL techniques have not achieved remarkable results. However, new ML/DL algorithms with advanced pattern recognition capabilities, such as attention-based transformers, could improve authentication performance based on hardware performance behavior.

As a summary of the open challenges, the literature suggests that while hardware behavior fingerprinting presents a promising solution, the field is still nascent, with considerable research gaps in methodology and practice for single-board devices. The individual identification of these devices requires innovative techniques to distinguish subtle manufacturing differences, as well as precise benchmarks and extensive datasets for validating new algorithms. The evolution of AI, particularly ML and DL, is shifting the paradigm for processing and evaluating device fingerprints, and new techniques are required in this area. Besides, there remains a critical need for the integration of hardware fingerprinting with broader network security measures to address various cybersecurity threats. Next, adversarial attacks pose a formidable risk, leveraging context-based tactics or ML/DL models to craft samples that mimic legitimate devices, necessitating a robust, adaptive security strategy that encompasses the comprehensive identification and mitigation of such threats. Finally, it is vital to explore the device authentication problem leveraging more complex ML/DL solutions. Authentication solutions have to generate advanced patterns in the fingerprint, generating unique models for each device based on its intrinsic characteristics.

Based on the previous considerations, this PhD Thesis explores the feasibility of individual device identification and authentication based on hardware performance behavior fingerprinting. It should investigate the methodology definition, its application in real frameworks and tools, their integration in real-world devices and scenarios, and the associated security threat analysis and mitigation. In this sense, several research questions arose from the previous challenges, guiding the research process of this PhD Thesis, and are presented as follows:

- RQ1: What is the current status of device behavior fingerprinting solutions applied in individual device identification and which data sources, techniques, application scenarios and datasets are present in the literature?

- RQ2: Which methodology should be followed to uniquely identify IoT devices in a scalable manner while leveraging on hardware behavior and ML/DL techniques?

- RQ3: Which hardware metrics can be extensively collected for measuring hardware performance and generating the required datasets for late behavior fingerprinting?

- RQ4: How can individual device identification be integrated with other behavior-based cybersecurity solutions deployed in real-world scenarios?

- RQ5: Which ML/DL techniques can achieve the best performance for individual device identification?

- RQ6: How can the resilience of hardware behavior-based individual device identification models be improved against context- and ML/DL-focused adversarial attacks?

- RQ7: Can attention-based ML/DL techniques, such as transformers, enable the individual identification of devices following an anomaly detection approach? Which resources are required?

II Objectives

The main goal of this PhD thesis is to advance the field of device behavior fingerprinting for improving the identification and security of IoT devices, with a particular focus on single-board computers and resource-constrained systems. The research aims to develop novel methodologies and frameworks for individual device identification, leveraging ML and DL techniques. By examining various application scenarios, including Smart Cities, Industry 4.0, and the Internet of Battlefield Things, the thesis aims to address the increasing cybersecurity threats, such as unauthorized device deployment, and other issues emerging due to the exponential growth of interconnected devices in the network-based computing world.

The thesis explores different aspects of device behavior fingerprinting, including data sources, techniques, application scenarios, and datasets. The research focuses on solving the unique problems faced when defining methodologies for identifying identical single-board computers based on hardware behavior fingerprinting and ML. Some of these open issues are the data availability and the solution integration in real-world scenarios. Therefore, they are explored as part of the research objectives. The research also investigates adversarial attacks and defenses in IoT fingerprinting, aiming to develop resilient architectures. Lastly, the work explores the authentication problem, where each device should be recognized without considering others and, therefore, more complex ML models are necessary. From the goal of covering the previous aspects, several specific objectives are derived as subsequently presented, indicating the research questions related to them:

O.1. Analyze the current state of the art regarding device behavior fingerprinting for individual identification through comprehensive review, analysis, and comparison of data sources, techniques, applications, and datasets to guide future research and solutions (RQ1).

O.2. Identify existing gaps in the literature in individual device identification based on hardware behavior fingerprinting and the required steps to cover them (RQ1).

O.3. Address single-board device identification challenges with a novel methodology leveraging hardware behavioral fingerprinting, ML/DL techniques, and essential properties for improved accuracy and robustness (RQ2).

O.4. Develop a low-level hardware benchmarking solution for Single-Board Computers to enable Edge-based AI solutions and versatile device management through extensive performance datasets (RQ3).

O.5. Integrate the single-board device identification solution into a security framework using behavioral fingerprinting and ML/DL techniques to accurately detect diverse cyber-attacks (RQ4).

O.6. Find the best ML/DL-based techniques for individual identification, iterating over the available datasets to achieve the best performance (RQ5).

O.7. Investigate the impact of context and ML/DL-focused attacks against hardware behavior-based device identification solutions leveraging ML/DL models (RQ6).

O.8. Implement and evaluate defense mechanisms to strengthen the solution resilience against adversarial attacks while maintaining its performance in the context of hardware behavior-based device identification (RQ6).

O.9. Explore the individual authentication problem to verify the capabilities of attention-based anomaly detection ML/DL models to detect advanced patterns in hardware behavior data (RQ7).

III Methodology

This PhD Thesis was conducted following a scientific approach based on the continuous study of the state of the art and the analysis of the results obtained during the different stages of the research. This thesis is defined as a set of six papers published in high-impact journals indexed in the Journal Citation Reports (JCR).

The first step was to solve RQ1 and provide a complete view of device behavior fingerprinting. A broad exploration of the existing literature was undertaken to provide a comprehensive understanding of the current status of device behavior fingerprinting solutions in cybersecurity. This exploration encompassed a variety of academic publications, spanning across journals and conference proceedings, ensuring a wide coverage of the topic. The gathered literature was meticulously categorized and analyzed based on specific criteria, such as the data sources used for fingerprinting, the techniques employed, the application scenarios addressed, and the datasets utilized. This granular categorization facilitated the identification of prevalent trends, commonly used methods, and potential gaps within the current body of knowledge. The analysis provided a nuanced understanding of how device behavior fingerprinting is being applied across different domains in cybersecurity, highlighting its versatility and the variety of ways it can be implemented. This extensive review culminated in a holistic view of the field, offering valuable insights and a solid foundation for future research endeavors in device behavior fingerprinting. All these considerations resulted in the first publication of this PhD Thesis, presented in the first chapter (A Survey on Device Behavior Fingerprinting: Data Sources, Techniques, Application Scenarios, and Datasets (Article 1–`IEEE_COMST`)), which covered the first two Objectives of the thesis.

After performing the state-of-the-art analysis and identifying the current gaps in IoT individual identification solutions, the focus moved to solve RQ2 and address Objective 3, which addresses the unique IoT device identification problem. In order to uniquely identify IoT devices based on their hardware behavior while leveraging ML and DL techniques, a structured approach was employed. The process commenced with the collection of a diverse and extensive set of hardware behavior data from a variety of RPi models. This data served as the foundation for the subsequent analysis. Following the data collection phase, pre-processing techniques were applied to refine and prepare the data for the ML and DL models. Various ML/DL models were then selected and trained on the pre-processed data, resulting in the development of identification algorithms. The performance of these algorithms was rigorously evaluated using common ML/DL classification metrics and real SBC devices, ensuring their accuracy and effectiveness in real-world scenarios. Afterward, the methodology was compared with other approaches existing in the literature, although they did not present a structured set of steps in order to develop an identification solution. In this comparison, the methodology performance improvement was contrasted practically. This comprehensive approach facilitated a clear and practical methodology for the unique identification of IoT devices based on their hardware behavior, leveraging the latest advancements in ML and DL. The proposal and validation of the individual device identification methodology resulted in the second chapter (A methodology to identify identical single-board computers based on hardware behavior fingerprinting (Article 2–`JNCA`)).

At this point, a limitation in the research line arose regarding data availability and how to collect it. A specialized benchmarking tool and dataset were developed as part

of this PhD thesis and utilized to address the challenge of extensive data collection for measuring hardware performance and generating the required datasets for later behavior fingerprinting. This tool was meticulously designed to measure and record the performance of various hardware components across a range of IoT devices, ensuring a standardized and comprehensive data collection process. The tool was executed then in some SBC devices to extract a complete and realistic dataset. The collected data was then subjected to rigorous validation and quality assurance processes, verifying its accuracy and reliability. In addition to the data collection, clear guidelines were provided on how to structure and store the data, facilitating its integration into comprehensive datasets. This meticulous approach ensured that the data collected was not only extensive but also of high quality, providing a solid foundation for subsequent behavior fingerprinting applications and analyses. The benchmarking application, together with the collected dataset, are publicly available for other researchers in the area. This work resulted in the third chapter of this thesis (LwHBench: A low-level hardware component benchmark and dataset for Single Board Computers (Article 3–IoT)), answering RQ3 and completing Objective 4.

A holistic and interoperable framework was developed once the individual identification viability was verified with the exhaustive dataset collected using the benchmark application. It sought to integrate individual device identification with other behavior-based cybersecurity solutions. This framework was based on an extensive analysis of device behavior fingerprinting and existing cybersecurity solutions, ensuring a comprehensive understanding of the necessary components and processes. The designed framework facilitated seamless integration, allowing for easy sharing and utilization of device fingerprints and behavior profiles across different security solutions. To achieve this, the framework was designed to be highly modular and scalable, allowing for easy integration with a variety of behavior-based cybersecurity solutions. The identification framework could work with various security systems, including intrusion detection systems, security information and event management solutions, and advanced threat protection tools. It could enhance their effectiveness and add an extra layer of security. To validate the framework suitability, extensive simulations, and real-world tests were conducted, evaluating its functionality in diverse scenarios and against various threat models. The real-world validation was performed considering an IoBT scenario based on the ElectroSense platform [39]. Then, a kernel event and syscall monitoring solution was integrated together with the individual identification solution to provide a complete security approach for the platform. The results of these evaluations confirmed the framework effectiveness, demonstrating its capability to integrate individual device identification with other behavior-based cybersecurity solutions and enhance the security posture of the network. This work solved RQ4 and Objective 5, available in the thesis's fourth chapter (SpecForce: A Framework to Secure IoT Spectrum Sensors in the Internet of Battlefield Things (Article 4–IEEE_COMMAG)).

The next work done in the PhD Thesis, presented in the fifth chapter of this document (Adversarial attacks and defenses on ML- and hardware-based IoT device fingerprinting and identification (Article 5–FGCS)) and aligned with the fifth and sixth research questions (Objectives 6, 7, and 8), tackled the challenge of security threats. It focused on finding the best ML/DL technique for identification and then enhancing its resilience for reliable hardware behavior-based individual device identification. This involved a thorough analysis of potential adversarial attacks, emphasizing understanding the nature and impact of these sophisticated attacks. The presented threat model addresses context- and ML/DL-focused adversarial attacks, ensuring that the identification models remain reliable and secure even in the face of sophisticated threats. The impact of the attacks proposed in the literature was verified, demonstrating the solution vulnerability to adversarial attacks.

Then, countermeasures and mitigation strategies were formulated and implemented to fortify the identification models against adversarial manipulations. Adversarial training and knowledge distillation techniques [40] were incorporated to enhance model resilience against adversarial attacks.

The last work of this compendium worked on the seventh research question (RQ7) and Objective 9 (Single-board Device Individual Authentication based on Hardware Performance and Autoencoder Transformer Models (Article 6–COSE)). It dealt with the authentication problem instead of with identification. As explained in the Introduction, the main difference between these problems is that for authentication, only data from the device being authenticated can be employed in the training process and a higher granularity is necessary in the data processing. This fact makes much more difficult the authentication problem, as the distribution of hardware performance data on devices from the same model usually overlaps in a great proportion. The methodology followed to solve this problem involved time series processing to train Transformer-based autoencoder models [41], with each model tailored to a specific device. The data collection and preprocessing were similar to the one applied in the identification-focused works, only varying the size of the employed time window. Then, attention-based transformer models were designed to function as outlier detection mechanisms, identifying any deviations from the expected hardware behavior that would indicate an unauthorized device. This work placed a strong emphasis on the practical application of their methodology in real-world scenarios. Therefore, it conducted extensive testing and validation to ensure that the approach was not only theoretically sound but also effective in practice.

This thesis creates a holistic storyline, addressing the critical aspects of device behavior fingerprinting from foundational knowledge and practical identification techniques to data collection, integration with broader security solutions, and resilience against adversarial attacks. Finally, the authentication problem is also explored using modern transformer-based approaches. This comprehensive approach ensures a thorough understanding and robust application of device behavior fingerprinting in the realm of cybersecurity. This methodology allowed for meeting the objectives defined in the thesis, previously presented in Section II.

IV Results

The first publication of the PhD Thesis, presented in (Article 1–IEEE_COMST), offered a comprehensive study on device behavior fingerprinting, providing an extensive analysis and synthesis of solutions applied in the realm of cybersecurity, along with valuable insights and findings. The results highlighted a broad spectrum of data sources, techniques, application scenarios, and datasets prevalent in the current literature, underlining the multifaceted nature of device behavior fingerprinting. One of the key findings revolves around the diversity of data sources used for device behavior fingerprinting. The results indicated that a wide variety of data types are employed, ranging from network traffic and system logs to hardware-specific attributes. This diversity underscores the versatility of device fingerprinting solutions, demonstrating their applicability across different layers of the system and network architecture.

In terms of fingerprinting techniques, the study revealed a rich landscape of methodologies, each with its unique strengths and capabilities. The results show that there is no one-size-fits-all solution, with different techniques catering to specific requirements and scenarios. This variety ensures that practitioners and researchers have a plethora of options to choose from, enabling them to tailor their device fingerprinting solutions to meet the

specific needs of their application. When it comes to application scenarios, the document outlined a wide range of contexts in which device behavior fingerprinting is employed. From enhancing network security and detecting unauthorized devices, to facilitating device authentication and integrity verification, the applications are vast and varied. This highlighted the crucial role that device behavior fingerprinting plays in bolstering cybersecurity measures, providing an additional layer of security and trust in digital environments. The examination of datasets revealed that while there are numerous datasets available for research and development purposes, there is still a need for more comprehensive and standardized datasets. The results point out that the existing datasets vary significantly in terms of size, quality, and relevance, indicating a gap that needs to be addressed to further advance the field. The availability of high-quality, standardized datasets is paramount for the validation and benchmarking of device fingerprinting solutions, ensuring their reliability and effectiveness in real-world scenarios.

In the last section of the document, a profound and insightful analysis was presented, encapsulating the lessons learned, identifying prevailing trends, and highlighting the challenges faced in the domain of device behavior fingerprinting within cybersecurity. This section serves as a critical reflection on the current state of the field, offering guidance for future research and practical implementations. One of the main challenges identified is associated with the lack of solutions dealing with individual device identification, compared to other topics such as behavior-based malware detection or network security. Moreover, there is a lack of datasets available to enable the design of individual identification solutions. Another important challenge is related to the attacks directly focused on disrupting the effectiveness of behavior-based security solutions. Besides, an obvious imbalance between solutions for identification and authentication of devices is detected, possibly because it is a more complex problem to solve.

Following the main challenges found in the literature, the second publication (Article 2–JNCA) provided a detailed exploration and analysis of the steps required to uniquely identify IoT devices based on their hardware behavior, with a particular emphasis on leveraging ML and DL techniques. The results gleaned from this investigation offer valuable insights into the intricacies of device fingerprinting and the practicalities of implementing such methodologies in real-world scenarios.

The research identified essential properties for single-board device identification, including uniqueness, stability, diversity, scalability, efficiency, robustness, and security. A novel methodology was then introduced, relying on behavioral fingerprinting to identify identical single-board devices while meeting the aforementioned properties. This methodology utilizes the different built-in components of the system, along with ML/DL techniques, to compare the internal behavior of devices and detect variations that occurred during the manufacturing processes. A key outcome of the study was the identification and validation of specific hardware attributes that can be used as reliable indicators for device fingerprinting. These attributes, when analyzed and processed correctly, have been shown to provide a unique signature for each device, facilitating accurate and efficient identification. The results underscore the importance of selecting the right combination of hardware attributes, as this choice significantly impacts the effectiveness of the fingerprinting process.

The document also highlighted the critical role of hardware isolation in the device fingerprinting workflow. The results demonstrated that careful handling and preparation of the hardware attributes are paramount, as this ensures the integrity of the fingerprinting process and enhances the accuracy of device identification. Besides, the application of various data preprocessing techniques, including normalization and dimensionality reduction, has been shown to contribute positively to the outcome of the fingerprinting process.

The integration of ML and DL techniques into the device fingerprinting process has also been shown to introduce an element of adaptability and learning, enabling the system to evolve and improve over time. Deep learning techniques, including various configurations of neural networks, were explored for their ability to learn and model the device fingerprints directly from raw hardware data. ML and DL classifiers were among the architectures tested, chosen for their prowess in handling sequential and time-series data, which is prevalent in hardware behavior information. Additionally, the document addressed the challenge of model overfitting, especially pertinent when employing complex models like deep neural networks. Strategies such as cross-validation and regularization were applied, ensuring that the models generalize well to unseen data and maintain high performance when deployed in real-world settings. The methodology is validated in a real environment, consisting of 15 identical Raspberry Pi 4 Model B and 10 Raspberry Pi 3 Model B+ devices whose CPU and GPU performance was analyzed. The results showcase a 91.9% average True Positive Rate (TPR) with an XGBoost model, achieving identification for all devices by setting a 50% threshold in the evaluation process. Furthermore, the proposed methodology was engaged in a critical discussion, comparing the proposed solution with related work, highlighting the fingerprint properties not met by other solutions, and providing valuable lessons learned and limitations of the presented methodology.

The main result in the third publication of the thesis (Article 3–IoT) was the development of a low-level hardware benchmarking application tailored for SBCs, addressing the need for lower-level benchmarking applications and datasets in the realm of IoT identification. The benchmark was named *LwHBench* [42] and focuses on measuring the performance of CPU, GPU, Memory, and Storage, while taking into account the component constraints inherent in SBCs. The application has been specifically implemented for Raspberry Pi devices. It was run for 100 days on a set of 45 devices to generate an extensive dataset containing 2386126 vectors in +4GB of data. This dataset paves the way for the application of AI techniques in scenarios where performance data can aid in the device management process. To showcase the inter-scenario capability of the dataset, the document also presented a series of AI-enabled use cases related to device identification and the impact of context on performance. In a practical setup, the benchmark application was adapted and applied to a scenario involving three RockPro64 devices, demonstrating its versatility and applicability in real-world settings.

In the fourth publication of the thesis (Article 4–IEEE_COMMAG), the research delves into the emerging and highly dynamic field of the Internet of Battlefield Things (IoBT). It focused particularly on the pivotal role of wireless communications within this sphere. In this intricate battlefield scenario, a myriad of devices, ranging from soldiers to diverse military equipment, interact in real time, exchanging information wirelessly and forming a complex network of interconnected entities. The proposed scenario delves into three primary use cases: IoT device identification, malware detection, and Spectrum Sensing Data Falsification (SSDF) attack detection.

To solve these use cases, an IoT behavior fingerprinting framework was introduced, namely *SpecForce*, which was meticulously designed to enhance the security of IoBT spectrum sensors, crucial components in monitoring the frequency spectrum, transmitting over unoccupied bands, intercepting enemy transmissions, and decoding valuable information. In this real-world scenario, individual device identification is essential to avoid possible attacks based on identity manipulations. *SpecForce* stands out as a robust solution, employing device behavioral fingerprinting alongside ML/DL techniques. The framework was adept at considering heterogeneous data sources, enhancing its capability to detect and mitigate a wide array of cyber threats effectively. The emphasis was placed on ensuring

the integrity and reliability of communications within the battlefield scenario, a critical aspect considering the spectrum scarcity and the burgeoning number of IoBT devices. The framework included an AI-based cybersecurity module that employs ML/DL classification algorithms for identifying different IoBT spectrum sensors based on RPi devices. The document provides a comprehensive analysis of various ML/DL models for classification. The results indicate that Random Forest and XGBoost are the best-performing models, achieving over 91% TPR. The document also discusses a use case demonstrating the capability of the system to uniquely identify 25 IoBT spectrum sensors, addressing identity-focused attacks and enhancing security.

As commented before, *SpecForce* was equipped with other cybersecurity approaches using kernel event and syscall behavior monitoring to detect higher-level cyberattacks. In the context of SSDF attack detection, the framework allows the syscall monitoring of IoBT spectrum sensors, aiming to detect various SSDF attacks. System calls are processed to generate feature vectors that model the spectrum sensing activities, with anomaly detection algorithms employed to distinguish between normal and malicious behaviors. The results showcase a high performance in recognizing normal behavior, with over 99% True Negative Rate (TNR), and a commendable TPR of over 92% for SSDF attack detection. Besides, regarding heterogeneous malware detection (botnets, backdoors, etc.), high performance was achieved by monitoring kernel events combined with ML/DL anomaly detection, with a 90% TPR and a 96% TNR.

In the context of adversarial attacks, the next publication of the PhD Thesis (Article 5–FGCS) shed light on the potential vulnerabilities and threats that can compromise the integrity of device fingerprinting and identification mechanisms. It discusses various attack vectors, illustrating how malicious entities could manipulate or bypass hardware-based identification mechanisms to achieve their nefarious goals. The paper underscores the need for comprehensive defense strategies capable of mitigating the risks associated with adversarial attacks, ensuring the robustness of device identification processes. The first main result of this work was the improvement in the identification results achieved in previous works. Using time series approaches combined with DL models, identification results were increased to +0.96 average TPR with a minimum 0.80 TPR in the 45 RPi devices used for validation by leveraging an LSTM+1D-CNN combined model. Regarding adversarial attacks, both context- and ML/DL-focused attacks are applied to evaluate the robustness of the device identification model. A specific mention is made of a temperature-based context attack, which, interestingly, was found to be ineffective in disrupting the device identification process, as the hardware isolation during data collection was already considering context impact mitigation. However, the document does acknowledge the success of certain state-of-the-art ML/DL evasion attacks, such as BIM, MIM, and JSMA.

On the defense side, the document provides an exhaustive exploration of various strategies and methodologies aimed at protecting IoT devices from adversarial threats. It delves into ML and context-based approaches, evaluating their effectiveness in enhancing the security and reliability of device fingerprinting and identification. Context-based attacks focused on temperature are ineffective due to the hardware stability and isolation measures taken during data collection. Regarding defenses on ML/DL evasion attacks, knowledge distillation and adversarial training are applied to reduce the impact of the attacks. The results highlight the critical role of these defense mechanisms in maintaining the integrity of IoT ecosystems, ensuring that devices are accurately identified and malicious entities are thwarted. In terms of performance, the success ratio of attacks was reduced from 0.88 to 0.17 in the worst-case scenario without causing a substantial degradation in performance. Finally, various security metrics are used to assess the resilience of neural networks

against adversarial perturbations and input variations. Metrics such as CLEVER score, Loss sensitivity, and Empirical robustness [40] are discussed, providing insights into how the robustness of ML models can be quantified and evaluated.

The last publication of the PhD thesis (Article 6–COSE) focused on the individual authentication problem. It proposed an authentication framework that utilized hardware performance data and transformer-based autoencoder models. The framework design is supported by a threat model that outlines the security challenges encountered in implementing hardware-based authentication in IoT contexts. As in previous works, key hardware components, such as CPU, GPU, RAM, and storage, were monitored for fingerprint data collection. These fingerprints were then utilized as time series data, applying time windows from 10 to 100 values. The generated time series are then used to train transformer models for outlier detection, tailored to each individual device, thereby aiming to represent and authenticate it accurately. The framework effectiveness was further demonstrated through its application in a spectrum crowdsensing system using Raspberry Pi devices. Transformer models were compared with LSTM and 1D-CNN approaches in terms of performance. Here, in a series of rigorous experiments involving 45 devices for validation, each device transformer model proved capable of accurately authenticating it. In contrast, other approaches were not capable of uniquely authenticating all the devices. The approach achieved an impressive average TPR of 0.74 ± 0.13 and maintained an average maximum FPR of 0.06 ± 0.09, underscoring its potential to significantly enhance authentication, security, and trustworthiness in crowdsensing applications. Moreover, the resource usage of the different approaches tested was also analyzed, confirming one of the main drawbacks of transformer models; this was the ML/DL technique using the most resources in terms of time and memory.

V Conclusions and future work

This thesis provides an in-depth and comprehensive examination of the field of device behavior fingerprinting, with a specific lens focused on its application within the realm of cybersecurity, and a nuanced emphasis on single-board and IoT device identification and authentication. The research begins with an extensive and systematic review of the current landscape, capturing the breadth and depth of device behavior fingerprinting solutions that have been explored and implemented within the cybersecurity domain. This initial phase of exploration serves to lay a robust foundation of knowledge, unraveling the complexities of various data sources, fingerprinting techniques, application scenarios, and the datasets that are prevalent within the academic and practical spheres of this field. From this exploration, a novel set of literature lessons and trends is derived, giving a holistic view of previous works. However, the main novelty from a research perspective is the list of challenges identified in the literature, which were not defined before and paved the way for future research.

As the storyline progresses, the spotlight shifts to the practical implementation of device identification, with the proposal of a clear and detailed methodology outlined for uniquely identifying IoT devices. This process is intricately tied to the capabilities afforded by ML and DL techniques. These advanced computational tools can be harnessed to decipher the subtle nuances of hardware behavior, ensuring a high level of authenticity and integrity for devices embedded within a network. The importance of this process cannot be overstated, as it plays a pivotal role in safeguarding the security and reliability of interconnected devices, forming a critical component of the broader cybersecurity infrastructure. To highlight this importance, the presented methodology is compared to other works in the

field, where no methodological set of steps was followed. Previous solutions were not able to perform full identification in the real device scenario used for validation. Therefore, the methodology becomes the state-of-the-art procedure for developing a functional individual identification solution.

Building upon the established identification methodology, the narrative delves deeper into the critical aspect of data collection, presenting a comprehensive approach for the systematic acquisition of hardware performance data. Generating the datasets required for later behavior fingerprinting is essential. These datasets help ensure that the identification models are trained and validated on data that is both extensive and accurate. The meticulous attention to detail in this process ensures the reliability of the data, setting a high standard for the quality of information used in device behavior fingerprinting. As a result of the proposed tool for data collection, an exhaustive dataset is generated to be applied in the following steps of the thesis. This dataset is published to be employed by the research community in the field, together with the benchmark application employed to generate the data. This is one of the first public datasets of hardware performance data that is focused on the identification problem.

With a solid foundation of data in place, the exploration then navigates toward the integration of individual device identification within the larger ecosystem of behavior-based cybersecurity solutions. This integration is paramount, as it ensures that the device identification procedures do not operate in isolation but are seamlessly linked with other general security frameworks. For this integration, the focus is strategically placed on the Internet of Battlefield Things. A real-world scenario is employed to integrate the hardware-based identification solution with higher-level behavior monitoring for the detection of malware and spectrum-sensing data falsification attacks. Specifically, the unified approach monitors hardware, kernel events, and system calls to provide a unified security solution based on behavior. This holistic approach enhances the environment resilience, fortifying its defenses against a myriad of cyber threats and vulnerabilities, and ensuring a robust and secure digital environment. The unified framework demonstrates experimentally its capabilities to detect different malware samples and spectrum data-focused attacks, as well as perform individual identification of the sensors deployed. This is, to the best of our knowledge, the first framework combining these cybersecurity capabilities together.

As the work reaches its culmination, the focus turns to the security of the device identification models themselves, specifically addressing the challenges posed by adversarial attacks. The landscape of adversarial attacks is analyzed, particularly focusing on context-aware and ML/DL-centric threats that pose significant risks to the integrity of device identification methodologies. Then, the focus delves into strategies and methodologies designed to enhance the resilience of hardware behavior-based identification models. A particular emphasis is placed on counteracting sophisticated adversarial threats, including context-based and ML/DL-focused attacks. Context-based attacks are ineffective against the solution, but evasion attacks targeting the identification models achieve high success rates. Therefore, defense techniques based on adversarial training and knowledge distillation are applied. This ensures that the device identification methodologies remain reliable, secure, and effective, even in the face of evolving and complex cyber threats. This is one of the first works demonstrating experimentally the effectiveness of adversarial attacks on ML models for IoT device identification, and also of state-of-the-art defense methods.

Finally, the research explores the more complex problem of individual device authentication, where only data from one device can be leveraged for model generation. The methodology applied for identification is tweaked for its application in authentication. The main difference is the change of the classifier model for an anomaly detection model

per device. However, more powerful ML/DL models are necessary to solve this problem as data distributions overlap between devices from the same model. To solve this issue, transformer models are employed. Using large time windows for data processing (100 values), the transformer-based approach improves the results achieved by previous state-of-the-art models such as LSTM, 1D-CNN, and their combination. In contrast, higher training time and memory are employed in the model generation process. This outcome confirms the effectiveness of the transformer architecture in a novel area. Unlike previous model architectures in existing literature, it successfully addresses the issue of individually authenticating IoT devices based on their hardware performance in the experimental scenario studied.

Looking ahead, there is a vast horizon of opportunities for future work building upon the foundations laid by this thesis in the domain of device behavior fingerprinting. The first iteration of future work could delve into the adversarial attack evaluation over the unsupervised transformer models employed to solve the individual identification problem, and the consequent application of defense techniques. These results will close the work on the authentication topic in a similar manner to the methodology applied for the individual identification problem.

Another promising avenue is expanding the scope of device behavior fingerprinting to encompass a broader array of devices and contexts. The current work has predominantly focused on single-board and IoT devices. However, the principles and methodologies developed could be adapted and applied to other types of devices and networks, such as industrial control systems, automotive systems, and smart home devices. Future research could explore the nuances and specific requirements of these different contexts, tailoring the fingerprinting techniques to suit the unique characteristics of each device category and usage scenario. One of the pivotal areas highlighted for future exploration is the continual quest for new and diverse data sources. The field stands to benefit significantly from broadening the scope of data collection, capturing a wider array of device behaviors, and ensuring a richer and more comprehensive dataset for analysis. This endeavor is not just limited to increasing the quantity of data but also emphasizes the importance of improving the quality and reliability of the data collected. Future work in this area could explore advanced data collection methodologies, innovative sensor technologies, and novel data preprocessing techniques, all aimed at ensuring that the data used for device behavior fingerprinting is of the highest caliber.

Building upon the theme of data, there is a clear call for the development and refinement of fingerprinting techniques. The future holds potential for the exploration of new algorithms, ML models, and DL architectures, each offering unique capabilities and advantages for device identification. Federated Learning (FL), and more concretely decentralized FL, is one of these promising areas worth exploring in the following years. The continuous evolution of computational power and ML/DL technologies opens up exciting possibilities for creating more sophisticated and accurate device behavior fingerprinting models, capable of discerning even the most subtle nuances in device behavior.

Addressing the challenge of adversarial attacks, there is a pressing need for the development of robust countermeasures and mitigation strategies. Future work in this area could explore innovative approaches to enhancing the resilience of device behavior fingerprinting models, with a specific focus on countering sophisticated adversarial threats, including context-based and ML/DL-focused attacks. This involves not only fortifying the identification models but also developing comprehensive threat detection and response mechanisms, ensuring the long-term reliability and security of device identification methodologies.

Another critical area for future work lies in enhancing the adaptability of device be-

havior fingerprinting models. With the rapid pace of technological advancement, devices are constantly evolving, and their behavior patterns may change over time due to software updates, hardware modifications, or changes in usage patterns. Future research could focus on developing fingerprinting models that are capable of adapting to these changes, ensuring that they remain accurate and reliable over the device lifecycle. This could involve the integration of online learning techniques, continual learning approaches, or transfer learning methodologies to enable the models to update and refine their fingerprinting profiles in response to observed changes in device behavior.

Finally, there is significant potential for future work in the integration of device behavior fingerprinting solutions with other cybersecurity tools and frameworks. This thesis has laid the groundwork for such integration, demonstrating the potential benefits of combining device behavior fingerprinting with other behavior-based security solutions. Future research could build upon this, exploring ways to further streamline the integration process, enhance interoperability, and maximize the synergies between different security tools. This could lead to the creation of more holistic and resilient cybersecurity frameworks, providing comprehensive protection against a diverse range of cyber threats.

Resumen

I Introducción y motivación

Para 2025, se espera que casi 64 mil millones de dispositivos de Internet de las Cosas (IoT) estén conectados en diversos entornos de vanguardia, como Ciudades Inteligentes, Industria 4.0 y crowdsensing [1]. El crecimiento del IoT promete no solo mejorar la calidad y accesibilidad de los servicios, sino también revolucionar la experiencia del usuario, ya que fomenta ecosistemas inteligentes que reducen significativamente la intervención humana, agilizando así los procesos, reduciendo costos y mitigando errores. Sin embargo, los objetivos únicos de estos entornos también aumentan la complejidad de optimizar el rendimiento de dispositivos y servicios.

Desde una perspectiva de escenario, la vasta gama de dispositivos IoT utilizados hoy en día incluye dispositivos de placa única (SBC por su nombre en inglés) como Raspberry Pi (RPi), que han ganado popularidad debido a su flexibilidad, asequibilidad, amplio soporte y periféricos disponibles [2]. Sin embargo, las limitaciones de conectividad y recursos de las SBC y los dispositivos IoT generan numerosas preocupaciones de ciberseguridad para diversas plataformas [3]. Un problema significativo es la presencia de dispositivos no autorizados con configuraciones de hardware y software idénticas a las de los nodos autorizados, lanzando ataques que impactan áreas de aplicación como Industria 4.0 [4], teléfonos inteligentes [5] o Internet de las Cosas del Campo de Batalla (IoBT) [6]. Estos dispositivos maliciosos pueden estar presentes como consecuencia de varias amenazas de ciberseguridad [7], incluyendo i) suplantación de dispositivos, donde un atacante reemplaza un dispositivo legítimo por uno malicioso usando la misma identidad; ii) despliegue no autorizado de dispositivos, que implica la instalación de un nuevo dispositivo con una identidad no registrada; y iii) ataque Sybil, donde un dispositivo malicioso utiliza múltiples identidades para simular numerosos dispositivos. En consecuencia, otras amenazas como la filtración de información sensible, envenenamiento de datos y escalada de privilegios y movimientos laterales pueden surgir de dispositivos suplantados.

Para resolver estos problemas emergentes de ciberseguridad, la ciencia de datos del comportamiento se ha expandido desde el estudio de comportamientos humanos [8] hacia la modelización de comportamientos de dispositivos [9], con un enfoque en la creación de patrones de comportamiento de dispositivos (*huellas digitales*) para optimizar el rendimiento y detectar problemas potenciales tempranamente [10]. Dos escenarios primarios de aplicación de IoT para la construcción de huellas digitales de dispositivos son i) identificación y autenticación de dispositivos en diferentes niveles de granularidad [11], y ii) detección de

ciberataques, malfuncionamiento o comportamiento indebido [12, 13]. Diversos estudios han aplicado el fingerprinting de dispositivos a ambos escenarios [14, 15]. En el escenario de identificación, la ciencia de datos del comportamiento ha mejorado significativamente las capacidades de identificación de dispositivos, superando las limitaciones asociadas con metodologías convencionales que predominantemente dependen del uso de nombres, identificadores, etiquetas o tags para el reconocimiento de dispositivos [16]. Una limitación crítica de las estrategias tradicionales es su susceptibilidad a alteraciones y duplicaciones. Estas técnicas convencionales a menudo carecen del dinamismo para adaptarse al paisaje rápidamente evolutivo de los escenarios IoT, donde muchos dispositivos se despliegan con alta movilidad. Estos problemas son especialmente relevantes en escenarios con un aumento exponencial en la cantidad de dispositivos, como en industrias inteligentes o la agricultura. Por lo tanto, la identificación de dispositivos basada en huellas digitales de comportamiento ha mejorado significativamente las soluciones tradicionales al enfocarse en niveles de granularidad de tipo [11], modelo [17], e individual [18]. Por otro lado, la detección de mal comportamiento o malfuncionamiento debido a problemas de ciberseguridad ha visto el surgimiento del fingerprinting de dispositivos como una solución prometedora. Numerosos trabajos crean huellas digitales de comportamiento "normal" para detectar cambios causados por problemas como ciberataques, ejecución de malware o malfuncionamiento de dispositivos [12, 19].

En el nivel de granularidad de identificación individual, el fingerprinting (generación de huellas) del comportamiento del hardware se destaca como una vía prometedora para identificar de manera única dispositivos idénticos, una necesidad en el paisaje tecnológico interconectado de hoy. A pesar de su potencial, este dominio todavía está en desarrollo, caracterizado por desafíos abiertos y una falta de investigación dedicada específicamente a la identificación de dispositivos de placa única idénticos [20]. La complejidad de esta tarea se amplifica por las similitudes inherentes entre dichos dispositivos, lo que requiere metodologías de identificación innovadoras y precisas. En el contexto más amplio, para dispositivos que no están limitados por componentes y recursos, la literatura existente aboga por la adopción del *fingerprinting del comportamiento del hardware* [21]. Esta técnica es fundamental para discernir pequeñas variaciones de rendimiento que son un subproducto de imperfecciones de fabricación [22]. Al analizar meticulosamente estas sutiles diferencias, el fingerprinting del comportamiento del hardware proporciona un nivel granular de caracterización del dispositivo, abriendo el camino para una identificación precisa del dispositivo. Sin embargo, es importante distinguir entre identificación y autenticación en este contexto. Mientras que la identificación implica reconocer un dispositivo de un conjunto basado en sus características únicas, la autenticación va un paso más allá verificando la legitimidad del dispositivo. La autenticación aborda la pregunta de si un dispositivo es quien dice ser, ofreciendo una capa de seguridad que la mera identificación no proporciona.

El análisis del comportamiento del hardware para la identificación de dispositivos se ha realizado parcialmente en la literatura pero sin seguir una metodología clara sobre los pasos a realizar para lograr una solución exitosa [20]. Por lo tanto, existe una necesidad de esta metodología específica, arraigada en la necesidad crítica de reforzar las medidas de ciberseguridad en infraestructuras de red que incorporan dispositivos de ordenador de placa única idénticos, como Raspberry Pi. El conjunto de estos dispositivos trabajando de manera coordinada es lo que permite ofrecer servicios de IoT basados en crowdsourcing o recolección/procesamiento de datos conjuntos. Estos dispositivos, a menudo desplegados extensamente y en entornos donde los recursos son escasos, son vulnerables a una miríada de amenazas de seguridad [23, 24] que afectan la fiabilidad de los servicios ofrecidos. Estos van desde intentos de suplantación de dispositivos maliciosos hasta la introducción de

dispositivos no autorizados en una red. Asegurar la capacidad de identificar y autenticar dispositivos de manera precisa se convierte en un requisito no negociable, fundamental para mantener la integridad y la resiliencia del servicio. Por lo tanto, un nivel de seguridad consistente es esencial para tener servicios confiables. La implementación del fingerprinting del comportamiento del hardware en estos contextos exige un nivel de precisión e innovación que trasciende los métodos convencionales de identificación. Los desafíos provienen de las sutiles diferencias entre dispositivos idénticos y las limitaciones impuestas por sus recursos limitados [25]. Abordar estos desafíos requiere un esfuerzo de investigación concertado, dirigido a desentrañar las complejidades del fingerprinting del comportamiento del hardware, seleccionar las características de rendimiento del hardware más prometedoras, desarrollar algoritmos de identificación robustos y establecer mejores prácticas para la aplicación práctica.

Además, tener una solución de identificación completa basada en hardware es crítico para la seguridad de IoT. Sin embargo, en escenarios del mundo real, las soluciones de identificación individual no funcionan aisladas de otras aplicaciones de ciberseguridad. Deben integrarse con otras soluciones de red y comportamiento que buscan cubrir problemas de ciberseguridad heterogéneos, como el malware. En este contexto, la integración del fingerprinting del comportamiento del hardware para dispositivos de placa única se convierte en un componente estratégico de un framework de seguridad más amplio. Identificar y verificar de manera única cada dispositivo permite una base sólida para la comunicación segura y la integridad de los datos. Sin embargo, aún se requieren otras herramientas para alcanzar un nivel de seguridad más completo. Este nivel de seguridad no se trata solo de proteger la información; se trata de garantizar la efectividad operativa y el éxito de los entornos IoT.

Para reforzar la efectividad del fingerprinting del comportamiento del hardware para dispositivos de placa única idénticos, es imperativo identificar las fuentes de datos más relevantes para resolver este problema. Luego, se requiere tener un punto de referencia integral y un conjunto de datos bien curado [26]. Estos recursos sirven como herramientas críticas en la verificación y validación de la metodología, asegurando su precisión, fiabilidad y aplicabilidad en escenarios del mundo real. Un punto de referencia preciso es necesario para evaluar sistemáticamente el rendimiento de los dispositivos, proporcionando una medida estandarizada que se puede usar para discernir las sutiles variaciones en el comportamiento del hardware [27]. Esto se vuelve particularmente crucial al tratar con dispositivos idénticos, donde las diferencias son mínimas pero significativas para una identificación precisa. Sin embargo, no hay aplicaciones de benchmarking disponibles en la literatura para el fingerprinting de comportamiento de hardware de bajo nivel [26]. Junto con un punto de referencia robusto, un conjunto de datos rico juega un papel vital en el proceso. Actúa como un repositorio de perfiles de comportamiento de dispositivos de hardware, capturando las complejidades y características únicas de cada dispositivo. Este conjunto de datos se convierte en un punto de referencia, ayudando en el entrenamiento de algoritmos y sirviendo como un punto de referencia para la validación. La combinación de un punto de referencia integral y un conjunto de datos detallado asegura un enfoque holístico para la identificación de dispositivos, fortaleciendo la metodología de fingerprinting del comportamiento del hardware. La motivación para establecer herramientas de verificación tan rigurosas surge de las necesidades cambiantes de los paradigmas de IoT y computación Edge, donde los dispositivos están cada vez más interconectados y la integridad del entorno es primordial. En estos escenarios, la capacidad de identificar y autenticar dispositivos con precisión no es solo una medida de seguridad; es una necesidad para garantizar el funcionamiento fluido de la red y la confiabilidad de los datos intercambiados. Al invertir en el desarrollo

de puntos de referencia precisos y conjuntos de datos integrales, la metodología de fingerprinting del comportamiento del hardware puede empoderarse, mejorando su precisión y fiabilidad. Esto, a su vez, allana el camino para un futuro donde dispositivos de placa única idénticos puedan integrarse sin problemas en redes complejas, operando de manera segura y eficiente, y contribuyendo a la solidez de la infraestructura digital.

Tras adquirir las huellas digitales, surge un dominio de investigación apasionante, centrado en emplear técnicas óptimas para procesarlas y evaluarlas. Mientras que los métodos estadísticos han mantenido una posición dominante en este campo durante décadas, la llegada de la Inteligencia Artificial (IA), específicamente el Machine Learning (ML) y el Deep Learning (DL), ha precipitado un cambio de paradigma, ganando mayor prominencia en las soluciones que se están desarrollando y desplegando hoy en día [28]. La complejidad y diversidad de los comportamientos de los dispositivos IoT requieren modelos de DL sofisticados capaces de capturar patrones y variaciones intrincadas. Técnicas como las Redes Neuronales Convolucionales (CNNs) se están adaptando para discernir relaciones espaciales en huellas digitales de hardware [29], mientras que las Redes Neuronales Recurrentes (RNNs), incluyendo su variante más sofisticada, las redes de Memoria a Corto y Largo Plazo (LSTM), se están desplegando para el análisis de datos temporales, crucial para comprender el comportamiento del dispositivo a lo largo del tiempo [17]. Los autoencoders también han creado un nicho en este sector, permitiendo la reducción de la dimensionalidad para datos de huellas digitales de gran volumen y la detección de anomalías mediante la reconstrucción del comportamiento normal del dispositivo y destacando desviaciones [30]. Además, los avances recientes en Redes Neuronales de Grafos (GNNs) permiten la incorporación de datos topológicos, facilitando la modelización de relaciones complejas dentro de redes de dispositivos IoT interconectados. Finalmente, los modelos basados en atención como los transformers están logrando un éxito extraordinario en herramientas centradas en el lenguaje, como ChatGPT [31], y también pueden aplicarse a otros datos de series temporales. Sin embargo, la aplicabilidad de todas estas técnicas modernas está condicionada por la falta de conjuntos de datos disponibles, ya que requieren la existencia de conjuntos de datos exhaustivos capaces de entrenar los modelos requeridos [32]. Por lo tanto, después de tener suficientes datos, el siguiente desafío es explorar los métodos de ML/DL disponibles para encontrar cuál podría proporcionar el mejor rendimiento en la identificación individual de dispositivos. Aquí, se pueden crear nuevas arquitecturas de red para modelar mejor el comportamiento del dispositivo y mejorar las capacidades de identificación y autenticación.

Además, una solución totalmente funcional y confiable debe considerar los posibles problemas de seguridad intrínsecos a la identificación individual de dispositivos basada en hardware [33]. Los ataques adversarios surgen como una de las principales amenazas dirigidas a soluciones de identificación en IoT. Estos ataques sofisticados presentan un desafío formidable, ya que están específicamente diseñados para manipular o evadir los mecanismos de seguridad existentes, lo que podría llevar a accesos no autorizados, violaciones de datos y compromiso de la integridad de la red [34, 35]. Los ataques adversarios en el contexto de la identificación de dispositivos de placa única explotan los mecanismos diseñados para garantizar la autenticidad del dispositivo y la seguridad de la red. Al alterar sutilmente el comportamiento del dispositivo o imitar las características de dispositivos legítimos, los adversarios pueden engañar a los sistemas de identificación, lo que resulta en la clasificación errónea o la aceptación falsa de dispositivos pícaros. Esto no solo socava la confiabilidad de la red, sino que también abre la puerta a actividades maliciosas adicionales, poniendo en peligro la confidencialidad, integridad y disponibilidad de los datos y servicios. Profundizando en la naturaleza de estos ataques adversarios, hay una tendencia

creciente en tácticas de evasión basadas en contexto y ML/DL. Los ataques basados en contexto manipulan inteligentemente las condiciones ambientales o el contexto operativo de los dispositivos, con el objetivo de distorsionar los datos utilizados para la identificación y, por lo tanto, engañar al sistema [36]. Estos ataques son particularmente insidiosos, ya que explotan la variabilidad natural en el comportamiento del dispositivo debido a cambios en factores externos, haciéndolos más difíciles de detectar y contrarrestar. Por otro lado, los ataques de evasión basados en ML/DL representan un asalto sofisticado y calculado a los sistemas de identificación de dispositivos de placa única. Los adversarios que emplean estas técnicas utilizan modelos avanzados de ML para aprender los patrones y características de dispositivos legítimos. Armados con este conocimiento, crean muestras adversarias que imitan de cerca a dispositivos auténticos, difuminando efectivamente las líneas entre dispositivos legítimos y pícaros [37]. Estas muestras se utilizan luego para sondear y engañar al sistema de identificación, llevando a clasificaciones erróneas y potencialmente otorgando acceso no autorizado a la red. Estos ataques explotan las complejidades inherentes y variabilidades en el comportamiento del dispositivo y las capacidades avanzadas de los modelos de ML/DL para engañar y comprometer la seguridad de la red. Abordar estos desafíos requiere una estrategia de seguridad integral y adaptable capaz de anticipar, detectar y mitigar la miríada de amenazas que representan los ataques adversarios.

Aparte de todo el trabajo por hacer en el contexto de identificación individual de dispositivos IoT, el problema de la autenticación es un desafío abierto por sí solo debido a sus requisitos más complejos [38]. Aunque el fingerprinting del comportamiento del hardware ofrece una base robusta para la identificación, extender este enfoque para incluir mecanismos de autenticación es vital para garantizar un grado más alto de seguridad en sistemas interconectados. Sin embargo, los dispositivos del mismo modelo pueden exhibir características de comportamiento muy similares o incluso superpuestas debido a procesos de producción estandarizados. Esta similitud en el comportamiento del hardware hace que sea desafiante distinguir entre dispositivos legítimos y no autorizados basándose únicamente en datos de comportamiento del hardware, ya que las distribuciones de datos se fusionan. Por lo tanto, la autenticación requiere una granularidad de análisis de datos mucho más fina. Aquí, los enfoques de procesamiento tradicionales y las técnicas de ML/DL no han logrado resultados notables. Sin embargo, nuevos algoritmos de ML/DL con capacidades avanzadas de reconocimiento de patrones, como los transformers basados en atención, podrían mejorar el rendimiento de la autenticación basada en el comportamiento del rendimiento del hardware.

Como resumen de los desafíos abiertos, la literatura sugiere que, aunque el fingerprinting del comportamiento del hardware presenta una solución prometedora, el campo aún es incipiente, con brechas de investigación considerables en metodología y práctica para dispositivos de placa única. La identificación individual de estos dispositivos requiere técnicas innovadoras para distinguir sutiles diferencias de fabricación, así como puntos de referencia precisos y conjuntos de datos extensos para validar nuevos algoritmos. La evolución de la IA, particularmente ML y DL, está cambiando el paradigma para procesar y evaluar huellas digitales de dispositivos, y se requieren nuevas técnicas en esta área. Además, sigue siendo una necesidad crítica la integración del fingerprinting del hardware con medidas de seguridad de red más amplias para abordar varias amenazas de ciberseguridad. A continuación, los ataques adversarios representan un riesgo formidable, aprovechando tácticas basadas en contexto o modelos de ML/DL para crear muestras que imitan dispositivos legítimos, lo que requiere una estrategia de seguridad robusta y adaptable que abarque la identificación y mitigación integral de tales amenazas. Finalmente, es vital explorar el problema de la autenticación de dispositivos aprovechando soluciones de ML/DL más complejas.

Las soluciones de autenticación tienen que generar patrones avanzados en la huella digital, generando modelos únicos para cada dispositivo basados en sus características intrínsecas.

Basado en las consideraciones anteriores, esta Tesis Doctoral explora la viabilidad de la identificación y autenticación de dispositivos individuales basada en el fingerprinting del comportamiento del rendimiento del hardware. Investiga la definición de la metodología, su aplicación en frameworks y herramientas reales, su integración en dispositivos y escenarios del mundo real, y el análisis y mitigación de amenazas de seguridad asociadas. En este sentido, varias preguntas de investigación surgieron de los desafíos anteriores, guiando el proceso de investigación de esta Tesis Doctoral, y se presentan de la siguiente manera:

- RQ1: ¿Cuál es el estado actual de las soluciones de fingerprinting del comportamiento de dispositivos aplicadas en la identificación individual de dispositivos y qué fuentes de datos, técnicas, escenarios de aplicación y conjuntos de datos están presentes en la literatura?
- RQ2: ¿Qué metodología se debe seguir para identificar de manera única dispositivos IoT de manera escalable mientras se aprovecha el comportamiento del hardware y las técnicas de ML/DL?
- RQ3: ¿Qué métricas de hardware se pueden recolectar extensivamente para medir el rendimiento del hardware y generar los conjuntos de datos requeridos para el fingerprinting del comportamiento posterior?
- RQ4: ¿Cómo se puede integrar la identificación individual de dispositivos con otras soluciones de ciberseguridad basadas en comportamiento desplegadas en escenarios del mundo real?
- RQ5: ¿Qué técnicas de ML/DL pueden lograr el mejor rendimiento para la identificación individual de dispositivos?
- RQ6: ¿Cómo se puede mejorar la resiliencia de los modelos de identificación individual de dispositivos basados en el comportamiento del hardware frente a ataques adversarios centrados en contexto y ML/DL?
- RQ7: ¿Pueden las técnicas de ML/DL basadas en atención, como los transformers, permitir la identificación individual de dispositivos siguiendo un enfoque de detección de anomalías? ¿Qué recursos se requieren?

II Objetivos

El objetivo principal de esta tesis doctoral es avanzar en el campo del fingerprinting del comportamiento de dispositivos para mejorar la identificación y seguridad de los dispositivos IoT, con un enfoque particular en computadoras de placa única y sistemas con recursos limitados. La investigación tiene como objetivo desarrollar metodologías y frameworks de trabajo novedosos para la identificación individual de dispositivos, aprovechando las técnicas de ML y DL. Al examinar varios escenarios de aplicación, incluyendo Ciudades Inteligentes, Industria 4.0 y el Internet de las Cosas del Campo de Batalla, la tesis tiene como objetivo abordar las crecientes amenazas de ciberseguridad, como el despliegue no autorizado de dispositivos, y otros problemas emergentes debido al crecimiento exponencial de dispositivos interconectados en el mundo computacional basado en redes.

La tesis explora diferentes aspectos del fingerprinting del comportamiento de dispositivos, incluyendo fuentes de datos, técnicas, escenarios de aplicación y conjuntos de

datos. La investigación se centra en resolver los problemas únicos enfrentados al definir metodologías para identificar computadoras de placa única idénticas basadas en el fingerprinting del comportamiento del hardware y ML. Algunos de estos problemas abiertos son la disponibilidad de datos y la integración de la solución en escenarios del mundo real. Por lo tanto, se exploran como parte de los objetivos de investigación. La investigación también investiga ataques adversarios y defensas en el fingerprinting de IoT, con el objetivo de desarrollar arquitecturas resilientes. Por último, el trabajo explora el problema de la autenticación, donde cada dispositivo debe ser reconocido sin considerar a otros y, por lo tanto, son necesarios modelos de ML más complejos. A partir del objetivo de cubrir los aspectos anteriores, se derivan varios objetivos específicos como se presenta a continuación, indicando las preguntas de investigación relacionadas con ellos:

O.1. Analizar el estado actual del arte respecto al fingerprinting del comportamiento de dispositivos para la identificación individual a través de una revisión completa, análisis y comparación de fuentes de datos, técnicas, aplicaciones y conjuntos de datos para guiar investigaciones y soluciones futuras (RQ1).

O.2. Identificar brechas existentes en la literatura en la identificación individual de dispositivos basada en el fingerprinting del comportamiento del hardware y los pasos requeridos para cubrirlas (RQ1).

O.3. Abordar los desafíos de identificación de dispositivos de placa única con una metodología novedosa aprovechando el fingerprinting del comportamiento del hardware, técnicas de ML/DL y propiedades esenciales para mejorar la precisión y robustez (RQ2).

O.4. Desarrollar una solución de benchmarking de hardware de bajo nivel para Computadoras de Placa Única para habilitar soluciones de IA basadas en Edge y gestión versátil de dispositivos a través de extensos conjuntos de datos de rendimiento (RQ3).

O.5. Integrar la solución de identificación de dispositivos de placa única en un framework de seguridad utilizando fingerprinting del comportamiento y técnicas de ML/DL para detectar con precisión diversos ciberataques (RQ4).

O.6. Encontrar las mejores técnicas basadas en ML/DL para la identificación individual, iterando sobre los conjuntos de datos disponibles para lograr el mejor rendimiento (RQ5).

O.7. Investigar el impacto de ataques contextuales y centrados en ML/DL contra soluciones de identificación de dispositivos basadas en el comportamiento del hardware aprovechando modelos de ML/DL (RQ6).

O.8. Implementar y evaluar mecanismos de defensa para fortalecer la resiliencia de la solución contra ataques adversarios mientras se mantiene su rendimiento en el contexto de la identificación de dispositivos basada en el comportamiento del hardware (RQ6).

O.9. Explorar el problema de la autenticación individual para verificar las capacidades de modelos de ML/DL basados en atención para la detección de anomalías en los datos de comportamiento del hardware (RQ7).

III Metodología

Esta Tesis Doctoral se llevó a cabo siguiendo un enfoque científico basado en el estudio continuo del estado del arte y el análisis de los resultados obtenidos durante las distintas

etapas de la investigación. Esta tesis se define como un conjunto de seis artículos publicados en revistas de alto impacto indexadas en los Journal Citation Reports (JCR).

El primer paso fue resolver la RQ1 y proporcionar una visión completa del fingerprinting del comportamiento de dispositivos. Se realizó una amplia exploración de la literatura existente para proporcionar una comprensión integral del estado actual de las soluciones de fingerprinting del comportamiento de dispositivos en ciberseguridad. Esta exploración abarcó una variedad de publicaciones académicas, que abarcan revistas y actas de conferencias, asegurando una amplia cobertura del tema. La literatura recopilada fue meticulosamente categorizada y analizada en función de criterios específicos, como las fuentes de datos utilizadas para el fingerprinting, las técnicas empleadas, los escenarios de aplicación abordados y los conjuntos de datos utilizados. Esta categorización granular facilitó la identificación de tendencias prevalentes, métodos comúnmente utilizados y posibles brechas dentro del cuerpo actual de conocimiento. El análisis proporcionó una comprensión matizada de cómo se aplica el fingerprinting del comportamiento de dispositivos en diferentes dominios de ciberseguridad, destacando su versatilidad y la variedad de formas en que se puede implementar. Esta revisión extensa culminó en una visión holística del campo, ofreciendo valiosas perspectivas y una base sólida para futuros esfuerzos de investigación en el fingerprinting del comportamiento de dispositivos. Todas estas consideraciones resultaron en la primera publicación de esta Tesis Doctoral, presentada en el primer capítulo (Article 1–IEEE_COMST), que cubrió los dos primeros Objetivos de la tesis.

Después de realizar el análisis del estado del arte e identificar las brechas actuales en las soluciones de identificación individual de IoT, el foco se trasladó a resolver la RQ2 y abordar el Objetivo 3, que trata el problema de identificación individual de dispositivos IoT. Para identificar de manera única dispositivos IoT basándose en su comportamiento de hardware mientras se aprovechan las técnicas de ML y DL, se empleó un enfoque estructurado. El proceso comenzó con la recopilación de un conjunto diverso y extenso de datos de comportamiento de hardware de una variedad de modelos de RPi. Estos datos sirvieron como base para el análisis posterior. Tras la fase de recopilación de datos, se aplicaron técnicas de preprocesamiento para refinar y preparar los datos para los modelos de ML y DL. Se seleccionaron y entrenaron varios modelos de ML/DL en los datos preprocesados, lo que resultó en el desarrollo de algoritmos de identificación. El rendimiento de estos algoritmos fue evaluado rigurosamente utilizando métricas comunes de clasificación de ML/DL y dispositivos reales de SBC, asegurando su precisión y efectividad en escenarios del mundo real. Posteriormente, se comparó la metodología con otros enfoques existentes en la literatura, aunque no presentaron un conjunto estructurado de pasos para desarrollar una solución de identificación. En esta comparación, se contrastó prácticamente la mejora del rendimiento de la metodología. Este enfoque integral facilitó una metodología clara y práctica para la identificación única de dispositivos IoT basada en su comportamiento de hardware, aprovechando los últimos avances en ML y DL. La propuesta y validación de la metodología de identificación de dispositivos individuales resultó en el segundo capítulo (Article 2–JNCA).

En este punto, surgió una limitación en la línea de investigación con respecto a la disponibilidad de datos y cómo recopilarlos. Como parte de esta Tesis Doctoral, se desarrolló una herramienta de benchmarking especializada y un conjunto de datos para abordar el desafío de la recopilación extensiva de datos para medir el rendimiento del hardware y generar los conjuntos de datos requeridos para el fingerprinting del comportamiento posterior. Esta herramienta fue meticulosamente diseñada para medir y registrar el rendimiento de varios componentes de hardware en una variedad de dispositivos IoT, asegurando un proceso de recopilación de datos estandarizado y completo. La herramienta se ejecutó luego

en algunos dispositivos SBC para extraer un conjunto de datos completo y realista. Los datos recopilados fueron sometidos a rigurosos procesos de validación y control de calidad, verificando su precisión y fiabilidad. Además de la recopilación de datos, se proporcionaron pautas claras sobre cómo estructurar y almacenar los datos, facilitando su integración en conjuntos de datos integrales. Este enfoque meticuloso aseguró que los datos recopilados no solo fueran extensos, sino también de alta calidad, proporcionando una base sólida para aplicaciones y análisis de fingerprinting del comportamiento posteriores. La aplicación de benchmarking, junto con el conjunto de datos recopilados, están disponibles públicamente para otros investigadores en el área. Este trabajo resultó en el tercer capítulo de esta tesis (Article 3–IoT), respondiendo a la RQ3 y completando el Objetivo 4.

Una vez verificada la viabilidad de la identificación individual con el conjunto de datos exhaustivos recopilados mediante la aplicación de benchmark, se desarrolló un framework holístico e interoperable. Este framework buscó integrar la identificación individual de dispositivos con otras soluciones de ciberseguridad basadas en comportamiento. Este framework se basó en un análisis extenso del fingerprinting del comportamiento de dispositivos y soluciones de ciberseguridad existentes, asegurando una comprensión integral de los componentes y procesos necesarios. El framework diseñado facilitó una integración sin problemas, permitiendo compartir y utilizar fácilmente las huellas digitales y los perfiles de comportamiento de los dispositivos en diferentes soluciones de seguridad. Para lograr esto, el framework fue diseñado para ser altamente modular y escalable, permitiendo una fácil integración con una variedad de soluciones de ciberseguridad basadas en el comportamiento. Ya fueran sistemas de detección de intrusiones, soluciones de gestión de información y eventos de seguridad, o herramientas avanzadas de protección contra amenazas, el framework era capaz de interactuar con ellas, mejorando sus capacidades y proporcionando una capa adicional de seguridad. Esta interoperabilidad fue crucial para mejorar la seguridad general de la red y asegurar la efectividad de las soluciones integradas. Para validar el rendimiento del framework, se realizaron extensas simulaciones y pruebas en el mundo real, evaluando su funcionalidad en diversos escenarios y contra varios modelos de amenazas. La validación en el mundo real se realizó considerando un escenario de Internet of Battlefield Things (IoBT) basado en la plataforma ElectroSense [39]. Luego, se integró una solución de monitoreo de eventos de kernel y llamadas al sistema junto con la solución de identificación individual para proporcionar un enfoque de seguridad completo para la plataforma. Los resultados de estas evaluaciones confirmaron la efectividad del framework, demostrando su capacidad para integrar la identificación individual de dispositivos con otras soluciones de ciberseguridad basadas en el comportamiento y mejorar la postura de seguridad de la red. Este trabajo resolvió RQ4 y el Objetivo 5, disponible en el cuarto capítulo de la tesis (Article 4–IEEE_COMMAG).

El siguiente trabajo realizado en la Tesis Doctoral, presentado en el quinto capítulo de este documento (Article 5–FGCS) y alineado con las quintas y sextas preguntas de investigación (Objetivos 6, 7 y 8), abordó el desafío de las amenazas de seguridad. Se centró en encontrar la mejor técnica de ML/DL para la identificación y luego mejorar su resiliencia para una identificación fiable de dispositivos individuales basada en el comportamiento del hardware. Esto involucró un análisis exhaustivo de los posibles ataques adversarios, enfatizando la comprensión de la naturaleza y el impacto de estos sofisticados ataques. El modelo de amenazas presentado aborda los ataques adversarios centrados en contexto y ML/DL, asegurando que los modelos de identificación sigan siendo confiables y seguros incluso frente a amenazas sofisticadas. Se verificó el impacto de los ataques propuestos en la literatura, demostrando la vulnerabilidad de la solución a los ataques adversarios. Luego, se formularon e implementaron contramedidas y estrategias de mitigación para fortalecer

los modelos de identificación contra manipulaciones adversarias. Se incorporaron técnicas de entrenamiento adversario y destilación de conocimiento [40] para mejorar la resiliencia del modelo contra ataques adversarios.

El último trabajo de este compendio se enfocó en la séptima pregunta de investigación (RQ7) y el Objetivo 9 (Article 6–COSE). Se ocupó del problema de autenticación en lugar de la identificación. Como se explicó en la Introducción, la principal diferencia entre estos problemas es que para la autenticación, solo se pueden emplear datos del dispositivo que se está autenticando en el proceso de entrenamiento y se necesita una mayor granularidad en el procesamiento de datos. Este hecho hace que el problema de autenticación sea mucho más difícil, ya que la distribución de datos de rendimiento del hardware en dispositivos del mismo modelo generalmente se superpone en gran proporción. La metodología seguida para resolver este problema implicó el procesamiento de series temporales para entrenar modelos autoencoder basados en transformers [41], con cada modelo adaptado a un dispositivo específico. La recolección y preprocesamiento de datos fueron similares a los aplicados en los trabajos centrados en identificación, variando solo el tamaño de la ventana de tiempo empleada. Luego, se diseñaron modelos transformers basados en atención para funcionar como mecanismos de detección de anomalías, identificando cualquier desviación del comportamiento del hardware esperado que indicaría un dispositivo no autorizado. Este trabajo puso un fuerte énfasis en la aplicación práctica de su metodología en escenarios del mundo real. Por lo tanto, realizó pruebas y validaciones extensas para asegurarse de que el enfoque no solo fuera teóricamente sólido sino también efectivo en la práctica.

Esta tesis crea una narrativa holística, abordando los aspectos críticos del fingerprinting del comportamiento de dispositivos desde el conocimiento fundamental y técnicas prácticas de identificación hasta la recolección de datos, integración con soluciones de seguridad más amplias y resiliencia frente a ataques adversarios. Finalmente, también se explora el problema de autenticación utilizando enfoques modernos basados en transformers. Este enfoque integral asegura una comprensión profunda y una aplicación robusta del fingerprinting del comportamiento de dispositivos en el ámbito de la ciberseguridad. Esta metodología permitió cumplir con los objetivos definidos en la tesis, previamente presentados en la Sección II.

IV Resultados

La primera publicación de la Tesis Doctoral, presentada en (Article 1–IEEE_COMST), ofreció un estudio exhaustivo sobre el modelado del comportamiento de dispositivos, proporcionando un análisis y síntesis amplios de soluciones aplicadas en el ámbito de la ciberseguridad, junto con valiosos conocimientos y hallazgos. Los resultados destacaron un amplio espectro de fuentes de datos, técnicas, escenarios de aplicación y conjuntos de datos prevalentes en la literatura actual, subrayando la naturaleza multifacética del modelado del comportamiento de dispositivos. Uno de los hallazgos clave gira en torno a la diversidad de fuentes de datos utilizadas para el modelado del comportamiento de dispositivos. Los resultados indicaron que se emplea una amplia variedad de tipos de datos, que van desde el tráfico de red y registros del sistema hasta atributos específicos del hardware. Esta diversidad subraya la versatilidad de las soluciones de modelado de dispositivos, demostrando su aplicabilidad en diferentes capas de la arquitectura del sistema y de la red.

En términos de técnicas de modelado, el estudio reveló un rico panorama de metodologías, cada una con sus fortalezas y capacidades únicas. Los resultados muestran que no hay una solución única para todos, con diferentes técnicas que atienden a requisitos y escenarios

específicos. Esta variedad asegura que los profesionales e investigadores tengan una plétora de opciones para elegir, permitiéndoles adaptar sus soluciones de modelado de dispositivos para satisfacer las necesidades específicas de su aplicación. En cuanto a los escenarios de aplicación, el documento esbozó una amplia gama de contextos en los que se emplea el modelado del comportamiento de dispositivos. Desde mejorar la seguridad de la red y detectar dispositivos no autorizados, hasta facilitar la autenticación de dispositivos y la verificación de integridad, las aplicaciones son vastas y variadas. Esto destacó el papel crucial que juega el modelado del comportamiento de dispositivos en el refuerzo de las medidas de ciberseguridad, proporcionando una capa adicional de seguridad y confianza en entornos digitales. El examen de los conjuntos de datos reveló que, aunque hay numerosos conjuntos de datos disponibles para fines de investigación y desarrollo, todavía se necesita más conjuntos de datos completos y estandarizados. Los resultados señalan que los conjuntos de datos existentes varían significativamente en términos de tamaño, calidad y relevancia, indicando una brecha que necesita ser abordada para avanzar aún más en el campo. La disponibilidad de conjuntos de datos estandarizados de alta calidad es fundamental para la validación y evaluación comparativa de soluciones de modelado de dispositivos, asegurando su fiabilidad y efectividad en escenarios del mundo real.

En la última sección del documento, se presentó un análisis profundo y perspicaz, encapsulando las lecciones aprendidas, identificando tendencias predominantes y destacando los desafíos enfrentados en el dominio del modelado del comportamiento de dispositivos dentro de la ciberseguridad. Esta sección sirve como una reflexión crítica sobre el estado actual del campo, ofreciendo orientación para futuras investigaciones e implementaciones prácticas, incluyendo los próximos trabajos en esta tesis doctoral.

La segunda publicación (Article 2–JNCA) proporcionó una exploración y análisis detallados de los pasos necesarios para identificar de manera única dispositivos IoT basados en su comportamiento de hardware, con un énfasis particular en el aprovechamiento de técnicas de ML y DL. Los resultados obtenidos de esta investigación ofrecen valiosas perspectivas sobre las complejidades del modelado de dispositivos y las prácticas de implementar tales metodologías en escenarios del mundo real.

La investigación identificó propiedades esenciales para la identificación de dispositivos de placa única, incluyendo la unicidad, estabilidad, diversidad, escalabilidad, eficiencia, robustez y seguridad. Luego se introdujo una metodología novedosa, basada en el modelado de comportamiento para identificar dispositivos de placa única idénticos mientras se cumplen las propiedades mencionadas. Esta metodología utiliza los diferentes componentes integrados del sistema, junto con técnicas de ML/DL, para comparar el comportamiento interno de los dispositivos y detectar variaciones que ocurrieron durante los procesos de fabricación. Un resultado clave del estudio fue la identificación y validación de atributos de hardware específicos que pueden usarse como indicadores confiables para el modelado de dispositivos. Estos atributos, cuando se analizan y procesan correctamente, han demostrado proporcionar una firma única para cada dispositivo, facilitando una identificación precisa y eficiente. Los resultados subrayan la importancia de seleccionar la combinación adecuada de atributos de hardware, ya que esta elección impacta significativamente en la efectividad del proceso de modelado.

El documento también destacó el papel crítico del aislamiento de hardware en el flujo de trabajo del modelado de dispositivos. Los resultados demostraron que el manejo y la preparación cuidadosos de los atributos de hardware son fundamentales, ya que esto asegura la integridad del proceso de modelado y mejora la precisión de la identificación de dispositivos. Además, se ha demostrado que la aplicación de varias técnicas de preprocesamiento de datos, incluyendo la normalización y reducción de dimensionalidad, contribuye

positivamente al resultado del proceso de modelado.

La integración de técnicas de ML y DL en el proceso de modelado de dispositivos también ha demostrado introducir un elemento de adaptabilidad y aprendizaje, permitiendo que el sistema evolucione y mejore con el tiempo. Se exploraron técnicas de Deep Learning, incluyendo varias configuraciones de redes neuronales, por su capacidad para aprender y modelar las huellas de dispositivos directamente a partir de datos brutos de hardware. Clasificadores de ML y DL estuvieron entre las arquitecturas probadas, elegidos por su habilidad en manejar datos secuenciales y de series temporales, que son prevalentes en la información de comportamiento del hardware. Además, el documento abordó el desafío del sobreajuste del modelo, especialmente pertinente al emplear modelos complejos como redes neuronales profundas. Se aplicaron estrategias como la validación cruzada y la regularización, asegurando que los modelos generalicen bien a datos no vistos y mantengan un alto rendimiento cuando se desplieguen en entornos del mundo real. La metodología se validó en un entorno real, consistiendo en 15 dispositivos Raspberry Pi 4 Model B y 10 Raspberry Pi 3 Model B+, cuyo rendimiento de CPU y GPU fue analizado. Los resultados muestran una Tasa de Positivos Verdaderos (TPR) promedio del 91.9% con un modelo XGBoost, logrando la identificación de todos los dispositivos estableciendo un umbral del 50% en el proceso de evaluación. Además, se entabló una discusión crítica sobre la metodología propuesta, comparando la solución propuesta con trabajos relacionados, destacando las propiedades de modelado no cumplidas por otras soluciones y proporcionando lecciones valiosas aprendidas y limitaciones de la metodología presentada.

El principal resultado en la tercera publicación de la tesis (Article 3–`IoT`) fue el desarrollo de una aplicación de evaluación comparativa de hardware de bajo nivel adaptada para SBCs, abordando la necesidad de aplicaciones y conjuntos de datos de evaluación comparativa de bajo nivel en el ámbito de IoT. La evaluación comparativa se denominó *LwHBench* y se centra en medir el rendimiento de CPU, GPU, Memoria y Almacenamiento, teniendo en cuenta las limitaciones de componentes inherentes en los SBCs. La aplicación se implementó específicamente para dispositivos Raspberry Pi. Se ejecutó durante 100 días en un conjunto de 45 dispositivos para generar un extenso conjunto de datos que contiene 2386126 vectores en más de 4GB de datos. Este conjunto de datos allana el camino para la aplicación de técnicas de AI en escenarios donde los datos de rendimiento pueden ayudar en el proceso de gestión de dispositivos. Para mostrar la capacidad inter-escenario del conjunto de datos, el documento también presentó una serie de casos de uso habilitados por AI relacionados con la identificación de dispositivos y el impacto del contexto en el rendimiento. En una configuración práctica, la aplicación de evaluación comparativa se adaptó y aplicó a un escenario que involucra tres dispositivos RockPro64, demostrando su versatilidad y aplicabilidad en entornos del mundo real.

En la cuarta publicación de la tesis (Article 4–`IEEE_COMMAG`), la investigación se adentra en el campo emergente y altamente dinámico del Internet of Battlefield Things (IoBT). Se centró particularmente en el papel crucial de las comunicaciones inalámbricas dentro de este ámbito. En este intrincado escenario de batalla, una miríada de dispositivos, que van desde soldados hasta diversos equipos militares, interactúan en tiempo real, intercambiando información de manera inalámbrica y formando una red compleja de entidades interconectadas. El escenario propuesto profundiza en tres casos de uso principales: identificación de dispositivos IoT, detección de malware y detección de ataques de Falsificación de Datos de Detección de Espectro (SSDF, por sus siglas en inglés).

Para resolver estos casos de uso, se introdujo un framework de modelado de comportamiento IoT, denominado *SpecForce*, que fue meticulosamente diseñado para mejorar la seguridad de los sensores de espectro IoBT, componentes cruciales en el monitoreo del

espectro de frecuencias, la transmisión sobre bandas desocupadas, la intercepción de transmisiones enemigas y la decodificación de información valiosa. En este escenario del mundo real, la identificación individual de dispositivos es esencial para evitar posibles ataques basados en manipulaciones de identidad. *SpecForce* se destaca como una solución robusta, empleando modelado de comportamiento de dispositivos junto con técnicas de ML/DL. El framework fue hábil al considerar fuentes de datos heterogéneas, mejorando su capacidad para detectar y mitigar eficazmente una amplia gama de amenazas cibernéticas. Se hizo hincapié en garantizar la integridad y fiabilidad de las comunicaciones dentro del escenario de batalla, un aspecto crítico considerando la escasez de espectro y el creciente número de dispositivos IoBT. El framework incluyó un módulo de ciberseguridad basado en AI que emplea algoritmos de clasificación de ML/DL para identificar diferentes sensores de espectro IoBT basados en dispositivos RPi. El documento proporciona un análisis exhaustivo de varios modelos de ML/DL para la clasificación. Los resultados indican que Random Forest y XGBoost son los modelos con mejor rendimiento, logrando más del 91% de TPR. El documento también discute un caso de uso que demuestra la capacidad del sistema para identificar de manera única 25 sensores de espectro IoBT, abordando ataques enfocados en la identidad y mejorando la seguridad.

Como se comentó anteriormente, *SpecForce* estaba equipado con otros enfoques de ciberseguridad utilizando monitoreo de eventos de kernel y comportamiento de llamadas al sistema para detectar ataques cibernéticos de nivel superior. En el contexto de la detección de ataques SSDF, el framework permite el monitoreo de llamadas al sistema de sensores de espectro IoBT, con el objetivo de detectar varios ataques SSDF. Las llamadas al sistema se procesan para generar vectores de características que modelan las actividades de detección de espectro, utilizando algoritmos de detección de anomalías para distinguir entre comportamientos normales y maliciosos. Los resultados muestran un alto rendimiento en el reconocimiento de comportamientos normales, con más del 99% de Tasa de Negativos Verdaderos (TNR) y un TPR loable de más del 92% para la detección de ataques SSDF. Además, en lo que respecta a la detección de malware heterogéneo (botnets, puertas traseras, etc.), se logró un alto rendimiento al monitorear eventos de kernel combinados con detección de anomalías de ML/DL, con un TPR del 90% y un TNR del 96%.

En el contexto de ataques adversariales, la siguiente publicación de la Tesis Doctoral (Article 5–FGCS) arrojó luz sobre las posibles vulnerabilidades y amenazas que pueden comprometer la integridad de los mecanismos de modelado e identificación de dispositivos. Discute varios vectores de ataque, ilustrando cómo las entidades maliciosas podrían manipular o eludir mecanismos de identificación basados en hardware para alcanzar sus nefastos objetivos. El artículo subraya la necesidad de estrategias de defensa integrales capaces de mitigar los riesgos asociados con ataques adversariales, asegurando la robustez de los procesos de identificación de dispositivos. El primer resultado principal de este trabajo fue la mejora en los resultados de identificación alcanzados en trabajos anteriores. Utilizando enfoques de series temporales combinados con modelos de DL, los resultados de identificación se incrementaron a un TPR promedio de +0.96 con un TPR mínimo de 0.80 en los 45 dispositivos RPi utilizados para la validación, aprovechando un modelo combinado de LSTM+1D-CNN. Con respecto a ataques adversariales, tanto ataques centrados en el contexto como en ML/DL se aplican para evaluar la robustez del modelo de identificación de dispositivos. Se hace una mención específica de un ataque de contexto basado en la temperatura, que, curiosamente, se encontró ineficaz para interrumpir el proceso de identificación de dispositivos, ya que el aislamiento de hardware durante la recolección de datos ya estaba considerando la mitigación del impacto del contexto. Sin embargo, el documento reconoce el éxito de ciertos ataques de evasión de ML/DL de última generación, como BIM,

MIM y JSMA.

En el lado de la defensa, el documento proporciona una exploración exhaustiva de varias estrategias y metodologías destinadas a proteger los dispositivos IoT de amenazas adversariales. Profundiza en enfoques basados en ML y contexto, evaluando su efectividad para mejorar la seguridad y fiabilidad del modelado e identificación de dispositivos. Los ataques basados en contexto centrados en la temperatura son ineficaces debido a las medidas de estabilidad y aislamiento de hardware tomadas durante la recolección de datos. En cuanto a las defensas contra ataques de evasión de ML/DL, se aplican la destilación de conocimientos y el entrenamiento adversarial para reducir el impacto de los ataques. Los resultados destacan el papel crítico de estos mecanismos de defensa para mantener la integridad de los ecosistemas IoT, asegurando que los dispositivos sean identificados con precisión y que las entidades maliciosas sean frustradas. En términos de rendimiento, la tasa de éxito de los ataques se redujo de 0.88 a 0.17 en el peor de los casos sin causar una degradación sustancial en el rendimiento. Finalmente, se utilizan varias métricas de seguridad para evaluar la resiliencia de las redes neuronales contra perturbaciones y variaciones adversarias en la entrada. Se discuten métricas como el puntaje CLEVER, la sensibilidad a la pérdida y la robustez empírica [40], proporcionando ideas sobre cómo se puede cuantificar y evaluar la robustez de los modelos de ML.

La última publicación de la tesis doctoral (Article 6–COSE) se centró en el problema de la autenticación individual. Propuso un framework de autenticación que utilizaba datos de rendimiento del hardware y modelos autoencoder basados en transformers. El diseño del framework está respaldado por un modelo de amenazas que describe los desafíos de seguridad encontrados al implementar la autenticación basada en hardware en contextos IoT. Como en trabajos anteriores, se monitorearon componentes clave del hardware, como la CPU, GPU, RAM y almacenamiento, para la recolección de datos de modelado. Estas huellas fueron utilizadas como datos de series temporales, aplicando ventanas de tiempo de 10 a 100 valores. Las series temporales generadas se utilizan entonces para entrenar modelos transformers para la detección de anomalías, adaptados a cada dispositivo individual, con el objetivo de representarlo y autenticarlo con precisión. La efectividad del framework se demostró además a través de su aplicación en un sistema de crowdsensing de espectro utilizando dispositivos Raspberry Pi. Los modelos transformers se compararon con enfoques LSTM y 1D-CNN en términos de rendimiento. Aquí, en una serie de experimentos rigurosos que involucraron 45 dispositivos para validación, cada modelo transformador del dispositivo demostró ser capaz de autenticarlo con precisión. En contraste, otros enfoques no fueron capaces de autenticar de manera única todos los dispositivos. El enfoque logró una impresionante TPR promedio de 0.74 ± 0.13 y mantuvo un FPR máximo promedio de 0.06 ± 0.09, subrayando su potencial para mejorar significativamente la autenticación, la seguridad y la confiabilidad en aplicaciones de crowdsensing. Además, también se analizó el uso de recursos de los diferentes enfoques probados, confirmando uno de los principales inconvenientes de los modelos transformers; esta fue la técnica de ML/DL que utilizó más recursos en términos de tiempo y memoria.

V Conclusiones y trabajo futuro

Esta tesis proporciona un examen exhaustivo y completo del campo del fingerprinting del comportamiento de dispositivos, con un enfoque específico en su aplicación dentro del ámbito de la ciberseguridad, y un énfasis matizado en la identificación y autenticación de dispositivos de placa única y IoT. La investigación comienza con una revisión extensa y sistemática del panorama actual, capturando la amplitud y profundidad de las soluciones

de fingerprinting del comportamiento de dispositivos que se han explorado e implementado dentro del dominio de la ciberseguridad. Esta fase inicial de exploración sirve para sentar una base sólida de conocimiento, desentrañando las complejidades de varias fuentes de datos, técnicas de fingerprinting, escenarios de aplicación y los conjuntos de datos que prevalecen dentro de las esferas académicas y prácticas de este campo. A partir de esta exploración, se deriva un nuevo conjunto de lecciones y tendencias de la literatura, dando una visión holística de trabajos anteriores. Sin embargo, la principal novedad desde una perspectiva de investigación es la lista de desafíos identificados en la literatura, que no estaban definidos antes y allanaron el camino para futuras investigaciones.

A medida que la historia avanza, el foco se desplaza a la implementación práctica de la identificación de dispositivos, con la propuesta de una metodología clara y detallada para identificar de manera única dispositivos IoT. Este proceso está intrínsecamente ligado a las capacidades proporcionadas por las técnicas de ML y DL. Estas herramientas computacionales avanzadas pueden ser aprovechadas para descifrar las sutiles matices del comportamiento del hardware, asegurando un alto nivel de autenticidad e integridad para dispositivos integrados dentro de una red. La importancia de este proceso no puede ser exagerada, ya que juega un papel fundamental en la salvaguarda de la seguridad y confiabilidad de dispositivos interconectados, formando un componente crítico de la infraestructura de ciberseguridad más amplia. Para resaltar esta importancia, la metodología presentada se compara con otros trabajos en el campo, donde no se siguió un conjunto metodológico de pasos. Las soluciones anteriores no fueron capaces de realizar una identificación completa en el escenario de dispositivo real utilizado para la validación. Por lo tanto, la metodología se convierte en el procedimiento de vanguardia para desarrollar una solución funcional de identificación individual.

Construyendo sobre la metodología de identificación establecida, la narrativa profundiza en el aspecto crítico de la recolección de datos, presentando un enfoque integral para la adquisición sistemática de datos de rendimiento del hardware. Generar los conjuntos de datos requeridos para el fingerprinting de comportamiento posterior es esencial. Estos conjuntos de datos ayudan a asegurar que los modelos de identificación se entrenen y validen con datos que son tanto extensos como precisos. La meticulosa atención al detalle en este proceso asegura la fiabilidad de los datos, estableciendo un alto estándar para la calidad de la información utilizada en el fingerprinting del comportamiento de dispositivos. Como resultado de la herramienta propuesta para la recolección de datos, se genera un conjunto de datos exhaustivo para ser aplicado en los siguientes pasos de la tesis. Este conjunto de datos se publica para ser utilizado por la comunidad de investigación en el campo, junto con la aplicación de benchmark empleada para generar los datos. Este es uno de los primeros conjuntos de datos públicos de datos de rendimiento del hardware que se centra en el problema de identificación.

Con una base sólida de datos en su lugar, la exploración luego navega hacia la integración de la identificación individual de dispositivos dentro del ecosistema más amplio de soluciones de ciberseguridad basadas en comportamiento. Esta integración es fundamental, ya que asegura que los procedimientos de identificación de dispositivos no operen de manera aislada, sino que estén vinculados sin problemas con otros frameworks de seguridad generales. Para esta integración, el enfoque se coloca estratégicamente en el Internet de las Cosas del Campo de Batalla. Se emplea un escenario del mundo real para integrar la solución de identificación basada en hardware con un monitoreo de comportamiento de nivel superior para la detección de malware y ataques de falsificación de datos de espectro. Específicamente, el enfoque unificado monitorea hardware, eventos de kernel y llamadas al sistema para proporcionar una solución de seguridad unificada basada en comportamiento.

Este enfoque holístico mejora la resiliencia del entorno, fortaleciendo sus defensas contra una miríada de amenazas y vulnerabilidades cibernéticas, y asegurando un entorno digital robusto y seguro. El framework unificado demuestra experimentalmente sus capacidades para detectar diferentes muestras de malware y ataques centrados en datos de espectro, así como realizar la identificación individual de los sensores desplegados. Hasta donde sabemos, este es el primer framework que combina estas capacidades de ciberseguridad juntas.

A medida que el trabajo alcanza su culminación, el enfoque se dirige a la seguridad de los modelos de identificación de dispositivos en sí mismos, abordando específicamente los desafíos planteados por los ataques adversarios. Se analiza el panorama de los ataques adversarios, centrándose particularmente en las amenazas conscientes del contexto y centradas en ML/DL que representan riesgos significativos para la integridad de las metodologías de identificación de dispositivos. Luego, el enfoque se adentra en estrategias y metodologías diseñadas para mejorar la resiliencia de los modelos de identificación basados en el comportamiento del hardware. Se pone un énfasis particular en contrarrestar amenazas adversarias sofisticadas, incluyendo ataques basados en contexto y centrados en ML/DL. Los ataques basados en contexto son ineficaces contra la solución, pero los ataques de evasión dirigidos a los modelos de identificación logran altas tasas de éxito. Por lo tanto, se aplican técnicas de defensa basadas en entrenamiento adversario y destilación de conocimiento. Esto asegura que las metodologías de identificación de dispositivos sigan siendo confiables, seguras y efectivas, incluso frente a amenazas cibernéticas en evolución y complejas. Este es uno de los primeros trabajos que demuestra experimentalmente la efectividad de los ataques adversarios en modelos de ML para la identificación de dispositivos IoT, y también de métodos de defensa de vanguardia.

Finalmente, la investigación explora el problema más complejo de la autenticación de dispositivos individuales, donde solo se pueden aprovechar los datos de un dispositivo para la generación del modelo. La metodología aplicada para la identificación se ajusta para su aplicación en la autenticación. La principal diferencia es el cambio del modelo clasificador por un modelo de detección de anomalías por dispositivo. Sin embargo, se necesitan modelos de ML/DL más potentes para resolver este problema, ya que las distribuciones de datos se superponen entre dispositivos del mismo modelo. Para resolver este problema, se emplean modelos transformers. Utilizando ventanas de tiempo grandes para el procesamiento de datos (100 valores), el enfoque basado en transformers mejora los resultados logrados por modelos de vanguardia anteriores, como LSTM, 1D-CNN y su combinación. En contraste, se emplea más tiempo de entrenamiento y memoria en el proceso de generación del modelo. Este resultado confirma la efectividad de la arquitectura transformadora en un área novedosa. A diferencia de las arquitecturas de modelos existentes en la literatura, aborda con éxito el problema de autenticar individualmente dispositivos IoT basados en su rendimiento de hardware en el escenario experimental estudiado.

Mirando hacia adelante, hay un vasto horizonte de oportunidades para trabajos futuros basados en los cimientos establecidos por esta tesis en el dominio del fingerprinting del comportamiento de dispositivos. La primera iteración de trabajos futuros podría profundizar en la evaluación de ataques adversarios sobre los modelos transformers no supervisados empleados para resolver el problema de identificación individual, y la consecuente aplicación de técnicas de defensa. Estos resultados cerrarán el trabajo sobre el tema de autenticación de manera similar a la metodología aplicada para el problema de identificación individual.

Otra avenida prometedora es expandir el alcance del fingerprinting del comportamiento de dispositivos para abarcar una gama más amplia de dispositivos y contextos. El trabajo actual se ha centrado predominantemente en dispositivos de placa única e IoT. Sin em-

bargo, los principios y metodologías desarrollados podrían adaptarse y aplicarse a otros tipos de dispositivos y redes, como sistemas de control industrial, sistemas automotrices y dispositivos de hogares inteligentes. Investigaciones futuras podrían explorar las sutilezas y requisitos específicos de estos diferentes contextos, adaptando las técnicas de fingerprinting para ajustarse a las características únicas de cada categoría de dispositivo y escenario de uso. Uno de los ámbitos cruciales destacados para futuras exploraciones es la búsqueda continua de nuevas y diversas fuentes de datos. El campo se beneficiaría significativamente de ampliar el alcance de la recolección de datos, capturando una gama más amplia de comportamientos de dispositivos y asegurando un conjunto de datos más rico y completo para análisis. Este esfuerzo no solo se limita a aumentar la cantidad de datos sino también enfatiza la importancia de mejorar la calidad y fiabilidad de los datos recopilados. Trabajos futuros en esta área podrían explorar metodologías avanzadas de recolección de datos, tecnologías innovadoras de sensores y técnicas novedosas de preprocesamiento de datos, todo con el objetivo de asegurar que los datos utilizados para el fingerprinting del comportamiento de dispositivos sean de la más alta calidad.

Construyendo sobre el tema de datos, hay un claro llamado para el desarrollo y refinamiento de técnicas de fingerprinting. El futuro tiene potencial para la exploración de nuevos algoritmos, modelos de ML y arquitecturas de DL, cada uno ofreciendo capacidades y ventajas únicas para la identificación de dispositivos. El Aprendizaje Federado (FL), y más concretamente el FL descentralizado, es una de estas áreas prometedoras que vale la pena explorar en los próximos años. La continua evolución del poder computacional y las tecnologías de ML/DL abre posibilidades emocionantes para crear modelos de fingerprinting del comportamiento de dispositivos más sofisticados y precisos, capaces de discernir incluso las sutilezas más sutiles en el comportamiento de dispositivos.

Abordando el desafío de los ataques adversarios, hay una necesidad apremiante de desarrollar contramedidas robustas y estrategias de mitigación. Trabajos futuros en esta área podrían explorar enfoques innovadores para mejorar la resiliencia de los modelos de fingerprinting del comportamiento de dispositivos, con un enfoque específico en contrarrestar amenazas adversarias sofisticadas, incluyendo ataques basados en contexto y centrados en ML/DL. Esto implica no solo fortalecer los modelos de identificación sino también desarrollar mecanismos integrales de detección y respuesta a amenazas, asegurando la fiabilidad y seguridad a largo plazo de las metodologías de identificación de dispositivos.

Otra área crítica para trabajos futuros radica en mejorar la adaptabilidad de los modelos de fingerprinting del comportamiento de dispositivos. Con el rápido ritmo de avance tecnológico, los dispositivos están en constante evolución, y sus patrones de comportamiento pueden cambiar con el tiempo debido a actualizaciones de software, modificaciones de hardware o cambios en los patrones de uso. Investigaciones futuras podrían centrarse en desarrollar modelos de fingerprinting capaces de adaptarse a estos cambios, asegurando que sigan siendo precisos y confiables a lo largo del ciclo de vida del dispositivo. Esto podría implicar la integración de técnicas de aprendizaje en línea, enfoques de aprendizaje continuo o metodologías de aprendizaje por transferencia para permitir que los modelos actualicen y refinan sus perfiles de fingerprinting en respuesta a los cambios observados en el comportamiento de los dispositivos.

Finalmente, hay un potencial significativo para trabajos futuros en la integración de soluciones de fingerprinting del comportamiento de dispositivos con otras herramientas y frameworks de ciberseguridad. Esta tesis ha sentado las bases para dicha integración, demostrando los beneficios potenciales de combinar el fingerprinting del comportamiento de dispositivos con otras soluciones de seguridad basadas en comportamiento. Investigaciones futuras podrían basarse en esto, explorando formas de optimizar aún más el proceso de

integración, mejorar la interoperabilidad y maximizar las sinergias entre diferentes herramientas de seguridad. Esto podría llevar a la creación de frameworks de ciberseguridad más holísticos y resilientes, proporcionando protección integral contra una amplia gama de amenazas cibernéticas.

Bibliography

[1] K. Riad, T. Huang, and L. Ke, "A dynamic and hierarchical access control for IoT in multi-authority cloud storage," *Journal of Network and Computer Applications*, vol. 160, p. 102633, 2020.

[2] R. Fayos-Jordan, S. Felici-Castell, J. Segura-Garcia, J. Lopez-Ballester, and M. Cobos, "Performance comparison of container orchestration platforms with low cost devices in the fog, assisting internet of things applications," *Journal of Network and Computer Applications*, vol. 169, p. 102788, 2020.

[3] A. L. Perales Gómez, L. Fernández Maimó, A. Huertas Celdran, F. J. García Clemente, C. Cadenas Sarmiento, C. J. Del Canto Masa, and R. Méndez Nistal, "On the generation of anomaly detection datasets in industrial control systems," *IEEE Access*, vol. 7, pp. 177 460–177 473, 2019.

[4] S. Jagdale, "The Role of Hardware Root of Trust in Edge Devices," https://www.eetimes.eu/the-role-of-hardware-root-of-trust-in-edge-devices/, 2022, [Online; accessed 21-June-2022].

[5] E. Montalbano, "Bluetooth Spoofing Bug Affects Billions of IoT Devices," https://threatpost.com/bluetooth-spoofing-bug-iot-devices/159291/, 2020, [Online; accessed 21-June-2022].

[6] R. Francese, M. Frasca, and M. Risi, "Are iobt services accessible to everyone?" *Pattern Recognition Letters*, vol. 147, pp. 71–77, 2021.

[7] Y. Liu, J. Wang, J. Li, H. Song, T. Yang, S. Niu, and Z. Ming, "Zero-bias deep learning for accurate identification of internet-of-things (iot) devices," *IEEE Internet of Things Journal*, vol. 8, no. 4, pp. 2627–2634, 2020.

[8] A. Fuentes, "Human niche, human behaviour, human nature," *Interface Focus*, vol. 7, no. 5, p. 20160136, 2017.

[9] N. Shone, Q. Shi, M. Merabti, and K. Kifayat, "Misbehaviour monitoring on system-of-systems components," in *2013 International Conference on Risks and Security of Internet and Systems*, Oct. 2013, pp. 1–6.

[10] A. Sivanathan, H. H. Gharakheili, F. Loi, A. Radford, C. Wijenayake, A. Vishwanath, and V. Sivaraman, "Classifying IoT devices in smart environments using network traffic characteristics," *IEEE Transactions on Mobile Computing*, vol. 18, no. 8, pp. 1745–1759, 2019. DOI: 10.1109/TMC.2018.2866249

[11] S. Marchal, M. Miettinen, T. D. Nguyen, A. Sadeghi, and N. Asokan, "AuDI: Toward autonomous IoT device-type identification using periodic communication," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1402–1412, 2019. DOI: 10.1109/JSAC.2019.2904364

[12] K. Haefner and I. Ray, "ComplexIoT: Behavior-based trust for IoT networks," in *1st IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications*, Dec. 2019, pp. 56–65. DOI: 10.1109/TPS-ISA48467.2019.00016

[13] G. Manco, E. Ritacco, P. Rullo, L. Gallucci, W. Astill, D. Kimber, and M. Antonelli, "Fault detection and explanation through big data analysis on sensor streams," *Expert Systems with Applications*, vol. 87, pp. 141–156, 2017.

[14] M. Miettinen, S. Marchal, I. Hafeez, T. Frassetto, N. Asokan, A. Sadeghi, and S. Tarkoma, "IoT Sentinel: Automated device-type identification for security enforcement in IoT," in *37th IEEE International Conference on Distributed Computing Systems*, June 2017, pp. 2511–2514.

[15] T. D. Nguyen, S. Marchal, M. Miettinen, H. Fereidooni, N. Asokan, and A. Sadeghi, "DÏot: A federated self-learning anomaly detection system for IoT," in *39th IEEE International Conference on Distributed Computing Systems*, July 2019. ISSN 1063-6927 pp. 756–767. DOI: 10.1109/ICDCS.2019.00080

[16] X. Liu, B. Xiao, S. Zhang, and K. Bu, "Unknown tag identification in large RFID systems: An efficient and complete solution," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 6, pp. 1775–1788, 2014.

[17] J. Ortiz, C. Crawford, and F. Le, "DeviceMien: Network device behavior modeling for identifying unknown IoT devices," in *International Conference on Internet of Things Design and Implementation*, Apr. 2019. ISBN 9781450362832 p. 106–117. [Online]. Available: https://doi.org/10.1145/3302505.3310073. DOI: 10.1145/3302505.3310073

[18] H. Jafari, O. Omotere, D. Adesina, H. Wu, and L. Qian, "IoT devices fingerprinting using deep learning," in *2018 IEEE Military Communications Conference*, Oct. 2018. ISSN 2155-7578 pp. 1–9. DOI: 10.1109/MILCOM.2018.8599826

[19] D. Li, D. Chen, B. Jin, L. Shi, J. Goh, and S. K. Ng, "MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks," in *28th International Conference on Artificial Neural Networks*, Sept. 2019, pp. 703–716.

[20] T. Sabhanayagam, "A comparative analysis to obtain unique device fingerprinting," in *Proceedings of International Conference on Deep Learning, Computing and Intelligence*. Springer, 2022, pp. 349–354.

[21] C. M. Ahmed and A. P. Mathur, "Hardware identification via sensor fingerprinting in a cyber physical system," in *2017 IEEE International Conference on Software Quality, Reliability and Security Companion*, July 2017, pp. 517–524.

[22] A. Al-Omary, A. Othman, H. M. AlSabbagh, and H. Al-Rizzo, "Survey of hardware-based security support for iot/cps systems," *KnE Engineering*, pp. 52–70, 2018.

[23] D. Marabissi, L. Mucchi, and A. Stomaci, "Iot nodes authentication and id spoofing detection based on joint use of physical layer security and machine learning," *Future Internet*, vol. 14, no. 2, p. 61, 2022.

[24] Z. Chen, J. Liu, Y. Shen, M. Simsek, B. Kantarci, H. T. Mouftah, and P. Djukic, "Machine learning-enabled iot security: Open issues and challenges under advanced persistent threats," *ACM Computing Surveys (CSUR)*, 2022.

[25] I. Sanchez-Rola, I. Santos, and D. Balzarotti, "Clock around the clock: Time-based device fingerprinting," in *2018 ACM SIGSAC Conference on Computer and Communications Security*, Jan. 2018, pp. 1502–1514.

[26] P. M. Sánchez Sánchez, J. M. Jorquera Valero, A. Huertas Celdrán, G. Bovet, M. Gil Pérez, and G. Martínez Pérez, "A survey on device behavior fingerprinting: Data sources, techniques, application scenarios, and datasets," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1048–1077, 2021. DOI: 10.1109/COMST.2021.3064259

[27] B. Varghese, N. Wang, D. Bermbach, C.-H. Hong, E. D. Lara, W. Shi, and C. Stewart, "A survey on edge performance benchmarking," *ACM Computing Surveys (CSUR)*, vol. 54, no. 3, pp. 1–33, 2021.

[28] B. Bezawada, M. Bachani, J. Peterson, H. Shirazi, I. Ray, and I. Ray, "Behavioral fingerprinting of iot devices," in *Proceedings of the 2018 workshop on attacks and solutions in hardware security*, 2018, pp. 41–50.

[29] S. Thouti, N. Venu, D. R. Rinku, A. Arora, and N. Rajeswaran, "Investigation on identify the multiple issues in iot devices using convolutional neural network," *Measurement: Sensors*, vol. 24, p. 100509, 2022.

[30] S. Zhang, Z. Wang, J. Yang, D. Bai, F. Li, Z. Li, J. Wu, and X. Liu, "Unsupervised iot fingerprinting method via variational auto-encoder and k-means," in *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021, pp. 1–6.

[31] J. Kocoń, I. Cichecki, O. Kaszyca, M. Kochanek, D. Szydło, J. Baran, J. Bielaniewicz, M. Gruza, A. Janz, K. Kanclerz *et al.*, "Chatgpt: Jack of all trades, master of none," *Information Fusion*, p. 101861, 2023.

[32] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[33] O. Ibitoye, R. Abou-Khamis, A. Matrawy, and M. O. Shafiq, "The threat of adversarial attacks on machine learning in network security–a survey," *arXiv preprint arXiv:1911.02621*, 2019.

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

[35] A. Namvar, C. Thapa, S. S. Kanhere, and S. Camtepe, "Evaluating the security of machine learning based iot device identification systems against adversarial examples,"

in *International Conference on Service-Oriented Computing*. Springer, 2021, pp. 800–810.

[36] T. Laor, N. Mehanna, A. Durey, V. Dyadyuk, P. Laperdrix, C. Maurice, Y. Oren, R. Rouvoy, W. Rudametkin, and Y. Yarom, "Drawnapart: A device identification technique based on remote gpu fingerprinting," *arXiv preprint arXiv:2201.09956*, 2022.

[37] K. Sadeghi, A. Banerjee, and S. K. Gupta, "A system-driven taxonomy of attacks and defenses in adversarial machine learning," *IEEE transactions on emerging topics in computational intelligence*, vol. 4, no. 4, pp. 450–467, 2020.

[38] I. Ali, S. Sabir, and Z. Ullah, "Internet of things security, device authentication and access control: a review," *arXiv preprint arXiv:1901.07309*, 2019.

[39] S. Rajendran, R. Calvo-Palomino, M. Fuchs, B. V. den Bergh, H. Cordobés, D. Giustiniano, S. Pollin, and V. Lenders, "Electrosense: Open and Big Spectrum Data," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 210–217, January 2018.

[40] I. Rosenberg, A. Shabtai, Y. Elovici, and L. Rokach, "Adversarial machine learning attacks and defense methods in the cyber security domain," *ACM Computing Surveys (CSUR)*, vol. 54, no. 5, pp. 1–36, 2021.

[41] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, and L. Sun, "Transformers in time series: A survey," *arXiv preprint arXiv:2202.07125*, 2022.

[42] P. M. S. Sánchez, J. M. J. Valero, A. H. Celdrán, G. Bovet, M. G. Pérez, and G. M. Pérez, "Lwhbench: A low-level hardware component benchmark and dataset for single board computers," *Internet of Things*, vol. 22, p. 100764, 2023.

Other publications/works

In addition to the main publications composing the PhD Thesis, several works have been published due to the research activities carried toward the PhD completion. Apart from the *six JCR articles* composing the present thesis, *sixteen other JCR journal articles*, *thirteen conference articles*, *four book chapters*, and *one 5G PPP technical report* have been co-authored. Besides, *three tutorial* sessions have been given at different conferences. All these publications can be framed in different collaborations and side-projects developed during the last years:

- Conference tutorial sessions directly based on the PhD thesis content:
  - (*Tutorial*) Alberto Huertas, **Pedro M. Sánchez Sánchez**, Muriel Franco, Bruno Rodrigues, Gérôme Bovet, Gregorio Martínez, Burkhard Stiller. (2021). Intelligent Behavioral Fingerprinting - From Theory to Practice. 17th IEEE International Conference on Network and Service Management (CNSM 2021).
  - (*Tutorial*) Alberto Huertas, **Pedro M. Sánchez Sánchez**, Muriel Franco, Gérôme Bovet, Gregorio Martínez, Burkhard Stiller. (2022). Theoretical and Practical Intelligent Behavioral Fingerprinting. IEEE/IFIP Network Operations and Management Symposium (NOMS 2022).
- Collaboration with the University of Zurich and armasuisse S&T in AI Trustworthiness, Federated Learning, and behavior fingerprinting for cyberattack detection:
  - (*Journal*) Rey, V., **Sánchez Sánchez, P. M.**, Huertas Celdrán, A., & Bovet, G. (2022). Federated learning for malware detection in iot devices. Computer Networks, 204, 108693.
  - (*Journal*) Huertas Celdrán, A., **Sánchez Sánchez, P. M.**, Castillo, M. A., Bovet, G., Martínez Pérez, G., & Stiller, B. (2022). Intelligent and behavioral-based detection of malware in IoT spectrum sensors. International Journal of Information Security, 1-21.
  - (*Journal*) **Sánchez Sánchez, P. M.**, Huertas Celdrán, A., Schenk, T., Iten, A. L. B., Bovet, G., Martínez Pérez, G., & Stiller, B. (2022). Studying the Robustness of Anti-Adversarial Federated Learning Models Detecting Cyberattacks in IoT Spectrum Sensors. IEEE Transactions on Dependable and Secure Computing. In press.

- (*Journal*) Huertas Celdrán, A., **Sánchez Sánchez, P. M**., Feng, C., Bovet, G., Martínez Pérez, G., & Stiller, B. (2023). Privacy-preserving and Syscall-based Intrusion Detection System for IoT Spectrum Sensors Affected by Data Falsification Attacks. IEEE Internet of Things Journal, 10 (10), 8408-8415.
- (*Journal*) **Sánchez Sánchez, P. M**., Huertas Celdrán, A., Buendía Rubio, J. R., Bovet, G., & Martínez Pérez, G. (2023). Robust Federated Learning for execution time-based device model identification under label-flipping attack. Cluster Computing, 1-12.
- (*Journal*) Huertas Celdrán, A., **Sánchez Sánchez, P. M**., Bovet, G., Martínez Pérez, G., & Stiller, B. (2023). CyberSpec: Behavioral Fingerprinting for Intelligent Attacks Detection on Crowdsensing Spectrum Sensors. IEEE Transactions on Dependable and Secure Computing. In press.
- (*Journal*) Huertas Celdrán, A., **Sánchez Sánchez, P. M.**, von der Assen, J., Shushack, D., Gómez, Á. L. P., Bovet, G., ... & Stiller, B. (2023). Behavioral Fingerprinting to Detect Ransomware in Resource-constrained Devices. Computers & Security, 103510.
- (*Journal*) **Sánchez Sánchez, P. M.**, Huertas Celdrán, A., Xie, N., Bovet, G., Martínez Pérez, G., & Stiller, B. (2024). FederatedTrust: A solution for trustworthy federated learning. Future Generation Computer Systems, 152, 83-98.
- (*Conference*) **Sánchez Sánchez, P. M**., Huertas Celdran, A., Bovet, G., Martínez Pérez, G., & Stiller, B. (2021). Secure Crowdsensing Platforms Through Device Behavior Fingerprinting. Jornadas Nacionales en Investigación en Ciberseguridad (JNIC).
- (*Conference*) Huertas Celdrán, A., **Sánchez Sánchez, P. M.**, Scheid, E. J., Besken, T., Bovet, G., Martínez Pérez, G., & Stiller, B. (2022, April). Policy-based and Behavioral Framework to Detect Ransomware Affecting Resource-constrained Sensors. In NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium (pp. 1-7). IEEE.
- (*Conference*) Huertas Celdrán, A., **Sánchez Sánchez, P. M**., Bovet, G., Martínez Pérez, G., & Stiller, B. (2022, May). Intelligent Fingerprinting to Detect Data Leakage Attacks on Spectrum Sensors. In ICC 2022-IEEE International Conference on Communications (pp. 4080-4085). IEEE.
- (*Conference*) Huertas Celdrán, A., Bauer, J., Demirci, M., Leupp, J., Franco, M. F., **Sánchez Sánchez, P. M**., ... & Stiller, B. (2022, October). RITUAL: a Platform Quantifying the Trustworthiness of Supervised Machine Learning. In 2022 18th International Conference on Network and Service Management (CNSM) (pp. 364-366). IEEE.
- (*Conference*) Huertas Celdrán, A., Kreischer, J., Demirci, M., Leupp, J., **Sanchez, P. M**., Franco, M. F., ... & Stiller, B. (2023). A Framework Quantifying Trustworthiness of Supervised Machine and Deep Learning Models. In SafeAI2023: The AAAI's Workshop on Artificial Intelligence Safety (pp. 2938-2948).
- (*Conference*) Huertas Celdrán, A., von der Assen, J., Moser, K., **Sánchez Sánchez, P. M.**, Bovet, G., Martínez Pérez, G., & Stiller, B. (2023, May). Early Detection of Cryptojacker Malicious Behaviors on IoT Crowdsensing Devices. In NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium (pp. 1-8). IEEE.

  - (*Conference*) **Sánchez Sánchez, P. M.**, Huertas Celdrán, A., Bovet, G., Martínez Pérez, G., & Stiller, B. (2023, June). A Trustworthy Federated Learning Framework for Individual Device Identification. In 2023 JNIC Cybersecurity Conference (JNIC) (pp. 1-8). IEEE.
  - (*Conference*) J. von der Assen, A. H. Celdrán, **P. M. S. Sánchez**, J. Cedeño, G. Bovet, G. M. Pérez, and B. Stiller, (2023) A Lightweight Moving Target Defense Framework for Multi-purpose Malware Affecting IoT Devices, in IEEE International Conference on Communications, ICC 2023.
  - (*Conference*) **Sánchez Sánchez, P. M.**, Huertas Celdrán, A., Beltrán, E. T. M., Wassink, R., Bovet, G., Martínez Pérez, G., & Stiller, B. (2023) Stealth Spectrum Sensing Data Falsification Attacks Affecting IoT Spectrum Monitors on the Battlefield, in 2023 IEEE Military Communications Conference (MILCOM 2023) (pp. 1-6). IEEE.
  - (*Conference*) J. von der Assen, A. H. Celdrán, J. Luechinger, **P. M. S. Sánchez**, G. Bovet, G. M. Pérez, and B. Stiller, (2023) RansomAI: AI-powered Ransomware for Stealthy Encryption, in IEEE Global Communications Conference (GLOBECOM 2023) (pp. 1-6). IEEE.
  - (*Chapter*) Huertas Celdrán, A., **Sánchez Sánchez, P. M**., Sisi, F., Bovet, G., Martínez Pérez, G., & Stiller, B. (2022). Creation of a Dataset Modeling the Behavior of Malware Affecting the Confidentiality of Data Managed by IoT Devices. In Robotics and AI for Cybersecurity and Critical Infrastructure in Smart Cities (pp. 193-225). Cham: Springer International Publishing.

- Collaboration in 5GZORRO (Zero-tOuch secuRity and tRust for ubiquitous cOmputing and connectivity in 5G networks) H2020 project and José María's PhD work:

  - (*Journal*) Jorquera Valero, J. M., **Sánchez Sánchez, P. M**., Lekidis, A., Hidalgo, J. F., Gil Pérez, M., Siddiqui, M. S., ... & Martínez Pérez, G. (2022). Design of a Security and Trust Framework for 5G Multi-domain Scenarios. J. Netw. Syst. Manag., 30(1), 7.
  - (*Journal*) Jorquera Valero, J. M., **Sánchez Sánchez, P. M**., Gil Pérez, M., Huertas Celdrán, A., & Martínez Pérez, G. (2022). Toward pre-standardization of reputation-based trust models beyond 5G. Computer Standards & Interfaces, 81, 103596.
  - (*Journal*) Jorquera Valero, J. M., **Sánchez Sánchez, P. M.**, Gil Pérez, M., Huertas Celdrán, A., & Martinez Perez, G. (2023). Cutting-Edge Assets for Trust in 5G and Beyond: Requirements, State of the Art, Trends, and Challenges. ACM Computing Surveys, 55(11), 1-36.
  - (*Journal*) Jorquera Valero, J. M., **Sánchez Sánchez, P. M.**, Gil Pérez, M., Huertas Celdrán, A., & Martínez Pérez, G. (2023). Trust-as-a-Service: A reputation-enabled trust framework for 5G network resource provisioning. Computer Communications, 211, 229-238.
  - (*Chapter*) Jorquera Valero, J. M., **Sánchez Sánchez, P. M**., Lekidis, A., Martins, P., Diogo, P., Gil Pérez, M., ... & Martínez Pérez, G. (2022). Trusted Execution Environment-enabled platform for 5G security and privacy enhancement. Security and Privacy Preserving for IoT and 5G Networks: Techniques, Challenges, and New Directions, 203-223.

  - (*Technical Report*) 5G PPP Technology Board. (2021). AI and ML–Enablers for beyond 5G Networks. `https://zenodo.org/record/4299895`

- Collaboration in EU-GUARDIAN (European framework and proofs-of-concept for the intelliGent aUtomAtion of cybeR Defence Incident mAnagemeNt) European Defence Fund (EDF) project:
  - (*Conference*) Cid, M. I. G., Gil Pérez, M., Jorquera Valero, J. M., Martínez, A. L., Vidal, J. M., Martínez Pérez, G., ..., **Sánchez Sánchez, P. M.**, & Monge, M. A. S. (2023, June). European framework and proofs-of-concept for the intelliGent aUtomAtion of cybeR Defence Incident mAnagemeNt. In 2023 JNIC Cybersecurity Conference (JNIC) (pp. 1-2). IEEE.

- Work in Decentralized Federated Learning and Enrique Martínez's PhD supervision:
  - (*Journal*) Beltrán, E. T. M., Pérez, M. Q., **Sánchez Sánchez, P. M.**, Bernal, S. L., Bovet, G., Gil Pérez, M., Martínez Pérez, G., & Huertas Celdrán, A. (2023). Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges. IEEE Communications Surveys & Tutorials, vol. 25, no. 4, pp. 2983-3013.
  - (*Journal*) Beltrán, E. T. M., Peráles Gómez, A. L., Feng, C., **Sánchez Sánchez, P. M.**, Bernal, S. L., Bovet, G., Gil Pérez, M., Martínez Pérez, G., & Huertas Celdrán, A. (2024) Fedstellar: A Platform for Decentralized Federated Learning. Expert Systems with Applications, 242, 122861.
  - (*Conference*) Beltrán, E. T. M., **Sánchez Sánchez, P. M.**, López, S., Bernal, G. B., Gil Pérez, M., Martínez Pérez, G., & Huertas Celdrán, A., (2023) Fedstellar: A platform for training models in a privacy-preserving and decentralized fashion, in International Joint Conference on Artificial Intelligence (IJCAI-23) Demo Track.
  - (*Tutorial*) Alberto Huertas Celdrán, Gregorio Martínez Pérez, Enrique Tomás Martínez Beltrán, **Pedro Miguel Sánchez Sánchez**, Gérôme Bovet, Burkhard Stiller. (2023). Decentralized Federated Learning: Enabling Collaborative AI With Enhanced Trust and Efficiency. 26th European Conference on Artificial Intelligence (ECAI 2023)

- Work in user behavior fingerprinting for authentication, an extension of my Master's thesis:
  - (*Journal*) **Sánchez Sánchez, P. M**., Jorquera Valero, J. M., Zago, M., Huertas Celdrán, A., Maimó, L. F., Bernal, E. L., ... & Martínez Pérez, G. (2020). BEHACOM-a dataset modelling users' behaviour in computers. Data in Brief, 31, 105767.
  - (*Journal*) **Sánchez Sánchez, P. M**., Fernández Maimó, L., Huertas Celdrán, A., & Martínez Pérez, G. (2021). AuthCODE: A privacy-preserving and multi-device continuous authentication architecture based on machine and deep learning. Computers & Security, 103, 102168.
  - (*Conference*) **Sánchez Sánchez, P. M**., Huertas Celdrán, A., Fernández Maimó, L., Martínez Pérez, G., & Wang, G. (2019). Securing smart offices through an intelligent and multi-device continuous authentication system. In Smart City and Informatization: 7th International Conference, iSCI 2019, Guangzhou, China, November 12–15, 2019, Proceedings 7 (pp. 73-85). Springer Singapore.

– (*Chapter*) **Sánchez Sánchez, P. M.**, Jorquera Valero, J. M., Huertas Celdrán, A., & Martínez Pérez, G. (2020). Intelligent User Profiling Based on Sensors and Location Data to Detect Intrusions on Mobile Devices. In Handbook of Research on Intrusion Detection Systems (pp. 1-25). IGI Global.

– (*Chapter*) Jorquera Valero, J. M., **Sánchez Sánchez, P. M.**, Huertas Celdrán, A., & Martínez Pérez, G. (2020). Machine Learning as an Enabler of Continuous and Adaptive Authentication in Multimedia Mobile Devices. In Handbook of Research on Multimedia Cyber Security (pp. 21-47). IGI Global.

Publications composing

the PhD Thesis

1

A Survey on Device Behavior Fingerprinting: Data Sources, Techniques, Application Scenarios, and Datasets

Title:	A Survey on Device Behavior Fingerprinting: Data Sources, Techniques, Application Scenarios, and Datasets
Authors:	Pedro Miguel Sánchez Sánchez, José María Jorquera Valero, Alberto Huertas Celdrán, Gérôme Bovet, Manuel Gil Pérez, Gregorio Martínez Pérez
Journal:	IEEE Communications Surveys & Tutorials
JIF:	33.84 D1 (2021)
Publisher:	IEEE
Volume:	23
Number:	2
Pages:	1048-1077
Year:	2021
Month:	Mar
DOI:	10.1109/COMST.2021.3064259
Status:	Published

Abstract

In the current network-based computing world, where the number of interconnected devices grows exponentially, their diversity, malfunctions, and cybersecurity threats are increasing at the same rate. To guarantee the correct functioning and performance of novel environments such as Smart Cities, Industry 4.0, or crowdsensing, it is crucial to identify the capabilities of their devices (e.g., sensors, actuators) and detect potential misbehavior that may arise due to cyberattacks, system faults, or misconfigurations. With this goal in mind, a promising research field emerged focusing on creating and managing fingerprints that model the behavior of both the device actions and its components. The article at hand studies the recent growth of the device behavior fingerprinting field in terms of application scenarios, behavioral sources, and processing and evaluation techniques. First, it performs a comprehensive review of the device types, behavioral data, and processing and evaluation techniques used by the most recent and representative research works dealing with two major scenarios: device identification and device misbehavior detection. After that, each work is deeply analyzed and compared, emphasizing its characteristics, advantages, and limitations. This article also provides researchers with a review of the most relevant characteristics of existing datasets as most of the novel processing techniques are based

on Machine Learning and Deep Learning. Finally, it studies the evolution of these two scenarios in recent years, providing lessons learned, current trends, and future research challenges to guide new solutions in the area.

Keywords

Behavioral data · cyberattack detection · device behavior datasets · device behavior fingerprinting · device identification · processing and evaluation techniques

2

A methodology to identify identical single-board computers based on hardware behavior fingerprinting

Title:	A methodology to identify identical single-board computers based on hardware behavior fingerprinting.
Authors:	Pedro Miguel Sánchez Sánchez, José María Jorquera Valero, Alberto Huertas Celdrán, Gérôme Bovet, Manuel Gil Pérez, Gregorio Martínez Pérez
Journal:	Journal of Network and Computer Applications
JIF:	8.7 D1 (2022)
Publisher:	Elsevier
Volume:	212
Number:	
Pages:	103579
Year:	2023
Month:	Mar
DOI:	10.1016/j.jnca.2022.103579
Status:	Published

Abstract

The connectivity and resource-constrained nature of single-board devices open the door to cybersecurity concerns affecting Internet of Things (IoT) scenarios. One of the most important issues is the presence of unauthorized IoT devices that want to impersonate legitimate ones by using identical hardware and software specifications. This situation can provoke sensitive information leakages, data poisoning, or privilege escalation in IoT scenarios. Combining behavioral fingerprinting and Machine/Deep Learning (ML/DL) techniques is a promising approach to identify these malicious spoofing devices by detecting minor performance differences generated by imperfections in manufacturing. However, existing solutions are not suitable for single-board devices since they do not consider their hardware and software limitations, underestimate critical aspects such as fingerprint stability or context changes, and do not explore the potential of ML/DL techniques. To improve it, this work first identifies the essential properties for single-board device identification: uniqueness, stability, diversity, scalability, efficiency, robustness, and security. Then, a novel methodology relies on behavioral fingerprinting to identify identical single-board devices and meet the previous properties. The methodology leverages the different built-in components of the system and ML/DL techniques, comparing the device internal

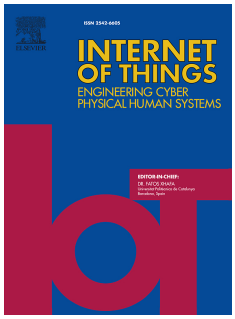
behavior with each other to detect variations that occurred in manufacturing processes. The methodology validation has been performed in a real environment composed of 15 identical Raspberry Pi 4 Model B and 10 Raspberry Pi 3 Model B+ devices, obtaining a 91.9% average TPR with an XGBoost model and achieving the identification for all devices by setting a 50% threshold in the evaluation process. Finally, a discussion compares the proposed solution with related work, highlighting the fingerprint properties not met, and provides important lessons learned and limitations.

Keywords

Device behavior fingerprinting · Device identification · Cyberattack detection · Behavioral data · Hardware fingerprinting

3

LwHBench: A low-level hardware component benchmark and dataset for Single Board Computers



Title:	LwHBench: A low-level hardware component benchmark and dataset for Single Board Computers
Authors:	Pedro Miguel Sánchez Sánchez, José María Jorquera Valero, Alberto Huertas Celdrán, Gérôme Bovet, Manuel Gil Pérez, Gregorio Martínez Pérez
Journal:	Internet of Things
JIF:	5.9 Q1 (2022)
Publisher:	Elsevier
Volume:	22
Number:	
Pages:	100764
Year:	2023
Month:	Jul
DOI:	10.1016/j.iot.2023.100764
Status:	Published

Abstract

In today's computing environment, where Artificial Intelligence (AI) and data processing are moving toward the Internet of Things (IoT) and Edge computing paradigms, benchmarking resource-constrained devices is a critical task to evaluate their suitability and performance. Between the employed devices, Single-Board Computers arise as multi-purpose and affordable systems. The literature has explored Single-Board Computers performance when running high-level benchmarks specialized in particular application scenarios, such as AI or medical applications. However, lower-level benchmarking applications and datasets are needed to enable new Edge-based AI solutions for network, system and service management based on device and component performance, such as individual device identification. Thus, this paper presents LwHBench, a low-level hardware benchmarking application for Single-Board Computers that measures the performance of CPU, GPU, Memory and Storage taking into account the component constraints in these types of devices. LwHBench has been implemented for Raspberry Pi devices and run for 100 days on a set of 45 devices to generate an extensive dataset that allows the usage of AI techniques in scenarios where performance data can help in the device management process. Besides, to demonstrate the inter-scenario capability of the dataset, a series of AI-enabled use cases about

device identification and context impact on performance are presented as exploration of the published data. Finally, the benchmark application has been adapted and applied to an agriculture-focused scenario where three RockPro64 devices are present.

Keywords

Hardware benchmarking · System performance · Dataset · IoT device · Identification

4

SpecForce: A Framework to Secure IoT Spectrum Sensors in the Internet of Battlefield Things

Title:	SpecForce: A Framework to Secure IoT Spectrum Sensors in the Internet of Battlefield Things
Authors:	Pedro Miguel Sánchez Sánchez, Alberto Huertas Celdrán, Gérôme Bovet, Gregorio Martínez Pérez, Burkhard Stiller
JIF:	11.2 D1 (2022)
Publisher:	IEEE
Volume:	61
Issue:	5
Pages:	174 - 180
Year:	2023
Month:	May
DOI:	10.1109/MCOM.001.2200349
Status:	Published

Abstract

The battlefield has evolved into a mobile and dynamic scenario where soldiers and heterogeneous military equipment exchange information in real-time and wirelessly. This fact brings to reality the Internet of Battlefield Things (IoBT). Wireless communications are key enablers for the IoBT, and their management is critical due to the spectrum scarcity and the increasing number of IoBT devices. In this sense, IoBT spectrum sensors are deployed on the battlefield to monitor the frequency spectrum, transmit over unoccupied bands, intercept enemy transmissions, or decode valuable information. However, IoBT spectrum sensors are vulnerable to heterogeneous cyber-attacks, and their accurate detection is an open challenge in the literature. Thus, this paper presents SpecForce, a security framework for IoBT spectrum sensors based on device behavioral fingerprinting and ML/DL techniques. SpecForce considers heterogeneous data sources to detect the most dangerous and recent cyber-attacks affecting IoBT spectrum sensors, such as impersonation, malware, and spectrum sensing data falsification attacks. To evaluate the SpecForce detection performance, it has been deployed on 25 real spectrum sensors, and results show almost perfect detection for the three cyber-attack families previously mentioned.

Keywords

IoT · Battlefield · Spectrum Monitoring · Fingerprinting · Cybersecurity · Identification

5

Adversarial attacks and defenses on ML- and hardware-based IoT device fingerprinting and identification

Title:	Adversarial attacks and defenses on ML- and hardware-based IoT device fingerprinting and identification
Authors:	Pedro Miguel Sánchez Sánchez, Alberto Huertas Celdrán, Gérôme Bovet, Gregorio Martínez Pérez
Journal:	Future Generation Computer Systems
JIF:	7.5 D1 (2022)
Publisher:	Elsevier
Volume:	152
Number:	
Pages:	30-42
Year:	2024
Month:	March
DOI:	10.1016/j.future.2023.10.011
Status:	Published

Abstract

In the last years, the number of IoT devices deployed has suffered an undoubted explosion, reaching the scale of billions. However, some new cybersecurity issues have appeared together with this development. Some of these issues are the deployment of unauthorized devices, malicious code modification, malware deployment, or vulnerability exploitation. This fact has motivated the requirement for new device identification mechanisms based on behavior monitoring. Besides, these solutions have recently leveraged Machine and Deep Learning (ML/DL) techniques due to the advances in this field and the increase in processing capabilities. In contrast, attackers do not stay stalled and have developed adversarial attacks focused on context modification and ML/DL evaluation evasion applied to IoT device identification solutions. However, literature has not yet analyzed in detail the impact of these attacks on individual identification solutions and their countermeasures. This work explores the performance of hardware behavior-based individual device identification, how it is affected by possible context- and ML/DL-focused attacks, and how its resilience can be improved using defense techniques. In this sense, it proposes an LSTM-CNN architecture based on hardware performance behavior for individual device identification. Then, the most usual ML/DL classification techniques have been compared with the proposed architecture using a hardware performance dataset collected from 45

Raspberry Pi devices running identical software. The LSTM-CNN improves previous solutions achieving a +0.96 average F1-Score and 0.8 minimum TPR for all devices. Afterward, context- and ML/DL-focused adversarial attacks were applied against the previous model to test its robustness. A temperature-based context attack was not able to disrupt the identification, but some ML/DL state-of-the-art evasion attacks were successful. Finally, adversarial training and model distillation defense techniques are selected to improve the model resilience to evasion attacks, improving its robustness from up to 0.88 attack success ratio to 0.17 in the worst attack case, without degrading its performance in an impactful manner.

Keywords

Adversarial attacks · Device Identification · Internet of Things (IoT) Security · Context Attack · Machine Learning

6

Single-board Device Individual Authentication based on Hardware Performance and Autoencoder Transformer Models

Title:	Single-board Device Individual Authentication based on Hardware Performance and Autoencoder Transformer Models
Authors:	Pedro Miguel Sánchez Sánchez, Alberto Huertas Celdrán, Gérôme Bovet, Gregorio Martínez Pérez
Journal:	Computers & Security
JIF:	5.6 Q2 (2022)
Publisher:	Elsevier
Volume:	137
Number:	
Pages:	103596
Year:	2024
Month:	February
DOI:	10.1016/j.cose.2023.103596
Status:	Published

Abstract

The proliferation of the Internet of Things (IoT) has led to the emergence of crowdsensing applications, where a multitude of interconnected devices collaboratively collect and analyze data. Ensuring the authenticity and integrity of the data collected by these devices is crucial for reliable decision-making and maintaining trust in the system. Traditional authentication methods are often vulnerable to attacks or can be easily duplicated, posing challenges to securing crowdsensing applications. Besides, current solutions leveraging device behavior are mostly focused on device identification, which is a simpler task than authentication. To address these issues, an individual IoT device authentication framework based on hardware behavior fingerprinting and Transformer autoencoders is proposed in this work. To support the design, a threat model details the security problems faced when performing hardware-based authentication in IoT. This solution leverages the inherent imperfections and variations in IoT device hardware to differentiate between devices with identical specifications. By monitoring and analyzing the behavior of key hardware components, such as the CPU, GPU, RAM, and Storage on devices, unique fingerprints for each device are created. The performance samples are considered as time series data and

used to train outlier detection transformer models, one per device and aiming to model its normal data distribution. Then, the framework is validated within a spectrum crowdsensing system leveraging Raspberry Pi devices. After a pool of experiments, the model from each device is able to individually authenticate it between the 45 devices employed for validation. An average True Positive Rate (TPR) of 0.74 ± 0.13 and an average maximum False Positive Rate (FPR) of 0.06 ± 0.09 demonstrate the effectiveness of this approach in enhancing authentication, security, and trust in crowdsensing applications.

Keywords

Device Behavior Fingerprinting · Device Authentication · Transformer · Behavioral Data · Hardware Fingerprinting · Autoencoder