Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.
Digital Object Identifier XX.XXXX/YYY.20XX.ZZZZZZZ

The Pains of Hardware Security: An Assessment Model of Real-World Hardware Security Attacks

SOFIA MARAGKOU[1], LUKAS RAPPEL[2], HENDRIK DETTMER[3], THILO SAUTER[4], AND AXEL JANTSCH[5]

[1]Institute of Computer Technology (ICT), TU Wien (e-mail: sofia.maragkou@tuwien.ac.at)
[2]TÜV AUSTRIA GMBH, Vienna Austria (e-mail: lukas.rappel@tuv.at)
[3]TÜV AUSTRIA GMBH, Vienna Austria (e-mail: hendrik.dettmer@tuv-austria.com)
[4]Institute of Computer Technology (ICT), TU Wien, and Dep. of Integrated Sensor Systems, Univ. of Continuing Education Krems (e-mail: thilo.sauter@tuwien.ac.at)
[5]Institute of Computer Technology (ICT), TU Wien (e-mail: axel.jantsch@tuwien.ac.at)

Corresponding author: Sofia Maragkou (e-mail: sofia.maragkou@tuwien.ac.at).

This work was supported by TÜV AUSTRIA's #SafeSecLab Research Lab for Safety and Security in Industry, a research collaboration between TU Wien and TÜV AUSTRIA. Additional support was provided in part by the TU Wien Bibliothek through its Open Access Funding Program.

ABSTRACT From military applications to everyday devices, hardware (HW) security is more relevant than ever before. The supply chain of integrated circuits is global and involves multiple actors, which facilitate the implementation of various attacks. Its complexity increases the attack surfaces, violating not only the privacy of the users or even national security but also endangering human life. We review some of the publicly known HW attacks that have occurred and propose an assessment scheme for the attacks and the defense on hardware. Using this scheme, we relate the costs of attacks and defense and provide a structured landscape of HW attacks. To illustrate the utility of our assessment scheme, we apply it to a number of real-world and synthetic research cases. We observe a gap between the research use cases and the real-world attacks and envision that the comprehensive assessment of the attacks will enable the development of more suitable countermeasures. Additionally, we revised the security policies for hardware devices, and we conclude that the complexity and obscurity of the supply chain are key parameters impacting hardware security, providing attack surfaces. Finally, we identify the demystification of the supply chain as the main strategy to mitigate this problem.

INDEX TERMS hardware attacks, hardware security, hardware Trojans, counterfeit chips, side-channel analysis, assessment scheme, supply chain analysis, hardware security policies, real-world hardware attacks

I. INTRODUCTION AND BACKGROUND

The security of electronic systems has been a concern ever since these systems were equipped with computing platforms, and even more so with the introduction of communication and networking interfaces. Traditionally, the focus is on software security, which has been challenged since the 1980s. Less attention is paid to hardware attacks, i.e., the malicious alteration of hardware functionality to enable attacks, possibly at a much later time. Obviously, this kind of attack requires significantly more effort than pure software changes. Consequently, hardware faced the first publicly known challenge only in 1996 after a timing attack was published that leaked key information [1]. Since then, only a few but a very diverse set of hardware-based attacks have become known. Their nature suggests a startlingly huge potential for security risks.

The motivation for attacks, in general, can differ. Very often, profit is seen as the key parameter for the application of security as well as the main motivation for attacks, especially in industrial applications. Based on the potential profit of the attack, the security investment is estimated to ensure that the investment for the attack will exceed its profit for the attacker. Although it is in many cases sufficient to consider only the cost of the defense and the profit of the attack, current market models have changed the attack surfaces and altered the profit model of hardware attacks in a more complex way. However, this does not necessarily apply to state and military applications, as well as critical infrastructures [2], [3]. While profit is a main driver for criminal actors, attacks against critical infrastructures are often motivated by maximizing damage, public uncertainty, or political sabotage [4].

An interesting example of how profit-seeking can incentivize security risks comes from the firmware industry. The low profit of firmware led developers to sell access to

user information [5]. Millions of Android phones have been infected with silent plugins by malicious firmware, which allows black-hat entities to gain access to the information of the devices. One of those plugins was renting out the device for up to five minutes at a time, giving access to key typing, geographical location, IP addresses, etc. The manufacturing of the devices was outsourced, exposing the attack surface for silent plugins, which could stay hidden. According to Dobberstein [5] similar attacks have happened in the past. Generally speaking, the availability of remote connections facilitates software and firmware updates. This is often a desired part of customer support, but can present also a significance security risk [6].

Contrary to software security, hardware security is still an under-illuminated aspect of security research. The complex design process of electronics and particularly integrated circuit (IC), the different entities involved, and the new potential profit areas lead to attack surfaces regardless of the end use case of the chips (military, governmental, critical infrastructure, industrial, or commercial applications). And, based on recent events, hardware security has become more important than ever. In an attack in September 2024, attack surfaces in the supply chain allowed the injection of explosive mechanisms in pagers and walkie-talkies, which ended up costing human lives [7], [8]. Even though the details of the attack have not been disclosed, it is obvious that the supply chain of the specific pagers and walkie-talkies had been compromised (an analysis of the attack is provided in section V).

The obvious problem with threat analysis of electronic systems is the sheer size and complexity of the possible attack surface which is difficult to grasp. Therefore, in this paper, we propose a security assessment scheme for hardware attacks based on the IC design flow and life cycle. Its purpose is to cover diverse aspects of attacks and make them comparable. The proposed assessment scheme includes six categories: Class, Resources, Difficulty/Security Level, Impact/Risk Acceptance, Identification, and Exploitation. In each category, we identify 4-5 levels of severity. This allows us to compare different research use cases and real-world scenarios in a structured way. Based on real-world attacks, we identify new security policies (secure handling of the IC, and transparency of the IC supply chain and processes) that will strengthen security on hardware (refer to section VIII). We compare real attacks with attacks developed for academic research and notice that the gap between these two is significant. Finally, in section VII we present our vision of how the field of hardware security will be involved and which actions should be considered.

Based on our current knowledge, this is the first assessment scheme for hardware attacks. Most related is the threat model for hardware attacks proposed by Halak [9], which includes hazards that the software-aimed threat models could not cover, e.g., supply chain sabotage, IC counterfeit, etc. It allows to map the threats of an application and identify corresponding countermeasures. Widening the scope, the focus of our assessment scheme is to compare and position attacks in terms of required competence and effort. Moreover, our scheme exhibits the divergence between the attacks used in research and those that occurred in the real world.

In section II, we present the current IC supply chain, the entities involved, and the accessible design assets per phase. In section III we list the types of hardware attacks and the motivation of the adversaries. Then, in section IV, we discuss the technical standards and regulatory requirements of the European Union. In section V, we present examples of real-world attacks. We introduce our assessment scheme in section VI. The costs of attack and defense are summarized and related to each other in section VII. We review the hardware security policies and present our perspective on the topic and possible future directions in the field of hardware security in section VIII. A summary and conclusion are provided in section IX.

II. ANATOMY OF HARDWARE SECURITY

Since hardware is a physical good, attacks on hardware often start with the infiltration and modification of the supply chain. To identify the parameters influencing hardware security, it is therefore important to have a detailed understanding of the supply chain. This includes the entities involved, their roles in the design process and the total life cycle, and the accessible assets that could create hardware attack surfaces. Based on this analysis, security policies and regulatory requirements or standards can be devised and enforced.

The design flow for the development of ICs is exceedingly complex and characterized by a tight interplay of different companies and stakeholders with varying business interests. Moreover, it keeps changing due to evolving technology and business models. The increasing complexity of designs and fabrication processes, together with the general cost sensitivity of the microelectronics field fosters further specialization. This in turn leads design houses to employ third parties for the core design and fabrication processes.

It should be noted that even though the supply chain and the entities involved do change over time, for a given project or design, the supply chain is static. Exchanging service providers *on the fly* during the process is hardly possibly because of the specific interfaces between the entities that do not allow for easy replacement. Even if, as in the case of FPGA designs, a partial dynamic reconfiguration is planned during the operation phase of the design (like for dedicated, task-specific hardware accelerators), the individual designs result from a static design flow.

A. IC SUPPLY CHAIN AND LIFE CYCLE

The IC life cycle, as illustrated in fig. 1, is divided into five levels: $\boxed{1}$ design, $\boxed{2.a}$ fabrication or $\boxed{2.b}$ FPGA configuration (depending on the technology), $\boxed{3}$ testing and integration, $\boxed{4}$ operation level, and $\boxed{5}$ disposal or recycle.

The first step is the definition of the design requirements and specifications, which are used as input for the first level of the IC life cycle: design $\boxed{1}$. At this level, the design is divided into functional blocks, some of which are realized

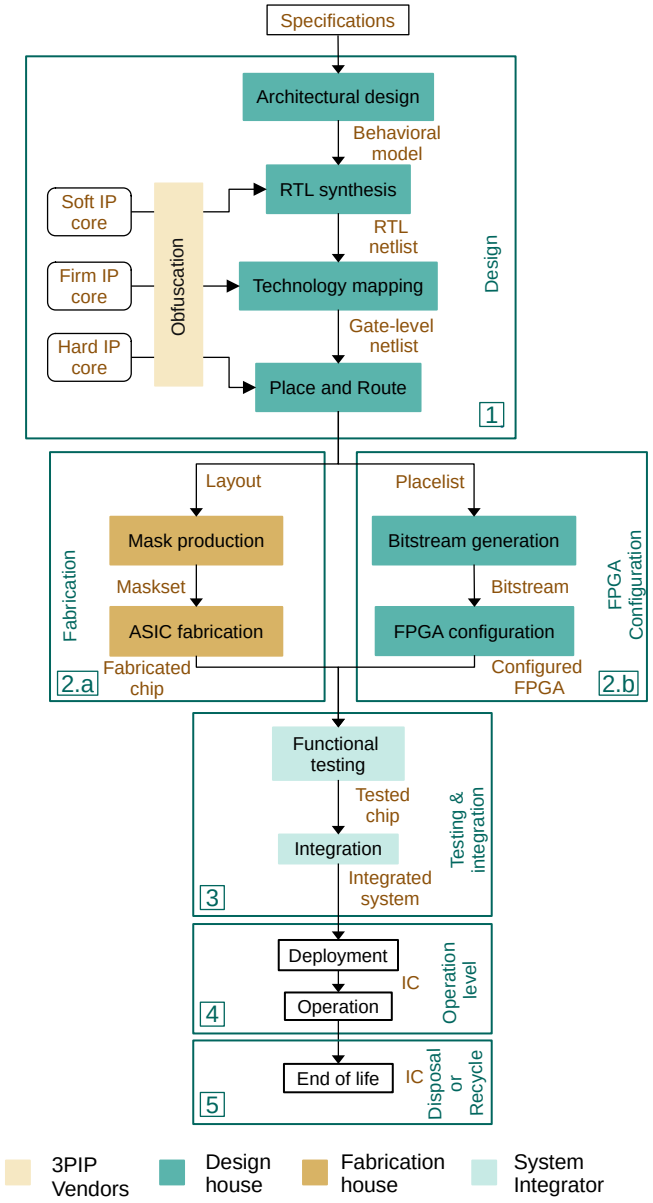

FIGURE 1. The life cycle of an IC. The different phases of the life cycle are enumerated inside the small boxes. Phase 2 differentiates per technology, and for this reason, we divide this phase into phase 2.a for ASIC and phase 2.b for FPGA. With brown letters, we name the assets that are exposed in each phase and exchanged between the phases. The different processes of the life cycle are colored based on the entity that executes them: 3PIP vendors, design houses, fabrication houses, and system integrators.

by well-defined intellectual property (IP) cores. Often IPs are purchased from third-party intellectual property (3PIP) vendors based on some pre-agreed specifications and standards. Design includes the processes of architectural design, Register-transfer Level (RTL) synthesis, technology mapping, and place and route. Through those processes IP cores from different vendors can be used as soft, firm, or hard IP cores depending on the abstraction level of the design in which they are integrated. The next step is the fabrication $\boxed{2.a}$ or the configuration $\boxed{2.b}$. Depending on the technology used, ASIC or FPGA, fabrication or configuration are the processes that give physical form to the hardware design. Typically, FPGA configuration is a process that is mostly performed in-house, while the fabrication is outsourced to IC manufacturers. Then, the ICs are tested and integrated $\boxed{3}$ into a complete system. This step is either done by the system house or again outsourced to a specialized company. Finally, the integrated system can be deployed and is ready for operation $\boxed{4}$. According to the guidelines for the scheduled replacement, or to possible failures from aging, the electronic components are sent for disposal or recycling $\boxed{5}$.

The steps of the supply chain can vary based on the different business models involved in the chain and their practices. In order to have a reference model, we identify the basic levels (or steps) in the supply chain. Their influence is important for the threat model.

B. ENTITIES INVOLVED IN THE DESIGN FLOW

To get a better understanding of the influence of each entity, we divide them into two categories: the direct- and the indirect-impact entities.

The direct-impact entities can only affect the design flow of a given design. Those entities are the IP vendors, the design house, the fabrication house, and the system integrator as featured in colored boxes of fig. 1. An attack deployed by these entities directly addresses a specific application and is, therefore, customized.

On the other hand, indirect-impact entities are the providers of the tools and libraries used during the IC design processes. The possible attacks deployed by these entities potentially have a larger impact on all applications developed using those tools or libraries. Specifically, back-doors in design tools can maliciously alter design behaviors without the hardware developers' knowledge [10] and [11].

C. ACCESSIBLE ASSETS

The term accessible assets describes the assets that are exposed during the IC design and manufacturing flow. They are considered attack surfaces.

The assets are identified in fig. 1 with brown letters. The available assets per phase of IC life cycle are listed in table 1. Essentially, they include design models and manufactured components. The design models include all models and files used during design and exchanged within and between the organizations contributing to the design process. They range from behavioral models to netlists of gates, configuration bitstreams, and mask sets. All these assets are potential targets for malicious manipulations or copying (theft).

A particularly interesting set of assets are design tools and libraries that are used not only in one but in many designs and products. If they are infected, the potential impact is much larger, extending to many very different and seemingly unrelated end-products.

III. TYPES OF HARDWARE ATTACKS AND MOTIVATION

The types of hardware attacks described in the following are motivated by the surveys in [12] and [1]. With these

TABLE 1. Critical assets per phase of IC life cycle.

Design time	FPGA conf. or Fabrication	Testing & Integration	Operation phase	Disposal phase	Recycle phase
IP rights	IP rights	IP rights	IP rights	IP rights	IP rights
Behavioral model	Layout or Placelist	Chip or FPGA	Chip or FPGA	Chip or FPGA	Chip or FPGA
RTL netlist	Maskset or Bitstream	-	-	-	-
Gate-level netlist	-	-	-	-	-

categories, we tried to cover a set of typical malicious actions, although we do not claim completeness. The categories are not strictly defined, and the distinction in research use cases and real-world attacks is not always clear. Very often, the adversaries also combine attacks to make their efforts more effective. For example, a hardware Trojan can be used to implement a side-channel attack. Nevertheless, the preparation of injecting the hardware Trojan is different from the planning of a side-channel attack.

Hardware Trojans are malicious sub-circuits designed to stay hidden during the verification process of the design under attack, and they deploy their malicious behavior in-field. The designs can be infected at any phase of the design flow.

Hardware Trojans consist mainly of two different parts: the trigger and the payload. The trigger is the mechanism that hides the malicious functionality during testing and the payload is the part of the design that deploys the actual attack. The possible attacks introduced by hardware Trojan-infected designs are information leakage [13], denial of service (DoS) [14], system degradation [15]–[17], change of functionality [18], [19], chip aging, computing exploitation, etc. The above attacks were identified from the body of literature from 2008 and later, further examples, in [11] and [17].

This list is not exhaustive since the possible attacks are based on the creativity of the attacker, and they are not restricted to the physical parameters of the attack surface.

Until 2016, the trigger was considered part of the design under attack (DuA), meaning that the trigger leaves a footprint in the original design (e.g., area-overhead, timing, functionality, not active logic under testing, etc.), which can be leveraged by detection methodologies. After 2016, a new attack vector was presented by [11] followed by [10], the malicious electronic design automation (EDA) . This attack vector provides a new aspect of hardware security analysis. Until this time, hardware Trojans were considered focused on single-application-attacks, but by tampering with the libraries of the EDA tools, multiple identical sub-circuits can be infected automatically in many different designs and products.

IP theft can include IP piracy and **reverse engineering**, the end goal of both is to get access to the information of the design. 3PIP vendors specialize in applications to create designs that can outperform competitors and gain dominance in the field. From the side of the design houses, the purchase and integration of the out-of-the-box cores facilitate the design process, helping the design houses to create IC according to the fast-changing demands of the market. Consecutively, IP rights are protected to ensure the profit of the companies developing them, while competitors try to steal successful designs in order to develop or sell them for their own profit.

IP piracy is not only gaining access to the design details without consent. This attack can also include the usage of an IP core without the knowledge of the vendor who developed it. An example can be that the core is used in more designs than what was agreed upon. The theft of the IP can be done in various ways and on different abstract levels during design.

By reverse engineering, the attacker is extracting the design details of the hardware core. It can be achieved in many different ways, depending on the equipment and the knowledge of the attacker. This threat exists throughout the whole IC supply chain. Nevertheless, it is a challenging task that requires high expertise from the side of the attacker and expensive equipment. This process can be applied at any phase of the IC life cycle.

Counterfeit chips are chips that have not been produced from the main supply chain and its specifications. Very often, they do not comply with any standards, or they do not even provide the promised functionality. Counterfeit chips can be produced and distributed in many different ways, such as grey market recycled chips, clone chips, and failed chips.

Overbuilding: Fabrication houses are usually third-party companies overseas. Thus, the complete control of the fabrication and the assembly process from the design house is not always possible. The yield information cannot be known. As Bhunia et al. [1] claim, fabrication houses leverage this in order to overbuild chips that will later sell for their own profit. Even though the direct impact is the loss of profit for the design house, indirectly, such attacks impact the consumers as well. The chips that have been fabricated additionally after the agreement with the design house are not tested for their correct functionality and reliability. Those chips have the name of the design house and can end up in different applications like military, safety-critical, etc. Thus, the reputation of the design house can be harmed.

Side channel attacks refer to attacks that extract information by monitoring the deviation of the physical parameters of the chips on the fly. Those channels can be voltage, temperature, and electromagnetic radiation.

On the other hand, **covert channels** are deliberate channels, made to leak information or to extract critical informa-

TABLE 2. Possible attacks on hardware for each phase of the supply chain.

IC life cycle	1	2.a 2.b	3	4	5
HW Trojan	x	x			
IP theft	x	x	x	x	x
Counterfeit					x
Overbuilding		x			
Side channel attacks			x	x	
Physical attacks			x	x	

tion from the existing channels of the chip.

Physical attacks include all the attacks that can be implemented by the user of the end device or with the final form of the IC. Those attacks include the physical injection of malicious devices or even the damage of the chip. For instance, one type of malicious device injection can be the injection of a USB flash drive or any other extra component with malicious functionality [20]. Additionally, chip damage can be any physical damage to the chip or the chip's housing.

Each phase of the supply chain gives access to specific assets. Those assets are the attack surfaces. As described in section II-C, those assets can be, e.g., design files, the final product, the chip for recycling, etc. Thus, different attacks can be implemented in different phases as seen in table 2.

The *motivation* for an attack may vary. Different kinds of motivation can be suitable for different attack vectors and thus require different threat models. As an example, an adversary who is motivated by profit ensures that the cost of the attack does not exceed the targeted profit of the attack. In contrast, a hacktivist may not feel constrained in the same way.

Possible attack motives can be the following.

(a) Profit: The interest of the adversary is only based on the financial profit. Such attacks are done by criminal teams and usually target sensitive information of the end user.
(b) Hacktivism: Teams with such motives target state organizations or private organizations in order to raise awareness, protest, or promote specific political or social ideologies.
(c) Sabotage and Disruption: Adversaries with such motives can be states or private companies. The result of such attacks can be the malfunction of specific infrastructure. Additionally, sabotage and disruption can also include the motive of political disorder. Terroristic organizations use security attacks to spread propaganda, enforce their ideology, or even take over the control of some technological applications. Such attacks usually target governmental or military applications and rarely individuals.
(d) Espionage: When the motivation of the adversaries is such, it can lead to the exploitation of sensitive information of the victim. This can happen on a governmental and military level or even on a private sector level. In this motivation category, we include attacks that target competitive advantage. In that case, the adversary steals the assets of a company or a country in order to reach the level of expertise of a competitor.
(e) Thrill-seeking and Notoriety: Such a motive only targets the rumor of the criminal. As security attacks can be challenging, executing some attacks can increase the reputation of the group or the individual who achieved them.

IV. REGULATORY REQUIREMENTS AND TECHNICAL STANDARDS

As the IC life cycle changes, the need to protect the exposed assets and restrict the payload of hardware attacks has increased. Attempts to protect industry, military, and critical infrastructures have been initiated by various countries and international organizations. The European Union and other bodies recently published the following regulatory requirements and technical standards. This list is not meant to be exhaustive but illustrates the types of relevant regulations.

A. NIS2

The NIS-2 Directive (also known as *The Network and Information Security (NIS) Directive*) [21] is an EU-wide legislation that aims to strengthen overall cybersecurity. NIS2 addresses the security of supply chains and supplier relationships by requiring individual companies to address cybersecurity risks in the supply chains and supplier relationships. At the European level, the Directive strengthens supply chain cybersecurity for key information and communication technologies. Member States, in cooperation with the European Commission and the European Cybersecurity Agency (ENISA), may carry out Union-level coordinated security risk assessments of critical supply chains, building on the successful approach taken in the context of the Commission Recommendation on Cybersecurity of 5G networks. This directive will have an indirect influence on products used in critical environments and their respective hardware security. The main focus of this regulation is the protection of EU-based companies and therefore, the supply chain protection focuses mostly on the resilience of the chain (removal of single points of failure in the supply chain) and not heavily on the detection of HW-Trojans or other cybersecurity threats.

B. CYBER RESILIENCE ACT (CRA)

The CRA [22] covers a wide range of products, in a nutshell, all digital products (hardware and software). For this paper, critical products with a Common Criteria or a European Union Common Criteria certification are relevant.

The Common Criteria (CC), or European Union Common Criteria (EUCC), is a globally recognized standard/certification (ISO/IEC 15408) that helps in choosing maximum security and assurance levels of hardware products. EUCC stands for European Union Common Criteria. It is an extension of the global CC standard, specifically tailored for use within the European Union. EUCC aligns with the international CC framework but includes additional

requirements and guidelines relevant to EU member states. Both standards use a standardized and well-defined attacker definition and different adversary levels against which the product, e.g., hardware, needs to protect itself. In these standards, the life cycle of the products is relevant and part of the evaluation, as well as an audit of the production facilities by the independent evaluation lab. Nevertheless, there are no spot checks or checks of finished products after the evaluation and certification are finished. The defined life cycle processes should protect all security policies, but for highly sophisticated hardware attacks the vendor must also have control over the whole supply chain. These requirements only allow a small number of vendors to achieve such a CC or EUCC certification on a high level.

C. EUROPEAN CHIPS ACT

The European Chips Act [22] aims to bolster Europe's competitiveness and resilience in semiconductor technologies and applications. It is a crucial step toward the EU's technological sovereignty. The CRA does not explicitly address cybersecurity. However, given the critical role of chips in various sectors, cybersecurity considerations are essential. Organizations involved in chip manufacturing and supply chains should prioritize robust security practices to protect against cyber threats.

D. IEC 62443

The IEC 62443 is an international series of standards focused on industrial communication networks and system security [23]. It addresses cybersecurity for operational technology (OT) in automation and control systems. The IEC 62443 standard includes Security Levels (SLs) that range from SL0 (no security) to SL4 (resistant against nation-state attacks). Security Levels 3 and 4 require hardware-based security due to the implementation of certain protections. Part 4-1 of the standard describes in the requirement "SM-9: Security requirements for externally provided components" that suppliers also implement security measures at the same level as the end product. However, these measures mostly cover a bill-of-material (BOM or SBOM) and an implemented vulnerability process. The standard also references ISO/IEC 27036-3 "Cybersecurity - Supplier relationships - Part 3: Guidelines for hardware, software, and services supply chain security". These standards go into more detail about which processes and regulations should be implemented and checked. Nevertheless, the described highly sophisticated attacks, described in this paper, which also include the producing party as an adversary, cannot be identified by these requirements and standards.

E. RADIO EQUIPMENT DIRECTIVE (RED)

The Radio Equipment Directive (RED) [24], which applies to wireless devices placed on the EU market, includes since 2022 cybersecurity requirements. The standard EN 18031 is a harmonized standard developed by the European Committee for Standardization (CEN) and the European Committee for Electrotechnical Standardization (CENELEC). This norm gives information about the processes and attack vectors for hardware devices. The requirements mainly focus on the implementation of standard cybersecurity mechanisms like multi-factor authentication. The standard wants to create the minimum level of security and not cover any medium or high-complexity attack vectors.

F. NIST HARDWARE SECURITY PROGRAM

The NIST Hardware Security Program focuses on identifying existing and emerging cybersecurity threats related to semiconductors. This Program is planning on performing the following activities grouped by topic area: Hardware Development Lifecycle, Metrology, Hardware/Silicon Testing, Vulnerability Management, and Standards. There are derived standards like NIST IR 8320 or NIST SP 1800-34, which specialize in certain hardware security topics.

V. EXAMPLES OF REAL-WORLD HARDWARE ATTACKS

Publicly known hardware attacks usually do not reveal all the technical details, making the gap between research and the real world hard to bridge. Nonetheless, to gain insight and understanding it is essential to carefully study real-world attacks. In the following, we present the best-known hardware attacks based on our current knowledge.

A. KILL SWITCH ATTACK

In 2008, a kill switch attack was reported in the public press [25]. In September 2007 a state-of-the-art jet radar failed to detect bombing jets with a flying duration of 4 hours. The result was that a nuclear installation in Syria was completely destroyed. Even though the details of the attack remain unknown, the known information about the attack suggests malicious behavior. As Adee suggests [25], the two possible scenarios are a backdoor or an incorporated kill switch in the off-the-shelf microprocessor, claiming that the scenario of the kill switch fits the provided details better. Such attacks enable a remote radar block without jeopardizing the rest of the functionality, giving the illusion of correct functionality. In this concrete case, malicious logic seems to be added at the hardware level in the form of a kill switch and additional logic. Similar microprocessors with kill switches were built at that time in Europe [25], targeting military applications.

B. SUPERMICRO - BIG HACK

The Big Hack (also called Supermicro) [20], [26] was another hardware-level attack that in 2018 targeted servers using Supermicro motherboards, by adding a microchip that was not specified in the original design. Those microchips seem to have been injected by the manufacturer in China and they were used as backdoors to every network used by the servers. The microchips included signal conditioning couplers, incorporated memory, network connection capability, and enough processing power to implement such an attack. In order to be used as backdoors, the chips had the capability

of altering the operating system's core functionality. The alteration was accepting the changes made by it and letting the microchip contact computers controlled by the adversary.

C. PAGER ATTACK

More recently, multiple simultaneous explosions got the world's attention. On September 17, 2024, multiple pagers exploded in Lebanon simultaneously, followed by the explosion of walkie-talkies the next day. The organization that used the pagers had turned to this low-tech solution to avoid the localization of its members being leaked to adversaries, as can happen with smartphones. Even though the two different explosion scenarios are considered the same attack, two different supply chains have been compromised for the pagers and the walkie-talkies. The targeted organization started using pagers in 2022 [27], and their orders increased in the months after February 2024. The supply chain remains unclear, but according to the most commonly referred scenario [8], the model was a Rugged Pager AR924, which originated from a Taiwanese company, Gold Apollo, produced by a consulting company in Europe that purchased the rights in 2022.

The explosion followed a notification for a message sent by a known sender [28]. Videos showed that only a few seconds after the messages were received, the pagers were detonated.

Pagers are part of a notification system where the paging controller receives the custom message and broadcasts it. The processor of the receiver decodes the message and, if the message is intended for this pager (the so-called capcode is known), it stores the message in the memory. In addition to the main components of the circuit, the pagers include some peripherals for interaction with the users. The pagers used in this attack have 4 buttons (up, down, enter, ok/power), possibly 4 LEDs, a buzzer, a display, and a USB-C port for charging and configuration. Conventional pagers run bare-metal applications, meaning that they include hardware and firmware, but no software.

As the details of the attacks are not known, we speculate how the attack could have been implemented. Starting from the known facts, we have the following information:

- The battery was different from the models produced by Gold Apollo.
- The pagers received a message from a known source.
- The pagers exploded a few seconds after the message was received.

As far as we know, there is only one description of a possible attack scenario, described in [29]. According to this scenario, the explosive charge was triggered by a detonator added in the protection circuit module (PCM) unit, a component used to prevent risks related to the charging of the batteries. The Gold Apollo pagers use AA-size batteries, but the models purchased in Lebanon were housing high-capacity Li-ion batteries, making the housing of the batteries larger. In addition, the used Li-ion batteries were manufactured in China, but the manufacturer is not known, which adds a new potentially malicious supply chain to the case. The exact route of the trigger from the message received by the RF antenna to the detonator is not known.

According to another source [30], experts claim that the amount of explosive material needed was very low and thus, it could have been hidden anywhere in the housing of the device. Additionally, they note that one of the possibilities is that the explosives were hidden in the battery housing together with an electronic detonating device. As there is evidence that the housing of the devices was different from the original device specifications, this seems to be a plausible scenario. Other speculations include the overheating of the batteries, which seems implausible as it would be too challenging to trigger so many concurrent explosions.

The detonator itself was likely electric, and it seems plausible that the trigger was mapped to the memory I/O of the pager. This way it could be activated by a specific message received over the air.

We know that the source of the message was known. This could imply the configuration of a dedicated group capcode (the identification code that checks if the message was addressed to this pager) just for the attack. Additionally, the paging controller also has to be manipulated in order to broadcast the message, or a malicious paging controller has to be added in a close location.

Thus, the trigger information of the attack could have been in the code or in the text. However, the pagers were functioning for at least a few months before the attack. In order to not detonate the pagers at random times, the most plausible scenario is that the capcode used was specifically dedicated to the attack. Storing a new capcode in the pager is a trivial process done via the USB-C port.

D. HARDWARE BACKDOORS

The injection of malicious functionality can take the form of a hardware Trojan or a backdoor. There are many reported cases of injected backdoors in hardware designs. The difference between them and a hardware Trojan attack is that the backdoor requires the action of the adversary even after the injection of the sub-circuit.

In [31], [32], the authors refer to a backdoor injected in cryptographic equipment from the company Crypto AG, delivering solutions to 120 countries for secure communication for governmental and diplomatic topics. In 2020 it was revealed that the company was controlled by the CIA, under the operations named "Thesaurus" and "Rubicon" from 1945 until 2018.

Even though further details of the attack have not been revealed in the sources, the encryption machines mentioned are the C-52 and CX-52, and they are both rotor-based cipher machines. As the functionality of a rotor-cipher machine is limited, we assume that the backdoor injected is a mechanism easing the extraction of the message or the key.

The long duration of the operations reveals that the compromised supply chain is a long-known problem.

S. Skorobogatov and C. Woods, in 2012, exposed a backdoor on a highly secure Actel/ Microsemi ProASIC3 FPGA, which was used in industrial and military applications [33] and even in aircrafts such as the Boeing 787 [34]. During security scanning, Skorobogatov and Woods found unknown JTAG commands which gave access to the configuration data. As the same backdoors exist in many of the FPGAs of the company, it is assumed that the backdoors were intentionally included in the device.

According to [35], the National Security Agency (NSA) was revealed in photos to implant backdoors (or hardware Trojans [2]) on Cisco routers by intervening in the supply chain. Details of the end users or the functionality implanted have not been published.

E. COUNTERFEIT CHIPS

Attacks on the hardware level also include counterfeit chips. According to [36], the US military bought 59,000 counterfeit microchips from China. The details of the attack, as well as the complete functionality of the chips, have not been officially revealed. Nevertheless, the incident raised awareness of possible malicious hardware alterations resulting from the outsourcing of manufacturing.

In 2010, Dell informed customers about spy malware injected into the servers PowerEdge R310, R410, R510, and T410, as mentioned in [37], [38], without specifying further details. Specifically, it is mentioned that the servers purchased directly from the Dell factory were not affected. As a result, the rest of the servers, and the stock in the supply chain had to be replaced, leading to major profit loss.

Military agencies and companies of critical infrastructures in the USA discovered 3500 counterfeit Cisco network components, as mentioned in [39]. The motive and the result of the attacks remain unknown. The known information suggests that the goal of the attack was just IP theft and no backdoor has been found. F.B.I. agents shared their concerns about potential results of counterfeit chips as remote jamming attacks or remote system control.

In more recent events, counterfeit Ryzen 7 9800X3D CPUs appeared in the market at the beginning of 2025 [40]. The fake CPUs cannot even boot and seem to be limited in the Chinese market so far. The differences from the original CPU are the information printed on the chip and the color of the board. As the fake CPUs cannot even boot, it seems that they are just fakes and not maliciously altered.

VI. ASSESSMENT SCHEME

To position a variety of severity and complexity levels depending on the perspective (attack or defense), an assessment scheme is required. The intention of this scheme is to provide a comprehensive description and to allow for an evaluation and comparison. To that end, we identify several categories that describe various aspects of attacks or defenses. These categories are enumerated in table 3. Each of these categorical variables can take on different severity levels. Most categories comprise four severity levels, only the categories related to human safety (impact and risk acceptance) have five levels. Arguably, the assignment of severity levels is to some extent subjective and depends on the concrete investigated case. Nevertheless, the assessment scheme is sufficient to shed light on attacks and defense mechanism from different angles.

In the last row of table 3, we translate the levels of the categories into numerical values which will later in section VII be used for display purposes. In the following, we discuss the individual assessment categories and the associated severity levels.

TABLE 3. Assessment categories and related severity levels for both attacks (A) and defense methods (D).

Categories	Validity	Severity Levels				
Class	A+D	Layman	Proficient	Expert	Multiple Experts	
Resources	A+D	Low	Moderate	Substantial	High	
Difficulty	A	Common Tool	Unusual Tool	Special Tool	Laboratory	
Security	D	Level 1	Level 2	Level 3	Level 4	
Impact	A	Low	Moderate	Substantial	High	Safety
Risk acceptance	D	High	Substantial	Moderate	Low	Zero
Identification	A+D	Low	Moderate	Substantial	High	
Exploitation	A+D	Low	Moderate	Substantial	High	
Numerical Value		1	2	3	4	5

A. CLASS

Motivated by the Common Criteria [41], the classes of attack and defense, based on their expertise, are four: layman, proficient, expert, and multiple experts.

Layman: The adversary/defender has no expertise on the topic.

Proficient: The adversary/defender is familiar with the specific application and the security perspective of this application.

Expert: The adversary/defender is a security expert with training on the topic.

Multiple experts: This class covers organized groups of experts specifically trained in the hardware security field.

A similar deviation of adversary classes has been provided by Hallak [9], according to which adversaries can be divided into small groups of hackers, academic research groups, organized criminal groups, and state-funded organizations.

B. ATTACK AND DEFENSE INVESTMENT AND RESOURCES

The topic of security is tightly coupled with the terms of profit and investment. Since both defenders and adversaries have very different budgets and capabilities, we formulate

the resources in terms of time, human effort, and equipment. To assess the resources of attack or defense, we define four different ranges named **low, moderate, substantial**, and **high** (table 3).

For all those parameters, the low and high scales are not expressed as absolute values but as a function of the motivation of the adversary. For instance, if the motivation of an attack is profit, the resources are restricted based on the possible profit of the attack. In contrast, in the case of hacktivism, the resources are not limited by a budget but can be limited by expertise.

C. ATTACK DIFFICULTY AND SECURITY LEVEL OF DEFENSE

The difficulty of the attack is coupled with the security level of the defense. An application that requires high security needs to cover more attack vectors, meaning that a potential attack has a higher difficulty level. According to Halak [9], the difficulty of the attack can be divided into five levels, based on the tools used. Those are **common tools, unusual tools, special tools, in a laboratory**, and **not feasible**. We do not consider the level not feasible as part of the assessment scheme as the purpose of the scheme is to position feasible attacks based on the technology existing.

The security level is a parameter defined to indicate the defense effort and depends exclusively on the application. That means that the higher the security level, the more difficult the attack.

D. ATTACK IMPACT AND DEFENSE RISK ACCEPTANCE

Two more coupled categories from the attack and defense perspectives are the impact of the attack and the risk acceptance of defense. The impact of the attack can be low, moderate, substantial, high, and safety-critical. Motivated by [42], we define the impact of the attack in a similar way as follows:

Low: The loss of the predefined security policies is expected to have a low to unnoticeable effect on the end users of the devices or the entities affected by the attack.

Moderate: The loss of the predefined security policies is expected to have serious effects on the end users of the devices or the entities offended by the attack.

Substantial: The loss of the predefined security policies is expected to have major or critical effects on the end users of the devices or the entities offended by the attack.

High: The loss of the predefined security policies is expected to have catastrophic effects on the end users or the entities offended by the attack.

Safety-critical: The loss of the predefined security policies is expected to have catastrophic effects on large groups of the population and their safety or the entities offended by the attack. As human safety is always prioritized in our assessment methodology, the severity level of safety-critical impact is noticeably higher than in other categories.

Risk acceptance is the strategy that each defender chooses based on the assets available in each use case or the likelihood of an attack and the loss tolerance. This strategy designates to what degree the defender is willing to accept the risk or mitigate it, and it is a function of the impact of a potential attack. Depending on the use case and the corresponding impact, we define different levels of risk acceptance.

Zero: Risk acceptance should be zero for applications where the impact of a potential attack is very high and safety-critical. Such applications can be critical infrastructures, military, and governmental applications where any risk is unacceptable under any circumstances. Nevertheless, in reality, zero risk is impossible, and a low-bound as low as reasonably practicable (ALARP) is considered.

Low: The risk acceptance should be low for applications where the impact of a potential attack can be high. The effort for mitigation should be immediate and significant. Such applications are not related to human safety but can still have catastrophic effects, e.g. critical infrastructures and governmental applications.

Moderate: The risk acceptance can be moderate for applications where the impact of a potential attack can be substantial. The effort for mitigation should be comprehensive. A potential attack has a high likelihood of taking place, and the impact would be noticeable for the end users.

Substantial: The risk acceptance can be substantial for applications where the impact of a potential attack can be moderate. The effort for mitigation should be limited to the basic standards. A potential attack has a low likelihood of taking place, and the impact would be limited to minor violations.

High: The risk acceptance can be high for applications where the impact of a potential attack can be low. No special mitigation strategy is required. A potential attack has a low likelihood of taking place, and the impact would be minor.

E. ATTACK AND DEFENSE IDENTIFICATION AND EXPLOITATION

Identification expresses the effort required to identify the attack vectors or to mitigate the attack. For every new attack, the adversary has to identify the assets required, technical details, and tools used, as well as the social parameters such as the profit, the targeted industry, etc. The initial effort for the attack planning can be much more intense than the attack repetition. Thus, this parameter should be considered separately.

It is considered a low identification effort when the attack surface is exposed and there is no additional overhead for the attack. The identification effort is high when the attack surfaces and the tools have to be identified by the adversary engineers. Very often, this includes attacks with high complexity that include reverse engineering or attacks targeting EDA tools.

Attack exploitation describes the effort of an existing attack to be repeated. Even though the initial research for the attack vectors can be challenging, the actual exploitation of the attack can also present its distinct challenges. The result of an attack can grow exponentially if it has a high exploita-

tion potential. Conversely, an attack that is challenging to repeat (exploit) will not yield the same total result. Thus, the low exploitation effort has an effect on the total result of the attack as well.

Likewise, from the perspective of defense, identification expresses the effort to identify and apply a countermeasure against an attack, and exploitation expresses the effort to apply that to different applications using the same technology.

The severity levels of the identification and exploitation of the attack and defense can vary from **low, moderate, substantial**, and **high**.

F. PARADIGMATIC EXAMPLES

To further explain the use of the model, we construct three fictional attack scenarios: Easy, Moderate, and Heavy, which can serve as a base for comparison. All assessed categories per scenario are summarized in Table 5 and illustrated in figure 2 top row.

1) Scenario Easy: IP Theft

The design house purchases an IP core from a vendor to incorporate it in its IC designs. The design house sells the IP core for its own profit to other companies or uses the core in multiple designs against the initial agreement with the vendor. Either of those scenarios falls into the category of IP piracy.

Attack: This attack takes place on the level 1 of the IC life cycle (at design time) at any abstraction level of the design and is considered easy in terms of the tools used. Specifically, there is no need for sophisticated tools or additional tools for the attack (Resources=1, Difficulty=1). The impact is low as there are no effects of the attack on the end users (Impact=1). The only entity harmed is the vendor, that has a profit loss. The attack surfaces are exposed as there is no way to quantify the use of the purchased cores, and manipulation of the design or the supply chain is not required. Thus, the effort required to identify the attack surface by the adversary (the design house) is low (Identification=1). The exploitation rate of the attack is high as the repetition of the attack requires minimum overhead (Exploitation=4). The class of the adversary can be described as a layman (Class=1). The adversary is located in the design house and can be a single individual with no special knowledge of the design or the design process.

Defense: To assess the resources for the mitigation of such an attack, we first define the security level of the assets and the risk acceptance. The security level of the mitigation strategy should be at least at level 2, as an overuse of the IP cores can be also intentional (Security level=2). The risk acceptance can be substantial as the impact of such an attack is considered low (Risk Acceptance=2). The minimum of the resources seems adequate to fulfill the requirements of such an attack, and the investment of more resources can be considered redundant (Resources=1). A sufficient countermeasure for this attack could be the IP watermarking [43] or the use of a license server, which means it can be handled by someone without special knowledge of the design (Class=1). As for the defense, identification is considered low, and the exploitation rate is low, too (Identification=1, Exploitation=1).

2) Scenario Moderate: Device Aging

We consider the physical tampering of the control system in a manufacturing plant. By physical access to the end device, the adversary manages to damage the end device, accelerating the aging process of the chips. This can be done by heating up the chip once or repeatedly. Since the maintenance of the industrial systems is done periodically based on the specifications, faster aging can create additional downtime, influencing the production of the factory and the profit. Additionally, it can reduce the quality of the production.

Attack: The attack difficulty is in the range of unusual tools (Difficulty=2). Its impact can be considered substantial, as there is no direct effect (functionality is unaffected), but in the long term, this attack can create a loss in profit (Impact=3). The identification of the attack is moderate as the attack does not include any sophisticated process (Identification=2), and the exploitation factor is moderate as well, as the same effort is required to repeat this attack (Exploitation=2). The adversary belongs in the class of the proficient, as knowledge is necessary to achieve the specific effect of aging on the chips (Class=2). Regarding the resources, the expense of time and equipment are moderate, but the human effort is substantial since physical presence in the targeted environment is required (Resources=3).

Defense: From the side of the defender, the security level of the control system in an industrial environment is Level 2 (Security=2). The defense should cover simple, intentional attacks with few resources. The risk acceptance is substantial (Risk Acceptance=3) as the impact of a possible attack is moderate and the effort for defense is limited to basic standards. The resources worth spending for defense for this system are moderate time and equipment and substantial human effort (Resources=2), and the required expertise is moderate (Class=2). As for the attacker, identification, and exploitation are considered moderate (Identification=2, Exploitation=2).

3) Scenario Heavy: HW Trojan

For the worst-case scenario, we consider a sophisticated attack with high complexity. This attack is the injection of hardware Trojans by the EDA tools. The reason we consider this attack being of high complexity is because both the identification and exploitation require high expertise and time.

The hardware Trojans are injected in design time and automatically by the EDA tools. This attack requires high expertise by EDA engineers as the changes of the malicious functionality should be sufficient to enable a complete attack but elegant enough so it is not detected by the area overhead. Such automatically malicious design injection can be triggered, e.g., by specific patterns in the original design, as in [11].

The payload is DoS, and the applications that are targeted are critical infrastructures. During the design time, the hardware Trojan stays inactive, passing all the test phases of the design flow. In the operation phase, the Trojan is triggered by a predefined event and it causes a DoS. In the case of critical infrastructure, the DoS can lead to safety hazards as the chip is no longer working as specified.

Attack: The attack difficulty is considered laboratory since good knowledge of hardware design, and the EDA tooling is required (Resources=4, Difficulty=4). The attack causes hazards in safety, which makes the impact safety-critical (Impact=5). Finally, the identification and the exploitation levels are both characterized as high (Identification=4, Exploitation=4). The adversary class is in the range of multiple experts as the complexity of this attack is considered high. Consequently, the resources of the attack are in the range of high (Class=4).

Defense: On the other hand, to defend critical infrastructures against this kind of attack is a very sophisticated and complex process. As a countermeasure for such an attack, sophisticated, exhausted detection methods should be incorporated during secure verification (Class=4, Resources=4). The security level is high as protection against intentional malicious behavior is required (Security =4). The risk acceptance in safety-relevant applications is zero, due to no risk can be acceptable (Risk Acceptance=5). As for the attacker, identification, and exploitation are considered high (Identification=4, Exploitation=4).

VII. PRACTICAL APPLICATION OF THE ASSESSMENT SCHEME

Based on the categories and the severity levels defined in section VI, we assess the "cost of hardware security" from the perspective of the attack and the defense, respectively. We consider the paradigmatic examples, the real-world attacks, and research use case attacks. The numerical values are summarized in the tables 4 and 5.

The severity level analysis for the **paradigmatic examples** has been given in section VI-F and is summarized in Table 5. The severity level analysis of the **real-world attack** follows the structure of section V and is presented below.

The **Kill Switch** attack [25] seems to be an attack designed ahead of time and includes the compromise of the supply chain of the processors. Thus, the difficulty of the attack is at the laboratory level. The impact of the attack was the destruction of a nuclear installation, which is considered safety-critical. The identification of the attack vector and the repetition of the attack require a lot of customization and they are not trivial. Many steps of the design process need to be compromised and therefore, the adversary can only be on the level of multiple experts, and the resources required are comparatively high. The pain of defending against such an attack requires all the possible resources in time, human effort, and equipment. As the attack comes from the military field, the security level of this attack is 4, and the risk acceptance should be zero, as no risk is acceptable.

To design the **Supermicro (Big Hack)** attack, the adversaries had to design microchips and integrate them into the servers of Supermicro citeRobertson2018,Mehta2020. Thus, the difficulty of the attack is at the laboratory level. The impact is not safety-critical, but it is still considered substantial as the microchips changed the core of the operating system and communicated with compromised computers. The identification of the attack vectors is high, but the exploitation is moderate, as the repetition of the attack does not require a high overhead of customization. The adversary's class is multiple experts and it requires substantial resources. The pain of defending against this attack requires security level 4 handling, against intentional malicious behavior using sophisticated means. The risk acceptance is moderate as the impact of the attack is substantial. The resources required are moderate.

The **Pager Attack** [7], [8] had a laboratory difficulty level, since multiple phases of the supply chain were compromised, and the expertise required was not only technological. The impact of the attack was safety-critical. As the attack seems to have been planned for at least 2 years, the identification of this attack vector is level 4, but the exploitation is level 2, as the repetition of the attack was not so challenging after the identification of the attack vectors. The order of the pagers was already done, and the processes of manufacturing and integration were already compromised. The pain of the adversary to implement such an attack is relatively high. The attack requires multiple experts and very high resources. The pain of defending such an attack requires all possible resources. The security level is high, and the risk acceptance is zero, as no risk should be accepted when human safety is compromised. At the same time, we assume that a visual inspection of the device's housing could have detected the explosion mechanism (as it happened by one of the users, leading to an earlier attack than planned). Thus, the class of the defender could have been a layman. Regarding the resources, for the scenario of the visual inspection, low resources are sufficient.

For the **hardware backdoors** [2], [31]–[35] described in the attacks in section V and based on the lack of detailed information, we formulated a common severity level analysis as seen in table 4. The class of the adversary is expert since multiple experts are required for the injection and concealing of a backdoor. The resources are substantial. The identification of a weak spot for a backdoor is substantial but the exploitation is low since no further effort is required for the repetition of the attack. The difficulty of the attack is based on each application, which cannot be assessed based on the unclear information provided in many attacks. The same applies to the defense categories of the assessment.

In the case of **counterfeit chips** [36]–[40], the class of the adversary can vary from layman to multiple experts, and the resources can vary from low to substantial. As an example, we can take the fake Ryzen CPUs, which cannot boot, and the military counterfeit chips, which may include malicious functionality. The level of identification can be moderate to

substantial and the exploitation remains, in this case, low.

Next, we present three **research use cases** from the literature and their severity level analysis: Malicious Routing, Power Virus, and Remote Power Analysis.

In the **Malicious Routing** attack [10] the difficulty requires special tools, but the impact of the attack cannot be assessed. The identification effort is substantial but the exploitation of the attack is uncomplicated as the EDA tool is expected to reproduce the attack. The class of the adversary is considered at the level of expert and the resources required are moderate. To defend against this attack, the required resources are moderate. We selected this attack for comparison as it is one of the latest attacks that changed the view of hardware Trojans. It included the trigger of the Trojan in the design flow instead of the original design, making the attack compatible with massive circuit infection. Thus, we consider it closer for comparison with real-world attacks.

Another example is the **Power Virus** attack implemented in [44]. For Power Virus attacks, the adversary includes in the original design, a sub-circuit with high switching activity, e.g., a ring oscillator. The power consumption is increased to the point of damage or denial-of-service failure. For safety-critical applications, denial of service or damage of any kind is not threatening human life as critical systems have safe failing procedures. Those attacks have low complexity as the circuits used for power-wasting are known even in the case of the power virus with a trigger. Nevertheless, it requires special tools, and the impact is substantial. The exploitation parameter of the attack is moderate and similar to the identification because customization may be required.

Finally, for the **Remote Power Analysis** [44], we consider that expert knowledge and moderate resources are required. The required tools are special and the impact can be moderate. The identification is substantial and the exploitation is moderate. From the side of defense, a proficient class of engineers is required and moderate resources. The identification and exploitation levels are moderate.

As is shown in our assessment, we consider the security level and the risk acceptance unknown due to the fact that these attacks were developed as use cases in research without a real-world environment of the attack. Thus, applicable regulations and standards are unknown for these applications.

In figure 2 we illustrate the assessments of attacks and defenses with a hardware security radar chart. By means of radar-diagrams, we can visualize the differences among the categories of attacks analyzed. From the real-world attacks, we included only those with a complete assessment scheme. Each radar has six dimensions for attack and six dimensions for defense, corresponding to the discussed categories. The red areas represent the attack assessments, and the green areas represent the defense assessments.

It is noticeable that the attacks observed in reality are more extensive and costly than the cases considered in research. Additionally, the research scenarios do not cover all the categories of defense. As the research attacks are not positioned in a specific environment, the security level and the risk acceptance cannot be assessed.

TABLE 4. Numerical values of the severity level analysis for the attack and defense of scenarios; KS=Kill Switch [25]; BH=Big Hack (Supermicro) [20], [26]; Pg=Pager [7], [8]; HB=Hardware Backdoors [2], [31]–[35]; CC=Counterfeit Chips [36]–[40]

	Numerical Values (Attack/Defense)				
	KS	BH	Pg	HB	CC
Class	4/4	4/4	4/1	4/-	1-4/-
Resources	4/4	3/3	4/1	3/-	1-3/-
Difficulty / Security Level	4/4	4/4	4/4	-/-	-/-
Impact /Risk Acceptance	5/5	3/3	5/5	-/-	-/-
Identification	4/4	4/3	4/2	3/-	2-3/-
Exploitation	4/4	2/2	1/1	1/-	1/-
	25/25	20/19	22/18	11/-	5-11/-

TABLE 5. Numerical values of the severity level analysis for the attack and defense of scenarios; MR=Malicious Routing [10]; PV= Power Virus [44]; RPA= Remote Power Analysis [44]; SE=Scenario Easy; SM=Scenario Moderate; SH=Scenario Heavy

	Numerical Values (Attack/Defense)					
	MR	PV	RPA	SE	SM	SH
Class	3/2	2/2	3/2	1/1	2/2	4/4
Resources	2/2	2/2	2/2	1/1	3/2	4/4
Difficulty / Security Level	3/0	3/0	3/0	1/2	2/2	4/4
Impact /Risk Acceptance	2/0	3/0	2/0	1/2	3/3	5/5
Identification	3/2	2/2	3/2	1/1	2/2	4/4
Exploitation	1/1	2/2	2/2	4/1	2/2	4/4
	14/15	14/8	15/8	9/8	14/13	25/25

VIII. DISCUSSION AND RECOMMENDATIONS

To restrict modifications that can lead to hardware attacks, it is vital to secure the hardware in any form throughout the IC supply chain. To define an asset as secure, it has to comply with some predefined policies. The main policies in the field of security are confidentiality, integrity, and availability (CIA) [45]. Even though these policies can cover a wide range of possible security threats, they are not sufficient when it comes to hardware security. Motivated by [12], we consider the following hardware security policies: *confidentiality, integrity, dependability, isolation, quantitative security properties, secure handling of the IC, transparency of the IC supply chain*.

Confidentiality: Confidential assets cannot be accessed by any means from entities that do not have the corresponding authorization.

Integrity ensures that the source of information, as well as the information itself has not been altered by any unauthorized entity.

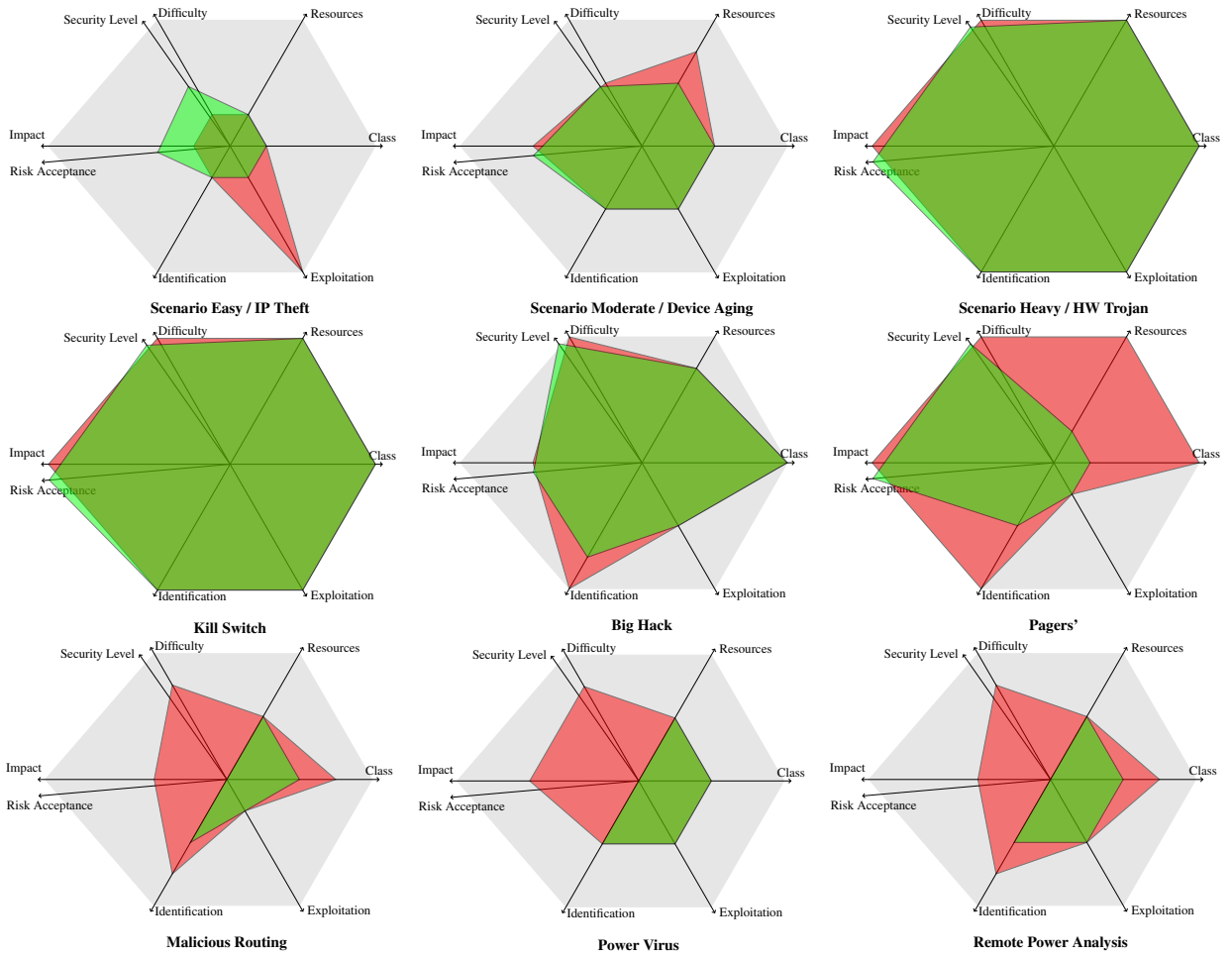

FIGURE 2. The Hardware Security radar for the three paradigmatic scenarios in the top row, the three real cases in the middle row, and the three research scenarios in the bottom row. Red areas represent the attacks, green areas the defense.

Dependability includes reliability, availability, and safety. A design should comply with those characteristics to be considered trustworthy. Reliability is the system's property of providing the correct and specified results even under the fluctuation of the correct function of several components. Availability is the ability of the system to behave as specified. Safety is the ability of the system to guarantee, that no human life is at risk at any time.

Isolation: Two different partitions of the design, a critical and a non-critical, should not communicate directly. As a critical partition, we consider the partition that handles confidential information. In the case of security level variations, two partitions with different security levels should not communicate directly.

Quantative security properties: This group of properties includes all the enforced security characteristics of the hardware designs that can be measured (e.g., enforcing constant time in a design such that the adversary cannot extract information by the invariants of the execution time). Some examples of such properties are the constant physical parameters of the designs, the randomness of the output of a crypto-core, etc. Those properties are the enforced protection of the circuit from side-channel analysis attacks.

In addition to those security policies, we identify two new security policies, addressing the threats rising from the IC supply chain.

Secure handling of the IC: Handling the chip without taking into account the specifications can lead to new attack surfaces and create attack vectors throughout the levels $3-5$ of the IC life cycle (e.g., the backyard industry of counterfeit chips). Secure IC handling can be secure verification, recycling, and disposal in a way that constricts the illegal IC market and prohibits unverified ICs from reaching the market.

Transparency of the IC supply chain and processes: All entities on the supply chain are known and provide clear information regarding the process they followed and their contribution to the IC supply chain. E.g., a verification entity should not only be known, but also the process and the result of the verification should be available.

From the severity level analysis of different attacks, the review of the security policies, and the regulations and standards enforced in the IC industry, we identify possible

future directions of research. One recommendation is to study more real-world attacks in detail and focus on the reasons why those attacks have not been mitigated on time. As an example, regarding the Pager attack, the visual inspection of the devices could have potentially revealed the explosive material and the detonator. Additionally, another question arises of how we can assess the trustworthiness not only of the chips used but also of the verification bodies. Taking the verification as part of the supply chain, we consider that the transparency of the supply chain can address this problem, too.

IX. CONCLUSION

Hardware security has become highly relevant. As the development process is accessible to more people, new attack surfaces are exposed. The supply chain, in the way it has formed, creates "blind spots" for several new kinds of attacks. The long list of entities involved makes it challenging to trace the origin of some components or some processes (e.g., integration). Additionally, as we noticed from the Pager Attack, the supply chain does not only include the technical processes but also the human factor. The pagers had to be altered maliciously before being delivered to specific users. This means that the vendors have been compromised or exchanged with fake ones. Human factors are a common weakness in security as it is one of the parameters that cannot be easily controlled.

At the same time, it is challenging to align research and real-world attacks. Research use cases are more technical and constructed often with a limited, laboratory-like scope. Therefore, they do not cover the complex environment of the application, leaving the regulations and the standards considered in real-world attacks out of consideration. On the other hand, information about real-world attacks is often concealed and does not become known in full detail. But from what is known, it is apparent that investing sufficient effort into analyzing them is both important and urgent. A good understanding of past attacks facilitates further development of tools and methodologies for defense.

Examples show that even simple devices that we carry with us could become a threat [27]. For us, this is enough motivation to look further into the field of hardware security and find ways to mitigate those attacks before they happen. A good step in this direction is taking into consideration the complete environment for research use cases including regulations and standards, creating more realistic study cases for mitigation methodologies.

Our analysis gives an insight into the supply chain of integrated circuits, their security policies, and relevant regulations and standards. In addition, we suggested an assessment method in order to position hardware attacks both from the real world and from research. This assessment and the visualization through radar plots helped to pinpoint the differences between research use cases and real-world attacks. Based on the analysis, we could formulate research gaps that should be addressed in the future. The most urgent recommendation is to include the environment of the attack in the threat model, rather than focusing on technical details alone. In this connection, we also argue that the transparency of the entities and the processes involved in the supply chain plays a vital role in securing hardware applications.

REFERENCES

[1] S. Bhunia, Hardware security : a hands-on learning approach. Cambridge, MA: Morgan Kaufmann, an imprint of Elsevier, 2019.

[2] S. Gallagner. Photos of an NSA "upgrade" factory show Cisco router getting implant. https://arstechnica.com/tech-policy/2014/05/photos-of-an-nsa-upgrade-factory-show-cisco-router-getting-implant/. (2014) Accessed on 2024-11-16.

[3] S. S. Hsu. U.S. charges Florida pair with selling counterfeit computer chips from China to the U.S. Navy and military. https://www.washingtonpost.com/wp-dyn/content/article/2010/09/14/AR2010091406468.html. (2010) Accessed on 2024-11-16.

[4] M. Bhole, T. Sauter, and W. Kastner, "Enhancing industrial cybersecurity: Insights from analyzing threat groups and strategies in operational technology environments," IEEE Open Journal of the Industrial Electronics Society, vol. 6, DOI 10.1109/OJIES.2025.3527585, pp. 145–157, 2025.

[5] L. Dobberstein. Millions of mobile phones come pre-infected with malware, say researchers. (2023, May) Accessed on 2025-01-31. [Online]. Available: https://www.theregister.com/2023/05/11/bh_asia_mobile_phones/

[6] T. Sauter and A. Treytl, "Iot-enabled sensors in automation systems and their security challenges," IEEE Sensors Letters, vol. 7, DOI 10.1109/LSENS.2023.3332404, no. 12, pp. 1–4, 2023.

[7] I. Thomson. Lebanon: At least nine dead, thousands hurt after hezbollah pagers explode. Accessed on 2025-01-31. [Online]. Available: https://www.theregister.com/2024/09/17/hezbollah_lebanon_explosive_pagers/

[8] C. Mallo, T. Qiblawi, J. Diamond, L. Kent, R. Picheta, C. Edwards, and H. Regan. Israel behind deadly pager explosions that targeted hezbollah and injured thousands in lebanon. [Online]. Available: https://edition.cnn.com/2024/09/17/middleeast/lebanon-hezbollah-pagers-explosions-intl/index.html

[9] B. Halak, CIST: A Threat Modelling Approach for Hardware Supply Chain Security, pp. 3–65. Cham: Springer International Publishing, 2021. [Online]. Available: https://doi.org/10.1007/978-3-030-62707-2_1

[10] Q. A. Ahmed, T. Wiersema, and M. Platzner, "Malicious routing: Circumventing bitstream-level verification for FPGAs," in 2021 Design, Automation Test in Europe Conference Exhibition (DATE), DOI 10.23919/DATE51398.2021.9474026, pp. 1490–1495, 2021.

[11] A. J. Christian Krieg, Clifford Wolf, "Malicious LUT: A stealthy FPGA Trojan Injected and Triggered by the Design Flow," in 2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD), 2016. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=7827620

[12] W. Hu, C.-H. Chang, A. Sengupta, S. Bhunia, R. Kastner, and H. Li, "An Overview of Hardware Security and Trust: Threats, Countermeasures, and Design Tools," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 40, DOI 10.1109/TCAD.2020.3047976, no. 6, pp. 1010–1038, 2021.

[13] N. Fern and K.-T. T. Cheng, "Detecting hardware trojans in unspecified functionality using mutation testing," in Proceedings of the IEEE/ACM International Conference on Computer-Aided Design, ser. ICCAD'15, p. 560–566. IEEE Press, 2015. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=7372619

[14] N. Fern, I. San, c. K. Koç, and K.-T. T. Cheng, "Hardware trojans in incompletely specified on-chip bus systems," in Proceedings of the 2016 Conference on Design, Automation amp; Test in Europe, ser. DATE '16, p. 527–530. San Jose, CA, USA: EDA Consortium. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7459366

[15] W. Hu, L. Zhang, A. Ardeshiricham, J. Blackstone, B. Hou, Y. Tai, and R. Kastner, "Why you should care about don't cares: Exploiting internal don't care conditions for hardware trojans," in Proceedings of the 36th International Conference on Computer-Aided Design, ser. ICCAD'17, p. 707–713. IEEE Press, 2017. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=8203846tag=1

[16] M. Kim, S. Kong, B. Hong, L. Xu, W. Shi, and T. Suh, "Evaluating coherence-exploiting hardware trojan," in Design, Automa-

tion Test in Europe Conference Exhibition (DATE), 2017, DOI 10.23919/DATE.2017.7926975, pp. 157–162, 2017.

[17] Y. Zhao, X. Hu, S. Li, J. Ye, L. Deng, Y. Ji, J. Xu, D. Wu, and Y. Xie, "Memory Trojan Attack on Neural Network Accelerators," in 2019 Design, Automation Test in Europe Conference Exhibition (DATE), DOI 10.23919/DATE.2019.8715027, pp. 1415–1420, 2019.

[18] A. P. Johnson, R. S. Chakraborty, and D. Mukhopadhyay, "A novel attack on a FPGA based true random number generator," in Proceedings of the WESS'15: Workshop on Embedded Systems Security, ser. WESS'15, DOI 10.1145/2818362.2818368. New York, NY, USA: Association for Computing Machinery, 2015. [Online]. Available: https://doi.org/10.1145/2818362.2818368

[19] X. Wang, T. Hoque, A. Basak, R. Karam, W. Hu, M. Qin, D. Mu, and S. Bhunia, "Hardware trojan attack in embedded memory," J. Emerg. Technol. Comput. Syst., vol. 17, DOI 10.1145/3422353, no. 1, Jan. 2021. [Online]. Available: https://doi.org/10.1145/3422353

[20] D. Mehta, H. Lu, O. P. Paradis, M. A. M. S., M. T. Rahman, Y. Iskander, P. Chawla, D. L. Woodard, M. Tehranipoor, and N. Asadizanjani, "The big hack explained: Detection and prevention of pcb supply chain implants," J. Emerg. Technol. Comput. Syst., vol. 16, DOI 10.1145/3401980, no. 4, Aug. 2020. [Online]. Available: https://doi.org/10.1145/3401980

[21] European Parlament, "Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 december 2022 on measures for a high common level of cybersecurity across the union," European Union Law, Dec. 2022.

[22] European Parlament, "Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 october 2024 on horizontal cybersecurity requirements for products with digital elements," European Union Law, Oct. 2024. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R2847&qid=1738150256240

[23] 2021-02-26 and E. Team, "Understanding iec 62443." [Online]. Available: https://www.iec.ch/blog/understanding-iec-62443

[24] European Commission, "Commission Delegated Regulation (EU) 2022/30 of 29 October 2021 Supplementing Directive 2014/53/EU of the European Parliament and of the Council with Regard to the Application of the Essential Requirements Referred to in Article 3(3), Points (d), (e) and (f), of that Directive (Text with EEA Relevance)," Online, Jan. 2022. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0030

[25] A. Adee, "The hunt for the kill switch," IEEE Spectrum, vol. 45, DOI 10.1109/MSPEC.2008.4505310, no. 5, pp. 34–39, 2008.

[26] J. Robertson and M. Riley. The big hack: How China used a tiny chip to infiltrate u.s. companies. [Online]. Available: https://www.bloomberg.com/news/features/2018-10-04/the-big-hack-how-china-used-a-tiny-chip-to-infiltrate-america-s-top-companies

[27] F. Sheera, B. Ronen, and S. Hwaida. How israel built a modern-day trojan horse: Exploding pagers. [Online]. Available: https://www.nytimes.com/2024/09/18/world/middleeast/israel-exploding-pagers-hezbollah.html

[28] M. Murphy and J. Tidy. What we know about the hezbollah device explosions. [Online]. Available: https://www.bbc.com/news/articles/cz04m913m49o

[29] Ar-924 rigged explosive radio pager - wanted item. [Online]. Available: https://www.cryptomuseum.com/covert/radio/apollo/ar924/

[30] J. Franks. How pagers and hand-held radios can be modified to remotely explode. [Online]. Available: https://news.sky.com/story/how-does-a-pager-explode-the-steps-needed-to-remotely-detonate-hezbollah-devices-13217335

[31] G. Miller. The intelligence coup of the century. [Online]. Available: https://www.washingtonpost.com/graphics/2020/world/national-security/cia-crypto-encryption-machines-espionage/

[32] J. Markoff. Old trick threatens the newest weapons. [Online]. Available: https://www.nytimes.com/2009/10/27/science/27trojan.html?pagewanted=all

[33] S. Skorobogatov and C. Woods, "Breakthrough silicon scanning discovers backdoor in military chip," in Proceedings of the 14th International Conference on Cryptographic Hardware and Embedded Systems, ser. CHES'12, DOI 10.1007/978-3-642-33027-8_2, p. 23–40. Berlin, Heidelberg: Springer-Verlag, 2012. [Online]. Available: https://doi.org/10.1007/978-3-642-33027-8_2

[34] C. Arthur. Cyber-attack concerns raised over boeing 787 chip's 'back door'. [Online]. Available: https://www.theguardian.com/technology/2012/may/29/cyber-attack-concerns-boeing-chip

[35] G. Greenwald. Glenn greenwald: how the nsa tampers with us-made internet routers. The guardian. [Online]. Available: https://www.theguardian.com/books/2014/may/12/glenn-greenwald-nsa-tampers-us-internet-routers-snowden

[36] A. Rawnsley. Fishy chips: spies want to hack-proof circuits. [Online]. Available: https://www.wired.com/2011/06/chips-oy-spies-want-to-hack-proof-circuits/

[37] Dell warns of hardware trojan. [Online]. Available: https://www.homelandsecuritynewswire.com/dell-warns-hardware-trojan

[38] J. Oates. Dell warns on spyware infected server motherboards. [Online]. Available: https://www.theregister.com/2010/07/21/dell_server_warning/

[39] J. Markoff. F.b.i. says the military had bogus computer gear. Accessed: 2025-01-31. [Online]. Available: https://www.nytimes.com/2008/05/09/technology/09cisco.html

[40] M. Tyson. Fake Ryzen 7 9800X3D CPUs are circulating in China — MSI China alerts buyers of new Zen 5 scam. [Online]. Available: https://www.tomshardware.com/pc-components/cpus/fake-ryzen-7-9800x3d-cpus-are-circulating-in-china-msi-china-alerts-buyers-of-new-zen-5-scam

[41] "Common criteria for information technology security evaluation," 2022, accessed: 2023-10-25. [Online]. Available: https://www.commoncriteriaportal.org/cc/index.cfm

[42] National Institute of Standards and Technology, "Standards for security categorization of federal information and information systems," U.S. Department of Commerce, Tech. Rep. FIPS PUB 199, Feb. 2004, accessed: 2024-11-11. [Online]. Available: https://csrc.nist.gov/publications/detail/fips/199/final

[43] N. N. Anandakumar, M. S. Rahman, M. M. M. Rahman, R. Kibria, U. Das, F. Farahmandi, F. Rahman, and M. M. Tehranipoor, "Rethinking watermark: Providing proof of ip ownership in modern socs," Cryptology ePrint Archive, 2022. [Online]. Available: https://eprint.iacr.org/2022/092

[44] J. Y. Koh and T. Nandha Kumar, "Review of side channel attacks and countermeasures of FPGA based systems," in 2021 IEEE 19th Student Conference on Research and Development (SCOReD), DOI 10.1109/SCOReD53546.2021.9652773, pp. 102–107, 2021.

[45] S. Snedaker and B. Cunningham, The best damn IT security management book period, 1st ed. Burlingont, Mass. : Oxford: Syngress ; Elsevier Science, 2007.

SOFIA MARAGKOU is a university assistant at the Institute of Computer Technology at TU Wien. She received her Dipl. Eng. degree from the Technical University of Crete in 2019. Since 2020, she has been a member of SafeSecLab and a PhD candidate. Her research interests include hardware security, specifically hardware Trojans and side-channel attacks. Additionally, she has recently started working on topics related to RISC-V.

LUKAS RAPPEL is a member of the TÜV Austria SafeSecLab Team. He received his B.Sc degree in Hard- and Software Design and M.Sc. Degree in Embedded Systems Design from the University of Applied Sciences Upper Austria, Campus Hagenberg, in 2017. He has been working as a technical expert at TÜV Austria GmbH since 2018. His field of activity includes the inspection of safety control systems (functional safety) in the industry and assessments of complex control systems such as programmable safety controls and integrated safety functions of electrical drives.

HENDRIK DETTMER was born in Bottrop, Germany in 1981. He received a diploma in IT-security engineering degrees in from the University of Bochum, in 2009. Since 2009, he has been working in the field of software and hardware security. First in the company SRC GmbH in Bonn as an evaluator and project manager of Common Criteria projects and in international workgroups to improve IT-security standards. In 2014, he joined Wincor Nixdorf in Paderborn to improve the software and hardware of payment terminals for world-wide use. In 2017, Hendrik joined TÜV TRUST IT GmbH TÜV AUSTRIA GROUP and created, among others, a certification scheme for IoT devices. He is also a technical expert in the OT-security domain and has several personal certifications from the German BSI govermental body.

THILO SAUTER (M'93, SM'09, F'14) holds a Ph.D. degree in electrical engineering and is professor for automation technology at TU Wien, Vienna, Austria, as well as senior scientist at the University of Continuing Education Krems, Wiener Neustadt, Austria. From 2004 to 2013, he also was the Founding Director of the Institute for Integrated Sensor Systems at the Austrian Academy of Sciences. His expertise and research interests include embedded systems and integrated circuit design, smart sensors, and automation and sensor networks with a focus on real-time, security, interconnection, and integration issues relevant to cyber-physical systems and the Internet of Things in various application domains such as industrial and building automation, smart manufacturing, or smart grids. Dr. Sauter is member of the Board of the Austrian Electrotechnical Association and Senior AdCom Member of the IEEE Industrial Electronics Society (IES). Moreover, he has been involved in the standardization of industrial communication systems for more than 25 years.

AXEL JANTSCH (Fellow, IEEE) received the Dipl.Ing. degree and the Ph.D. degree in computer science from TU Wien, Vienna, Austria, in 1987 and 1992, respectively. From 1997 to 2002, he was an Associate Professor with KTH Royal Institute of Technology, Stockholm. From 2002 to 2014, he was a Full Professor in electronic systems design at KTH. Since 2014, he has been a Professor of systems on chips with the Institute of Computer Technology, TU Wien. His current research interests include systems on chips and embedded machine learning. He has published five books as an editor and two as single author and over 400 peer-reviewed contributions in journals, books, and conference proceedings. He has given over 130 invited presentations at conferences, universities, and companies.