In an increasingly digitized world, the systems that sustain our daily lives – such as energy grids, water supplies, and manufacturing plants – are more interconnected than ever. However, this progress comes at a price. The rise of industrial Internet of Things (IoT) applications has significantly expanded attack surfaces, leaving critical infrastructures exposed to sophisticated cyber threats. Even with strong security controls, the risk of incidents persists. This dissertation tackles these challenges by focusing on incident response in industrial environments. It examines how to prepare for, detect, and respond to cybersecurity threats and incidents through the use of digital twins – virtual replicas of physical systems – and structured cybersecurity playbooks. Specifically, it demonstrates how digital twins support security operations when preparing and detecting attacks across hosts and networks. Additionally, it highlights how cybersecurity playbooks provide clear, actionable steps for proactive and reactive responses, helping to mitigate vulnerabilities and respond to incidents.

Universität Regensburg
UNIVERSITÄTSBIBLIOTHEK

ISBN 978-3-88246-521-1
9 783882 465211

Incident Response for the Internet of Things

Philip Empl

Philip Empl

Incident Response for the Internet of Things

Eine Publikation der
Universitätsbibliothek Regensburg

Incident Response for the Internet of Things

Dissertation zur Erlangung des Grades eines
Doktors der Wirtschaftswissenschaft

eingereicht an der
Fakultät für Wirtschaftswissenschaften
der Universität Regensburg

vorgelegt von:

Philip Maurice Empl

Berichterstatter
Prof. Dr. Günther Pernul
Prof. Dr. Stefan Schönig

Tag der Disputation: 04. November 2024

Erstgutachter (Betreuer): Prof. Dr. Günther Pernul

Zweitgutachter: Prof. Dr. Stefan Schönig

Tag der Prüfung: 04. November 2024

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie. Detaillierte bibliografische Daten sind im Internet unter http://dnb.dnb.de abrufbar.

ISBN (Print): 978-3-88246-521-1

ISBN (PDF): 978-3-88246-522-8

DOI: 10.5283/epub.59795

Link zur Online-Version: https://epub.uni-regensburg.de/59795

Satz und Layout: Philip Empl

Titelbild: Generiert mit Midjourney, Inc. am 14.10.2024

Umschlaggestaltung: Peter Brünsteiner

Herstellung: Universitätsbibliothek Regensburg

Erscheinungsort: Regensburg, 2025

Druck und Bindung: Digital Print Group o. Schimek GmbH, Nürnberg

To my wonderful parents, Birgit and Herbert,
whose love and support sustains me
since the beginning of my life.

Acknowledgement

Undoubtedly, succeeding in your PhD requires immense dedication. Yet it is mainly about the people who push you beyond your limits and catch you when you fall. I am fortunate to have many such outstanding individuals. Some are constants in my life, while others became invaluable parts of my story and friends along the way.

I extend my heartfelt thanks to Prof. Dr. Günther Pernul, an exemplary supervisor whose open-door policy and care provided me with guidance, freedom, and support when I needed it the most. His wisdom made my journey not only possible but enjoyable, and I will never forget my first IFS seminar. Prof. Dr. Stefan Schönig also deserves special mention for his keen interest in my work and support, enriching my research.

The journey was filled with many who left an indelible mark. Many thanks to Daniel; I really started to enjoy research with you. I will always reflect on our insightful discussions, the almost-missed award ceremony, and after-work sports. Lena, the 127.0.0.1 office colleague, brought laughter, scientific debate, and a shared joy in RWS 105 and beyond. Markus stood by me as a soul mate, a steadfast and reliable support, and a future business ally. Manfred brought interdisciplinary discussions, memorable lunch breaks, and priceless off-work fun. Thanks to Bene for the smooth onboarding during tough COVID-19 times and for taking me much further than Meteora. Also, thanks to all those other fellow IFS travelers on this journey, those who shared the same boat. Thanks to Sabrina for opening up the IoT topic together, Johannes for sharing the same passion for WGMs, and Tobi, who you can literally always ring. Thanks to Fabian, Ludwig, Marietheres, Daniel, and Mathis, with whom I had the honor of having time at the chair. Thanks, Thomas and Sascha; I will never forget our scientific pool talks. Also, thanks to Daniel, Henric, Lukas, David, and Timo, whose contributions as students were valuable. I sincerely thank Petra for her organizational support, which kept the gears turning smoothly, and Werner for the engaging technical ones. Thanks to Gerit, a true beacon of scientific mentorship, who advised me in GTD.

I am grateful that in Jonas I have a constant and reliable friend who I can confide in and always rely on. Thanks to my parents and sister Isabel for their strong emotional support, allowing me to be myself and pursue my dreams, always and in every way. Lisa, this moment shines for you. More than a partner, Lisa is that friend whose boundless love and joy have profoundly enriched my life. You always understand me. In closing, my appreciation extends to all who have believed in and supported me, but also to those who wrote me off at the time. You spurred me on toward greater heights.

Abstract

With the increasing digitization of critical infrastructures, highly organized threat actors pose significant threats to the confidentiality, integrity, and availability of *Internet of Things (IoT)* systems. While implementing security controls enhances cybersecurity, a residual risk persists. Driven by legal mandates, critical infrastructure operators must now address these residual risks through incident response measures. Given the heterogeneity and resource constraints of IoT systems, there is an urgent need for innovative incident response strategies. The *National Institute of Standards and Technology (NIST)* Incident Response Lifecycle provides a framework for this dissertation, which explores the application of digital twins in incident response for IoT systems. By adhering to the four phases of the NIST Incident Response Lifecycle, this dissertation advances three key areas: Digital Twins, Preparation & Detection, and Response & Learning.

First, this dissertation extends the role of digital twins beyond mere incident response. Second, it addresses the *Preparation & Detection* phases of the NIST Incident Response Lifecycle, focusing on integrating smaller IoT devices and considering situational awareness in intrusion detection systems both in hosts and networks. Subsequently, this dissertation delves into the *Containment, Eradication, Recovery, & Post-Incident* phases, tailoring incident response processes to organizational IoT systems' nuances and automating reactive and proactive playbooks for security orchestration, automation, and response. With the rise of new legal requirements and highly organized threat groups, this dissertation strives to contribute to more secure critical infrastructures with incident responses for IoT systems.

Contents

List of Tables

List of Figures

Abbreviations and Acronyms

AI	Artificial Intelligence
APT	Advanced Persistent Threat
CACAO	Collaborative Automated Course of Action Operations
CAPEC	Common Attack Pattern Enumeration and Classification
CERT	Computer Emergency Response Team
CIS	Center for Internet Security
CISO	Chief Information Security Officer
CORE	Computing Research & Education
CPE	Common Platform Enumeration
CRISP-DM	Cross Industry Standard Process for Data Mining
CSAF	Common Security Advisory Framework
CSF	Cybersecurity Framework
CSIRT	Computer Security Incident Response Team
CSMS	Cybersecurity Management System
CTI	Cyber Threat Intelligence
CVE	Common Vulnerabilities and Exposures
CVRF	Common Vulnerability Reporting Framework
CVSS	Common Vulnerability Scoring System
CWE	Common Weakness Enumeration
DApps	Decentralized Application
DSR	Design Science Research
DSRM	Design Science Research Methodology
EPSS	Exploit Prediction Scoring System
EU	European Union
EU-CyCLONe	European Cyber Crisis Liaison Organisation Network
FTP	File Transfer Protocol
GB	Gigabyte
HIDS	Host Intrusion Detection System

HMI	Human Machine Interface
ICS	Industrial Control System
ID	Identifier
IDS	Intrusion Detection System
IEC	International Electrotechnical Commission
IF	Impact Factor
IIRA	Industrial Internet Reference Architecture
IoT	Internet of Things
IP	Internet Protocol
IPFIX	Internet Protocol Flow Information Export
IS	Information Systems
ISA	International Society of Automation
ISO	International Organization for Standardization
IT	Information Technology
LED	Light-Emitting Diode
LLM	Large Language Models
MISP	Malware Information Sharing Platform
MITRE	Massachusetts Institute of Technology Research and Engineering
MQTT	Message Queuing Telemetry Transport
NIDS	Network Intrusion Detection System
NIS2	Network and Information Security 2
NIS2UmsuCG	NIS2-Umsetzungsgesetz
NIST	National Institute of Standards and Technology
NISTIR	NIST Interagency or Internal Reports
NoSQL	Not only SQL
OODA	Observe, Orient, Decide, and Act
OPC UA	Open Platform Communications United Architecture
OpenC2	Open Command and Control
OT	Operational Technology
OVAL	Open Vulnerability and Assessment Language
PCB	Printed Circuit Board
PPD-21	Presidential Policy Directive 21
PSIRT	Product Security Incident Response Team
RAG	Retrieval Augmented Generation

RAM	Random-Access Memory
RAMI 4.0	Referenzarchitekturmodell Industrie 4.0
RQ	Research Question
SIEM	Security Information and Event Management
SME	Small and Medium-sized Enterprises
SOAR	Security Orchestration, Automation, and Response
SOC	Security Operations Center
SSH	Secure Shell
SWID	Software Identification
TCP	Transmission Control Protocol
UI	User Interface
URL	Uniform Resource Locator

Disclaimer. This dissertation has benefited from using AI tools (e.g., Grammarly, ChatGPT, and DeepL) for language improvement and editing. While these tools have assisted in refining the text's clarity, coherence, and overall readability, all intellectual content and thoughts are solely those of the author.

Part I

Dissertation Overview

The first part of this dissertation outlines its structure and rationales.

1 Motivation

Information Technology (IT) is becoming ever more ubiquitous and, thus, increasingly shifting conflicts to the digital space. Cyberattacks are often orchestrated by organized actors known as *Advanced Persistent Threat (APT)* groups [32]. MITRE, a non-profit organization, has identified 152 APT groups, a significant number of which are state-sponsored or financially motivated (e.g., Dragonfly, APT33, and PLA Unit 61398) [71]. These threat actors have prompted a swift and decisive response from international and national bodies. Following other nations' lead, Germany will establish a cybersecurity command [57]. Of course, this command defends IT systems, but what truly demands attention are our critical infrastructures, and here is why:

Critical infrastructures ensure the well-being of our society, with essential sectors like water and wastewater systems, energy, and critical manufacturing [75]. They are built of *Operational Technology (OT)* systems that control machines and physical processes. Compared to commodity IT systems, OT systems are used for machine automation and engineered for long-term operations [14]. As a result, systems put into operation years back consist of non-upgradeable legacy hardware and software (e.g., Windows 3.11), communicate via proprietary automation protocols (e.g., S7comm), and are insecure by design (e.g., communication in plain text) [64]. This poses many security challenges when considering all critical vulnerabilities, exploits, and potential outcomes. The Ukraine blackouts [56] or the cyberattacks on nuclear facilities in Iran [66] serve as stark reminders of security incidents. Nevertheless, when those vulnerable critical infrastructures undergo steady digitization, they integrate more and more IT systems [58]. This opens up the prior isolated OT systems to the internet, shaping the industrial *Internet of Things (IoT)*. The IoT describes the interconnection of identifiable objects – ranging from sensors to actuators – that broadcast their services to semantic-rich applications (e.g., predictive maintenance) [2]. In short, critical infrastructures shift from OT systems to more complex IoT systems, expanding their attack surfaces.

To defend critical infrastructures, the United States (e.g., PPD-21 [75]) and China (e.g., Cybersecurity Law [62]) set clear examples. In line with those, the *European Union (EU)* has announced its cybersecurity strategy for the digital decade spanning from 2020 to 2030 [22]. This strategy aims at improving security resilience and includes three legislative acts specifically designed to protect critical infrastructures: the NIS2 Directive [25], the Cyber Resilience Act [23], and the Cyber Solidarity Act [24]. While the NIS2 Directive defines a global cybersecurity benchmark that EU member states must achieve by 2024[1], the other two legal acts are binding and await adoption. The Cyber Resilience Act addresses manufacturers and promotes proactive security controls by embedding security-by-design principles in digital products. In contrast, the NIS2 Directive and the Cyber Solidarity Act target operators and aim at reactive *incident response* management, including intrusion detection and cross-border collaboration.

[1] German legal authorities are drafting the NIS2UmsuCG [16] (successor to: "IT-Sicherheitsgesetz 2.0").

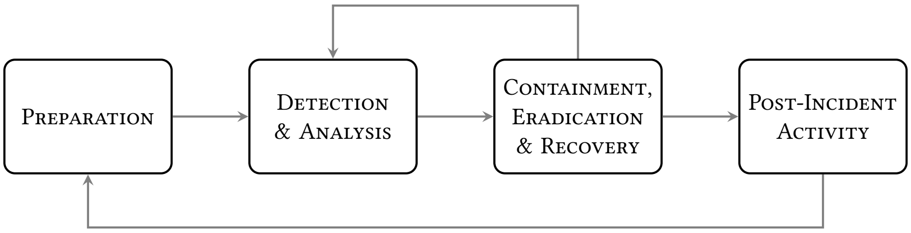


Figure 1: *The NIST Incident Response Lifecycle [12] handles residual security risks.*

While embedding cybersecurity into future products is essential, there remains an inherent residual risk [35] – essentially the risk that exists after implementing security controls (e.g., ISA/IEC 62443 [38] or NIST *Cybersecurity Framework (CSF)* [47]) and after obtaining security insurances. This residual risk opens a doorway for attackers, where incident response comes into play. The NIST Incident Response Lifecycle [12] is an established, iterative approach for managing security incidents (see Figure 1). During the Preparation phase, organizations develop policies, plans, and teams to ensure readiness. The Detection and Analysis phase involves monitoring to identify incidents. Upon detection, the Containment, Eradication, and Recovery phase aims to halt incidents, neutralize threats, and restore normal operations. The Post-Incident Activity phase conducts thorough reviews to refine future responses.

Managing threats and incidents within the industrial IoT proves challenging, as it diverges significantly from traditional IT systems and corresponding strategies (e.g., with rare patching windows and poor asset visibility [13]). Today's critical infrastructures are complex, with a mix of heterogeneous, resource-constrained, and legacy assets communicating over proprietary protocols [40, 44]. Coupled with the ever-evolving cyber threats, organized APT groups, and increasing legislative requirements (i.e., EU NIS2 Directive), the demand for organization-specific incident response has never been bigger. Since digital twins are organization-specific and asset-centric, they emerge as a cutting-edge solution, abstracting operations and introducing an additional defense-in-depth layer [4]. By offering dynamic virtual representations of physical assets [10], digital twins provide a decoupled, centralized, and unified platform for managing and securing critical infrastructures without risking operations [18, 19]. To date, incident response and related security orchestration in IoT are scarcely researched [39], no less with the backing of digital twins [5]. This dissertation investigates digital twins for incident response in the IoT. It contributes to solving common IoT security challenges, fulfilling legal requirements, and making critical infrastructures more responsive.

This dissertation is structured as follows: Section 2 delves into influential related work and relevant background information. Section 3 delineates the research questions, and Section 4 describes the research process relying on the design science research methodology and the research setting. Section 5 presents a detailed compilation of the research papers contributing to this cumulative dissertation. Section 6 concludes with summaries of the findings and implications for future research. This dissertation is complemented by Part II, which encompasses the original scientific publications.

2 Related Work

Each publication within this cumulative dissertation includes detailed related work, so this section provides a broader perspective. Incident response for IoT systems is highly regulated and shaped by industry advancements. Considering those aspects impacting this research is essential before delving into influential research.

Legal Requirements for Incident Response. In its strategy for the digital decade, the EU approaches secure critical infrastructures throughout their lifecycles [22]. This strategy includes secure product design under the Cyber Resilience Act [23], the assurance of secure operations through the NIS2 Directive [25], and enhanced EU-wide cybersecurity collaboration via the Cyber Solidarity Act [24]. The Cyber Resilience Act establishes stringent requirements to enhance incident response capabilities for products with digital elements. It mandates that such products be supplied free of known exploitable vulnerabilities, ensuring a baseline level of cybersecurity at delivery (*Annex I, 1.2*). Additionally, it requires mechanisms for addressing potential vulnerabilities through timely security updates (*Annex I, 1.3k*) and mandates malicious activity monitoring that may indicate security issues (*Annex I, 1.3j*). Manufacturers are also obligated to facilitate the sharing of information (e.g., vulnerabilities) to enhance collaborative cybersecurity efforts (*Annex I, 2.6*). The EU's NIS2 Directive strengthens cybersecurity across critical infrastructure operators by expanding coverage to include essential and important entities. It requires each member state to establish a *Computer Security Incident Response Team (CSIRT)*, a single point of contact for cybersecurity (*Art. 10*), and comprehensive incident response plans (*Art. 9*). The directive emphasizes operational readiness through incident handling, business continuity, including disaster recovery, and vulnerability management (*Art. 21*). It also imposes stringent and timely incident reporting obligations to ensure rapid communication of cyber threats (*Art. 23*). Networks of CSIRTs and the EU-CyCLONe, a joint EU cybersecurity unit, foster cooperation among EU member states and enable informed responses to incidents (*Art. 16*). The Cyber Solidarity Act further enhances the EU's capacity to collaboratively prepare for, detect, and respond to security threats. It aims to protect critical infrastructures by establishing the European Cyber Shield and a cybersecurity emergency mechanism. Integrating national and cross-border *Security Operations Center (SOC)* improves EU-wide coordination and strengthens the resilience against cyberattacks. In summary, the EU has put effort into enhancing critical infrastructure incident response capabilities.

Cybersecurity Standards for Incident Response. Beyond legal requirements, compliance with industry standards and certifications is crucial for accessing new markets, enhancing competitiveness, and earning customer trust. While the NIST CSF [47] is an open standard to start, organizations may opt for industry-specific standards like the OT-centric IEC 62443 [38] or the more IT-focused ISO 27001 [36] to obtain certificates and secure their critical infrastructures. The NIST CSF, along with standards such as NIST 800-82 [64], NISTIR 8428 [28], IEC 61508 [33], and IEC 61511 [34], details the management of security incidents through detection, response, and recovery functions. It highlights continuous monitoring (*DE.CM*) and the analysis of potential threats (*DE.AE*) to detect security incidents swiftly. Upon detection, the NIST CSF advises organizations on managing and analyzing incidents, ensuring thorough incident response planning (*RS.RP*), effective communication (*RS.CO*), and mitigation efforts to address and eliminate the threats (*RS.MI*). The recovery phase is marked by executing a detailed plan to restore affected systems and services (*RC.RP*), coupled with ongoing communication with stakeholders (*RC.CO*). The IEC 62443:2019 standard series caters to different roles, such as asset owners, product suppliers, and service providers. Asset owners and service providers must develop a *Cybersecurity Management System (CSMS)* that includes risk assessment as well as policies and procedures for incident response (*2-1*). Product suppliers are expected to meet organizational and procedural requirements (*4-1)*, including continuous monitoring (*4-2, CR 6.2*). System integrators need to be aware of vulnerabilities and implement measures for control system recovery and reconstitution (*3-3, SR 7.4*). The IT-oriented ISO 27001:2022 standard covers event monitoring (*A.8.16*) and the technical management of vulnerabilities (*A.8.8*). It expects a comprehensive incident response plan (*A.5.24*), delineates roles and responsibilities (*A.5.2*) related to security incidents responses (*A.5.26*), and encourages learning from past incidents to enhance future responses (*A.5.27*). Standards and frameworks overlap, indicating the necessity of incident response in organizations.

Threat Information for Incident Response. The success of IoT incident response management hinges significantly on an organization's ability to detect and manage incidents. For the former, incident detection is about consuming and standardizing threat information (e.g., MISP [45], OVAL [74], and STIX [50]), thereby building on open standard-driven processes. With more structured data about hardware and software (e.g., CPE [68], SWID [37], and CycloneDX [53]), organizations can identify associated vulnerabilities (e.g., CVE [69], and CWE [70]), their criticality (e.g., CVSS [26] and EPSS [27]), and advisories (e.g., CSAF [49]), as well as potential APT behavior and techniques (e.g., MITRE ATT&CK for ICS [72], or CAPEC [67]). With these data formats, organizations can proactively monitor threats and secure their IoT systems. For reactive incident detection, *Intrusion Detection System (IDS)* are designed to monitor malicious events on hosts (e.g., HIDS) and networks (e.g., NIDS) [7, 15]. On the incident response management side, the focus shifts to a synergy of people, processes, and technologies. The rise of *Security Orchestration, Automation, and Response (SOAR)* platforms has brought the adoption of playbooks. These playbooks are structured processes that enhance incident handling by coordinating people (e.g., communication with analysts) and integrating technologies (e.g., firewall), promoting a standardized approach to incident response with organizational specifics. Central to these efforts are open standards like CACAO [48], OpenC2 [51], and D3FEND [73]. In summary, threat intelligence is about creating and sharing data and knowledge around incident response.

Influential Research and Domain Problems. Although it may initially appear that incident response is primarily legal and technical, it also involves organizational and social aspects [46]. At its core, incident response is about processes and making decisions, particularly transitioning the phases of the *Observe, Orient, Decide, and Act (OODA)* loop [31], a widely recognized decision-making framework. Deciding the response to incidents requires a high level of situational awareness, which involves the continuous cycle of sensing, understanding, and predicting environmental changes [21] – capabilities critical in IoT systems. Moreover, as incident response processes are currently generic, they lack organizational contexts, complicating incident response management [61]. Digital twins may solve these IoT incident response challenges.

The digital twin paradigm is a highly discussed topic [65] and gives a set of existing technologies a new name. Simply put, a digital twin is a virtual representation of a real-world entity, set apart by its bidirectional communication relationship with its physical counterpart [41]. Beyond the clear benefits of synchronization, a digital twin serves as a technological framework – akin to a middleware – that allows on-building applications to access real-time and descriptive data (e.g., states), semantics (e.g., behavior), as well as services (e.g., simulation) [10]. This allows for realizing various use cases like predictive maintenance. In cybersecurity, the digital twin not only presents inherent security challenges but also offers a defense-in-depth layer for IoT systems [4]. It acts as a decoupled, centralized, and unified platform for securing critical infrastructures without compromising operational availability [18, 19]. At the same time, the current research on digital twins for security operations is extensive [19], while their potential for *incident response remains untapped* (*domain problem #1*).

Additionally, attackers are becoming increasingly aware of operational processes. APTs are on the rise [32] and organize their expertise to infiltrate, monitor, and disrupt IoT systems [72]. Detecting such process-aware attacks necessitates high situational awareness on the defenders' side. Research has already identified the digital twin's potential to raise situational awareness in cybersecurity [20, 76]. Indeed, hidden IoT devices by network scans and high manual effort in deriving operational processes complicate situational awareness and effective intrusion detection. It requires elaborations as currently *incident response lacks situational awareness* (*domain problem #2*).

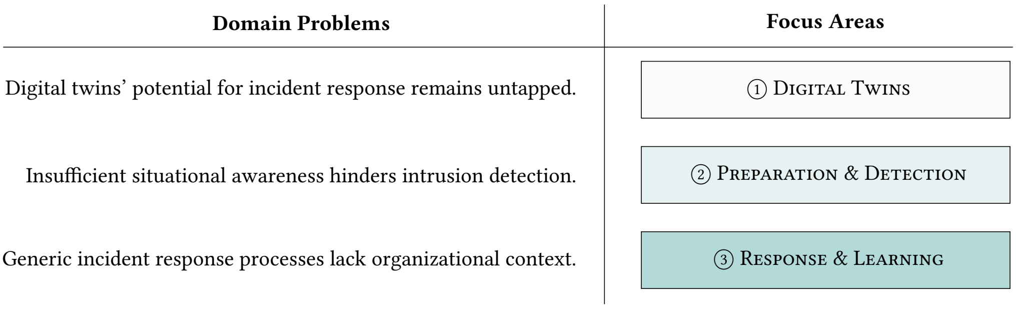
As situational awareness might indicate, IoT systems are organization-specific. For instance, each IoT system uses different devices with different hardware and software and runs different operational processes in different sectors with different legal requirements. In his 2014 work [60], Bruce Schneier articulated his vision for future incident response by integrating people, technologies, and processes. A decade later, this integration is evident in adopting playbooks, reinterpretations of traditional business processes, now retooled for cybersecurity. Playbooks indicate a trend to envision organization-specific incident response. Stevens et al. [63] identified that playbooks simplify incident response. Tough, it leaves open research and potentials in playbooks for IoT systems as currently *incident response processes are generic* (*domain problem #3*).

3 Research Questions

In light of stringent legal requirements, the ever-evolving threat landscape, and rapid technological advancements, incident response within critical infrastructures, particularly those incorporating IoT systems, demands a novel perspective. Leveraging digital twins enables organizations to enhance their IoT systems' readiness for incident response. This cumulative dissertation explores three domain problems of incident response for IoT systems, with the following overarching *Research Question (RQ)*:

Main Research Question.
How can organizations streamline incident response for IoT systems?

The main research question comprehensively addresses incident response in IoT systems. To thoroughly target the domain problems from the previous section, this thesis is structured around three key focus areas: ① Digital Twins, ② Preparation & Detection, and ③ Response & Learning. The organizational structure of this dissertation and the corresponding focus areas are shown in Figure 2.



Domain Problems	Focus Areas
Digital twins' potential for incident response remains untapped.	① Digital Twins
Insufficient situational awareness hinders intrusion detection.	② Preparation & Detection
Generic incident response processes lack organizational context.	③ Response & Learning

Figure 2: *Dissertation structure including domain problems and focus areas.*

In alignment with cybersecurity standards and frameworks, this dissertation follows the NIST Incident Response Lifecycle [12]. The ① Digital Twins area delves into the role of digital twins within broader security operations, including incident response. It serves as the technological backbone for all focus areas. The ② Preparation & Detection area addresses the lack of situational awareness in intrusion detection for hosts and networks with digital twins. It aligns with the first two phases of the NIST Incident Response Lifecycle, focusing on enhancing IoT asset visibility and investigating situational awareness in HIDS and NIDS. The ③ Response & Learning area examines the final two phases of the NIST Incident Response Lifecycle, discussing the intricacies of incident response processes, including organizational factors and the automation of organization-specific reactive and proactive playbooks with digital twins.

Consequently, the main research question is divided into three distinct research questions RQ1, RQ2, and RQ3, each corresponding to one of the three focus areas outlined above. This cumulative dissertation builds upon these research questions, organizing and presenting the findings through various scientific publications.

Research Question 1. Digital Twins
How can organizations use digital twins for security operations?

While the digital twin paradigm attracted the scientific community's interest decades ago, it was not until 2017 that the industry and research community began to take an interest in security operations [18, 19]. Yet, the full potential of its capabilities still needs to be explored. While the digital twin is commonly associated with intrusion detection, there are various other essential security operations it can be beneficial to, including educational training, risk assessment, and access control. A systematic literature review is crucial to effectively organize and assess the existing body of knowledge around the digital twin's use in incident response and other security operations. Such a review will identify prevailing research trends and clarify how organizations can use digital twins for security operations.

Research Question 2. Preparation & Detection
How can digital twins assist organizations in creating situational awareness?

Situational awareness is crucial in IoT systems for sensing the environment, understanding its implications, and projecting future scenarios [21]. As many IoT assets communicate via proprietary communication protocols, they remain hidden from TCP/IP network scans. Compiling a comprehensive asset inventory, including smaller and resource-constrained IoT devices, is essential for effective incident response. Additionally, complex systems impede asset-centric intrusion detection, requiring different approaches for intrusion detection on hosts. Approaching HIDS for IoT systems, especially systems-of-systems, is crucial. Last, as attackers become increasingly process-aware, NIDS should incorporate IoT systems' processes. Using digital twins to map out organizational processes creates situational awareness in IDS for IoT systems.

Research Question 3. Response & Learning
How can digital twins automate organizational incident response processes?

Organizations across various sectors differ significantly in their people, processes, and technologies [60]. These differences heavily influence their incident response processes. For instance, different legal requirements imply different reporting obligations. With the advent of SOAR platforms, incident response management is entering a new era, fortified by cybersecurity playbooks. Investigating the factors that shape an incident response playbook is a prerequisite and crucial for considering organizational specifics in playbook design. This includes examining reactive playbooks designed to counteract incidents and proactive playbooks aimed at managing the vast number of vulnerabilities in critical infrastructures and IoT systems. Understanding these influencing factors, and aligning those playbooks with digital twins, allows the automation of highly organization-specific incident response processes and informed learning.

4 Methodology

In order to ensure consistent, transparent, and rigorous results, this dissertation follows a well-defined set of methods. It was undertaken at the *Chair of Information Systems* at the University of Regensburg, aligning with the principles of information systems research – abbreviated as IS and known as *Wirtschaftsinformatik* in Germany. Information systems is a research domain that leverages IT to address domain problems.

4.1 Information Systems Research

At its core, information systems research is about the convergence of people, technologies, and processes. It is divided into two principal streams: Behavioral Science and *Design Science Research (DSR)* [30]. Behavioral science aims to understand and explain human behaviors and interactions with technologies, building theories to predict and explain phenomena [30]. DSR, in contrast, is more than just the design [8] and focuses on creating and evaluating IT artifacts, thereby solving practical problems through artifact design and assessment [52]. In DSR, design theory conceptualizes principles, patterns, and models guiding future artifact development (e.g., knowledge processes) [43]. In short, behavioral science aims to derive human behavior, DSR supports human behavior with artifacts, and design theory explores the principles guiding artifacts. As Hevner et al. [30] succinctly put it: *"Truth informs design and utility informs theory"*.

However, this dissertation primarily follows the DSR methodology (DSRM) as it is related to the IS domain and creates IT artifacts solving real-world domain problems around incident response for IoT systems. It is complemented by design theory and behavioral science when required. Most of the publications within this cumulative dissertation are about developing IT artifacts, especially prototypical implementations. Additionally, design theory was, for example, used to identify common knowledge about incident response processes in SOAR platforms and behavioral science to identify organizational influences on incident response processes. Individual publications correspond to the DSRM, while the structure of this dissertation and the corresponding focus areas – Digital Twins, Preparation & Detection, and Response & Learning – are about classifying and bringing them into a larger context.

4.2 Research Process

Peffers et al. [54] outline the six key steps of the DSRM as shown in Figure 3. Below, these steps are briefly described, highlighting their relevance in addressing domain problems related to incident response in IoT systems.

Step 1 – Problem Identification. As APT groups become highly organized, legal authorities are prompting critical infrastructure operators to implement stringent incident response measures. These measures are designed to secure the convergence of IT and OT systems (= IoT systems).

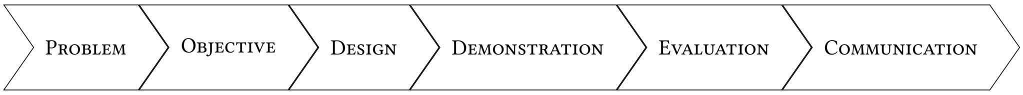


Figure 3: *DSRM according to Peffers et al. [54].*

Step 2 – Objective Definition. The primary objective of this dissertation is to manage incident response for critical infrastructures, focusing on IoT systems. To address this, RQ1 investigates the use of digital twins for incident response and other security operations. Additionally, RQ2 centers around creating situational awareness in HIDS and NIDS for IoT systems, and RQ3 concerns bringing organizational context into incident response processes for automating reactive and proactive playbooks.

Step 3 – Design & Development. In this step, artifacts are developed based on the objectives outlined in Step 2. Design and development are guided by systematic and scientific approaches such as literature reviews [77], data mining techniques (e.g., CRISP-DM [78]), knowledge processes (e.g., network management [6], or sense-making loop [59]), and thematic analysis [11]. Artifacts are implemented using software development patterns (e.g., "backend-for-frontend" and microservices) and best practices.

Step 4 – Demonstration. The artifacts developed address relevant domain problems, demonstrating that they provide a more effective and/or efficient solution than existing approaches. This leads to the next step, evaluation.

Step 5 – Evaluation. The evaluation step assesses how well the artifacts meet the objectives defined in Step 2. This cumulative dissertation employs a range of evaluation methods that are consistent with DSR practices [30], including:

- *Observable* methods like user studies ($N = 147$) and expert interviews ($N = 9$) are conducted, allowing for industrial insights.
- *Analytical* methods such as collecting and analyzing over 1,200 playbooks and manually assessing their contents foster generalization and rigor.
- *Experimental* methods like implementing a multi-tenant MQTT simulation with over 50 devices showcase the applicability of artifacts.
- *Descriptive* methods using industrial case studies (e.g., project partners) validate the artifacts' use in real-world settings and research projects.

Step 6: Communication. The final step is to share the research findings with the appropriate audiences. This is accomplished through (open access) academic publications and presentations at scientific conferences, ensuring the research is accessible to academic and industry stakeholders. Additionally, when artifacts are implemented, the source code and software are always publicly available (e.g., GitHub or DockerHub) to encourage transparency and further research as well as development in this field.

4.3 Research Setting

Since this dissertation is associated with the Chair of Information Systems at the University of Regensburg, it is deeply rooted in both previous dissertations and research projects. While previous dissertations inspire this one, research projects foster insightful discussions and knowledge transfers.

Research Environment. Digital twins in cybersecurity have already been explored in a previous dissertation at the Chair of Information Systems, University of Regensburg. The dissertation "*A two-fold Perspective on Enterprise Security in the Digital Twin Context*" [17] by Dr. Marietheres Dietz deals with the digital twin as an enabler and a problem. On the one hand, a digital twin can be used for security operations (e.g., intrusion detection), as in this dissertation. On the other hand, digital twins can also open up new gateways into industrial IoT systems. However, Dr. Marietheres Dietz was among the first researchers to study the convergence of digital twins and cybersecurity, which undoubtedly influences this dissertation. In contrast, this dissertation focuses on a single domain, examining the role of digital twins in incident response. It stands out in particular by i) systematically analyzing the state of the art of digital twins in security operations, ii) addressing situational awareness in HIDS and NIDS, and iii) investigating incident response processes and playbooks.

SISSeC Project. Since the beginning of the dissertation, the Chair of Information Systems has been involved in a research project called SISSeC – Secure Industrial Semantic Sensor Cloud – funded by the Federal Ministry of Economics and Climate Protection (No.: 16KN085725). This research project accompanied the development of this dissertation until the end of 2022. The research project involved collecting industrial sensor, actuator, and machine data from a PCB manufacturer's production plant, transferring it to a cloud environment, where it is stored and made available. The objective was to identify critical parameters for predictive maintenance that can lead to a general optimization of the machine processes and, thus, to reduce waste. The Chair of Information Systems's role was to create secure data processing, storage, and analysis pipelines. To this end, a digital twin was developed that semantically represents a drilling and milling machine, which was continuously fed with real-time data. Intrusion detection was carried out using the digital twin model. Productive project meetings and shared operational data contributed significantly to the outcomes of this dissertation.

INSIST Project. Various research projects started during this dissertation, including INSIST – Industrial IoT Security Operations Center – funded by the Bavarian Ministry of Economic Affairs, Regional Development, and Energy (No.: DIK0338/01). INSIST uses digital twins within an industrial SOC. In recent years, there has been close collaboration and knowledge transfer with the colleagues involved at the Chair of Information Systems and the Professorship of IoT-based Information Systems.

5 Results

With a strong IS research influence and the domain problems involved, this dissertation explores three main focus areas to address the research questions defined in Section 3: Digital Twins, Preparation & Detection, and Response & Learning. Each focus area is mapped to a domain problem, often requiring independent investigations of subproblems. The following sections provide a brief overview of the publications and a detailed summary of the primary findings for each focus area. Aside from the domain problem, contribution, and methodology, each section outlines several findings and implications that arose while solving them.

5.1 Overview of Research Papers

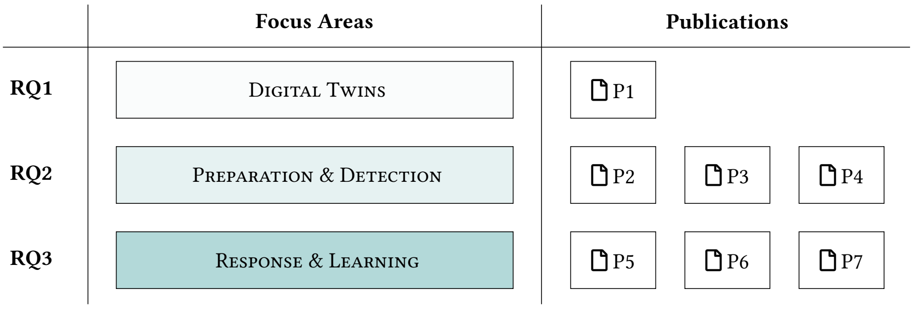
This cumulative dissertation has led to seven publications, either already published or forthcoming, in scientific conferences and journals. Each publication addresses a specific domain (sub-)problem, contributing to incident response for IoT systems. Table 2 provides an overview of these publications, listing the authors, titles, and venues, along with their respective rankings. Journals are ranked by their *Impact Factor (IF)*, while conferences are classified according to the CORE rankings[2] if available.

Table 2: *Publications overview with bibliometric data and corresponding rankings.*

No.	Publication	Ranking
P1	Empl, P., Koch, D., Dietz, M., & Pernul, G. (2024). Digital Twins in Security Operations: State of the Art and Future Perspectives.	–
P2	Empl, P., Hager, H., & Pernul, G. (2023). Digital Twins for IoT Security Management. In *Proceedings of the XXXVII Data and Applications Security and Privacy* (pp. 141–149).	B core
P3	Empl, P., & Pernul, G. (2023). Digital-Twin-Based Security Analytics for the Internet of Things. *Information 14*(2), Article 95.	3.1 if
P4	Empl, P., Böhm, F., & Pernul, G. (2024). Process-Aware Intrusion Detection in MQTT Networks. In *Proceedings of the 14th ACM Conference on Data and Application Security and Privacy* (pp. 91–102).	–
P5	Schlette, D., Empl, P., Caselli, M., Schreck, T., & Pernul, G. (2024). Do You Play It by the Books? A Study on Incident Response Playbooks and Influencing Factors. In *Proceedings of the 45th IEEE Symposium on Security and Privacy* (pp. 59:1–59:19).	A* core
P6	Empl, P., Schlette, D., Zupfer, D., & Pernul, G. (2022). SOAR4IoT: Securing IoT Assets with Digital Twins. In *Proceedings of the 17th International Conference on Availability, Reliability and Security* (pp. 4:1-4:10). **Best Paper Runner-Up Award**.	B core
P7	Empl, P., Schlette, D., Stöger, L., & Pernul, G. (2024). Generating ICS Vulnerability Playbooks with Open Standards. *International Journal of Information Security 23*(2), 1215-1230.	3.2 if

[2] https://portal.core.edu.au/conf-ranks

Figure 4 shows the respective RQs, the corresponding focus areas, and the assigned publications targeting specific domain problems within the focus areas. While the focus area Digital Twins is addressed with one publication, the remaining two focus areas are targeted via three publications, respectively. Note that all publications have been published except for *Publication P1*, which is currently under review. Each publication is shortly described and brought into the larger context in the following.



	Focus Areas	**Publications**		
RQ1	Digital Twins	P1		
RQ2	Preparation & Detection	P2	P3	P4
RQ3	Response & Learning	P5	P6	P7

Figure 4: *Focus areas and related publications solving domain problems.*

In the focus area of Digital Twins, *Publication P1* provides the state of the art of digital twins in corporate security operations by conducting a systematic literature review. It provides insights into the digital twin paradigm and organizational use. In the Preparation & Detection focus area, *Publication P2* enables intrusion detection and incident response for resource-constrained and smaller IoT devices, increasing their visibility. Building on this, *Publication P3* explores digital twins and HIDS, analyzing different security analytics techniques with various IoT data sources. Moreover, *Publication P4* delves into NIDS, with a process-aware intrusion detection within an MQTT network, combating APT groups. Transitioning to the Response & Learning focus area, *Publication P5* demystifies incident response playbooks, providing insights for creating, sharing, and using them in IoT systems while highlighting their intricacies and organizational factors in design. Taking this further, *Publication P6* explores the automation of reactive incident response playbooks in IoT systems, demonstrating how they can detect Sybil node attacks. Complementing this, *Publication P7* examines the automation of proactive playbooks, detailing effective strategies to automatically mitigate vulnerabilities in IoT systems by considering security advisories. In the following sections, the results of the individual focus areas are presented in more detail.

5.2 Focus Area 1: Digital Twins

The first focus area, Digital Twins, is about digital twins and their use in security operations. Research on this focus area attempts to address the domain problem that the digital twins' potential for incident response remains partially untapped. Thereby, *Publication P1* targets this domain problem and answers RQ1 ("*How can organizations use digital twins for security operations?*"). It assesses the state of the art of the digital twins in security operations by following a systematic literature review process.

Publication P1: Digital Twins in Security Operations: State of the Art and Future Perspectives

Domain Problem. Digital twins are increasingly gaining attention in industry and research. Despite their growing adoption, summarizing the convergence of digital twins and security operations remains a relatively unexplored area. Past surveys have primarily focused on specific security operations or technologies like augmented reality [9, 55]. *Publication P1* addresses this gap and systematically examines the state-of-the-art. Based on 201 research papers, it answers the following guiding questions:

- *What constitutes a digital twin within the realm of cybersecurity?*
- *How is the digital twin harnessed to fortify cybersecurity?*
- *How can digital twin components be instantiated?*
- *What pivotal challenges remain unaddressed, and what promising avenues beckon for future research?*

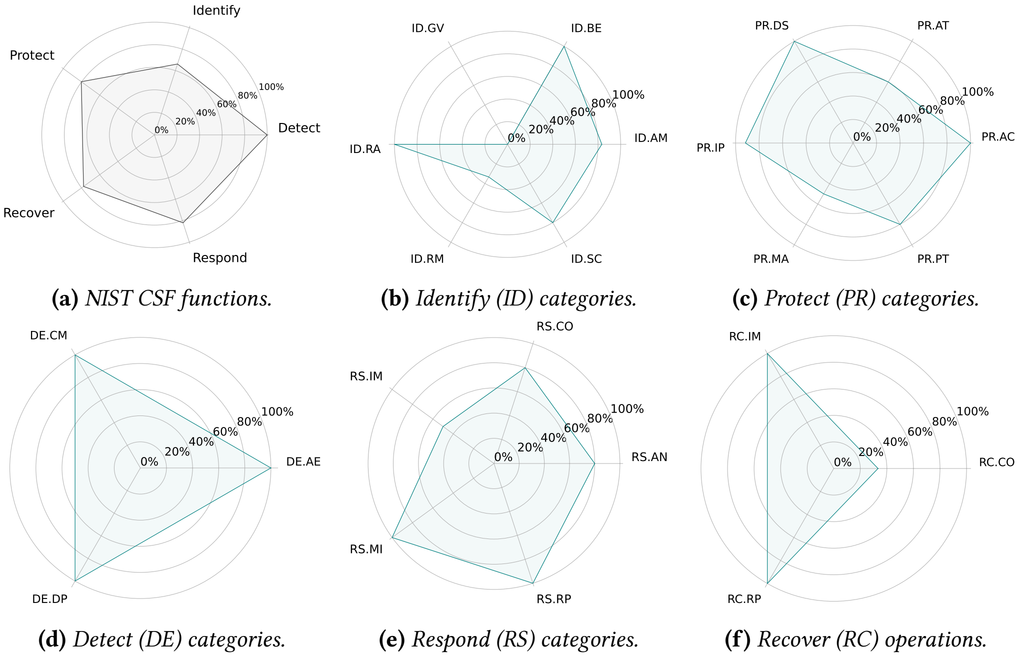
Contribution. *Publication P1* summarizes the body of knowledge and offers the first comprehensive overview of digital twins in security operations. This publication draws insights from 201 research papers to explore digital twins, their applications, and technologies. By going beyond those insights, this paper identifies challenges and outlines a research agenda, providing a guide for future research.

Methodology. This publication is based on the DSRM combined with a structured literature review process [77]. This involves querying academic databases with search terms to gather relevant literature. Those results are then screened and filtered according to predefined criteria applied to titles, abstracts, and full texts. Thematic analysis [11] is used to extract key insights from the selected papers. All data is publicly available to make decision-making and extraction processes transparent.

Finding 1: Digital twins are more than just simulations. Digital twins, a concept receiving various interpretations such as middleware, simulation, and virtual replicas, are more than just that. That is why they comprise multiple components and characteristics. A central element is the *virtual representation* that can be either structural, functional, or behavioral, offering a detailed capture of the real-world entity. *Data management* combines collecting, transforming, and processing data to create semantics for security operations. *User interaction* varies from reporting to virtual reality experiences. *Simulation capabilities* allow for advanced testing, forecasting, and prediction without impacting the physical systems. *Communication and synchronization* provide continuous and seamless data flows between digital twins and their real-world counterparts. Lastly, *security operations* adhere to the functions of the NIST CSF. These components collectively describe the digital twin in organizational security operations.

What is a digital twin in security operations?
A digital twin in the realm of cybersecurity is a dynamic *virtual representation* of an entity that seamlessly integrates *real-time and historical data*, covering the entire *lifecycle* of its counterpart. Thereby, it can *simulate* and replicate its behavior and states. A digital twin engages users in the *interactions* and *security operations*.

Finding 2: Digital twins are mostly used in intrusion detection operations. Digital twins offer several benefits for security operations. Their *asset centrality* allows a one-to-one mapping with physical entities and enables comprehensive data collection and management. Their *high fidelity* ensures advanced analytics, machine learning, and continuous monitoring, providing deep insights. *Virtual decoupling* ensures that digital twins operate without affecting physical systems, allowing resource-intensive tasks and parallel simulations without risk. *Physical/virtual intertwining* allows seamless communication, facilitating a defense-in-depth strategy and direct interaction with real-world assets. *Throughout the asset lifecycle*, digital twins support cybersecurity, from design to production and service.



(a) *NIST CSF functions.* **(b)** *Identify (ID) categories.* **(c)** *Protect (PR) categories.*

(d) *Detect (DE) categories.* **(e)** *Respond (RS) categories.* **(f)** *Recover (RC) operations.*

Figure 5: *Detailed security operations of digital twins in relation to the NIST CSF for critical infrastructures if they have been addressed at least once.*

Digital twins are crucial in enhancing security operations across critical infrastructure sectors. By mapping literature on digital twins in security operations, we can understand their application in this context according to the NIST CSF. Figure 5 shows the findings, revealing that intrusion detection is well-researched and saturated with a significant body of literature. However, the Identify, Respond, and Recover functions are yet to be explored, indicating gaps where more research is needed.

Finding 3: Digital twins are either created by data or domain experts. A digital twin is built out of several technologies (e.g., NoSQL) and protocols (e.g., MQTT). At its core, models provide semantics. Figure 6 shows different digital twin models.

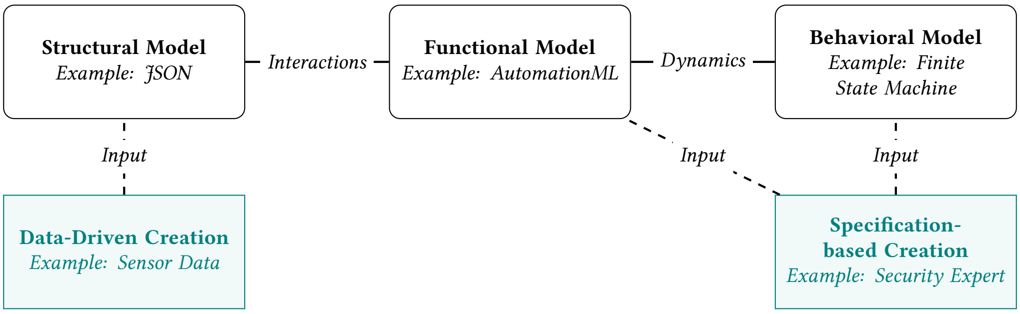


Figure 6: *Identified digital twin models, along with their creation.*

Structural models focus on the physical aspects of a system, providing a blueprint for the digital twin, often in formats like JSON or XML. *Functional models* delve into how the system operates, detailing interactions and processes, commonly using formats like AutomationML. *Behavioral models*, on the other hand, capture the dynamics and state changes over time, frequently using finite state machines or Petri nets to simulate system behavior. These models are created through either data-driven or specification-based approaches. *Data-driven creation* uses real-time or historical data to reflect the system's current state, offering continuity and constant updates. In contrast, *specification-based creation* involves predefined functional or behavioral specifications, often using expert input to reflect an asset's behavior accurately.

Result 4: Future research should target digital twins and incident response. Digital twins face challenges in standardization, data quality, security, and interoperability, necessitating ongoing research and cross-industry collaboration. In addition to technical challenges like communication delays, model attacks, and vulnerabilities, research should explore explainable digital twin models to maintain transparency in IoT systems. Future research should also focus on open standards and investigate less-explored security operations like incident response and risk management.

5.3 Focus Area 2: Preparation & Detection

Focus area 2 is about incident response preparation and incident detection. Research on this focus area attempts to address the domain problem of insufficient situational awareness. Three publications target this domain problem and answer RQ2 ("*How can digital twins assist organizations in creating situational awareness?*"). *Publication P2* contributes to situational awareness for IDS by integrating smaller IoT devices in the incident response process. *Publication P3* creates situational awareness for HIDS by aligning digital twins and intrusion detection, transitioning from IoT data to actionable insights. Last, *Publication P4* generates situational awareness for NIDS by automatically deriving operational processes from IoT networks to be used for intrusion detection.

Publication P2: Digital Twins for IoT Security Management

Domain Problem. IoT devices' diversity and resource constraints, combined with their insecure configurations and communication, make traditional security measures inadequate, e.g., devices in proprietary networks may be hidden from TCP/IP network scans [1]. Digital twins can emerge as a solution to these challenges. By creating digital twins of IoT networks, security analysts can monitor and analyze IoT networks proactively, detecting and addressing threats without disrupting physical systems. *Publication 2* explores how digital twins can create situational awareness and how they manage and secure complex IoT networks. It answers the following guiding question:

- *How can digital twins enable IoT security management?*

Contribution. *Publication 2* opens the scientific discussion on digital twins and IoT network management. It introduces a novel concept that uses digital twins for proactive IoT security management. Additionally, it offers a proof of concept that showcases the practical application of this concept across four distinct security use cases.

Methodology. This publication adheres to the principles of the DSRM and incorporates a domain-specific knowledge process. The resulting artifact is built according to the network management cycle [6], maintains an open-source approach, and is evaluated descriptively across four distinct use cases.

Finding 1: Incident response in IoT systems involves two distinct device classes. The proposed artifact employs digital twins to manage IoT networks, addressing the complexities posed by the heterogeneity of IoT devices (see Figure 7). It comprises three main components: the real world, virtual representation, and network management.

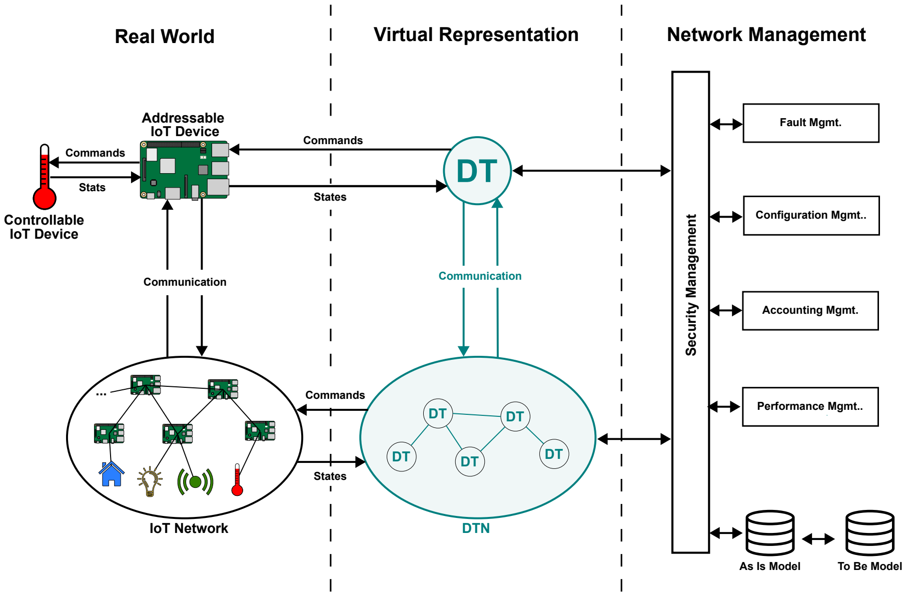


Figure 7: *Managing and securing IoT networks with digital twins.*

The *real world* encompasses a variety of IoT devices, ranging from actuators to sensors. These devices fall into two primary categories: addressable and controllable. Addressable devices function as coordinators and possess IP addresses, while controllable devices operate within their specific subnets. Both types are represented and modeled by digital twins. The digital twin's *virtual representation* accurately models the real-world IoT network, holding devices and communication links. This allows data collection as well as the management and configuration of IoT devices through the bidirectional communication relationship. For example, within *network management*, digital twins can be used to monitor and improve the security of IoT networks, leveraging the FCAPS framework (Fault, Configuration, Accounting, Performance, Security). The digital twin continuously oversees the IoT network, comparing the existing structural model (the As-Is state) with an optimal (To-Be) state. Consequently, the digital twin provides security recommendations for the IoT network. These recommendations are then forwarded to security analysts, who have the ultimate authority on execution.

Finding 2: It requires remote access to the coordinator to manage IoT subnets. The concept is validated through a prototypical artifact evaluated against four use cases: open ports, malicious USB dongles, weak access controls, and insecure configurations. We construct an experimental setup comprising a Raspberry Pi 3B+ model, a CC2531 ZigBee dongle, and a MacOS machine serving as the server (see Figure 8).

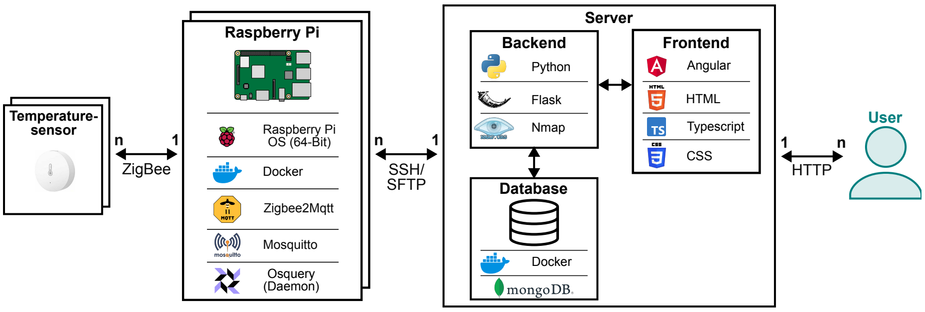


Figure 8: *Experimental setting and technologies to test the artifact.*

The artifact under consideration is a prototypical user interface developed using Angular, complemented by a Python-based backend. It integrates a structural digital twin model stored in a document-oriented database, meticulously capturing all details of the IoT networks. We employ nmap to scan the IP network to identify addressable IoT devices. Additionally, acquiring information about the subnets necessitates remote access via SSH or FTP to retrieve configuration files via osquery (e.g., MQTT brokers). The intelligence within this proof of concept is rule-based, relying on specific thresholds, such as the presence of more than five open ports. Based on the data collected and subsequent analysis, the prototype successfully suggests targeted recommendations to enhance the security of the IoT networks, such as closing unnecessary ports. Moreover, we measure the high performance of each use case, focusing mainly on the duration required to collect and analyze data and to suggest recommendations.

Publication 3: Digital-Twin-Based Security Analytics for the Internet of Things

Domain Problem. Nowadays, critical infrastructures are complex and composed of heterogeneous IoT components, collectively forming systems-of-systems. These systems feature long product lifecycles and operate within opaque supply chains. Consequently, organizations must understand their IoT systems (situational awareness) to manage cybersecurity effectively. Security analytics, simply big data analytics in cybersecurity, is a data-driven approach that extends beyond intrusion detection and synthesizes historical and real-time data to generate knowledge [42]. As digital twins hold essential semantics, they can assist in converting data to actions. However, digital twins and security analytics have yet to be explored. *Publication 3* delves into the role of digital twins in security analytics, focusing on generating and sharing actionable knowledge across the supply chain. The following question drives it:

- *How can organizations align security analytics and digital twins?*

Contribution. *Publication 3* conceptually explores how security knowledge is generated and shared within security analytics, leveraging digital twins to transition from raw data to actionable insights about incidents. By drawing on a framework, an artifact has been developed within the research project SISSeC. This artifact, named TWINSIGHT, is open-source and employs digital twins to run detective analytics for HIDS.

Methodology. This publication embodies the DSRM, initially deriving requirements for digital twins, security analytics, and the generation and sharing of knowledge, which is finalized in an entity-relationship model. Subsequently, a formal model is designed to define security analytics operations with digital twins, detailing the necessary data and knowledge. Afterward, a framework is introduced and instantiated through the TWINSIGHT microservice artifact. TWINSIGHT is evaluated in an experimental setup focused on real-time intrusion detection with HIDS-based agents.

Finding 1: Security analytics centers around five distinct analytical operations. At the core of the framework are digital twins and security analytics. This publication defines security analytics as *a repertoire of capabilities and technologies to systematically process and analyze data, detecting threats and imminent incidents.* Raising awareness among organizations about the intricacies of security analytics is paramount. Organizations can select from five distinct operations to cultivate security knowledge, as depicted in Figure 9. *Descriptive analytics* identifies trends and patterns by summarizing historical data. *Diagnostic analytics* delves into data to uncover the causes of past events and incidents. *Detective analytics* actively monitors data in near real-time to detect malicious behaviors or signatures signaling potential incidents. *Predictive analytics* leverages semantic models to forecast future events and incidents. Lastly, *prescriptive analytics* not only predicts but also offers recommendations on responding to incidents.

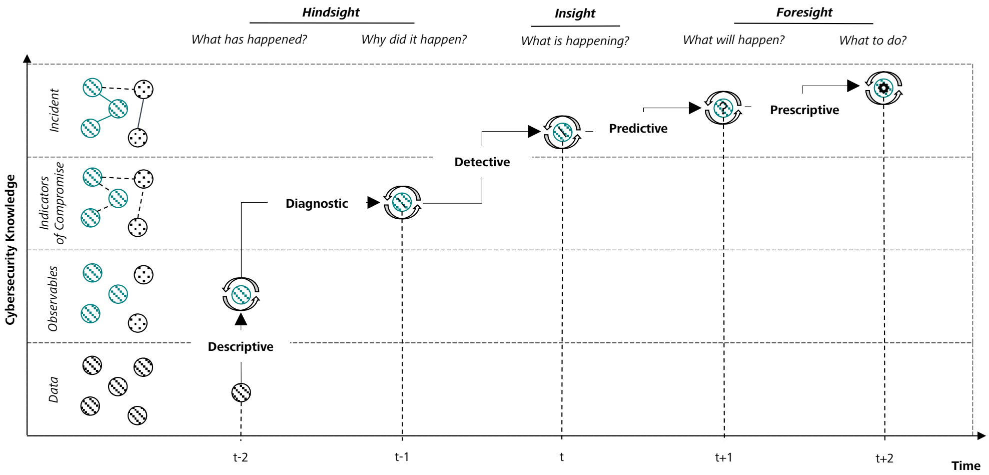


Figure 9: *Five distinct operations and their contribution to knowledge generation.*

Finding 2: Agent-based HIDS should follow asset centrality in IoT systems. We conduct a proof-of-concept, employing the Twinsight artifact within the SISSeC research project. SISSeC's objective is to securely integrate machine and sensor data from a PCB manufacturer with the cloud for real-time intrusion detection. We instantiate this use case with our framework: detective security analytics centering around a structural digital twin model. Consequently, Twinsight employs a set of different microservices, as depicted in Figure 10.

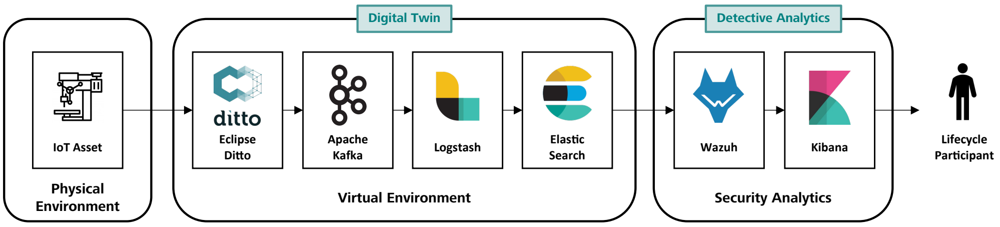


Figure 10: *Implemented Twinsight microservice architecture.*

The *physical environment* is managed by a Raspberry Pi 3B+ model (running Raspbian GNU/Linux 11), equipped with an MQTT client and a Wazuh agent that functions like a HIDS. The *virtual environment* gathers real-time data to populate the structural digital twin model within the Eclipse Ditto framework. This data then gets transformed and stored in a NoSQL database after passing through a Kafka messaging broker. For *detective analytics*, the data is sent to the Wazuh manager, which acts like a SIEM tool. The data is available to users via a dashboard. However, our results reveal significant limitations in the Wazuh agents, which rely on software to collect asset data. These agents view the system as a whole, overlooking the complex relationships between individual components. Detective analytics could model and incorporate their relationships if each component possessed its own agent. However, grouping and modeling Wazuh agents would enhance the effectiveness of intrusion detection in complex IoT systems.

Publication 4: Process-Aware Intrusion Detection in MQTT Networks

Domain Problem. IDS are essential for monitoring the industrial IoT, especially as organizations increasingly integrate IoT devices into their OT systems. Traditional IDS approaches, including NIDS, often fail to protect against sophisticated APTs, as they do not go beyond the transport layer data in the ISO/OSI model [29]. As APTs become increasingly process-aware, NIDS should consider IoT messaging protocols such as MQTT, CoAP, and OPC-UA. It should integrate contextual information from the application layer of the ISO/OSI model, such as MQTT topics, and leverage operational processes. *Publication P4* employs distributed tracing and process mining to extract operational processes from the application layer of IoT messaging protocols, enhancing NIDS. The following question guides this publication:

- *How can distributed tracing be utilized to automatically mine IoT processes for NIDS?*

Contribution. *Publication P4* introduces Mission, an explainable and distributed framework for NIDS in IoT systems. More than just a framework, the Mission on MQTT artifact is open source and accessible to both the scientific and industrial communities through GitHub and DockerHub. Integrating process mining and distributed tracing techniques enhances IoT process awareness in NIDS.

Methodology. To develop Mission, *Publication P4* employs the DSRM alongside the widely recognized data mining standard CRISP-DM [78]. Initially, the paper categorizes various NIDS techniques and establishes clear objectives. Subsequently, it delves into an in-depth analysis of the MQTT protocol, interprets relevant MQTT data, and models IPFIX flows from MQTT network traffic packets. The Mission framework is shown in Figure 11, detailing the mining of processes from IoT networks.

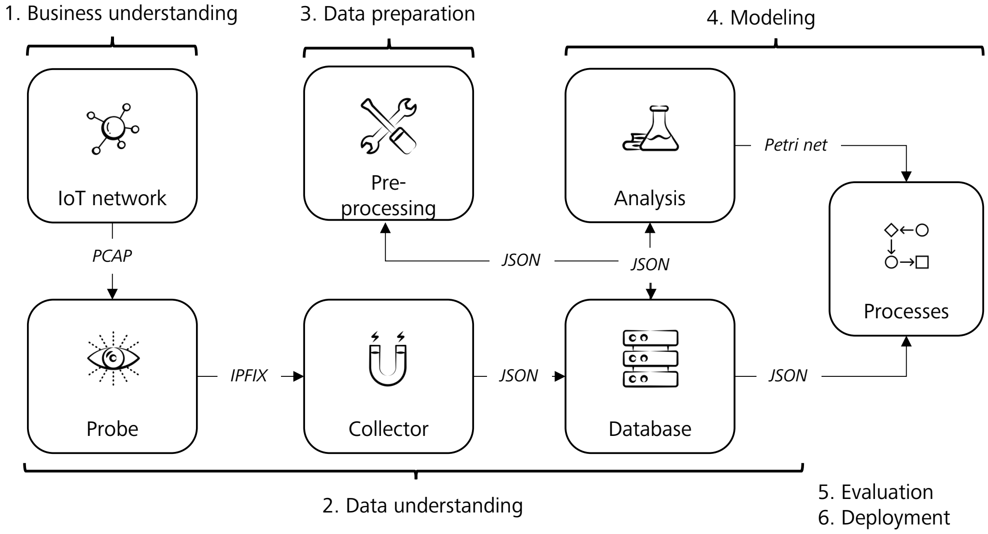
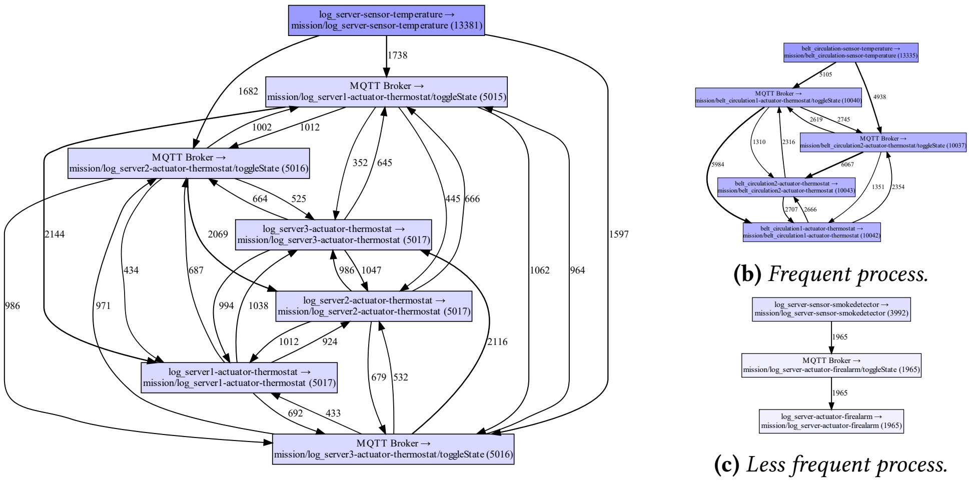


Figure 11: *MISSION framework based on CRISP-DM [78].*

Finding 1: IoT network flows should reflect activities instead of sessions. While the Mission framework is agnostic and versatile enough to adapt to any IoT messaging protocol, our initial implementation concentrates on the MQTT protocol. We begin by setting up the framework to thoroughly explore the features of MQTT 5.0. Using the protocol's control packet types, such as Connect, Publish, and Subscribe, we establish IPFIX flows. IPFIX, a standard widely recognized in network monitoring, traditionally constructs flows from TCP sessions (for example, a client maintaining a session with a web server). However, adhering strictly to TCP sessions in IoT systems risks losing valuable application layer data. To capture as much information as possible for the NIDS, we redefine a flow based on activities identified by packet control types – such as a sensor publishing a value to a specific topic. This approach enables a more nuanced capture of the IoT network, thereby enhancing situational awareness.

Finding 2: IoT networks must hand over trace identifiers to derive processes. After collecting IPFIX flows, we extensively evaluate our artifact against the outputs' quality, performance, and Mission's efficacy in detecting two attacks on MQTT networks: topic flooding and unauthorized subscriptions. We learn that IoT networks should incorporate trace identifiers (ID) to identify traces across IoT networks uniquely. For example, when devices publish data, they should attach unique IDs to be reused by other devices. Figure 12 shows the discovered processes (directly-follows graphs).


log_server-sensor-temperature →
mission/log_server-sensor-temperature (13381)
1738
1682
MQTT Broker →
mission/log_server1-actuator-thermostat/toggleState (5015)
1002
1012
MQTT Broker →
mission/log_server2-actuator-thermostat/toggleState (5016)
352
645
664
525
445
666
2144
2069
1597
log_server3-actuator-thermostat →
mission/log_server3-actuator-thermostat (5017)
434
687
986
1047
1062
964
986
971
994
1038
log_server2-actuator-thermostat →
mission/log_server2-actuator-thermostat (5017)
1012
924
2116
log_server1-actuator-thermostat →
mission/log_server1-actuator-thermostat (5017)
679
532
692
433
MQTT Broker →
mission/log_server3-actuator-thermostat/toggleState (5016)
belt_circulation-sensor-temperature →
mission/belt_circulation-sensor-temperature (13335)
5105
4938
MQTT Broker →
mission/belt_circulation1-actuator-thermostat/toggleState (10040)
2619
2745
MQTT Broker →
mission/belt_circulation2-actuator-thermostat/toggleState (10037)
1310
2316
5984
6067
belt_circulation2-actuator-thermostat →
mission/belt_circulation2-actuator-thermostat (10043)
1351
2354
2707
2666
belt_circulation1-actuator-thermostat →
mission/belt_circulation1-actuator-thermostat (10042)
log_server-sensor-smokedetector →
mission/log_server-sensor-smokedetector (3992)
1965
MQTT Broker →
mission/log_server-actuator-firealarm/toggleState (1965)
1965
log_server-actuator-firealarm →
mission/log_server-actuator-firealarm (1965)

(b) *Frequent process.*

(c) *Less frequent process.*

(a) *Complex process.*

Figure 12: *Discovery of three MQTT processes using the Mission framework.*

Mission is operated on a virtual machine equipped with Ubuntu 20.04.3 LTS (16GB RAM, eight cores, and 80GB storage) and consistently processes up to 200 packets per second. It empowers organizations to be aware of their IoT networks' behavior, making it a valuable input for NIDS. We successfully reuse the discovered processes for intrusion detection. Thereby, we analyze each flow, including preceding and subsequent activities as well as the devices involved. As Mission incorporates IoT processes in NIDS, it detects process-aware attacks that were previously unknown.

5.4 Focus Area 3: Response & Learning

Focus area 3 concerns the incident phases, including containment, recovery, and post-incident activity. Research on this focus area attempts to address the domain problem of needing more organizational context in generic incident response processes. Three publications target this domain problem and answer RQ3 ("*How can digital twins automate organizational incident response processes?*"). *Publication P5* contributes to an organizational context by empirically investigating playbooks' content and organizational influencing factors. *Publication P6* contributes to an organizational context in reactive playbooks by orchestrating, automating, and responding to incidents in IoT systems. Last, *Publication P7* incorporates organizational context in proactive playbooks by automatically remedying critical vulnerabilities in IoT systems.

Publication 5: Do You Play It by the Books? A Study on Incident Response Playbooks and Influencing Factors

Domain Problem. Organizations must establish and document standard operating procedures to ensure consistent security operations and incident response. Incident response playbooks offer a structured approach for managing security incidents, complementing proactive measures such as system hardening and vulnerability handling with reactive ones like blocking an IP address. Demand for these playbooks has surged as the threat landscape evolves due to their potential for consistency and automation. Nevertheless, challenges persist as empirical playbook research is limited (e.g., [63]). However, playbooks developed and shared by SOAR vendors and open-source communities offer organizations a valuable starting point and require customization for organizational use. *Publication P5* seeks to demystify incident response playbooks and organizational drivers to modify playbooks (*influencing factors*) by analyzing over 1,200 community playbooks and surveying over 140 participants about what influences their incident response processes. The following questions guide it:

- *What are characteristics of community playbooks made available by trusted sources?*
- *Which influencing factors shape incident response processes and organization-specific playbooks?*

Contribution. *Publication P5* delivers valuable insights into incident response playbooks, leveraging three data sources: community playbooks, online surveys, and expert interviews. This paper emphasizes the adaptation and use of playbooks within organizations. Based on the first data source, we closely examine community playbooks, providing a thorough analysis of their structure and content. The playbooks and related analyses are publicly available, inviting researchers to build upon and challenge the paper's findings. These findings are complemented by drawing on insights from the remaining data sources, survey participants, and security experts. Indeed, playbooks are influenced by internal and external factors in organizations.

Methodology. This publication uses a mixed-methods approach by combining quantitative and qualitative research to study incident response playbooks. It analyzes data from 1,217 selected playbooks (out of 1,347) sourced from 14 sources to address the first question, preprocessing them with Python and manually categorizing steps from 8,623 actuator-action-artifact triplets. An online survey of 147 participants and nine semi-structured interviews explore influencing factors to answer the second question, which guides in building a model that categorizes these factors into internal and external groups. Ethical considerations include anonymizing participant data.

Finding 1: Incident response playbooks are not exclusively tied to SOAR tools. In detail, this paper initially explores the structure and content of community playbooks, categorizing steps and identifying multi-category playbooks. Across vendors, playbooks have a significant structural consistency, distinguished by nuanced implementation details. However, playbooks typically include meta-information (e.g., name or description), workflows (process), parameters (e.g., inputs and outputs), and additional features like sharing or ownership. They are mainly represented by JSON, YAML, or XML formats, with PDF and BPMN formats as the exception. This shows that playbooks are not exclusively tied to SOAR tools; they are also used for documentation. However, the workflow details the incident response process at the playbook's core. A workflow contains typically automated workflow steps that encompass meta-information, logic, and actuator-action-artifact triplets. An example of an actuator-action-artifact triplet is a firewall (*actuator*) that blocks (*action*) an IP address (*artifact*). Figure 13 shows that the median number of steps per playbook is 10.

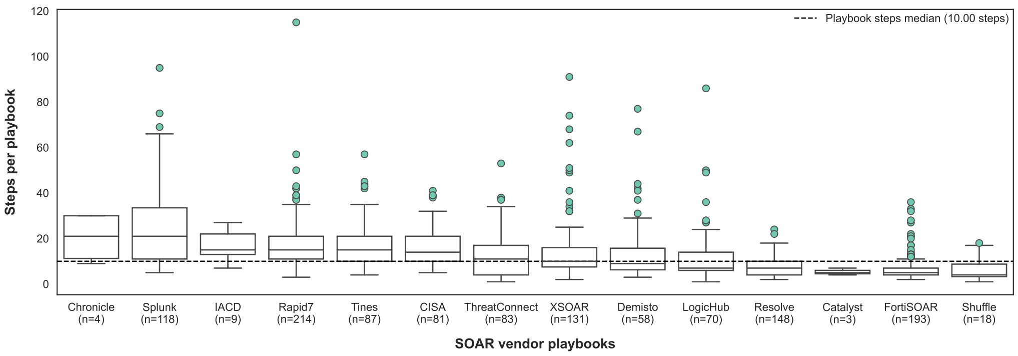


Figure 13: *Exploring 1,217 playbooks based on the number of steps.*

Categorizing workflow steps unveils five main categories: logic, utility, ticketing, investigation, and countermeasure. Logic steps organize the workflow (e.g., gateways), utility steps support operations (e.g., cleanse data), ticketing steps manage incidents (e.g., create ticket), investigation steps analyze threats (e.g., lookup IP), and countermeasure steps mitigate them (e.g., block URL). Besides, this paper investigates whether those community playbooks are built from multiple categories. Results indicate that 91.4% of those playbooks are built at least out of two categories, where investigation steps are prevalent, particularly in incident-specific playbooks (e.g., Ransomware).

Finding 2: Organizational playbooks are mainly influenced by internal factors. In the second finding, this paper draws from online survey results (with a sample size of $N = 147$) and insights gathered from nine security professionals. The survey participants, predominantly from the IT sector, disclose robust incident response capabilities and widespread utilization of playbooks, primarily for documentation and automation purposes. Notably, the playbooks exhibit variations in step count, with participants identifying multiple factors influencing their playbook design. These insights are further illuminated through detailed interviews, as summarized in Table 3.

Table 3: *Survey beliefs on influencing factor importance vs. interview findings.*

	Influencing Factors and Beliefs	Interview Findings
External	Attacker characteristics – *50%*	*Attacker behavior and motivation beat location.*
	Industry standards – *55.6%*	*Industry standards provide broad guidance only.*
	Laws and regulations – *52.1%*	*Laws and regulations apply throughout.*
	– Business structure	*Business requires compliance, but the influence is vague.*
	– Location	*Location-based influence is about privacy and data protection.*
	– Sector	*Sector-based influence has implications on security operations.*
	Supply chain – *32.4%*	*When in doubt, collaborate with business partners.*
Internal	Incident response directives – *69%*	*Directives address the finer points of incident response.*
	– Data operations	*Organizations with storage constraints perform data operations.*
	– Targets	*Organizations adapt response when targets are business critical.*
	People – *66.2%*	*More people imply more tasks and coordination.*
	– Security culture	*Security culture is about management support.*
	– Security team	*Multi-national organizations cover CTI sharing.*
	Technology – *78.9%*	*Organizations stand on the shoulders of technology.*
	– IT infrastructure	*Organizations strive for redundant response infrastructure.*
	– Security tools	*Technology drives communications and collaborations.*

Internal factors largely dominate, with 78.9% of participants citing technology as the primary influence on their playbooks, followed by incident response directives and people. Yet, organizations often rely on community playbooks initially, despite needing modifications to align with influencing factors. For instance, modification is required if the community playbook integrates Microsoft Teams for communication, but the organization uses Slack. However, organizations retain flexibility regarding the level of detail of their playbooks, opting for strict specificity with little leeway or leaving room for analyst discretion. Balancing creativity and specificity is essential to mitigate bias and maintain consistency in incident response.

Finding 3: Organizations should maintain playbooks and influencing factors. Playbooks serve beyond automation, enhancing documentation, reporting, and onboarding. Organizations must uphold playbooks and monitor influencing factors regardless of the use case. Whether modifications are prompted by external triggers (such as legal authorities adopting new laws) or established through regular reviews (like weekly sprints), maintaining playbooks is essential. There remains an unexplored potential in incident response playbook maintenance, where using playbook repositories and agile methods is highly recommended.

Publication 6: SOAR4IoT: Securing IoT Assets with Digital Twins

Domain Problem. Security analysts grapple with thousands of security alerts from various security tools, including many false positives [3]. Organizations are turning to SOAR platforms to manage this flood, which offer security tool integration and automated processes (playbooks) for more efficient incident response. With the proliferation of IoT devices, heterogeneous and resource-constrained IoT assets contribute to industrial digitization but currently overlook cybersecurity. The inadequate or absent security measures often resulting from "set-it-and-forget-it" manner lead to critical vulnerabilities. There is a need for SOAR platforms to extend their capabilities to include security operations for IoT assets. *Publication P6* explores this unexplored potential by addressing the following question:

- *How to use Security Orchestration, Automation and Response for the Internet of Things?*

Contribution. *Publication P6* contributes by shedding light on SOAR platform features and introduces the SOAR4IoT framework. This framework is designed to orchestrate IoT assets and automate incident response with playbooks. Additionally, this paper instantiates the framework with an open-source implementation using digital twins. It designs and implements two playbooks for orchestration and automated response by addressing two security threats of IoT assets.

Methodology. This publication follows the DSRM to develop an artifact that integrates IoT assets into SOAR platforms. Initially, it extensively researches existing SOAR platforms and common IoT security challenges. Subsequently, the paper adapts SOAR specifically for the IoT by proposing a new framework, SOAR4IoT, which centrally incorporates digital twins. The SOAR4IoT framework is open-source and designed to be extendable, promoting ongoing development and customization by both the research community and industry practitioners.

Finding 1: Orchestration collects IoT data, while incident response modifies it. The SOAR4IoT framework strengthens IoT incident response capabilities by incorporating SOAR and playbooks with digital twins. The framework is shown in Figure 14.

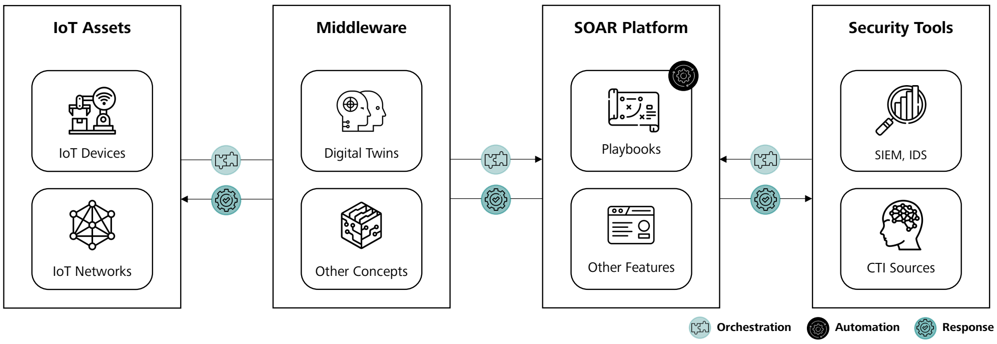


Figure 14: *SOAR4IoT framework with orchestration, automation, and response.*

The core elements of the SOAR4IoT framework include IoT assets, digital twins as middleware, the SOAR platform itself, and various security tools. *IoT assets* encompass all devices and networks within IoT systems, forming the primary focus for incident response efforts. *Digital twins*, serving as middleware, offer a virtual representation of these assets for precise management and control. The framework seamlessly integrates an array of *security tools* and intelligence, such as *Cyber Threat Intelligence (CTI)* and SIEM, enhancing the platform's security capabilities. At its core, the *SOAR platform* incorporates digital twins with these security tools to orchestrate operations, using automated playbooks to execute security operations and responses that bolster the IoT security posture. These playbooks minimize human error and automate routine tasks, pushing the boundaries of security hyper-automation.

Finding 2: Playbook execution depends on matching IoT assets and playbooks. The *SOAR4IoT* framework is successfully validated through a practical experimental setup by implementing a digital twin-based SOAR platform to secure IoT assets (devices and networks). The experiments involve two tailored reactive playbooks addressing two security threats on IoT devices and networks: the Mirai botnet and the Sybil attack. The experimental setup uses several technologies aligned to the four main components of the *SOAR4IoT* framework. It is shown in Figure 15.

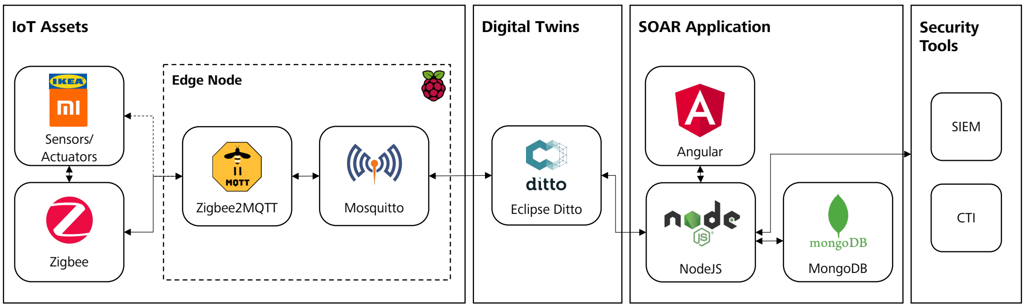


Figure 15: *SOAR4IoT experimental setting in IoT systems.*

The experimental setup uses IoT devices including Xiaomi Aqara temperature sensors and IKEA Tradfri LED bulb actuators, controlled via a Raspberry Pi 3B+ with 1GB RAM, and managed using Eclipse Ditto digital twins on a microservices architecture on a Ubuntu virtual machine (16GB RAM, eight kernels, and 80GB storage). The *User Interface (UI)* of the SOAR platform, developed using the Angular framework and TypeScript, provides a streamlined, minimalistic design that facilitates the easy configuration of playbooks and management of IoT assets. Suppose the SOAR platform captures an event using a security tool. In that case, it matches the event parameters with assets and playbook to choose the correct one. In the first case, the Mirai playbook preemptively reacts and updates the firmware on vulnerable devices to guard against the Mirai botnet. In contrast, the Sybil playbook reacts to fake identity creations by monitoring and removing suspicious nodes from the IoT network. Afterward, analysts can assess the playbook's success by checking its status (success, timeout, or failure).

Publication 7: Generating ICS Vulnerability Playbooks with Open Standards

Domain Problem. There are significant challenges organizations face in managing vulnerabilities within ICS and industrial IoT systems. Cyberattacks increasingly target these systems due to numerous vulnerabilities and the complexity of maintaining cybersecurity across diverse and interconnected IoT assets. Traditional methods of patching or updating to new software versions are often not feasible in industrial environments due to system availability requirements or device heterogeneity. Thus, there is a pressing need for more structured and automated approaches to vulnerability management, shifting focus to workarounds. *Publication P7* investigates proactive vulnerability playbooks based on open standards, offering a systematic and automated approach to handling vulnerabilities. The following question guides it:

- *Is it possible to generate ICS vulnerability playbooks?*

Contribution. *Publication P7* introduces a novel process model for generating ICS vulnerability playbooks. The key contribution is the process model of transforming security advisories into structured, actionable playbooks that automate and enhance vulnerability management. By integrating CSAF for advisory input and CACAO for playbook output, this paper results in a prototypical artifact that promotes a more efficient and consistent approach to vulnerability handling in industrial environments.

Methodology. This publication employs DSRM to develop an innovative process model artifact. The paper's methodology involves identifying the problem of inefficient vulnerability management in ICS, followed by iterative development, testing, and refinement based on a real-world use case at DEHN SE. Based on the model, the open-source prototype is designed to demonstrate practical applicability and playbook quality, fostering widespread adoption and customization by hinting at open challenges.

Finding 1: Standardizing security advisory exchanges would reduce efforts. The process model for generating ICS vulnerability playbooks streamlines vulnerability handling through a series of phases designed to automate the process (see Figure 16).

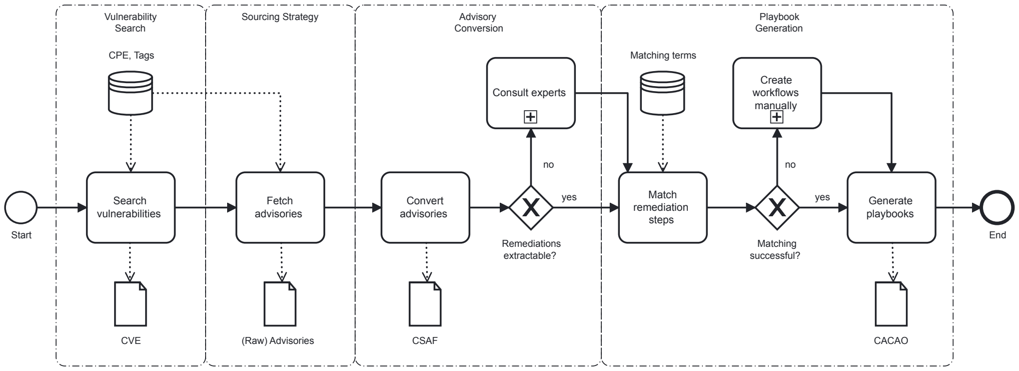
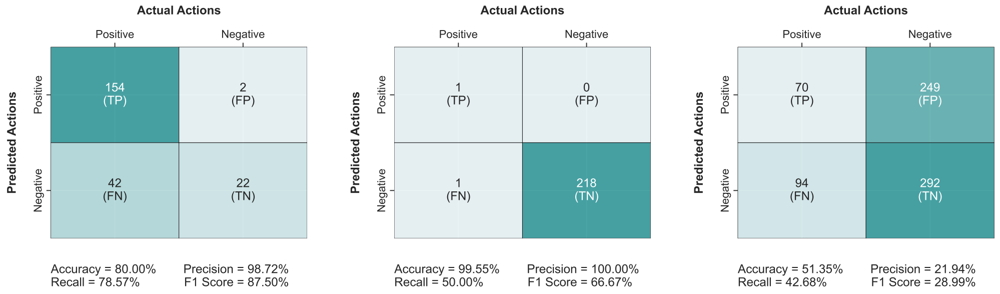


Figure 16: *Process model from security advisories to vulnerability playbooks.*

It begins with the *vulnerability search*, using structured asset information and CPE identifiers to identify at-risk components. Next, in the *sourcing strategy* phase, security advisories must be queried from different standardized (e.g., CVRF and CSAF) and unstructured (e.g., Tweets, RSS feeds, and email notifications) data sources. These advisories are then transformed in the *advisory conversion* phase using the CSAF standard to unify the data format, facilitating automated processing. The final phase, *playbook generation*, translates these advisories into actionable playbooks using the CACAO standard, thus guiding the automated response to vulnerabilities. Human involvement is crucial in overseeing and refining automated processes, particularly verifying data accuracy, configuring sources, and reviewing generated playbooks.

Finding 2: Structured asset representations support vulnerability handling. The process model is evaluated by implementing a prototype to demonstrate the feasibility of automated ICS vulnerability playbook generation. The evaluation is conducted on an experimental setup that features a dedicated virtual machine running Ubuntu 22.04 LTS, equipped with 8GB RAM and 80GB of storage. The system architecture includes a Vue.js front-end for user interactions and a Node.js back-end for processing logic, with MongoDB utilized for data storage. The application's functionality is tested using security advisories from two primary ICS vendors – Siemens ProductCERT and Cisco PSIRT– and from the CISA, which vary in format from structured (CSAF and CVRF) to unstructured (HTML). With two asset representations (structural digital twin models), our prototype automatically identifies vulnerabilities. It queries and transforms 79 security advisories in CSAF format across the three data sources, with 71 generated CACAO playbooks. Afterward, a manual assessment of the generated playbooks allows for quality insights (see Figure 17).



(a) *Siemens ProductCERT (n=220).* **(b)** *Cisco PSIRT (n=136)* **(c)** *CISA ICS CERT (n=705).*

Figure 17: *CACAO playbook quality (n equals the number of workflow actions).*

Results show the accuracy and completeness of the playbooks by analyzing the number of workflow actions and calculating rates of mistaken acceptance (*type I errors*) and mistaken rejection (*type II errors*). Results indicate that playbooks generated from structured advisories demonstrate higher precision and recall, confirming their greater efficacy in automation than those from unstructured ones. Lastly, the artifact processes five years of historical security advisories in under nine minutes.

5.5 Complementary Publications

In addition to the publications P1 to P7 in this cumulative dissertation, Table 4 lists complementary papers that contribute to IS research, IoT systems, and cybersecurity.

Table 4: *Overview of complementary publications.*

No.	Publication	Ranking
C1	Wagner, G., Empl, P., & Schryen, G. (2020). Designing a novel strategy for exploring literature corpora. In *Proceedings of the 28th European Conference on Information Systems* (pp. 44:1-44:17).	A core
C2	Empl, P., & Pernul, G. (2021). A Flexible Security Analytics Service for the Industrial IoT. In *Proceedings of the 2021 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems* (pp. 23–32).	–
C3	Putz, B., Dietz, M., Empl, P., & Pernul, G. (2021). EtherTwin: Blockchain-based Secure Digital Twin Information Management. In *Information Processing & Management, 58*(1), 102425.	8.6 if
C4	Hornsteiner, M., Empl, P., Bunghardt, T., & Schönig, S. (2024). Reading Between the Lines: Process Mining in OPC UA Network Data. *Sensors, 24*(14), 4497.	3.4 if

Publication C1. *Publication C1* explores a strategic approach for reviewing literature in IS. This strategy prioritizes reading influential papers first, underpinned by schemata theory, to enhance comprehension by linking new to known information. It develops a prototypical artifact, ENLIT, to automate reading sequences, validated through a case study in the IT business value domain. The findings highlight the strategy's potential to streamline literature reviews, mainly benefiting junior scholars.

Publication C2. *Publication C2* introduces a security analytics service for the industrial IoT tailored to *Small and Medium-sized Enterprises (SME).* This service integrates with known frameworks (e.g., RAMI 4.0 and IIRA), using security analytics to enhance IoT security. Designed for cost-efficiency and scalability, it overcomes financial and knowledge barriers for SMEs and is proven within the project SISSeC.

Publication C3. *Publication C3* presents EtherTwin, a blockchain-based solution for managing digital twin data within Industry 4.0. This approach utilizes distributed ledger technology to ensure data availability, integrity, and confidentiality, supporting operations across organizational boundaries. It promotes trustless data sharing with *Decentralized Application (DApps).* The framework is evaluated through expert interviews and an industry use case, demonstrating practical applicability and scalability.

Publication C4. *Publication C4* elaborates on a method for using IoT network data for process mining to improve operational efficiency. This innovative methodology captures and analyzes network traffic data, mainly focusing on the OPC UA protocol. Process mining algorithms and a rule-based event log generation technique enhance transparency and adaptability. By demonstrating it through a real-world industrial case study, this method effectively identifies and optimizes operational processes.

6 Conclusion and Future Work

The dissertation addresses cybersecurity standards and domain problems in IoT systems, particularly digital-twin-based incident response. Central to this dissertation is the research question: *"How can organizations streamline incident response for IoT systems?"*. Subsequently, it delves into three RQs stemming from this question. The resulting publications, ranging from P1 to P7, significantly answer these RQs with intersecting boundaries and contributions. As evidenced by the findings, Table 5 delineates the primary and secondary contributions of the publications within this cumulative dissertation. These publications contribute to the digital twins' organizational use and situational awareness in HIDS and NIDS, as well as in automating reactive and proactive organization-specific playbooks in IoT systems.

Table 5: *Primary (●) and secondary (○) contributions of publications.*

Publications	**RQ1** – Digital Twins	**RQ2** – Ids		**RQ3** – Playbooks	
	Organizational Use	*Host*	*Network*	*Reactive*	*Proactive*
Publication P1	●				
Publication P2	○	●	●		○
Publication P3	○	●			
Publication P4	○		●		
Publication P5				●	●
Publication P6	○			●	○
Publication P7	○				●

RQ1 and the focus area Digital Twins delve into the untapped potentials of digital twins in incident response. Publications address this domain problem by raising awareness of digital twins and their organizational use. In this context, *Publication P1* elaborate on the state of the art of digital twins' use, transcending their conventional role as mere simulations for IDS. It provides insightful guidance to organizations, including digital twin models and their creation. Furthermore, it anticipates future research directions, particularly emphasizing the use of digital twins in less-investigated research areas like incident response and risk assessment – whereas incident response is explored in this dissertation. Primarily through *Publication P1*, organizations can enhance their understanding of digital twins and gain actionable insights on leveraging them for their security operations in IoT systems.

RQ2 and the respective focus area Preparation & Detection address the escalating threats posed by APT groups and the pressing need for enhanced organizational situational awareness in IoT systems. Three key publications enrich this area, each making substantial contributions. *Publication P2* elevates situational awareness by preparing IoT systems for incident response, particularly emphasizing the identification of controllable and addressable IoT assets. *Publication P3* delves into the intricacies of HIDS, revealing the complexities within system-of-systems that hinder situational awareness. It advocates for more structured approaches to address these challenges, offering insights into five analytical operations. *Publication P4* unravels operational processes and introduces a process-aware NIDS tailored for MQTT networks. Focusing on activities rather than sessions and employing trace identifiers enhances situational awareness. Collectively, publications P2 to P4 empower organizational incident response capabilities in preparing their IoT systems, bolstering situational awareness, and effectively thwarting sophisticated APT groups.

RQ3 and the focus area Response & Learning dive into the lack of organization context in incident response processes by investigating playbooks. Three publications contribute to this area by investigating organization-specific influencing factors as well as reactive and proactive playbooks. *Publication P5* sheds light on playbooks' use, like automation, documentation, and onboarding, by investigating over 1,200 community playbooks. It further identifies internal and external influences along a playbook's lifecycle by conducting an online survey and expert interviews, with internal factors being prevalent, opening up broad playbook discussions. *Publication P6* investigates SOAR platforms for IoT systems, defining the tasks of orchestration, automation, and response. It showcases reactive playbook executions on IoT devices and networks, paving the way for more streamlined and structured incident response. Furthermore, *Publication P7* delves into the execution of proactive playbooks for vulnerability handling in high-availability IoT systems, advocating for standardized exchange to security advisories and device representations. With publications P5 to P7, organizations can leverage community playbooks tailored to their IoT systems, proactively mitigating the risks posed by open vulnerabilities and responding to incidents.

In essence, this dissertation advances incident response for IoT systems by leveraging digital twins and adhering to the NIST Incident Response Lifecycle. It addresses three pivotal focus areas, introducing various artifacts such as prototypical implementations and effectively demonstrating their efficiency and effectiveness in securing the critical infrastructures of tomorrow. Notably, the individual publications stemming from this cumulative dissertation have undergone rigorous peer review and have garnered significant recognition within the academic community, as evidenced by citations. Moreover, the research impact is documented by research projects (e.g., SISSeC and INSIST), which build upon the findings of this dissertation. Acknowledging the expansive scope of incident response in IoT systems, it is evident that further research is imperative. This dissertation serves as a foundational cornerstone for future research and discussions in incident response. With the escalating demand for compliance with legal requirements in IoT systems, technological advances, and the ever-evolving threat landscape, there are noteworthy future research areas in incident response and beyond, which are discussed in the following.

Organizations should not always use large language models to crack nuts. What started a few years ago with natural language processing (e.g., stop words, stemming, and part of speech tagging) of security information has taken a decisive turn with generative *Artificial Intelligence (AI)* especially with *Large Language Models (LLM).* Research and the industry are racing to improve and challenge models in every use case. At the same time, it sometimes leaves the impression that they use the sledgehammer to crack a nut. As with every technology, disillusionment sets in at some point. Of course, LLMs undoubtedly have their purpose and are superficial in understanding and generating human languages. However, to date, outputs are too imprecise and misleading in a security context. More promising are LLM chains and agents. LLM chains realize use cases by sequentially arranging LLMs, knowledge bases (cf. embedding and RAG), and interfaces (e.g., APIs or security tools). These chains are deterministic and reliably run in the same manner. More autonomous LLM agents personalize security roles (e.g., CISO), deciding which tool or knowledge base is helpful in a given context.

Cybersecurity hyper-automation is there for those familiar with data formats and standards. Not only because of playbooks but with more structured data formats available, the security research and industry face hyper-automation not restricted to incident response. Nowadays, we find not only data formats but also open data that are not limited to vulnerabilities (e.g., CVE) or advisories (e.g., CSAF). Nations, research, and the industry are eager to map data formats (e.g., CVE with MITRE ATT&CK) and security standards (e.g., CIS controls and NIST CSF) to foster automation and standard interdependencies. The data, formats, and standards are out there and ready to be used to envision security hyper-automation and crosswalks. It is interesting to follow automation trends and see if humans-in-the-loop will be replaced by agents-in-the-loop.

Legal requirements shift focus to IoT risk assessment. With current legal developments and research trends, risk assessment has a revival in OT and IoT systems after being a trending topic in IT decades ago. While the EU NIS2 Directive calls critical infrastructure operators to assess their security risks, the EU Cyber Resilience Act does the same for products (e.g., HMI software). Given parallel developments in security hyper-automation and LLM usage, it is evident that risk assessment will be automated, too. For instance, automatically creating risk assessment reports based on software repositories and IoT firmware analysis might be an exciting field of research.

IoT security made for everyone. In closing, the developments of incident response in IoT systems by governments, industries, and research are tremendous and represent a quantum leap in the right direction. The technological prerequisites are in place to secure our critical infrastructures of tomorrow. Of course, it depends on skilled professionals as critical infrastructures converge in engineering, information science, IT, and cybersecurity. However, the organizational component rests with the operators, but it is up to us to simplify IoT security further.

References

[1] AboubDakar, M., Kellil, M., and Roux, P. A review of iot network management: Current status and perspectives. *Journal of King Saud University Computer and Information Sciences 34*, 7 (2022), 4163–4176.

[2] Al-Fuqaha, A. I., Guizani, M., Mohammadi, M., Aledhari, M., and Ayyash, M. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Communications Surveys & Tutorials 17*, 4 (2015), 2347–2376.

[3] AlAhmadi, B. A., Axon, L., and Martinovic, I. 99% false positives: A qualitative study of SOC analysts' perspectives on security alarms. In *Proceedings of the 31st USENIX Security Symposium* (2022), K. R. B. Butler and K. Thomas, Eds., USENIX Association, pp. 2783–2800.

[4] Alcaraz, C., and Lopez, J. Digital Twin: A Comprehensive Survey of Security Threats. *IEEE Communications Surveys & Tutorials 24*, 3 (2022), 1475–1503.

[5] Allison, D., Smith, P., and Mclaughlin, K. Digital Twin-Enhanced Incident Response for Cyber-Physical Systems. In *Proceedings of the 18th International Conference on Availability, Reliability and Security* (2023), ACM, pp. 28:1–28:10.

[6] Arzo, S. T., Naiga, C., Granelli, F., Bassoli, R., Devetsikiotis, M., and Fitzek, F. H. P. A Theoretical Discussion and Survey of Network Automation for IoT: Challenges and Opportunity. *IEEE Internet of Things Journal 8*, 15 (2021), 12021–12045.

[7] Axelsson, S. Intrusion detection systems: A survey and taxonomy. White paper, Citeseer, 2000.

[8] Baskerville, R. What design science is not. *European Journal of Information Systems 17*, 5 (2008), 441–443.

[9] Böhm, F., Dietz, M., Preindl, T., and Pernul, G. Augmented reality and the digital twin: State-of-the-art and perspectives for cybersecurity. *Journal of Cybersecurity and Privacy 1*, 3 (2021), 519–538.

[10] Boschert, S., Heinrich, C., and Rosen, R. Next Generation Digital Twin. In *Proceedings of the 12th International Symposium on Tools and Methods of Competitive Engineering* (2018), TMCE, pp. 209–217.

[11] Braun, V., and Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology 3*, 2 (2006), 77–101.

[12] Cichonski, P., Millar, T., Grance, T., and Scarfone, K. Computer Security Incident Handling Guide. White Paper SP 800-61r2, National Institute of Standards and Technology, 2012.

[13] Claroty Ltd. The Global State of Industrial Cybersecurity 2023: New Technologies, Persistent Threats, and Maturing Defenses. Survey, 2023.

[14] Conklin, W. A. IT vs. OT Security: A Time to Consider a Change in CIA to Include Resilience. In *Proceedings of the 49th Hawaii International Conference on System Sciences* (2016), T. X. Bui and R. H. S. Jr., Eds., IEEE, pp. 2642–2647.

[15] Debar, H., Dacier, M., and Wespi, A. Towards a taxonomy of intrusion-detection systems. *Computer Networks 31*, 8 (1999), 805–822.

[16] Deutscher Bundestag. Gesetz zur Umsetzung der Richtlinie (EU) 2016/1148 des Europäischen Parlaments und des Rates vom 6. Juli 2016 über Maßnahmen zur Gewährleistung eines hohen gemeinsamen Sicherheitsniveaus von Netz- und Informationssystemen in der Union. Legislation, 2017.

[17] Dietz, M. *A Two-fold Perspective on Enterprise Security in the Digital Twin Context.* PhD thesis, University of Regensburg, 2022.

[18] Dietz, M., and Pernul, G. Unleashing the Digital Twin's Potential for ICS Security. *IEEE Security & Privacy 18*, 4 (2020), 20–27.

[19] Eckhart, M., and Ekelhart, A. Digital Twins for Cyber-Physical Systems Security: State of the Art and Outlook. In *Security and Quality in Cyber-Physical Systems Engineering*, S. Biffl, M. Eckhart, A. Lüder, and E. R. Weippl, Eds. Springer, 2019, pp. 383–412.

[20] Eckhart, M., Ekelhart, A., and Weippl, E. Enhancing Cyber Situational Awareness for Cyber-Physical Systems through Digital Twins. In *Proceedings of the 24th IEEE International Conference on Emerging Technologies and Factory Automation* (2019), IEEE, pp. 1222–1225.

[21] Endsley, M. R. Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society Annual Meeting 32*, 2 (1988), 97–101.

[22] European Commission. New eu cybersecurity strategy and new rules to make physical and digital critical entities more resilient. https://ec.europa.eu/commission/presscorner/detail/en/IP_20_2391, 2020. Accessed: June 27, 2024.

[23] European Commission. Proposal for a Regulation of the European Parliament and of the Council on horizontal cybersecurity requirements for products with

digital elements and amending Regulation (EU) 2019/1020. Legislation COM(2022) 454 final, 2022.

[24] European Commission. Measures to strengthen solidarity and capacities in the Union to detect, prepare for, and respond to cybersecurity threats and incidents. Legislation 2023/0109(COD), 2023.

[25] European Parliament and the Council of the European Union. 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. Legislation (EU) 2022/2555, 2022.

[26] Forum of Incident Response and Security Teams. CVSS – Common Vulnerability Scoring System. https://www.first.org/cvss, 2024. Accessed: June 27, 2024.

[27] Forum of Incident Response and Security Teams. EPSS – Exploit Prediction Scoring System. https://www.first.org/epss, 2024. Accessed: June 27, 2024.

[28] Frazier, G., and Stouffer, K. Digital Forensics and Incident Response (DFIR) Framework for Operational Technology (OT). White Paper NISTIR 8428, National Institute of Standards and Technology, 2022.

[29] Gemalto. Gemalto: State of IoT Security. *Network Security 2019*, 2 (2019), 4.

[30] Hevner, A. R., March, S. T., Park, J., and Ram, S. Design Science in Information Systems Research. *MIS Quarterly 28*, 1 (2004), 75–105.

[31] Husák, M., Sadlek, L., Špaček, S., Laštovička, M., Javorník, M., and Komárková, J. CRUSOE: A toolset for cyber situational awareness and decision support in incident handling. *Computers & Security 115* (2022), 102609.

[32] Ifigeneia Lella and Eleni Tsekmezoglou and Marianthi Theocharidou and Erika Magonara and Apostolos Malatras and Rossen Svetozarov Naydenov and Cosmin Ciobanu. ENISA Threat Landscape 2023. Survey, European Union Agency for Cybersecurity, 2023.

[33] International Electrotechnical Commission. Functional safety of electrical/electronic/programmable electronic safety-related systems. Standard IEC 61508-1:2010, 2010.

[34] International Electrotechnical Commission. Functional safety - Safety instrumented systems for the process industry sector. Standard IEC 61511:2016, 2016.

[35] International Organization for Standardization. Information technology – Security techniques – Information security management systems – Overview and vocabulary. Standard ISO/IEC 27000:2018, 2018.

[36] International Organization for Standardization. Information Security Management Systems - Requirements. Standard ISO/IEC 27001:2022, 2022.

[37] International Organization for Standardization and International Electrotechnical Commission. Information Technology – Software Asset Management - Part 2: Software Identification Tag. Standard ISO/IEC 19770-2:2015, 2015.

[38] International Society of Automation. ISA/IEC 62443 Series of Standards: Automation and Control Systems Cybersecurity. Standard ISA/IEC 62443:2019, 2019.

[39] Islam, C., Babar, M. A., and Nepal, S. A Multi-Vocal Review of Security Orchestration. *ACM Computing Surveys 52*, 2 (2019), 1–45.

[40] Khan, M. A., and Salah, K. IoT security: Review, blockchain solutions, and open challenges. *Future Generation Computer Syststems 82* (2018), 395–411.

[41] Kritzinger, W., Karner, M., Traar, G., Henjes, J., and Sihn, W. Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine 51*, 11 (2018), 1016–1022.

[42] Mahmood, T., and Afzal, U. Security analytics: Big data analytics for cybersecurity: A review of trends, techniques and tools. In *Proceedings of the 2nd National Conference on Information Assurance* (2013), pp. 129–134.

[43] Markus, M. L., Majchrzak, A., and Gasser, L. A Design Theory for Systems That Support Emergent Knowledge Processes. *MIS Quarterly 26*, 3 (2002), 179–212.

[44] Mekala, S. H., Baig, Z., Anwar, A., and Zeadally, S. Cybersecurity for Industrial IoT (IIoT): Threats, countermeasures, challenges and future directions. *Computer Communications 208* (2023), 294–320.

[45] MISP Project. MISP Standard. https://www.misp-standard.org, 2024. Accessed: June 27, 2024.

[46] Mitropoulos, S., Patsos, D., and Douligeris, C. On Incident Handling and Response: A state-of-the-art approach. *Computers & Security 25*, 5 (2006), 351–370.

[47] National Institute of Standards and Technology. NIST Cybersecurity Framework 2.0. White Paper NIST CSWP 29, 2023.

[48] OASIS. CACAO Security Playbooks Version 2.0. https://docs.oasis-open.org/cacao/security-playbooks/v2.0/security-playbooks-v2.0.html, 2024. Accessed: June 27, 2024.

[49] OASIS. CSAF – Common Security Advisory Framework Version 2.0. https://docs.oasis-open.org/csaf/csaf/v2.0/os/csaf-v2.0-os.html, 2024. Accessed: June 27, 2024.

[50] OASIS Cyber Threat Intelligence (CTI) Technical Committee. Introduction to STIX. https://oasis-open.github.io/cti-documentation/stix/intro.html, 2024. Accessed: June 27, 2024.

[51] OASIS OpenC2 Technical Committee. OpenC2 – Open Command and Control. https://openc2.org, 2024. Accessed: June 27, 2024.

[52] Orlikowski, W. J., and Iacono, C. S. Research Commentary: Desperately Seeking the "IT" in IT Research—A Call to Theorizing the IT Artifact. *Information Systems Research 12*, 2 (2001), 121–134.

[53] OWASP Foundation. OWASP CycloneDX. https://owasp.org/www-project-cyclonedx, 2024. Accessed: June 27, 2024.

[54] Peffers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems 24*, 3 (2007), 45–77.

[55] Pokhrel, A., Katta, V., and Colomo-Palacios, R. Digital twin for cybersecurity incident prediction: A multivocal literature review. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops* (2020), p. 671–678.

[56] Reuters. Russian spies behind cyberattack on Ukrainian power grid in 2022. https://www.reuters.com/technology/cybersecurity/russian-spies-behind-cyberattack-ukrainian-power-grid-2022-researchers-2023-11-09, 2023. Accessed: June 27, 2024.

[57] Reuters. Germany announces military overhaul with eye on cyber threats. https://www.reuters.com/world/europe/germany-announces-military-overhaul-with-eye-cyber-threats-2024-04-04, 2024. Accessed: June 27, 2024.

[58] Rubio, J. E., Alcaraz, C., Roman, R., and Lopez, J. Current cyber-defense trends in industrial control systems. *Computers & Security 87* (2019), 101561.

[59] Sacha, D., Stoffel, A., Stoffel, F., Kwon, B. C., Ellis, G., and Keim, D. A. Knowledge Generation Model for Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1604–1613.

[60] Schneier, B. The Future of Incident Response. *IEEE Security & Privacy 12*, 5 (Sept. 2014), 96–96.

[61] Smith, R., Janicke, H., He, Y., Ferra, F., and Albakri, A. The Agile Incident Response for Industrial Control Systems (AIR4ICS) framework. *Computers & Security 109* (Oct. 2021), 102398.

[62] Standing Committee of the National People's Congress. Cybersecurity Law of the People's Republic of China. Legislation, 2017.

[63] Stevens, R., Votipka, D., Dykstra, J., Tomlinson, F., Quartararo, E., Ahern, C., and Mazurek, M. L. How Ready is Your Ready? Assessing the Usability of Incident Response Playbook Frameworks. In *Proceedings of the 2022 Conference on Human Factors in Computing Systems* (2022), S. D. J. Barbosa, C. Lampe, C. Appert, D. A. Shamma, S. M. Drucker, J. R. Williamson, and K. Yatani, Eds., ACM, pp. 589:1–589:18.

[64] Stouffer, K., Lightman, S., Pillitteri, V., Abrams, M., and Hahn, A. Guide to Industrial Control Systems (ICS) Security. White Paper SP 800-82 Revision 3, National Institute of Standards and Technology, 2023.

[65] Tao, F., Zhang, H., Liu, A., and Nee, A. Y. C. Digital Twin in Industry: State-of-the-Art. *IEEE Transactions on Industrial Informatics 15*, 4 (2019), 2405–2415.

[66] The Guardian. Israel appears to confirm cyberattack on Iran nuclear facility. https://www.theguardian.com/world/2021/apr/11/israel-appears-confirm-cyberattack-iran-nuclear-facility, 2021. Accessed: June 27, 2024.

[67] The MITRE Corporation. CAPEC – Common Attack Pattern Enumeration and Classification. https://capec.mitre.org, 2024. Accessed: June 27, 2024.

[68] The MITRE Corporation. CPE – Common Platform Enumeration. https://cpe.mitre.org/specification, 2024. Accessed: June 27, 2024.

[69] The MITRE Corporation. CVE – Common Vulnerabilities and Exposures. https://www.cve.org, 2024. Accessed: June 27, 2024.

[70] The MITRE Corporation. CWE – Common Weakness Enumeration. https://cwe.mitre.org, 2024. Accessed: June 27, 2024.

[71] The MITRE Corporation. Groups – MITRE ATT&CK. https://attack.mitre.org/groups, 2024. Accessed: June 27, 2024.

[72] The MITRE Corporation. MITRE ATT&CK Matrix for Industrial Control Systems. https://attack.mitre.org/matrices/ics, 2024. Accessed: June 27, 2024.

[73] The MITRE Corporation. MITRE D3DEND – A Knowledge Graph for Cybersecurity Countermeasures. https://d3fend.mitre.org, 2024. Accessed: June 27, 2024.

[74] THE MITRE CORPORATION. OVAL – Open Vulnerability and Assessment Language. https://oval.mitre.org, 2024. Accessed: June 27, 2024.

[75] THE WHITE HOUSE. Presidential Policy Directive – Critical Infrastructure Security and Resilience. Legislation PPD-21, 2013.

[76] WANG, Y., SU, Z., GUO, S., DAI, M., LUAN, T. H., AND LIU, Y. A Survey on Digital Twins: Architecture, Enabling Technologies, Security and Privacy, and Future Prospects. *IEEE Internet of Things Journal 10*, 17 (2023), 14965–14987.

[77] WEBSTER, J., AND WATSON, R. T. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly 26*, 2 (2002), xiii–xxiii.

[78] WIRTH, R., AND HIPP, J. CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining* (2000), pp. 29–40.

Part II

Research Papers

The second part of this dissertation features publications P1 through P7.

P1 Digital Twins in Security Operations: State of the Art and Future Perspectives

Status:	Submitted
Date of Submission:	15 February 2024
Date of Acceptance:	–
Date of Publication:	–
Journal:	ACM Computing Surveys
Authors' Contributions:	

Philip Empl	40%
David Koch	30%
Marietheres Dietz	20%
Günther Pernul	10%

Full Citation:	Empl, P., Koch, D., Dietz, M., & Pernul, G. (2024). Digital Twins in Security Operations: State of the Art and Future Perspectives. *ACM Computing Surveys.* Association for Computing Machinery.
DOI:	–
Artifact:	github.com/philipempl/DT4SEC

Journal description: These comprehensive, readable surveys and tutorial papers give guided tours through the literature and explain topics to those who seek to learn the basics of areas outside their specialties. The carefully planned and presented introductions in Computing Surveys help researchers and professionals gain perspectives on and identify trends in complex technologies. Contributions that bridge existing and emerging technologies, such as machine learning, with a variety of science and engineering domains in novel and interesting ways are welcomed. The journal aims to be accessible to a broad audience, featuring clear exposition, a lively tutorial style, and pointers to further literature.

Digital Twins in Security Operations: State of the Art and Future Perspectives

PHILIP EMPL, University of Regensburg, Germany
DAVID KOCH, Bundeswehr University Munich, Germany
MARIETHERES DIETZ, University of Regensburg, Germany
GÜNTHER PERNUL, University of Regensburg, Germany

In an era of rapid technological advancements, digital twins are gaining attention in industry and research. These virtual representations of real-world entities, enabled by the Internet of Things (IoT), offer advanced simulation and analysis capabilities. Their application spans various sectors, from smart manufacturing to healthcare, highlighting their versatility. However, the rise of digital technologies has also escalated cybersecurity concerns. Historical cyberattacks underscore the urgency for enhanced security operations. In this context, digital twins represent a novel approach to cybersecurity. Industry and academic research are increasingly exploring their potential to protect their assets. Despite growing interest and applications, more comprehensive research synthesis needs to be done, particularly in security operations based on digital twins. Our paper aims to fill this gap through a structured literature review aggregating knowledge from 201 publications. We focus on defining the digital twin in cybersecurity, exploring its applications, and outlining implementations and challenges. To maintain transparency, our data is documented and is publicly available. This survey serves as a crucial guide for academic and industry stakeholders, fostering digital twins in security operations.

CCS Concepts: • **General and reference** → **Surveys and overviews**; • **Networks** → **Cyber-physical networks**; • **Information systems**; • **Computer systems organization** → **Embedded and cyber-physical systems**; • **Security and privacy**;

Additional Key Words and Phrases: Digital Twin, Security Operations, Cybersecurity, Survey, Literature Review

ACM Reference Format:
Philip Empl, David Koch, Marietheres Dietz, and Günther Pernul. 2024. Digital Twins in Security Operations: State of the Art and Future Perspectives. In *ACM Computing Surveys, Vol. 1, No. 1, Article 1, Submission date: 15. February 2024.* ACM, New York, NY, USA, 35 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Offering opportunities like no technology before, digital twins are gaining significant attention in industry and science alike. Digital twins virtually represent real-world objects (e.g., products, systems, or even processes) and are thus able to forecast and simulate potential future happenings as well as to replicate current ones. Thereby, a digital twin bases its computations on data from the real-world object, adds context by semantically linking data, and produces results tailored to the asset in question. The extent of these results matches no current alternative regarding in-depth analysis of the investigated object and offers high degrees of individualization. While the notion of digital twins has been floating around for years, they only became realizable with the widespread adoption of the Internet of Things (IoT).

Thanks to these powerful features, the application scenarios of digital twins are manifold. For instance, the engineering company Bosch is instrumenting the digital twin mainly in the sphere of the Industrial IoT to envision a smart factory [90]:

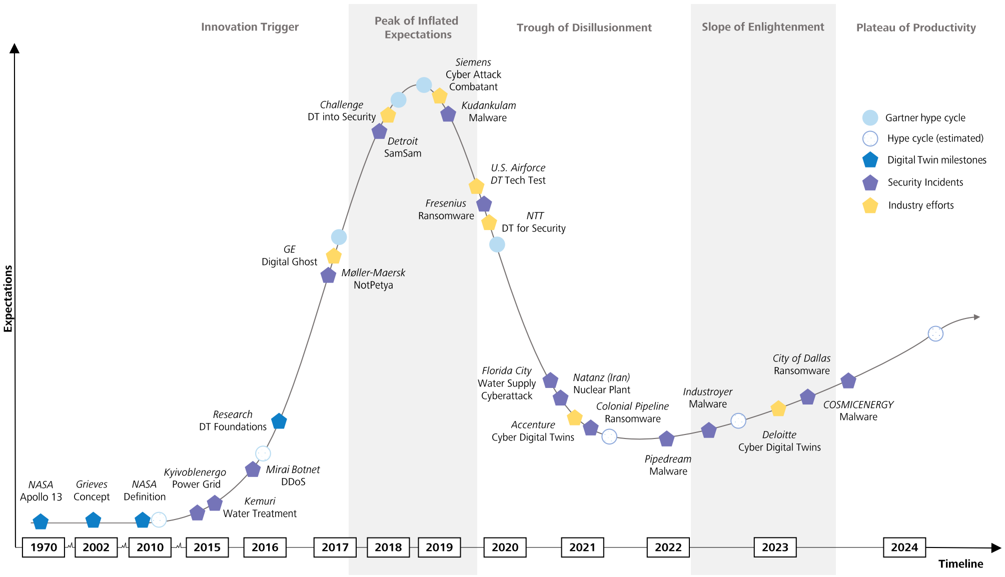


Fig. 1. Chronological development of digital twin (in security operations) research and industry efforts.

They begin with collecting raw data of their products, add semantics to generate contextual information and provide a digital representation of the individual product. On the other hand, Siemens is computing digital twins of patients' hearts to comprehend the divergent reactions to medicine [107] or, in times of SARS-CoV-2, to vaccines. On a greater level, the European Union is planning the *Destination Earth* initiative, which focuses on creating a digital twin of the whole world to estimate probabilities and impacts of future scenarios, such as in climate change. While the world is enhanced with digitization, crime is, too. Cybersecurity is turning into a major concern for companies these days. In the last decades, several cyberattacks (see Figure 1) greatly impacted the world, including:

- Ransomware attacks (e.g., *NotPetya* in 2017 [5], the Fresenius cyber attack in 2020 [217], and the Colonial Pipeline attack [81] in 2021)
- Malware attacks (e.g., Kudankulam power plant in 2019 [52])
- Worms (e.g., *Stuxnet* attack in 2010 [5])
- Exploits of unpatched vulnerabilities (e.g., the cyberattack at Kemuri water treatment plant in 2015 [5])
- Botnets (e.g., *Mirai* in 2016, exploiting IoT devices using default credentials for DDoS attacks [10])

One of the latest events highlights that more significant efforts to enhance cybersecurity are required. In 2023, the *Vulkan files* were published. These files implicate the Russian government of actively developing frameworks (e.g., *Amezit* until 2018 and *Krystal-2B* until 2020) to attack industrial systems. Figure 1 illustrates critical cybersecurity incidents while also informing on milestones of digital twins in research. Thereto, it adapts the Gartner hype cycle to show the (estimated) progression of the digital twin concept from 2017 to 2024. It also provides advances in practice concerning using digital twins for security. For instance, General Electric employs digital twins of machines to identify anomalies in their behavior. Their so-called "digital ghost" maintains cybersecurity via detection, localization, prediction, and even neutralization of harmful activities [87]. Meanwhile, Siemens also starts exploiting the potential of digital twins for detecting cyberattacks [67]. Others follow their lead (e.g., [120, 125, 153]).

Likewise, academic research increasingly focuses on the digital twin's potential for cybersecurity. Dietz & Pernul [58] describe the possibilities of a digital twin to enhance cybersecurity in industrial systems, identifying three modes of operation: simulation, replication, and analytics. Eckhart et al. [66] developed a prototypical digital twin to allow assumptions about the security state of industrial systems. Olivares-Rojas et al. [165] used the digital twin technology to test cybersecurity attacks on smart meters, while the real-world smart meters remain unaffected. Many more security-related digital twin applications have been introduced in research. However, they rarely rely on one another. This is why providing an overview of the existing body of knowledge concerning enhancing security through digital twins is vital. Over the last years, there have been some surveys on digital twins in general (e.g., [162]). However, those literature reviews commonly do not consider the security perspective – despite its apparent adoption in science and practice. Thus, one of the scientific challenges remains to provide a comprehensible overview of how digital twins can be used for security. Our structured literature review will close this gap and follows the principles of open data[1].

1.1 Research Questions & Contribution

Regarding the current development of the digital twin turning into a cybersecurity tool, it is vital to classify and present the state of the art. Such a survey allows us to study existing works and expand the body of knowledge by building on the already published research. Doing so, we investigate and aggregate 201 research paper. Next, to provide an overview of existing literature, the need for further research will be identified, showing the opportunities and diversity of the digital twin in cybersecurity. This survey not only endeavors to summarize previous research findings but also maps the landscape of digital twins for security operations following the NIST Cybersecurity Framework (CSF) [18]. Therefore, we focus on a dynamic and critical convergence point – where technology, cybersecurity, and industrial requirements intersect. The following *research questions* help to determine the status quo of digital twins for security operations:

RQ1 What constitutes a digital twin within the realm of cybersecurity?
RQ2 How is the digital twin harnessed to fortify cybersecurity?
RQ3 How can digital twin components be instantiated?
RQ4 What pivotal challenges remain unaddressed, and what promising avenues beckon for future research?

To answer **RQ1**, we will study the digital twin's components and how researchers understand the paradigm. Subsequently, exploring the applications of using the digital twin in security operations will respond to **RQ2**. This includes the benefits of using digital twins for cybersecurity. The answer to **RQ3** will include how digital twin components are set up, what data the digital twin consumes, and which tools and communication protocols are used. Finally, we will answer **RQ4** by extracting the authors' views from their publications to synthesize in which direction the digital twin in security operations needs to be explored as well as how to overcome challenges for designing and implementing digital twins in cybersecurity settings. In a nutshell, the *contribution* of this paper is as follows:

(1) It confirms the components of the digital twin from a cybersecurity perspective,
(2) re-defines the paradigm digital twin in cybersecurity,
(3) describes the application domains and security operations based on digital twins,
(4) provides an overview of implementation techniques and tools,
(5) shapes future research directions and open issues.

[1] https://github.com/philipempl/DT4Sec

1.2 Remainder

The remainder of this work is organized as follows. Section 2 will explore the theoretical background to fill in readers on the digital twin paradigm and to describe related work by other authors. The methodology of this survey is detailed in Section 3. Section 4 approaches the digital twin paradigm and essential components. Section 5 describes the application benefits, domains, and security operations. Section 6 highlights how implementing the digital twin components can be managed, and Section 7 pictures open issues and future research. In Section 8, we will conclude our paper.

2 BACKGROUND AND RELATED WORK

2.1 Digital Twin

Although the digital twin seems like a newly emerged concept, the notion of it has been around for quite some time. For instance, when Apollo 13 suffered an explosion inside its oxygen tanks in outer space in 1970, NASA modified the conditions of their simulators to match the specifications of Apollo 13 and identified strategies to solve this hazardous situation [19]. NASA first created something resembling an early-stage digital twin by customizing its generic simulator to an individual counterpart. Around 2012, NASA and the U.S. Air Force applied digital twins further to examine the life cycle of their physical assets and to diagnose as well as predict an asset's particular behavior [88]. Nowadays, digital twins go beyond mere simulations. In science as well as in practice, the digital twin only gained a foothold with the maturity and deployment of the IoT: Data generated thanks to the IoT enable on-building statistics, artificial intelligence, machine learning, deep learning, and data visualizations – all valuable features of digital twins [80]. So, only with the technology of the IoT have organizations been able to provide the twins with enough crucial data quickly. Essential data is usually stored and enriched with semantics to provide context [225]. On this basis, the digital twin can virtually represent its real-world counterpart, e.g., by using the data to create a digital model or simulation [183]. Also, current advances in digital twin technology attempt to create a bidirectional connection between the twin and its counterpart [124]. This enables the digital twin not only to monitor but also to control its real-world counterpart.

2.2 Related Work

In the past, researchers conducted a plethora of surveys concerning the digital twin [19, 115, 124, 136, 138, 162, 218, 223]. Most of them concentrated on certain application domains (e.g., smart manufacturing). However, a literature overview on digital twins regarding security operations has yet to be conducted. While the digital twin holds vital and often sensitive data, research mainly developed in the direction of provisioning cybersecurity for the digital twin – e.g., through the combination of digital twins and distributed ledger technology [112, 175, 231]. However, surveys scarcely focus on the application of the digital twin to enhance cybersecurity through security operations. Böhm et al. [24] systematically deal with combining digital twins and Augmented Reality (AR) by taking a cybersecurity perspective. Eckhart et al. [64] systematized the body of knowledge in an unstructured way and identified several security operations of the digital twin for cybersecurity in 2019. Pokhrel et al. [173] conducted a Multi-vocal Literature Review (MLR) on the application of the digital twin for incident prediction. Thalpage and Nisansala [208] systematically analyzed the usage of digital twins for industrial intrusion detection. In contrast, Wang et al. [220] do not apply a structured method when focusing on digital twins and security for IoT. Table 1 compares these works to our study. The most remarkable difference to our survey lies in the specialization of these. For example, some authors [64, 208, 220] focus on one specific application domain. Pokhrel et al. [173] as well as Thalpage and Nisansala [208] only target one specific security operation (intrusion detection), while Böhm et al. [24] review a given set of core technologies (e.g., AR). Our survey is

Aspects	[24]	[64]	[173]	[208]	[220]
Methodology	*SLR*	*n/s*	*MLR*	*SLR*	*n/s*
Publication year	*2021*	*2019*	*2020*	*2023*	*2023*
Security perspective	Augmented Reality (technology-res.)	Industry (app-res.)	Intrusion detection (ops-res.)	Ind. intr. det. (app-&ops-res.)	IoT (app.-res.)
Studies considered	355	n.a.	751	158	n.a.
Studies included	33	n.a.	17	10	n.a.
Applications	●	◑	◑	◑	●
Implementation	◑	○	○	○	◑
Challenges/ future research	○	●	◑	○	●

○ = not covered; ◑ = partially covered; ● = covered; app = application; ops =operations; res = restricted; n/s = not specified; n.a. = not available

Table 1. Comparison of our survey to existing surveys in terms of the content and scope

intended to be all-encompassing and to combine all perspectives of previously written overviews. Moreover, we base our literature review on a profound methodology to gather the results in a structured way, fostering comprehensibility and transparency. It is also important to note that the research of digital twins in security operations only started around 2018, turning our literature review into a 7-year summary including more than 200 relevant articles on this issue in contrast to earlier reviews, where the maximum of relevant articles is 33 (see Table 1). Surely, there are still other works do provide an overview on digital twin security [65, 74, 116, 202, 205]. However, none of these works provided a method. Also, those works only cover a very specific part of the literature, offering little basis for comparison. In contrast, the two studies without a method [64, 220] reference a multitude of research to provide very comprehensive overviews. In summary, our paper takes a comprehensive and methodological approach to investigate the multifaceted role of digital twins in security operations by filling a significant gap through a wide-ranging synthesis, in contrast to previous work that are either outdated, lack methodological rigor, or focus only on specific application domains, security operations, and technologies.

3 METHODOLOGY

Our survey strictly adheres to the methodology of a qualitative systematic review, a widely recognized approach in information systems research [167]. We adopt the guidelines by Schryen [191], which structure our review into four key phases: framing, searching, screening, and synthesis. This approach ensures a comprehensive and well-organized examination of our subject matter. For an in-depth understanding of our methodology and the specific outcomes of each phase, more details can be found in our GitHub repository[2]. The following sections provide a detailed account of our search, assessment, extraction, and synthesis processes.

Search & Assessment. Our literature search process begins with queries in ten different scientific databases: ACM, AISeL, arXiv, dblp, IEEE CSDL, IEEE Xplore, ScienceDirect, SpringerLink, Wiley Online, and WoS (see Table 2). We also conduct backward searches by automatically parsing PDFs and forward searches using Google Scholar to enhance our search (> 3,000 results). It is worth noting that the search term "*digital twins AND security*" produces an exceptionally large number of results in a three databases where searches cannot be constricted to data fields. To manage this, we

[2] https://github.com/philipempl/DT4SEC

Database Name	**URL**	**Results**
ACM	dl.acm.org	565
AISeL	aisel.aisnet.org	163
Arxiv	arxiv.org	77
dblp	dblp.org	64
IEEE CS	computer.org	504
IEEE Xplore	ieeexplore.ieee.org	1,075
Science Direct	sciencedirect.com	954
Springer Link	link.springer.com	1,631
Wiley	wiley.com	209
WoS	webofscience.com	318
Total		**5,560**

Table 2. Databases and initial search results.

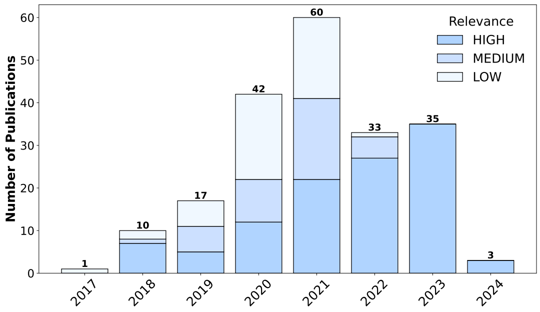


Fig. 2. Selected digital twin papers by year and relevance (*Total: 201*).

refine our search results by including additional keywords[3]. Our literature review process is started by cleaning the corpus to eliminate duplicate entries. Then, we assess each publication's relevance by considering the title (*screen 1*) and afterward the full-text (*screen 2*), where we apply specific inclusion and exclusion criteria established in advance. These criteria involve evaluating the publication's quality, language (English), relevance to digital twins, and cybersecurity topics. If a publication does not meet these criteria, it is excluded immediately. Through this rigorous process, we could identify a total of 201 publications that align with the scope of our survey, which is illustrated in Figure 2. We classified the relevance of the papers into three categories: high, medium, and low. A paper was deemed of high relevance (111 papers) if it directly addressed digital twins for security operations. Medium relevance (41 papers) was assigned to papers that partially mentioned the topic. Low relevance (49 papers) was attributed to papers that mentioned the topic in passing. Moreover, this figure indicates a noticeable increase in highly relevant academic research on digital twins in security operations. Please note that our search only covers January 2024.

Extraction & Synthesis. In this phase, we focus on extracting and organizing the required data. This part is crucial because it supports answering our research questions. To enable data extraction, we create a structured template applied to each relevant research paper. This template contains 18 fields and ensures consistent handling of each paper. In addition to the data related to our research questions, we gathered bibliographic information like author names, publication years, and titles. We have created an additional template for the NIST CSF mapping. Before applying the extraction template to all the papers, it is tested on a small sample to ensure it works effectively. Once this step is completed, the data is synthesized using the thematic analysis method [47]. Thematic analysis helps identify common patterns in the data and group them into specific themes (see Figure 3). We first define central themes for each research question in this thematic analysis process. Afterwards, sub-themes are linked to corresponding central themes. These central themes act as summaries for the sub-themes. Our approach to defining these themes combines inductive and deductive methods. While the sub-themes *Application Domains* and *Security Operations* are derived deductively, drawing from the NIST CSF and the critical infrastructure classification established by the Cybersecurity & Infrastructure Security Agency (CISA), the remaining sub-themes are generated inductively from our extracted data. Our primary objective is to gain a comprehensive understanding of the digital twin paradigm in the context of cybersecurity (*what is*

[3] *"digital twin"* ***AND*** *("internet of things"* ***OR*** *"IoT"* ***OR*** *"CPS"* ***OR*** *"cyber-physical systems"* ***OR*** *"cyber physical systems")* ***AND*** *("cybersecurity"* ***OR*** *"cyber-security"* ***OR*** *"information security")*

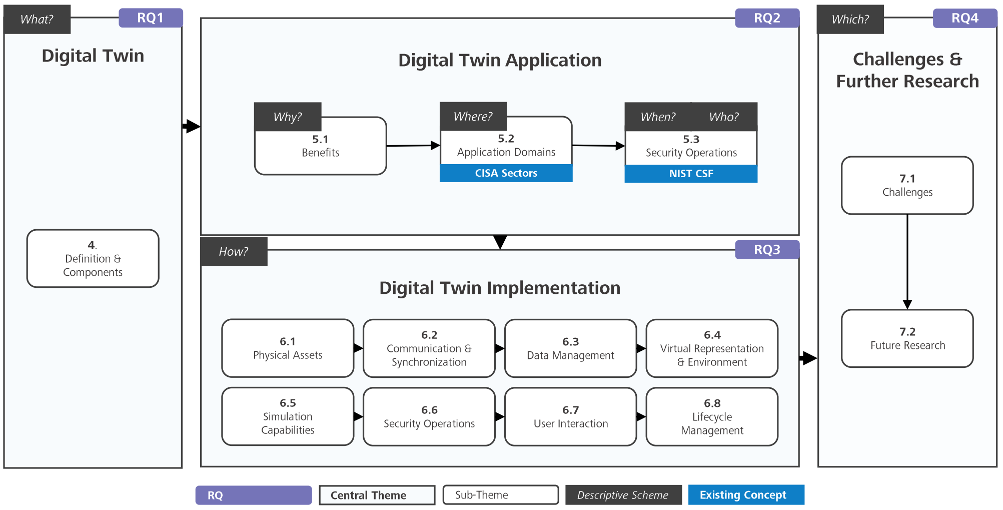


Fig. 3. Thematic classification of the survey's topic.

a digital twin?), with details about its applications (*"why/where/when to use a digital twin?"*, or *who applies a digital twin?*), implementations (*how to use a digital twin?*) and future research (*which open challenges and future research exist?*). Please note that we extract data from publications of all relevancy levels to address conceptual schemes. However, for the analysis of security operations aligned with the NIST CSF, we specifically focus on papers classified as highly relevant. These papers are selected because they provide detailed insights into the subject matter. In the following sections (Sections 4-7), we delve into each of these central themes and their corresponding sub-themes in more detail.

4 APPROACHING THE DIGITAL TWIN PARADIGM

In this section, we delve into the paradigm of the digital twin. We face different perceptions of a digital twin ranging from "*middleware*" [72] to "*holistic simulation*" [63] and "*virtual replica*" [84]. Additionally, there are differing viewpoints on what should be considered an integral part of the digital twin and what should be seen as external but still utilized by the digital twin. For instance, while data storage could externally provide data to a digital twin [26], it could also be a vital functional component [60, 118, 163] holding machine learning models [27, 40, 50, 227, 230]. We illustrate the digital twin by weaving together the functional components identified from academic literature (see Table 3). This lets us draw a definition of the digital twin paradigm for security operations:

> **What is a digital twin in the realm of cybersecurity?**
> A digital twin in the realm of cybersecurity is a dynamic *virtual representation* of an entity that seamlessly integrates *real-time and historical data*, covering the entire *lifecycle* of its counterpart. Thereby, it can *simulate* and replicate its behavior and states. A digital twin engages users in the *interactions* and *security operations.*

Component	References
Virtual Representation	[2, 3, 7, 11, 13, 16, 21, 28, 34, 39, 45, 49, 53–55, 62, 64, 65, 70, 72, 73, 76, 77, 82, 84–86, 92, 99, 101, 103, 117, 121–123, 133, 135, 137, 141, 144, 148, 150, 151, 154, 155, 160, 161, 164, 170–172, 178, 179, 184–187, 189, 190, 193, 194, 198, 202, 204, 205, 208–214, 221, 227–230, 233, 234]
Data Management	[4, 7, 15, 16, 20, 45, 49, 64, 65, 135, 164, 204, 208]
User Interaction	[16, 25, 45, 65, 70, 73, 77, 83, 113, 135, 164, 165, 170, 212, 229, 234]
Simulation Capabilities	[12, 15, 21, 23, 31, 35, 46, 49, 50, 53, 58, 60, 63, 65, 66, 69, 70, 75, 78, 85, 92, 95, 96, 101, 106, 111, 113, 118, 122, 128, 131, 132, 140, 148, 152, 166, 190, 193, 194, 198, 199, 202, 205, 233, 234]
Lifecycle Management	[32, 49, 55–60, 64, 72, 95, 135, 152, 164, 170, 211, 234]
Communication & Synchronization	[1, 2, 4, 11, 23, 27, 28, 57, 61, 63, 72, 76, 95, 96, 103, 103, 106, 113, 113, 118, 126, 132, 139, 141–143, 152, 159, 164, 177, 201, 202, 210, 214]
Security Operations	[8, 9, 12, 21, 46, 59, 61, 64, 69, 75, 164, 171, 199, 205]

Table 3. Functional components of a digital twin derived from academic definitions.

Virtual Representation. A digital twin is a virtual representation or replica of a real-world entity. The virtual representation is close to the replication mode of a digital twin [58], where the stimuli triggering states in the real-world counterpart are reproduced in the digital twin to mimic system states. The virtual representation encompasses either a structural (e.g., JSON [196] or RDF [40]), functional (e.g., AutomationML [32]), or behavioral (e.g., finite-state machine [62] or Kalman filter [1]) aspect of the entity or a combination thereof, to capture its nuances and intricacies. The representation's fidelity determines the accuracy of the security operation.

Data Management. Central to the digital twin concept is its ability to manage data. This starts with the collection of data, where different types might be required (e.g., device [70] or network [60]). It also involves data transformation and its corresponding tools and methods (e.g., using Logstash as proposed by [214]). Dependent on the security operation, a digital twin can consume real-time (e.g., [131]) and historical data (like relational [129], document-oriented [216], graph-based [40], or distributed [172]) to provide a comprehensive data-backed view.

User Interaction. Digital twins should dynamically allow user interaction. This component can also be referred to as the analytics mode of a digital twin [58], where cybersecurity professionals can monitor and correlate variables making the virtual representation more effective. In addition, cybersecurity novices can be introduced to the topic in a practical way [216]. Besides reporting (e.g., Kibana [60] or Prometheus [2]), cybersecurity professionals can interact with digital twins in different ways (like via virtual reality [132] or 3D models [234]).

Simulation Capabilities. Beyond being mere replicas, digital twins offer advanced simulation capabilities (*the process of generating data*). This is also named simulation mode [58]. Depending on the virtual representation, digital twins can simulate different scenarios and predict future states based on real-time and historical data. Digital twins are decoupled from the real-world entity and do not affect their physical counterparts [62]. Note that running a digital twin in the simulation mode requires behavioral models as virtual representations and appropriate software (e.g., MathWorks Simulink as shown by [103] or Cooja like in [118]).

Communication & Synchronization. The value of a digital twin is amplified by its connectivity. Seamless communication protocols ensure that there is a continuous flow of data between the digital twin and its physical counterpart (e.g., [13] focused on MQTT, [38] on Modbus, [222] on OPC UA). This connectivity is crucial for real-time monitoring, synchronization, and remote control. Regarding synchronization, a digital twin is characterized by bidirectional communication, where the entity and the digital twin share the same state.

Security Operations. Security operations, as outlined in the NIST CSF, refer to the structured processes and activities carried out by organizations or security teams to safeguard their assets, data, systems, and operations from various threats. Digital twins assist security operations given the framework's five phases:

- *Identifying* cybersecurity risks to establish a security baseline [61].
- *Protecting* assets through security measures, e.g., policies [46], access controls [33], and awareness [66].
- *Detecting* security incidents using monitoring, intrusion detection, and cyber threat intelligence [48].
- *Responding* to security incidents by defining and executing incident response processes [72].
- *Recovering* in the event of security incidents, emphasizing business continuity and resilience [2].

Considering these functional components, a digital twin in security operations is similar to one in the industrial context. The difference is that a digital twin in security operations addresses a different context, user group, and operations. We find different relevant data and different security operations. Regarding user groups, cybersecurity professionals are the main ones interacting with the digital twin. In the next section, we detail these security operations where digital twins are currently used.

5 UNDERSTANDING THE USE OF DIGITAL TWINS

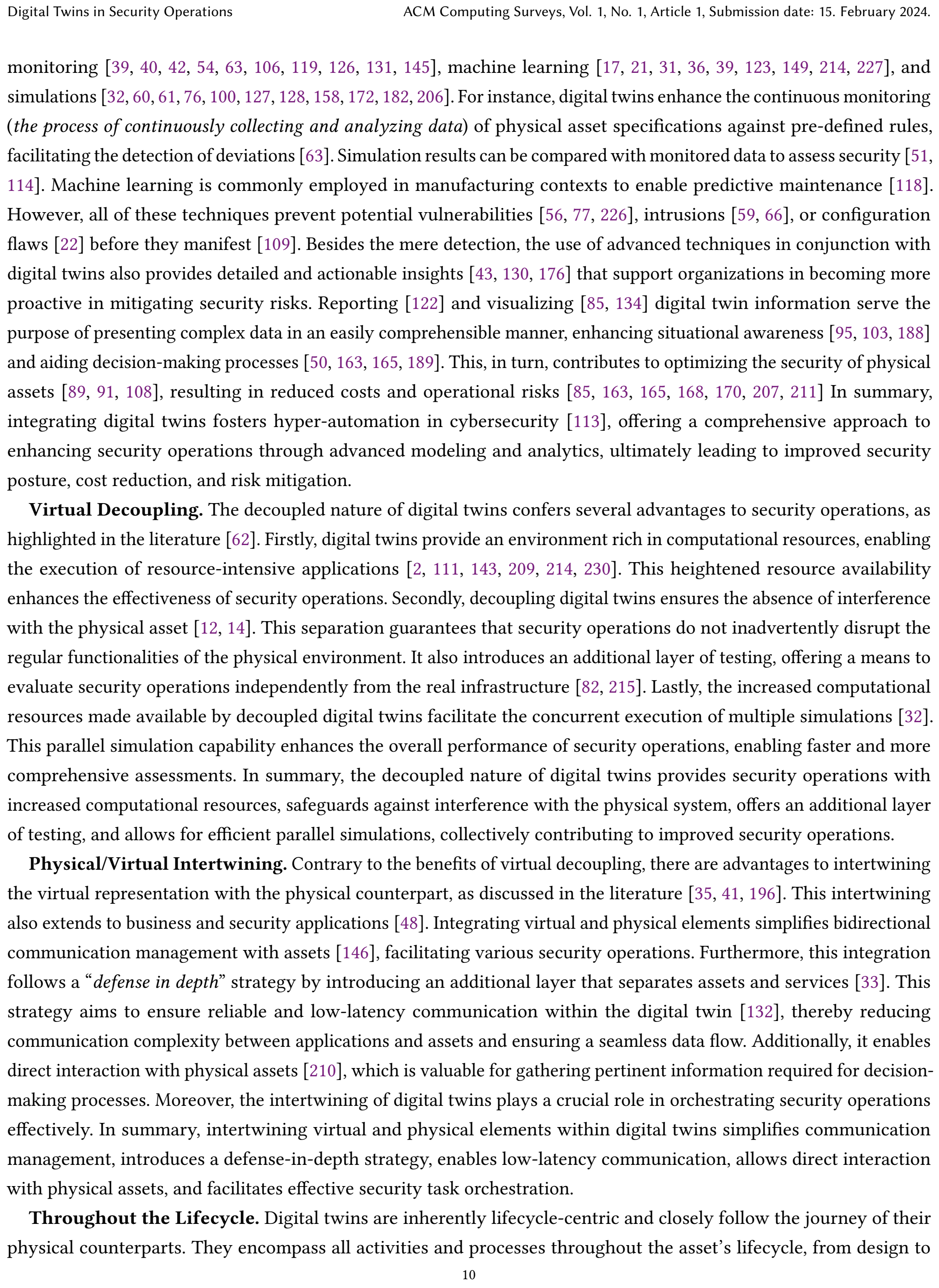
In this section, we delve into why researchers increasingly turn to digital twins to enhance their security operations. We will examine the application domains where digital twins are applied. Additionally, we will explore security operations to illustrate how digital twins operate to improve the cybersecurity posture.

5.1 Benefits of Digital Twins

A digital twin encompasses a virtual representation at its core and culminates in security operations. In the following, we delve into the rationale behind researchers' choice of digital twins as the cornerstone of their security operations.

Asset Centrality. Security operations can significantly benefit from asset centrality, which is central to digital twins. Digital twins are representations of physical assets, including devices, networks, products, or processes, and they maintain a one-to-one relationship with their physical counterparts. These digital twins collect and manage vast amounts of historical and real-time data [84, 98, 138, 157, 161, 180, 186, 197] focused on their respective assets' states and history [57, 58, 141, 142]. Their strength lies in collecting and aggregating data even if distributed across various storage systems [113, 200]. To achieve this, digital twins extract data related to assets, such as firmware binaries [46, 78], and specifications [55]. Additionally, digital twins incorporate expert knowledge about the functions and behaviors of the assets they represent [79]. Digital twin models equip their counterparts with intelligence [174]. To ensure the security of the assets' data, digital twins manage access for their associated physical counterparts [29], typically through mechanisms like access control [142]. Beyond the physical counterpart, digital twins define sharing policies to regulate data access for relevant assets and stakeholders [30, 158, 179], including scenarios like training [140]. However, combining data and knowledge within digital twins necessitates high documentation standards akin to those used in digital forensics [216]. Asset centrality is about maintaining a one-to-one relationship, collecting and managing historical and real-time data, incorporating expert knowledge, and ensuring secure data access.

(High) Fidelity. Security operations derive significant benefits from using digital twins, particularly in terms of their models and semantics. These models are designed to capture the real world with a high degree of fidelity [105, 129, 154, 160, 166, 198, 234], ensuring a reliable reflection of reality when compared to testbeds [23, 31, 37]. Additionally, these models enable more comprehensive analytics [25, 72], when used with more advanced techniques [201], including



monitoring [39, 40, 42, 54, 63, 106, 119, 126, 131, 145], machine learning [17, 21, 31, 36, 39, 123, 149, 214, 227], and simulations [32, 60, 61, 76, 100, 127, 128, 158, 172, 182, 206]. For instance, digital twins enhance the continuous monitoring (*the process of continuously collecting and analyzing data*) of physical asset specifications against pre-defined rules, facilitating the detection of deviations [63]. Simulation results can be compared with monitored data to assess security [51, 114]. Machine learning is commonly employed in manufacturing contexts to enable predictive maintenance [118]. However, all of these techniques prevent potential vulnerabilities [56, 77, 226], intrusions [59, 66], or configuration flaws [22] before they manifest [109]. Besides the mere detection, the use of advanced techniques in conjunction with digital twins also provides detailed and actionable insights [43, 130, 176] that support organizations in becoming more proactive in mitigating security risks. Reporting [122] and visualizing [85, 134] digital twin information serve the purpose of presenting complex data in an easily comprehensible manner, enhancing situational awareness [95, 103, 188] and aiding decision-making processes [50, 163, 165, 189]. This, in turn, contributes to optimizing the security of physical assets [89, 91, 108], resulting in reduced costs and operational risks [85, 163, 165, 168, 170, 207, 211] In summary, integrating digital twins fosters hyper-automation in cybersecurity [113], offering a comprehensive approach to enhancing security operations through advanced modeling and analytics, ultimately leading to improved security posture, cost reduction, and risk mitigation.

Virtual Decoupling. The decoupled nature of digital twins confers several advantages to security operations, as highlighted in the literature [62]. Firstly, digital twins provide an environment rich in computational resources, enabling the execution of resource-intensive applications [2, 111, 143, 209, 214, 230]. This heightened resource availability enhances the effectiveness of security operations. Secondly, decoupling digital twins ensures the absence of interference with the physical asset [12, 14]. This separation guarantees that security operations do not inadvertently disrupt the regular functionalities of the physical environment. It also introduces an additional layer of testing, offering a means to evaluate security operations independently from the real infrastructure [82, 215]. Lastly, the increased computational resources made available by decoupled digital twins facilitate the concurrent execution of multiple simulations [32]. This parallel simulation capability enhances the overall performance of security operations, enabling faster and more comprehensive assessments. In summary, the decoupled nature of digital twins provides security operations with increased computational resources, safeguards against interference with the physical system, offers an additional layer of testing, and allows for efficient parallel simulations, collectively contributing to improved security operations.

Physical/Virtual Intertwining. Contrary to the benefits of virtual decoupling, there are advantages to intertwining the virtual representation with the physical counterpart, as discussed in the literature [35, 41, 196]. This intertwining also extends to business and security applications [48]. Integrating virtual and physical elements simplifies bidirectional communication management with assets [146], facilitating various security operations. Furthermore, this integration follows a "*defense in depth*" strategy by introducing an additional layer that separates assets and services [33]. This strategy aims to ensure reliable and low-latency communication within the digital twin [132], thereby reducing communication complexity between applications and assets and ensuring a seamless data flow. Additionally, it enables direct interaction with physical assets [210], which is valuable for gathering pertinent information required for decision-making processes. Moreover, the intertwining of digital twins plays a crucial role in orchestrating security operations effectively. In summary, intertwining virtual and physical elements within digital twins simplifies communication management, introduces a defense-in-depth strategy, enables low-latency communication, allows direct interaction with physical assets, and facilitates effective security task orchestration.

Throughout the Lifecycle. Digital twins are inherently lifecycle-centric and closely follow the journey of their physical counterparts. They encompass all activities and processes throughout the asset's lifecycle, from design to production and service phases [96, 177, 185, 222]. Remarkably, even if the physical asset has yet to be realized [199, 212], digital twins can offer valuable support, such as system configuration [27] and aid in making security by design decisions to harden the physical assets [20, 216]. Moreover, digital twins persist throughout the product lifecycle, remaining prepared to accompany identical future assets. In summary, digital twins are inherently lifecycle-centric, accompanying physical assets from design to service, even before their physical existence, ensuring robust security at every stage.

Overall, digital twins represent a sophisticated approach to enhancing security operations, offering benefits such as centralized asset management, advanced modeling, a decoupled environment intertwined with physical elements, and comprehensive lifecycle coverage. Collectively, these benefits aim to enhance the overall security posture. Subsequent sections of this paper examine the application domains and security operations benefiting from digital twins.

5.2 Application Domains

Digital twins are applied to enhance security operations across various sectors. Our review of the application domains is based on the critical infrastructure classification established by the CISA[4].

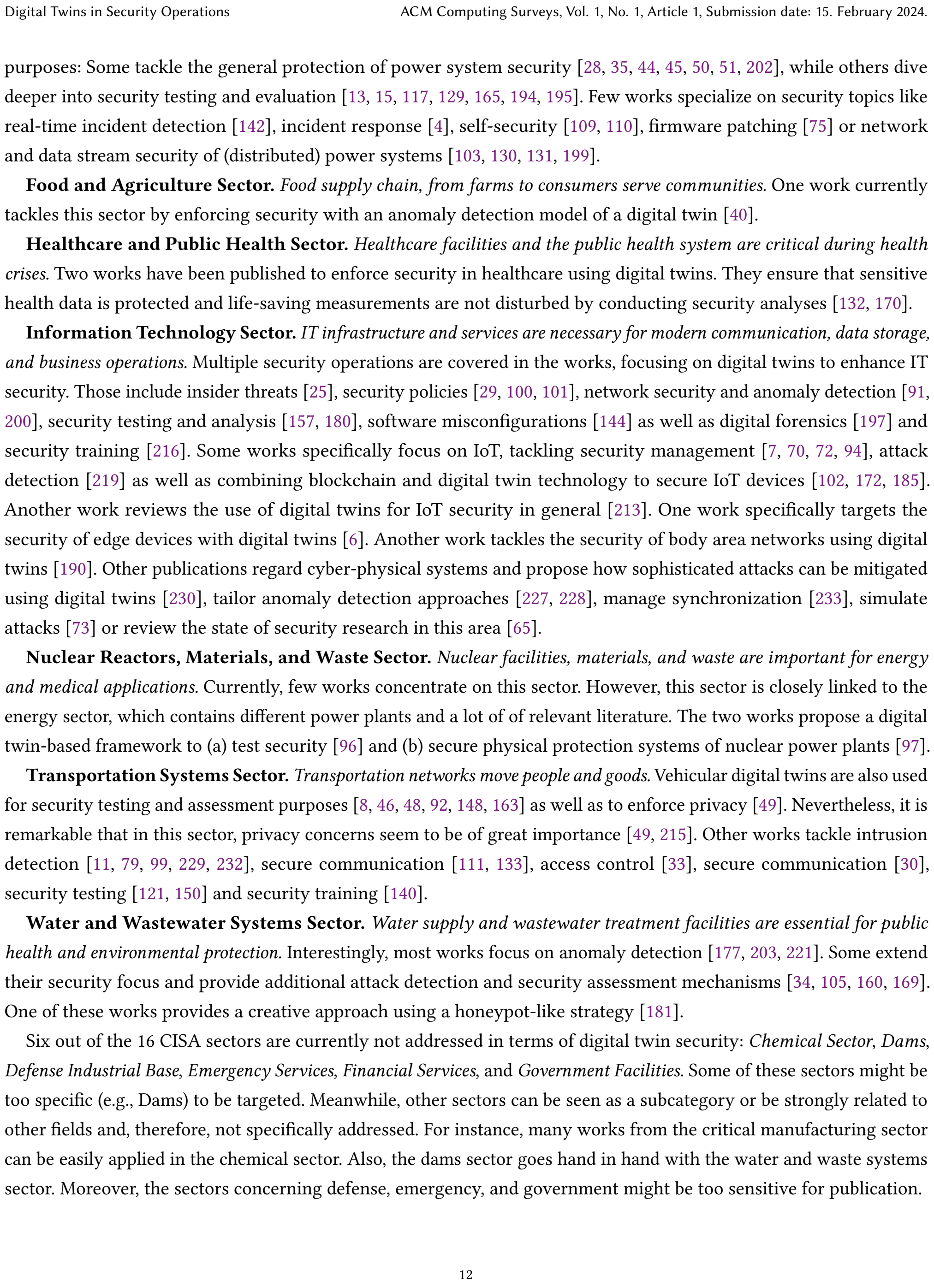
Commercial Facilities Sector. *Public spaces such as shopping centers, sports stadiums, and hotels play a key role in public safety and economic activity.* Current research focuses on malicious attack detection to protect smart cities using digital twins [188]. One work proposes security with digital twins for the IoT-enhanced theater [135]. While we could only find this one digital twin security work on a smaller scale that concerns only commercial or public facilities, we included smart homes in this sector. Despite not being strictly commercial, smart homes have economic characteristics and require safety (albeit more for the individuals than the public). For instance, digital twins are used to detect and check for anomalies via security and safety rules [17] or use other attack detection mechanisms [156] to prevent attacks on smart systems for home automation [224]. One work proposes the use of honeypots in combination with digital twins to detect attacks in smart homes [164], while another tackles security testing [171].

Communications Sector. *Communication networks are essential for everyday communication and emergency responses.* Two works propose a training based on digital twin technology to enhance skills in cybersecurity for IoT and 5G networks [104, 113]. A similar work uses a digital twin to test 5G and attack models [53]. 6G technology enabled by digital twins is regarded in terms of digital forensics [210]. Another work uses a digital twin to verify a security decoupling approach for an IoT network [54]. Some papers focus on communication (network) security with digital twins. However, those are mainly industry-oriented, so we categorized them into the respective manufacturing sector.

Critical Manufacturing Sector. *Manufacturing processes are vital to national security, including aerospace and automotive.* Most works using digital twins for security can be associated with smart manufacturing. In a nutshell, certain aspects are concentrated on when dealing with digital twins for industrial security. Those comprise attack detection [1, 3, 16, 60, 62, 63, 86, 128, 137, 168, 209, 214] including anomaly detection [31, 145], attack mitigation [2, 184, 187, 201] as well as security testing [21, 23, 27, 77, 139, 192, 193] and security analysis or evaluation [37, 58, 61, 66, 93, 106, 114, 122, 147, 186, 226]. Few works specifically regard one security operation – such as security-by-design [56] or security-in-depth [108]. Quite a lot of publications propose the combination with other security operations, including cyber ranges [20, 151, 204, 212, 234], forensics [55], cyber threat intelligence [59] or cybersecurity playbooks [68]. At last, some offer a research overview [64, 173] or deal with security in general [155, 158, 159, 198, 205, 207, 208].

Energy Sector. *Infrastructure for electricity, oil, and natural gas are essential for the functioning of other sectors.* Being the second-largest sector where digital twins are discussed for security, the papers show a manifold of application

[4] https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors



purposes: Some tackle the general protection of power system security [28, 35, 44, 45, 50, 51, 202], while others dive deeper into security testing and evaluation [13, 15, 117, 129, 165, 194, 195]. Few works specialize on security topics like real-time incident detection [142], incident response [4], self-security [109, 110], firmware patching [75] or network and data stream security of (distributed) power systems [103, 130, 131, 199].

Food and Agriculture Sector. *Food supply chain, from farms to consumers serve communities.* One work currently tackles this sector by enforcing security with an anomaly detection model of a digital twin [40].

Healthcare and Public Health Sector. *Healthcare facilities and the public health system are critical during health crises.* Two works have been published to enforce security in healthcare using digital twins. They ensure that sensitive health data is protected and life-saving measurements are not disturbed by conducting security analyses [132, 170].

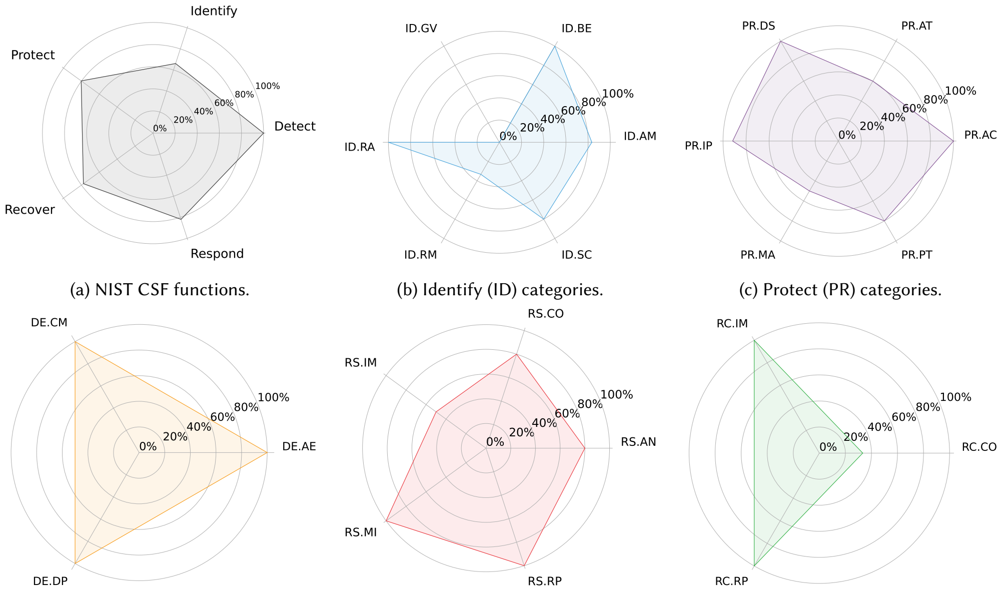
Information Technology Sector. *IT infrastructure and services are necessary for modern communication, data storage, and business operations.* Multiple security operations are covered in the works, focusing on digital twins to enhance IT security. Those include insider threats [25], security policies [29, 100, 101], network security and anomaly detection [91, 200], security testing and analysis [157, 180], software misconfigurations [144] as well as digital forensics [197] and security training [216]. Some works specifically focus on IoT, tackling security management [7, 70, 72, 94], attack detection [219] as well as combining blockchain and digital twin technology to secure IoT devices [102, 172, 185]. Another work reviews the use of digital twins for IoT security in general [213]. One work specifically targets the security of edge devices with digital twins [6]. Another work tackles the security of body area networks using digital twins [190]. Other publications regard cyber-physical systems and propose how sophisticated attacks can be mitigated using digital twins [230], tailor anomaly detection approaches [227, 228], manage synchronization [233], simulate attacks [73] or review the state of security research in this area [65].

Nuclear Reactors, Materials, and Waste Sector. *Nuclear facilities, materials, and waste are important for energy and medical applications.* Currently, few works concentrate on this sector. However, this sector is closely linked to the energy sector, which contains different power plants and a lot of of relevant literature. The two works propose a digital twin-based framework to (a) test security [96] and (b) secure physical protection systems of nuclear power plants [97].

Transportation Systems Sector. *Transportation networks move people and goods.* Vehicular digital twins are also used for security testing and assessment purposes [8, 46, 48, 92, 148, 163] as well as to enforce privacy [49]. Nevertheless, it is remarkable that in this sector, privacy concerns seem to be of great importance [49, 215]. Other works tackle intrusion detection [11, 79, 99, 229, 232], secure communication [111, 133], access control [33], secure communication [30], security testing [121, 150] and security training [140].

Water and Wastewater Systems Sector. *Water supply and wastewater treatment facilities are essential for public health and environmental protection.* Interestingly, most works focus on anomaly detection [177, 203, 221]. Some extend their security focus and provide additional attack detection and security assessment mechanisms [34, 105, 160, 169]. One of these works provides a creative approach using a honeypot-like strategy [181].

Six out of the 16 CISA sectors are currently not addressed in terms of digital twin security: *Chemical Sector*, *Dams*, *Defense Industrial Base*, *Emergency Services*, *Financial Services*, and *Government Facilities*. Some of these sectors might be too specific (e.g., Dams) to be targeted. Meanwhile, other sectors can be seen as a subcategory or be strongly related to other fields and, therefore, not specifically addressed. For instance, many works from the critical manufacturing sector can be easily applied in the chemical sector. Also, the dams sector goes hand in hand with the water and waste systems sector. Moreover, the sectors concerning defense, emergency, and government might be too sensitive for publication.



(a) NIST CSF functions.

(b) Identify (ID) categories.

(c) Protect (PR) categories.

(d) Detect (DE) categories.

(e) Respond (RS) categories.

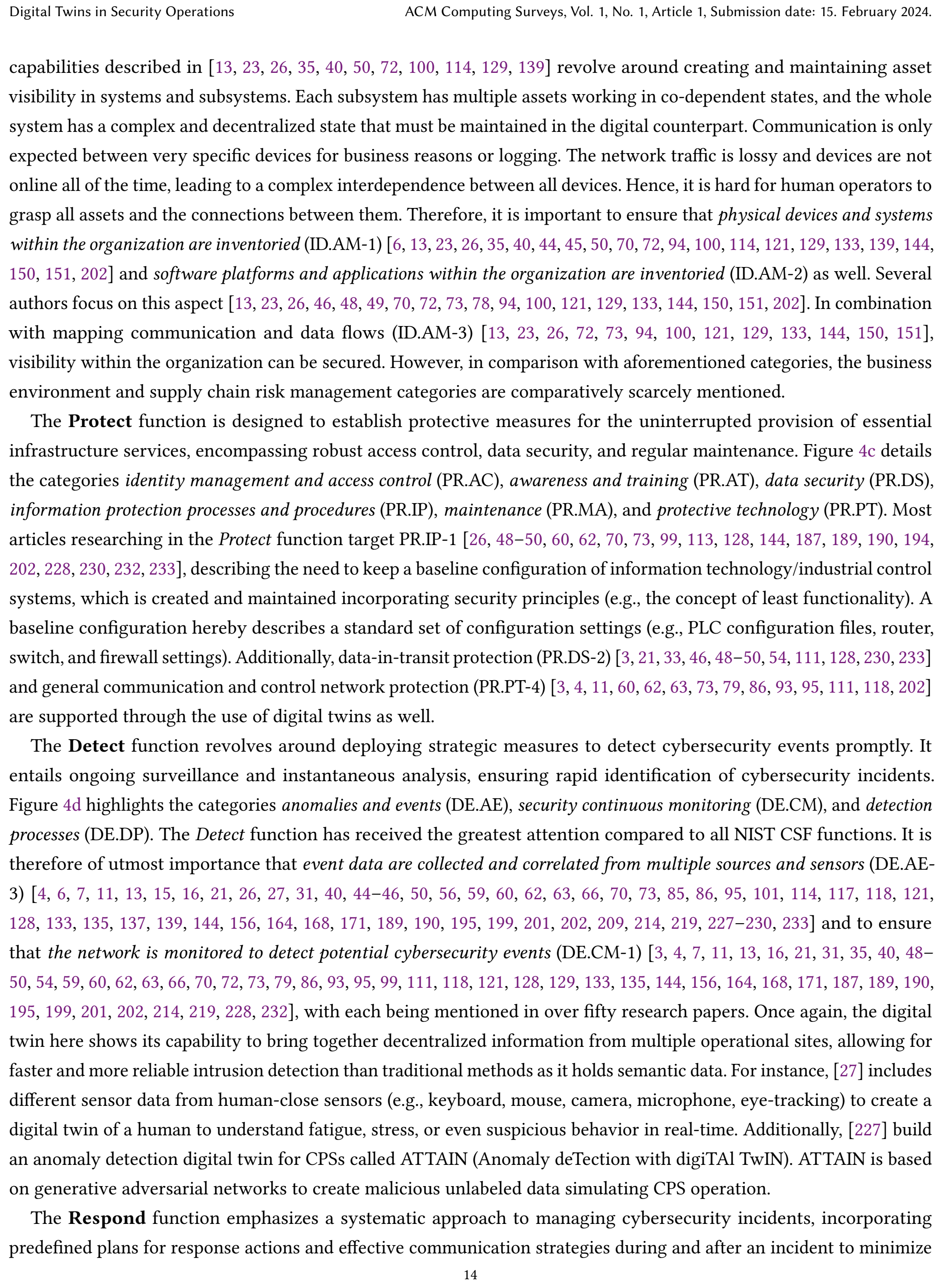
(f) Recover (RC) operations.

Fig. 4. Detailed security operations of digital twins in relation to the NIST Cybersecurity Framwork (CSF) for critical infrastructures if they have been addressed at least once.

5.3 Security Operations

In this section, we broaden our discussion on how digital twins can support security operations in alignment with the NIST CSF towards the diverse range of security operations facilitated by digital twins. Therefore, we conduct a comprehensive analysis and correlate the 111 research papers categorized as highly relevant to our survey with specific NIST CSF subcategories for improving cybersecurity in critical infrastructures. The detailed correlation can be found in Appendix A. The mapping results are illustrated in Figure 4, where Figure 4a provides an abstract view at the CSF function level. Meanwhile, Figures 4b-4f delve into more granular details at the CSF category level. For each category, we summarize whether the subcategories within this category have been addressed, denoting this with a simple binary (yes or no) indicator. For instance, we check whether for the category ID.GV (Identify – Governance), the research covers its subcategories ID.GV-1 to ID.GV-4. Figure 4a allows us to discern, for instance, that all categories of the *Detect* function have been covered in existing research, while not all categories of the *Identify* function have been equally addressed. In the following sections, we will delve into the specifics of the NIST CSF functions: *Identify*, *Protect*, *Detect*, *Respond*, and *Recover* to provide detailed insights how digital twins contribute to each of these security functions.

The **Identify** function involves developing an organizational understanding to manage cybersecurity risk to systems, assets, data, and capabilities. It ensures businesses know and understand their resources, risks, and overall cybersecurity posture. Figure 4b shows its categories *asset management* (ID.AM), *business environment* (ID.BE), *governance* (ID.GV), *risk assessment* (ID.RA), *risk management strategy* (ID.RM), and *supply chain risk management* (ID.SC). The asset management



capabilities described in [13, 23, 26, 35, 40, 50, 72, 100, 114, 129, 139] revolve around creating and maintaining asset visibility in systems and subsystems. Each subsystem has multiple assets working in co-dependent states, and the whole system has a complex and decentralized state that must be maintained in the digital counterpart. Communication is only expected between very specific devices for business reasons or logging. The network traffic is lossy and devices are not online all of the time, leading to a complex interdependence between all devices. Hence, it is hard for human operators to grasp all assets and the connections between them. Therefore, it is important to ensure that *physical devices and systems within the organization are inventoried* (ID.AM-1) [6, 13, 23, 26, 35, 40, 44, 45, 50, 70, 72, 94, 100, 114, 121, 129, 133, 139, 144, 150, 151, 202] and *software platforms and applications within the organization are inventoried* (ID.AM-2) as well. Several authors focus on this aspect [13, 23, 26, 46, 48, 49, 70, 72, 73, 78, 94, 100, 121, 129, 133, 144, 150, 151, 202]. In combination with mapping communication and data flows (ID.AM-3) [13, 23, 26, 72, 73, 94, 100, 121, 129, 133, 144, 150, 151], visibility within the organization can be secured. However, in comparison with aforementioned categories, the business environment and supply chain risk management categories are comparatively scarcely mentioned.

The **Protect** function is designed to establish protective measures for the uninterrupted provision of essential infrastructure services, encompassing robust access control, data security, and regular maintenance. Figure 4c details the categories *identity management and access control* (PR.AC), *awareness and training* (PR.AT), *data security* (PR.DS), *information protection processes and procedures* (PR.IP), *maintenance* (PR.MA), and *protective technology* (PR.PT). Most articles researching in the *Protect* function target PR.IP-1 [26, 48–50, 60, 62, 70, 73, 99, 113, 128, 144, 187, 189, 190, 194, 202, 228, 230, 232, 233], describing the need to keep a baseline configuration of information technology/industrial control systems, which is created and maintained incorporating security principles (e.g., the concept of least functionality). A baseline configuration hereby describes a standard set of configuration settings (e.g., PLC configuration files, router, switch, and firewall settings). Additionally, data-in-transit protection (PR.DS-2) [3, 21, 33, 46, 48–50, 54, 111, 128, 230, 233] and general communication and control network protection (PR.PT-4) [3, 4, 11, 60, 62, 63, 73, 79, 86, 93, 95, 111, 118, 202] are supported through the use of digital twins as well.

The **Detect** function revolves around deploying strategic measures to detect cybersecurity events promptly. It entails ongoing surveillance and instantaneous analysis, ensuring rapid identification of cybersecurity incidents. Figure 4d highlights the categories *anomalies and events* (DE.AE), *security continuous monitoring* (DE.CM), and *detection processes* (DE.DP). The *Detect* function has received the greatest attention compared to all NIST CSF functions. It is therefore of utmost importance that *event data are collected and correlated from multiple sources and sensors* (DE.AE-3) [4, 6, 7, 11, 13, 15, 16, 21, 26, 27, 31, 40, 44–46, 50, 56, 59, 60, 62, 63, 66, 70, 73, 85, 86, 95, 101, 114, 117, 118, 121, 128, 133, 135, 137, 139, 144, 156, 164, 168, 171, 189, 190, 195, 199, 201, 202, 209, 214, 219, 227–230, 233] and to ensure that *the network is monitored to detect potential cybersecurity events* (DE.CM-1) [3, 4, 7, 11, 13, 16, 21, 31, 35, 40, 48–50, 54, 59, 60, 62, 63, 66, 70, 72, 73, 79, 86, 93, 95, 99, 111, 118, 121, 128, 129, 133, 135, 144, 156, 164, 168, 171, 187, 189, 190, 195, 199, 201, 202, 214, 219, 228, 232], with each being mentioned in over fifty research papers. Once again, the digital twin here shows its capability to bring together decentralized information from multiple operational sites, allowing for faster and more reliable intrusion detection than traditional methods as it holds semantic data. For instance, [27] includes different sensor data from human-close sensors (e.g., keyboard, mouse, camera, microphone, eye-tracking) to create a digital twin of a human to understand fatigue, stress, or even suspicious behavior in real-time. Additionally, [227] build an anomaly detection digital twin for CPSs called ATTAIN (Anomaly deTection with digiTAl TwIN). ATTAIN is based on generative adversarial networks to create malicious unlabeled data simulating CPS operation.

The **Respond** function emphasizes a systematic approach to managing cybersecurity incidents, incorporating predefined plans for response actions and effective communication strategies during and after an incident to minimize its impact. Figure 4e showcases the categories *response planning* (RS.RP), *communications* (RS.CO), *analysis* (RS.AN), *mitigation* (RS.MI), and *improvements* (RS.IM). The *Response* function focuses mainly on the control *incidents are reported consistent with established criteria* (RS.CO-2) [1, 2, 6, 7, 13, 16, 48–50, 54, 59, 60, 62, 63, 72, 73, 93, 99, 101, 128, 168, 171, 187, 202, 214, 232], but some papers also aim to highlight the impact of an incident and enrich its context so that the *impact of the incident is understood* (RS.AN-2) [1, 2, 7, 35, 46, 72, 73, 128, 144, 194, 202, 216]. This allows a defender to conduct more in-depth root-cause analyses and respond to the incident faster and more precisely.

The **Recover** function focuses on formulating and executing suitable strategies to uphold resilience plans and rehabilitate any functionalities or services disrupted by a cybersecurity incident. It underscores the criticality of recovery planning, enhancements, and communication efforts to efficiently reinstate services and operations post-incident. Figure 4f illustrates the categories *recovery planing* (RC.RP), *improvements* (RC.IM), and *communications* (RC.CO). Within the *Recover* function, only a few researchers detail digital twin functionalities at all. These functionalities concentrate on the support of the following controls: *Recovery plan is executed during or after a cybersecurity incident* (RC.RP-1) [7, 72, 73, 144, 151, 187, 216], *Recovery strategies are updated* (RC.IM-2) [72, 216], *recovery plans incorporate lessons learned* (RC.IM-1) [73, 144, 216], and *recovery activities are communicated to internal and external stakeholders* as well as *executive and management teams* (RC.CO-3) [7, 72, 73]. However, compared to other functions of the NIST CSF, the *Recover* function is not yet as well researched, and more benefits might be found in the application of digital twins for recovery purposes, e.g., in case of and in response to a ransomware attack.

6 IMPLEMENTING DIGITAL TWINS

In this section, we explore the components of digital twins in security operations in detail. To this end, we provide an overview on how to set up digital twins and discuss supportive tools.

6.1 Physical Assets and Data

Digital twins are virtual representations of physical assets. While physical assets may be resource-efficient devices like sensors or actuators, they could also have more computing resources (e.g., CPS). The following details different data used within a digital twin in security operations.

Sensor and Measurement Data. This category centers data gathered from various sensors, ranging from simple temperature sensors to complex multi-spectral cameras. Research [8, 9, 17, 29, 31, 85, 103, 132, 188, 230] delves into the nuances of sensor data, including its precision, reliability, and the challenges of integrating data from multiple sensors into coherent systems. They explore innovative ways to enhance sensor data accuracy and utility, such as advanced calibration techniques, noise reduction algorithms, and integration strategies for combining data from disparate sources. Additional research [54, 110, 123, 165, 180] investigates the application of sensor data in specific technologies like transformers, inverters, and IoT devices, addressing challenges like real-time data processing and the integration of sensor data into existing systems for monitoring, control, and predictive maintenance.

CPS Data. In the context of CPS, the data primarily revolves around operational parameters, performance metrics, and environmental interactions. The works of [30, 76, 78] represent significant advancements in utilizing data for enhancing vehicle performance and machinery efficiency. This includes the analysis of engine telemetry, fuel efficiency, emission levels, and machinery wear and tear. The data also encompasses feedback from control systems and user inputs, providing a comprehensive view of machine performance in real-world conditions. This contributes to developing more innovative, efficient, safer vehicles and machinery through data-driven optimizations and predictive maintenance.

Table 4. Classification and description of protocols by ISO/OSI layers.

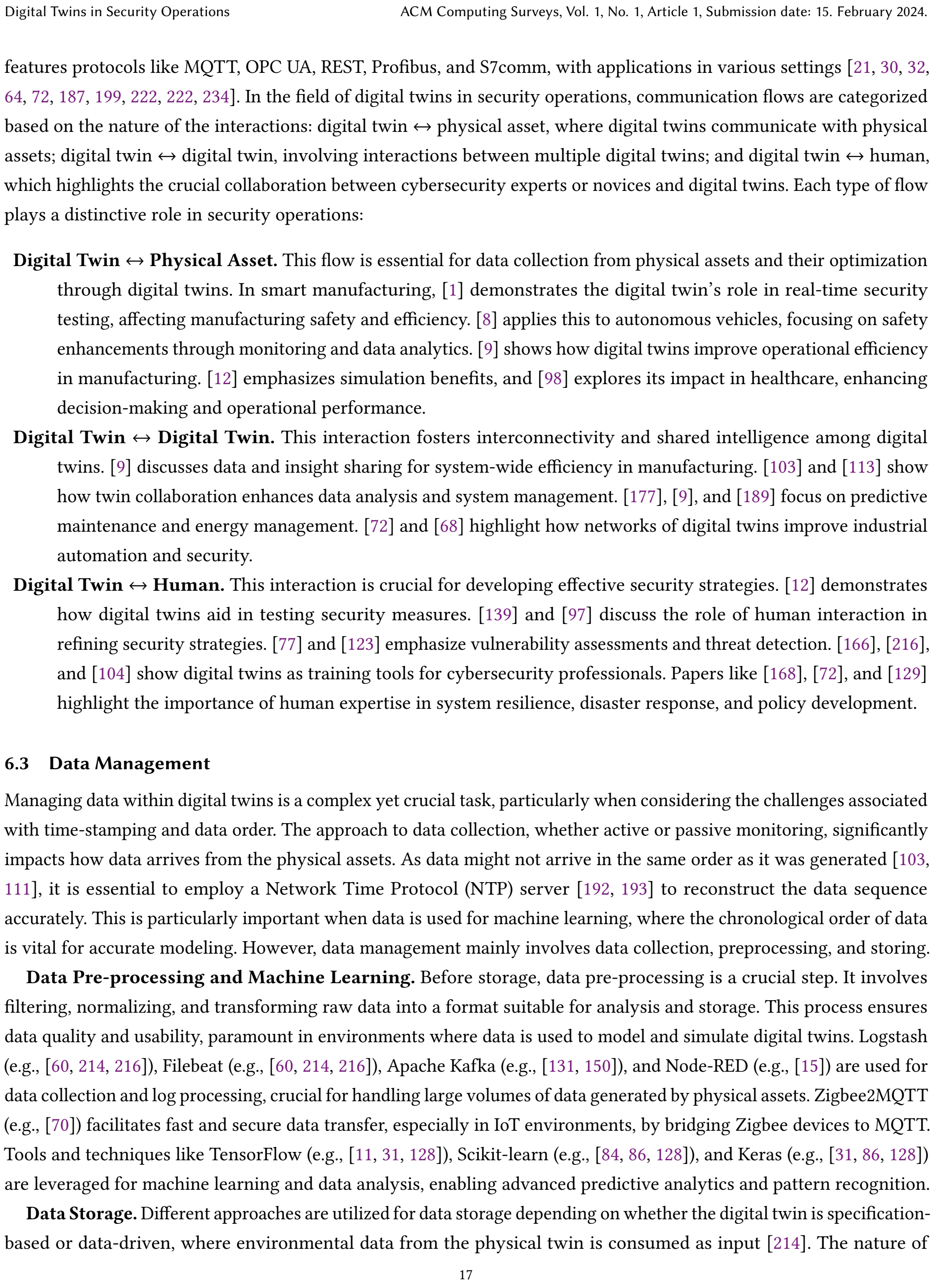
ISO/OSI Layer	Protocol	Description	References
L1 – Physical	Ethernet	Data framing & physical network transmission	[15, 59, 77, 117, 221, 229]
L2 – Data Link	ARP	IP address to physical address resolution	[216]
	WiFi	Wireless LAN communication	[34, 66, 224]
	Bluetooth	Short-range wireless communication	[224]
	Zigbee	Low-power wireless networking protocol	[70, 72, 121, 139]
L3 – Network	LoRaWAN	Long Range WAN protocol	[114, 232]
	DNP3	Distributed Network Protocol for utilities	[28, 77, 202]
L4 – Transport	TCP	Reliable, ordered data transmission	[13, 30, 32, 34, 55, 114, 132, 194, 195, 199, 201, 216]
	UDP	Fast, connectionless communication	[17, 103, 132]
L6 – Presentation	XML	Data format for structured documents	[189, 226]
	JSON	Lightweight data-interchange format	[32, 33, 60, 148, 165, 186, 187, 189]
	AutomationML	Data exchange in automation and engineering	[32]
	CAEX	Computer-Aided Engineering Exchange format	[226]
	SysML	Systems Modeling Language	[65]
	MSDL	Manufacturing Service Description Language	[32]
L7 – Application	AMQP	Messaging protocol for interoperability	[114]
	CoAP	Web transfer protocol for constrained nodes	[30, 32, 54, 185]
	HTTP/HTTPS	HyperText Transfer Protocol (Secure)	[44, 45, 70, 148, 185]
	MQTT	Lightweight messaging protocol for IoT	[13, 15, 30, 32–34, 70, 72, 114, 121, 129, 152, 154, 186, 187, 199, 222, 234]
	OPC UA	Machine to machine communication protocol	[64, 139, 222, 226, 234]
	REST	Representational state transfer	[32, 185, 216]
	SCADA	Supervisory control and data acquisition	[50, 58]
	SPARQL	Query language for graph databases	[40]
	SQLite	High-reliability database engine	[160]
	CIP	Common Industrial Protocol	[221]
	Modbus	Communication protocol for PLCs	[15, 28, 93, 117, 129, 194, 195, 199]
	S7comm	Communication protocol by Siemens	[21, 23, 234]
	Profinet	Industrial Ethernet standard for automation	[21]

Network Traffic Data. Research in this category [1, 13, 25, 37, 54, 55, 59, 62, 66, 103, 118, 128, 160, 166] focuses on the vast and complex datasets generated by network traffic. This data includes packet information, bandwidth usage, network topology, and security logs. The papers aim to optimize network performance, enhance data throughput, and ensure robust security against cyber threats. They leverage data analytics to understand network behavior, identify patterns, and predict potential issues. Their research is pivotal in designing more efficient and secure network systems capable of handling the increasing demands of modern digital communication.

Other Data. Data in this category [132, 140, 166, 216], includes a broad spectrum of information. This might involve user interaction data with devices, providing insights into user behavior, preferences, and use patterns. Additionally, specification sales documentation is used as a form of data, which could include information on product specifications, performance metrics, and customer feedback. Such data is invaluable in understanding how products are used in real-world environments and can guide design, functionality, and user experience improvements.

6.2 Communication & Synchronization

Digital twins communicate on the ISO/OSI stack (see Table 4), while they can interact with their physical counterparts, other digital twins, or human actors. In the physical layer (*ISO/OSI layer 1*), Ethernet is often applied in industrial settings [59, 77, 221]. At the data link layer (*ISO/OSI layer 2*), ARP, Wifi, Bluetooth, and Zigbee are noted for their diverse roles in systems [72, 139, 216, 224]. In the network layer (*ISO/OSI layer 3*), the importance of IP and LoRaWAN is highlighted [13] and [114]. The transport layer (*ISO/OSI layer 4*) sees applications of TCP and UDP [13, 17]. In the presentation layer (*ISO/OSI layer 6*), XML and JSON's roles are mentioned [32]. The application layer (*ISO/OSI layer 7*)



features protocols like MQTT, OPC UA, REST, Profibus, and S7comm, with applications in various settings [21, 30, 32, 64, 72, 187, 199, 222, 222, 234]. In the field of digital twins in security operations, communication flows are categorized based on the nature of the interactions: digital twin ↔ physical asset, where digital twins communicate with physical assets; digital twin ↔ digital twin, involving interactions between multiple digital twins; and digital twin ↔ human, which highlights the crucial collaboration between cybersecurity experts or novices and digital twins. Each type of flow plays a distinctive role in security operations:

Digital Twin ↔ Physical Asset. This flow is essential for data collection from physical assets and their optimization through digital twins. In smart manufacturing, [1] demonstrates the digital twin's role in real-time security testing, affecting manufacturing safety and efficiency. [8] applies this to autonomous vehicles, focusing on safety enhancements through monitoring and data analytics. [9] shows how digital twins improve operational efficiency in manufacturing. [12] emphasizes simulation benefits, and [98] explores its impact in healthcare, enhancing decision-making and operational performance.

Digital Twin ↔ Digital Twin. This interaction fosters interconnectivity and shared intelligence among digital twins. [9] discusses data and insight sharing for system-wide efficiency in manufacturing. [103] and [113] show how twin collaboration enhances data analysis and system management. [177], [9], and [189] focus on predictive maintenance and energy management. [72] and [68] highlight how networks of digital twins improve industrial automation and security.

Digital Twin ↔ Human. This interaction is crucial for developing effective security strategies. [12] demonstrates how digital twins aid in testing security measures. [139] and [97] discuss the role of human interaction in refining security strategies. [77] and [123] emphasize vulnerability assessments and threat detection. [166], [216], and [104] show digital twins as training tools for cybersecurity professionals. Papers like [168], [72], and [129] highlight the importance of human expertise in system resilience, disaster response, and policy development.

6.3 Data Management

Managing data within digital twins is a complex yet crucial task, particularly when considering the challenges associated with time-stamping and data order. The approach to data collection, whether active or passive monitoring, significantly impacts how data arrives from the physical assets. As data might not arrive in the same order as it was generated [103, 111], it is essential to employ a Network Time Protocol (NTP) server [192, 193] to reconstruct the data sequence accurately. This is particularly important when data is used for machine learning, where the chronological order of data is vital for accurate modeling. However, data management mainly involves data collection, preprocessing, and storing.

Data Pre-processing and Machine Learning. Before storage, data pre-processing is a crucial step. It involves filtering, normalizing, and transforming raw data into a format suitable for analysis and storage. This process ensures data quality and usability, paramount in environments where data is used to model and simulate digital twins. Logstash (e.g., [60, 214, 216]), Filebeat (e.g., [60, 214, 216]), Apache Kafka (e.g., [131, 150]), and Node-RED (e.g., [15]) are used for data collection and log processing, crucial for handling large volumes of data generated by physical assets. Zigbee2MQTT (e.g., [70]) facilitates fast and secure data transfer, especially in IoT environments, by bridging Zigbee devices to MQTT. Tools and techniques like TensorFlow (e.g., [11, 31, 128]), Scikit-learn (e.g., [84, 86, 128]), and Keras (e.g., [31, 86, 128]) are leveraged for machine learning and data analysis, enabling advanced predictive analytics and pattern recognition.

Data Storage. Different approaches are utilized for data storage depending on whether the digital twin is specification-based or data-driven, where environmental data from the physical twin is consumed as input [214]. The nature of collected data varies across papers, depending real world asset and the twin's intended purpose. Beside SQL and NoSQL databases like MongoDB (e.g., [70, 71]) or PostgreSQL (e.g., [101]), Elasticsearch (e.g., [27, 60, 214, 216]) or OpenSearch (e.g., [44, 45]) are employed for their powerful search and data indexing capabilities, ideal for the large-scale data storage needs of digital twins. Ethereum (e.g., [128, 172]) and Hyperledger Fabric (e.g., [172]) provide distributed ledger technologies for secure and immutable data storage, offering reliability and transparency in data management.

Supportive Software. Platforms provide architectural patterns and simplify digital twin (data) management. For instance, cloud platforms like AWS (IoT) (e.g., [32, 33, 187, 199]), Azure (IoT) (e.g., [32, 40]), Azure Digital Twins (e.g., [4]), and IBM Watson IoT (e.g., [32]) provide robust infrastructure for hosting and managing digital twins, offering scalability and high availability. Open source tools such as Eclipse Ditto (e.g., [32, 72, 158, 196]), OpenStack (e.g., [93, 166]), Kubernetes (e.g., [2, 3, 144]), and FIWARE (e.g., [44, 45]) facilitate the deployment and orchestration of digital twin applications, ensuring efficient and flexible management of digital twin environments. In addition to these, other tools such as Google Functions, AWS Lambda, and Amazon SageMaker (e.g., [73]) can be integrated to extend capabilities for data processing and analytics. These tools and platforms collectively enable the comprehensive collection, storage, processing, and analysis of digital twin data.

Deployment Strategy. The location of digital twin deployment dramatically affects their performance and utility. On-device or embedded digital twins, as noted in [29], [54], and [57], offer low latency and immediate responsiveness, crucial for rapid decision-making. However, they face limitations in processing power and storage capacity, affecting the complexity of tasks they can manage. Edge computing, discussed in [84] and [85], mitigates some constraints by processing data near its source. This approach offers lower latency than cloud solutions and significant computational resources, proving beneficial in industrial settings for real-time data processing. Cloud-based digital twins, referenced in [1] and [8], provide expansive computational resources and storage. This deployment fits complex simulations and large-scale data analytics, which is ideal for in-depth analysis and strategic planning where latency is less critical. Each deployment model — embedded, edge, or cloud — has unique benefits and limitations. The selection depends on the security operation's specific needs, considering latency, computational power, and data processing scope.

6.4 Virtual Representation & Environment

Creating virtual models in digital twin research starts with a deep understanding of the systems involved [1, 9]. This process involves data collection and analysis to develop digital twins that accurately reflect their physical counterparts [14, 20]. As Figure 5 illustrates, digital twin research encompasses three core models: structural, functional, and behavioral. Structural models, using formats like JSON, XML, OWL, or RDF, represent the physical aspects of systems [1, 32, 33, 40, 60, 148, 165, 185–187, 189, 226]. These models descriptively represent a system. Functional models, often employing AutomationML, detail the operational processes and logic [30, 58, 62–64, 66, 84, 160, 179, 222]. They illustrate the interactions within a system. Behavioral models focus on dynamic behaviors over time, using approaches like finite state machines [21, 32, 32, 43, 59, 62, 122, 161, 172, 185, 206]. Figure 5 also shows that these models can be created either data-driven, like from sensor data [30, 63, 72, 84, 179], or specification-based, as in the case of security expert inputs [1, 9, 17]. Please note that some papers may have ambiguities and can be assigned to multiple categories.

Data-Driven Creation. In the data-driven approach to creating digital twins, the emphasis is on leveraging real-time data to construct and refine the digital twin. This method is highlighted in [30], [84], or [72] and primarily relies on the integration of IoT technologies for continuous state updates. The real-time data collected from IoT devices provides a dynamic and evolving representation of the physical counterpart. For instance, [63, 179] further demonstrate how advanced data analytics are employed to process and interpret this vast stream of information, facilitating the creation

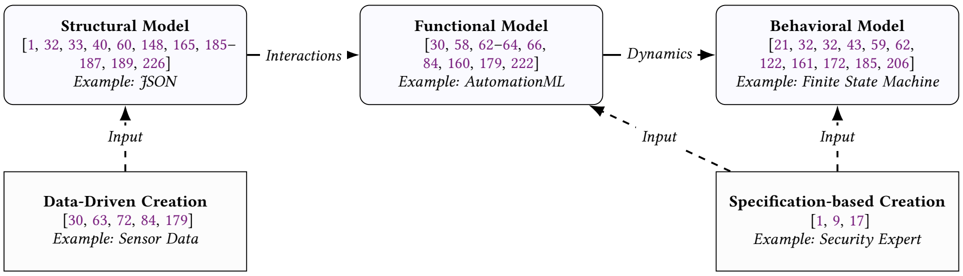


Fig. 5. Illustration of structural, functional, and behavioral models in digital twin research.

of a digital twin that is responsive and up-to-date. This approach is efficient in environments where operational data is abundant and can be used to update and improve the digital twin's accuracy continuously. The data-driven model excels in adaptability, reflecting the ever-changing state of the physical system it represents.

Specification-based Creation. Conversely, the specification-based approach to digital twin creation focuses on constructing a detailed and predefined functional or behavioral model of a system. This method is exemplified by [1], [9], or [17] and involves the use of simulation tools such as Simulink or Anylogic and security expert input to craft a virtual model that mirrors the physical system meticulously. The approach is grounded in deeply understanding the system's processes and operations. Furthermore, applying algorithms like the Kalman Filter and SVM, as seen in [1], enhances the model's capability to simulate complex environments. The specification-based approach is particularly suited to scenarios where the physical system's behavior is well-understood or documented and can be accurately represented through established models. This approach ensures control and predictability in the digital twin's behavior.

6.5 Simulation Capabilities

Simulations can be either in real-time or offline. Real-time simulation runs at the same pace as the real-world system, offering immediate feedback and interaction, while offline simulation operates at variable speeds for in-depth analysis, planning, and optimization without real-time processing constraints. The choice between them depends on the desired security operation. These simulations are supported through various tools.

Real-time Simulation. This type of simulation operates synchronously with system time. It supports *Detect* (DE) and *Respond* (RS) functions of the NIST CSF. In the *Detect* function, particularly in categories like anomalies and events (DE.AE) and security continuous monitoring (DE.CM), real-time simulations are crucial for the immediate collection and correlation of event data from diverse sources and sensors. This is further crucial in the prompt detection of unusual activities or cybersecurity incidents, thereby enhancing an organization's readiness and response capabilities, as demonstrated in papers like [60], [17], and [63]. In the *Respond* function, real-time simulations support the rapid execution of response plans (RS.RP-1) and consistent reporting of incidents (RS.CO-2), facilitating swift and efficient actions against cybersecurity threats. This minimizes potential damage and aids in quicker recovery [55, 59].

Offline Simulation. Offline simulations run at different (lower or higher) system speeds compared to the real world, like a Monte Carlo simulation. They significantly contribute to the *Identify* (ID) and *Protect* (PR) functions of the

Table 5. Classification of simulation technologies in digital twins.

Domain	Simulation Technology	Open Source	Simulation Type	References
General	Anylogic		Real-time/Offline	[9]
	ANSYS Twin Builder		Offline	[179]
Cyber-Physical Systems	CPS Twinning	●	Real-time	[62–64, 66]
	MATLAB		Real-time/Offline	[6, 17, 84, 129, 131, 169, 186, 199, 234]
	Simulink		Real-time/Offline	[1, 15, 17, 50, 103, 169, 186, 199]
Automation Systems	ScadaBR	●	Real-time	[56]
	OpenPLC	●	Real-time	[55], [56], [77], [226]
Robotics	Gazebo	●	Real-time/Offline	[27]
	LabVIEW		Real-time/Offline	[75]
Networks	Mininet	●	Real-time	[60, 62–64, 66, 160]
	Cooja	●	Real-time	[118]
Water Systems	EPANET	●	Real-time/Offline	[160, 169]
	DHALSIM	●	Real-time	[160]
Industrial Simulations	SIMIT		Real-time	[168]
	Tecnomatrix Simulation		Offline	[234]
Electrical and Power Systems	SimpowerSystems		Real-time/Offline	[50]

NIST CSF. In the "Identify" function, offline simulations assist in comprehensively understanding and documenting an organization's assets, such as in ID.AM-1 and ID.AM-2, where physical devices and software platforms are inventoried. This role is underscored in papers like [13] and [23], highlighting the importance of offline simulations in asset management and risk assessment (ID.RA). These simulations provide a detailed view of the organization's resources and vulnerabilities, enabling better strategic planning and prioritization. In the *Protect* function, offline simulations aid in establishing safeguards and maintaining baseline configurations of IT and control systems (PR.IP-1). They offer a controlled environment to test and validate security operations without impacting operations [66].

Simulation Tools. To support these simulations, software is pivotal, catering to specific domains and types (see Table 5). General-domain tools like Anylogic (e.g., [9]) and ANSYS Twin Builder (e.g., [9, 179]) offer wide-ranging capabilities, from simulating physical systems to electronic. For cyber-physical systems, CPS Twinning integrates physical and computational models [63], complemented by MATLAB (e.g., [17]) and Simulink (e.g., [1]) for computation and multi-domain simulation. Automation systems benefit from ScadaBR's focus on supervisory control [56] and OpenPLC's open-source approach [55]. Robotics simulations are adeptly handled by Gazebo (e.g., [27]) and LabVIEW (e.g., [75]), while network environments are emulated by Mininet (e.g., [60]) and Cooja (e.g., [118]). EPANET (e.g., [118]) and DHALSIM (e.g., [160]) offer specialized solutions for water systems analysis, and industrial simulations are advanced by SIMIT (e.g., [168]) and Tecnomatrix Simulation (e.g., [234]) for process optimization. Lastly, electrical and power systems find a niche in SimpowerSystems [50], underlining the comprehensive scope of simulation technologies.

6.6 Security Operations

In the realm of digital twins, security operations vary significantly based on their specific functions, particularly within the NIST CSF. For example, intrusion detection often involves unidirectional communication from the physical asset to the digital twin, focusing on data analysis for attack detection without immediate system response. The digital twin serves as a data processor, identifying potential threats without directly altering the physical system. In contrast, like those managed by SOAR platforms, incident response operations require communication from the digital twin to the physical asset. Here, the digital twin actively orchestrates and implements security responses to threats, playing a crucial role in timely mitigating security incidents. The diversity of these security operations highlights the complexity of cybersecurity in digital twins. There is no universal best practice for designing security operation applications due to the variety of available tools and distinct system requirements. Each function of the NIST CSF – Identify, Protect, Detect, Respond, and Recover – demands specific tools and methodologies tailored to particular security needs. Therefore, selecting tools for digital twin-based security operations must be customized to each specific function's unique challenges and organizational specifics.

However, in developing security operations based on digital twins, various programming languages, frameworks, and tools strategically support security operations by optimizing functionality, enhancing user experience, and bolstering security. Programming languages like C++ [84, 224], known for its performance efficiency, and Python [33, 129, 148, 199, 221, 224], celebrated for its versatility and user-friendliness, play a pivotal role in the backbone of digital twin development. Javascript [216] is commonly used for its effectiveness in web-based applications, while Docker [70, 129] ensures application consistency across various environments. Frameworks such as Flask [70, 148, 216] enable the creation of robust web applications, while MATLAB App Designer [50] is instrumental in designing interactive apps. Nginx [2, 3] serves as a reliable web server and reverse proxy. In the realm of tools, SDN-Wise [118] is utilized for software-defined networking, enhancing network management, and security. Ettercap [37, 60, 216] is a crucial resource for network security, specializing in packet sniffing and protocol analysis. Tools like Nmap [14, 56, 70] and Nessus [14] are key for network mapping and vulnerability scanning, respectively. Pfsense [37, 121] offers firewall and routing capabilities, whereas Scapy [216], Dsiem [60, 214, 216], Beacon [179], and KDiff3 [37] provide various functionalities ranging from packet manipulation to security information and event management. Additional security-focused tools such as Metasploit [121] for penetration testing, Security Onion [121] for network security monitoring, and OSSIM AlienVault [121] for security information and event management, also contribute to security operations.

6.7 User Interaction

In the ecosystem of digital twins, user interaction often transcends direct communication with the digital twin itself. Instead, users typically engage with digital twins through specialized applications designed for security operations, such as cyber ranges or SOAR platforms, as discussed in papers [17, 72, 104, 166]. These applications act as intermediaries, facilitating user interaction with the underlying digital twin. A cyber range, for instance, is an interactive, simulated representation of an OT environment used for cybersecurity training. In this context, the digital twin is the foundational (simulation) model, mimicking real-world Operational Technology (OT) systems and networks, as elaborated in [212] and [216]. Users interact with the cyber range, not directly with the digital twin, but through the scenarios and challenges presented within the cyber range environment, providing realistic, hands-on experience in a controlled setting. Similarly, SOAR platforms integrate with digital twins to enhance cybersecurity incident response and management [72]. In these platforms, the digital twin provides a dynamic system representation using structural models, enabling users to visualize, analyze, and respond to security threats more effectively, as detailed in [43] and [72]. Users interact with the SOAR platform's interface, leveraging the insights the digital twin offers to make informed decisions and automate responses to security incidents through playbooks.

Different specialized technologies and tools simplify the access to digital twins, each adding a unique dimension to its functionality ("*analytics mode*"). Polygon Cruncher, as referenced in [139], plays a crucial role in optimizing 3D models, enhancing the efficiency of digital twin representations. Kibana, highlighted in [27, 60, 214], serves as an invaluable tool for data visualization, enabling users to analyze and interpret complex data streams effectively. Prometheus, cited in [2], offers critical capabilities in monitoring and alerting, ensuring the digital twin systems perform optimally.

InTouch software, as seen in [37], is pivotal for industrial automation, providing robust management and visualization of operational data within digital twins. Lastly, WinCC, mentioned in [21], aids in process visualization for automated environments, playing a vital role in the control and oversight of digital twin systems. Together, these technologies and tools underscore the complexity and technological sophistication of accessing digital twins, highlighting their versatility and the necessity of a diverse toolkit to unlock their full potential.

6.8 Lifecycle Management

When discussing digital twins, it is necessary to consider lifecycle management, focusing on designing, operating, and integrating digital twins in security operations. In the design phase, the "security by design" principle can be integrated the foster cybersecurity at a foundational level [56, 185]. This approach embeds security into the system's architecture, making it inherent rather than supplementary, addressing vulnerabilities and establishing a robust security baseline. As the system evolves into production and operation, the digital twin becomes vital for real-time security management, enabling swift threat identification and mitigation [93, 207]. Moreover, maintaining a secure digital thread throughout the system's lifecycle [64] enables continuity and integrity of cybersecurity information. This thread provides a comprehensive, real-time view of the system's security, crucial for adapting to evolving cybersecurity challenges. Cybersecurity is a lifecycle mirroring all NIST CSF phases. Of course, research like [1] and [12] can only cover one security operation in a particular phase, e.g., intrusion detection. Nevertheless, we must understand that digital twins ensure proactive cybersecurity at all life cycle phases, from initial design to ongoing maintenance, response strategies, and recovery [13, 64]. Recognizing that digital twins offer a sophisticated approach to security operations across an asset's entire lifecycle allows us to adapt and counter the ever-evolving landscape of cybersecurity.

7 CHALLENGES AND FUTURE RESEARCH

This section explores the challenges in digital twins in security operations. While the challenges point to future research, we additionally discuss future research beyond these challenges.

7.1 Challenges

Lack of Standards. Digital twins face challenges due to the absence of standards, especially in diverse real-world OT environments. These environments, often built upon existing infrastructures, vary significantly in design and technology, creating hurdles for developing standardized digital twins. Each system is unique, and consequently, digital twins are typically custom-made and implemented as an afterthought, making widespread adoption challenging [178, 207, 211].

Automated Vulnerability Response. The dynamic nature of industrial environments makes manual vulnerability assessments impractical, highlighting the need for effective automated methods. Digital twins could play a crucial role in continuously monitoring and identifying vulnerabilities. The challenge lies in developing systems that can quickly detect weaknesses and automatically respond to the complexities and dependencies within these systems.

High-Level Guidance and Access Management. Establishing high-level guidance and managing access to digital twins are vital challenges. Defining strategic frameworks that align with organizational objectives is complex due to the diverse applications of digital twins. Moreover, effective access management is critical to prevent unauthorized use, requiring a delicate balance between collaborative openness and strict control of access, emphasizing the need for clear policies and robust governance which is especially challenging within heterogeneous IoT environments [147, 159].

Knowledge Requirements. Developing digital twins in security operations requires a unique blend of OT systems and cybersecurity expertise. There is a noticeable gap between the knowledge of cybersecurity experts and OT engineers, with the former often needing more insight into OT systems and the latter focusing more on operational excellence than security. This creates a demand for specialists proficient in both domains to develop effective digital twins for complex environments. Failing to adhere to such goals will limit the implementation success for creation and adoption of digital twins as the vast majority of companies and users will not be able to deploy an army of engineers to create custom digital twins and therefore the adoption of digital twins may be restricted to an industrial oligopoly [200].

OT/IoT Network Assets Visibility. Gaining comprehensive visibility into OT and IoT networks is challenging, especially with devices that are not continuously connected. The first step in OT asset management is constantly gaining clear visibility, guided by the principle that effective management starts with thorough visibility of the network's assets [49, 103]. Note that field devices or sensors are hidden in IP networks (see Purdue model).

Communication Delays and Model Attacks. Digital twins in OT environments are susceptible to communication delays and model attacks. When applied in real-world scenarios, these issues can lead to performance problems like network congestion or computational overload, underscoring the importance of addressing these challenges for the successful implementation of digital twins [111, 139].

Cybersecurity Concerns. Security vulnerabilities are significant concerns in digital twin development. Synchronization security, resilience against DoS attacks, and data privacy are critical. The insecurity of IoT devices and the lack of standards for securing digital twins add to these concerns, making them potential targets for cyber attacks [108, 168, 176].

Data Management and Quality. Ensuring high-quality data management is a crucial challenge in digital twins, especially when using machine learning algorithms. Handling data scarcity or biases and ensuring the quality of the data are pivotal for the accuracy and effectiveness of digital twins.

Fidelity and Accuracy. Maintaining high fidelity and accuracy in digital twin models is crucial but challenging. The diversity of OT environments and data quality issues make it difficult to achieve the desired level of model accuracy. Moreover, OT engineers prioritize operational continuity, making them cautious about adopting new technologies that might disrupt manufacturing processes [79, 161, 198].

Interoperability and Integration. Achieving interoperability and seamless integration with existing systems is a significant hurdle. Digital twins must be able to adapt to various protocols and integrate with different management tools, which is often challenging due to the proprietary nature of many OT protocols [145].

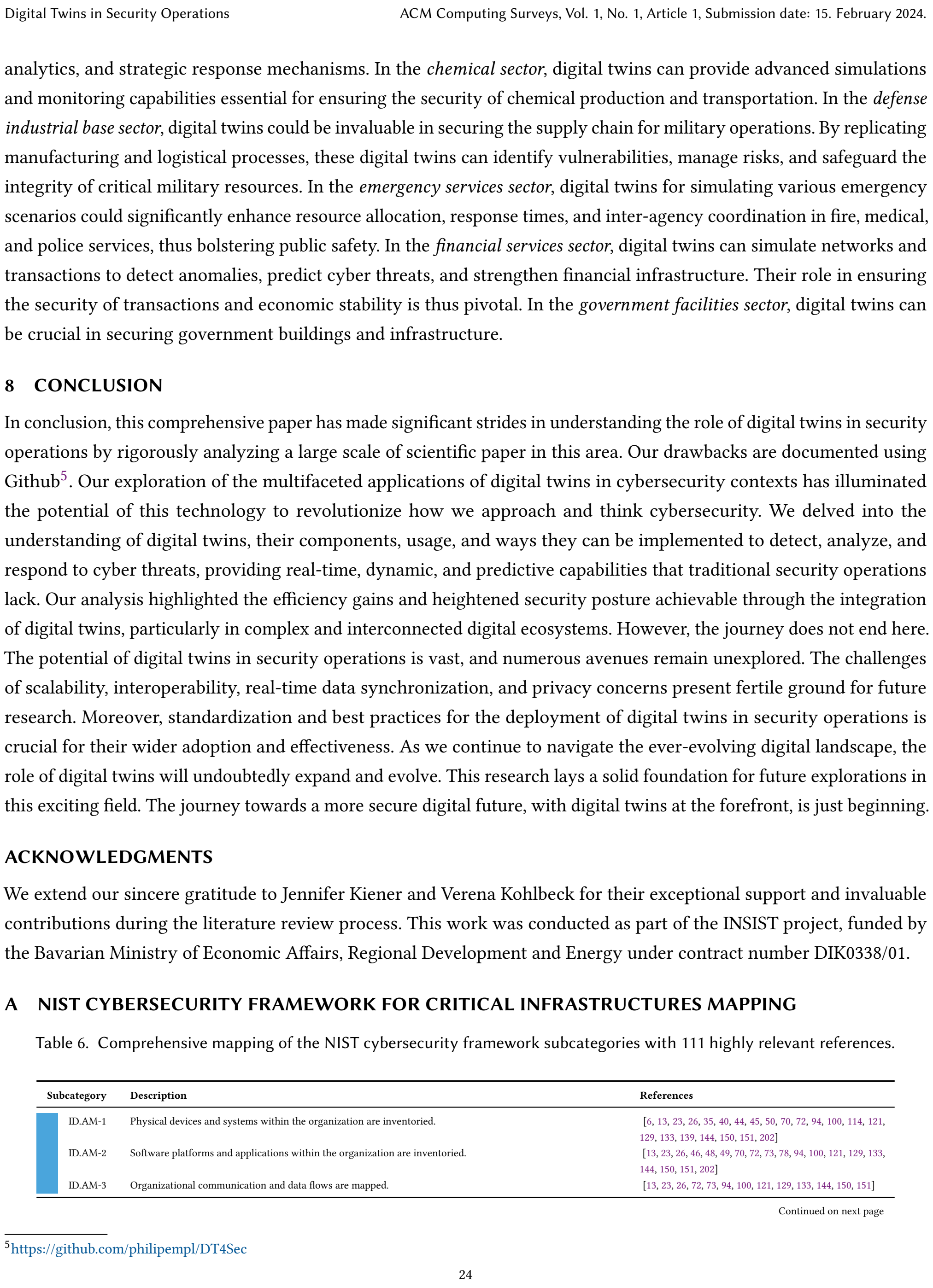
Explainable AI. The use of explainable AI in digital twins is essential for transparency and trust. This is especially important in safety-critical systems, where understanding AI decisions is crucial. Challenges include addressing the black-box nature of AI, ensuring model transparency, and protecting against attacks on the AI model itself [95, 168, 214].

7.2 Future Research

Future research endeavors should address the previously outlined challenges and explore the untapped potential of digital twins in lesser-investigated NIST CSF functions and application domains.

Governance, Risk Management, Incident Response, and Recovery. Within the NIST CSF framework, several subcategories are yet to be fully explored, presenting ripe opportunities for future research. These include governance and risk management (ID.GV-1 to ID.GV-4), remote maintenance (PR.MA-2), coordination with stakeholders (RS.CO-4), incorporation of lessons learned in response plans (RS.IM-1), and recovery and public relations management (RC.CO-1 to RC.CO-2). These areas lack extensive references and could significantly benefit from digital twins. By delving into these subcategories, researchers can uncover new ways to enhance security operations.

Exploring Application Domains for Digital Twins in Security Operations. The application of digital twins in various sectors promises to revolutionize security operations through enhanced situational awareness, predictive



analytics, and strategic response mechanisms. In the *chemical sector*, digital twins can provide advanced simulations and monitoring capabilities essential for ensuring the security of chemical production and transportation. In the *defense industrial base sector*, digital twins could be invaluable in securing the supply chain for military operations. By replicating manufacturing and logistical processes, these digital twins can identify vulnerabilities, manage risks, and safeguard the integrity of critical military resources. In the *emergency services sector*, digital twins for simulating various emergency scenarios could significantly enhance resource allocation, response times, and inter-agency coordination in fire, medical, and police services, thus bolstering public safety. In the *financial services sector*, digital twins can simulate networks and transactions to detect anomalies, predict cyber threats, and strengthen financial infrastructure. Their role in ensuring the security of transactions and economic stability is thus pivotal. In the *government facilities sector*, digital twins can be crucial in securing government buildings and infrastructure.

8 CONCLUSION

In conclusion, this comprehensive paper has made significant strides in understanding the role of digital twins in security operations by rigorously analyzing a large scale of scientific paper in this area. Our drawbacks are documented using Github[5]. Our exploration of the multifaceted applications of digital twins in cybersecurity contexts has illuminated the potential of this technology to revolutionize how we approach and think cybersecurity. We delved into the understanding of digital twins, their components, usage, and ways they can be implemented to detect, analyze, and respond to cyber threats, providing real-time, dynamic, and predictive capabilities that traditional security operations lack. Our analysis highlighted the efficiency gains and heightened security posture achievable through the integration of digital twins, particularly in complex and interconnected digital ecosystems. However, the journey does not end here. The potential of digital twins in security operations is vast, and numerous avenues remain unexplored. The challenges of scalability, interoperability, real-time data synchronization, and privacy concerns present fertile ground for future research. Moreover, standardization and best practices for the deployment of digital twins in security operations is crucial for their wider adoption and effectiveness. As we continue to navigate the ever-evolving digital landscape, the role of digital twins will undoubtedly expand and evolve. This research lays a solid foundation for future explorations in this exciting field. The journey towards a more secure digital future, with digital twins at the forefront, is just beginning.

ACKNOWLEDGMENTS

We extend our sincere gratitude to Jennifer Kiener and Verena Kohlbeck for their exceptional support and invaluable contributions during the literature review process. This work was conducted as part of the INSIST project, funded by the Bavarian Ministry of Economic Affairs, Regional Development and Energy under contract number DIK0338/01.

A NIST CYBERSECURITY FRAMEWORK FOR CRITICAL INFRASTRUCTURES MAPPING

Table 6. Comprehensive mapping of the NIST cybersecurity framework subcategories with 111 highly relevant references.

Subcategory	Description	References
ID.AM-1	Physical devices and systems within the organization are inventoried.	[6, 13, 23, 26, 35, 40, 44, 45, 50, 70, 72, 94, 100, 114, 121, 129, 133, 139, 144, 150, 151, 202]
ID.AM-2	Software platforms and applications within the organization are inventoried.	[13, 23, 26, 46, 48, 49, 70, 72, 73, 78, 94, 100, 121, 129, 133, 144, 150, 151, 202]
ID.AM-3	Organizational communication and data flows are mapped.	[13, 23, 26, 72, 73, 94, 100, 121, 129, 133, 144, 150, 151]

Continued on next page

[5] https://github.com/philipempl/DT4Sec

Table 6 – continued

Subcategory	Description	References
ID.AM-4	External information systems are catalogued.	[23, 72, 73, 94, 150]
ID.AM-5	Resources (e.g., hardware, devices, data, time, personnel, and software) are prioritized based on their classification, criticality, and business value.	[23, 48, 49, 73, 144, 150]
ID.AM-6	Cybersecurity roles and responsibilities for the entire workforce and third-party stakeholders (e.g., suppliers, customers, partners) are established.	
ID.BE-1	The organization's role in the supply chain is identified and communicated.	[73, 78, 114]
ID.BE-2	The organization's place in critical infrastructure and its industry sector is identified and communicated.	[35, 73, 114]
ID.BE-3	Priorities for organizational mission, objectives, and activities are established and communicated.	[35]
ID.BE-4	Dependencies and critical functions for delivery of critical services are established.	[35, 78, 94, 144]
ID.BE-5	Resilience requirements to support delivery of critical services are established for all operating states (e.g. under duress/attack, during recovery, normal operations).	[35, 78, 144, 187, 201]
ID.GV-1	Organizational cybersecurity policy is established and communicated.	
ID.GV-2	Cybersecurity roles and responsibilities are coordinated and aligned with internal roles and external partners.	
ID.GV-3	Legal and regulatory requirements regarding cybersecurity, including privacy and civil liberties obligations, are understood and managed.	
ID.GV-4	Governance and risk management processes address cybersecurity risks.	
ID.RA-1	Asset vulnerabilities are identified and documented.	[13, 23, 35, 46, 53, 72, 73, 78, 144]
ID.RA-2	Cyber threat intelligence is received from information sharing forums and sources.	[13, 72, 73, 78, 100, 144]
ID.RA-3	Threats, both internal and external, are identified and documented.	[13, 35, 70, 73, 78, 94, 100, 144, 202]
ID.RA-4	Potential business impacts and likelihoods are identified.	[73, 78, 94, 114, 144, 202]
ID.RA-5	Threats, vulnerabilities, likelihoods, and impacts are used to determine risk.	[46, 73, 94, 144]
ID.RA-6	Risk responses are identified and prioritized.	[73, 94, 187]
ID.RM-1	Risk management processes are established, managed, and agreed to by organizational stakeholders.	[73, 100]
ID.RM-2	Organizational risk tolerance is determined and clearly expressed.	
ID.RM-3	The organization's determination of risk tolerance is informed by its role in critical infrastructure and sector specific risk analysis.	
ID.SC-1	Cyber supply chain risk management processes are identified, established, assessed, managed, and agreed to by organizational stakeholders.	[46]
ID.SC-2	Suppliers and third party partners of information systems, components, and services are identified, prioritized, and assessed using a cyber supply chain risk assessment process.	[46]
ID.SC-3	Contracts with suppliers and third-party partners are used to implement appropriate measures designed to meet the objectives of an organization's cybersecurity program and Cyber Supply Chain Risk Management Plan.	[46]
ID.SC-4	Suppliers and third-party partners are routinely assessed using audits, test results, or other forms of evaluations to confirm they are meeting their contractual obligations.	[46, 78]
ID.SC-5	Response and recovery planning and testing are conducted with suppliers and third-party providers	
PR.AC-1	Identities and credentials are issued, managed, verified, revoked, and audited for authorized devices, users and processes.	[21, 128, 172]
PR.AC-2	Physical access to assets is managed and protected	[21, 73, 114, 180]
PR.AC-3	Remote access is managed	[21, 33, 54, 73, 113, 172]
PR.AC-4	Access permissions and authorizations are managed, incorporating the principles of least privilege and separation of duties.	[33, 48, 49, 54, 73, 113, 128, 172]
PR.AC-5	Network integrity is protected (e.g., network segregation, network segmentation).	[48, 49, 54, 73, 128]
PR.AC-6	Identities are proofed and bound to credentials and asserted in interactions.	[33, 48, 49, 73, 113, 128, 172]
PR.AC-7	Users, devices, and other assets are authenticated (e.g., single-factor, multi-factor) commensurate with the risk of the transaction (e.g., individuals' security and privacy risks and other organizational risks).	[33, 48, 49, 73, 114, 172]
PR.AT-1	All users are informed and trained.	[20, 34, 66, 92, 151, 202, 204, 216]
PR.AT-2	Privileged users understand their roles and responsibilities.	[20, 34, 79, 151, 202, 216]
PR.AT-3	Third-party stakeholders (e.g., suppliers, customers, partners) understand their roles and responsibilities.	
PR.AT-4	Senior executives understand their roles and responsibilities.	
PR.AT-5	Physical and cybersecurity personnel understand their roles and responsibilities.	[20, 34, 60, 79, 151, 202, 204, 216]
PR.DS-1	Data-at-rest is protected.	[21, 33, 46, 48, 49, 99, 128]
PR.DS-2	Data-in-transit is protected.	[3, 21, 33, 46, 48–50, 54, 111, 128, 230, 233]
PR.DS-3	Assets are formally managed throughout removal, transfers, and disposition.	[26, 48, 49, 73, 93]
PR.DS-4	Adequate capacity to ensure availability is maintained.	[50, 73]
PR.DS-5	Protections against data leaks are implemented.	[46, 48, 49, 54, 70, 73]
PR.DS-6	Integrity checking mechanisms are used to verify software, firmware, and information integrity.	[46, 50, 62, 63, 73, 78, 128]
PR.DS-7	The development and testing environment(s) are separate from the production environment.	[1, 2, 13, 46, 50, 63, 194, 202]
PR.DS-8	Integrity checking mechanisms are used to verify hardware integrity.	[6, 21]
PR.IP-1	A baseline configuration of information technology/industrial control systems is created and maintained incorporating security principles (e.g. concept of least functionality).	[26, 48–50, 60, 62, 70, 73, 99, 113, 128, 144, 187, 189, 190, 194, 202, 228, 230, 232, 233]
PR.IP-2	A System Development Life Cycle to manage systems is implemented.	[46, 144]
PR.IP-3	Configuration change control processes are in place.	[46, 60, 93, 144, 195, 202]
PR.IP-4	Backups of information are conducted, maintained, and tested.	[60, 144]
PR.IP-5	Policy and regulations regarding the physical operating environment for organizational assets are met.	[35, 48, 49, 63, 144]
PR.IP-6	Data is destroyed according to policy.	[48, 49, 128]
PR.IP-7	Protection processes are improved.	[7, 46, 51, 56, 63, 72, 144, 164, 194, 202]
PR.IP-8	Effectiveness of protection technologies is shared.	
PR.IP-9	Response plans (Incident Response and Business Continuity) and recovery plans (Incident Recovery and Disaster Recovery) are in place and managed.	[7, 72, 73, 144, 202]
PR.IP-10	Response and recovery plans are tested.	[7, 72, 73, 144, 180, 202, 216]

Continued on next page

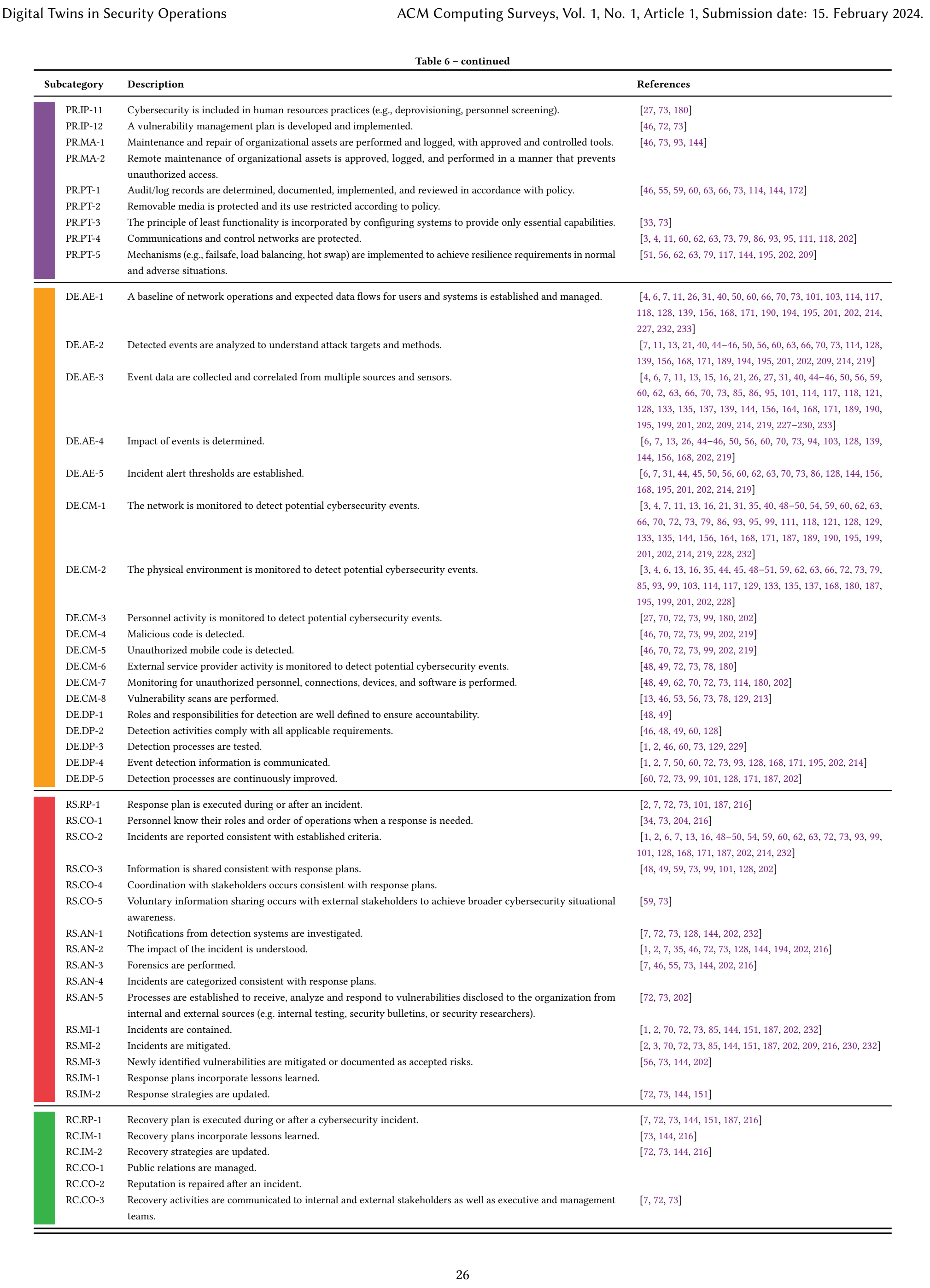


Table 6 – continued

	Subcategory	Description	References
🟪	PR.IP-11	Cybersecurity is included in human resources practices (e.g., deprovisioning, personnel screening).	[27, 73, 180]
	PR.IP-12	A vulnerability management plan is developed and implemented.	[46, 72, 73]
	PR.MA-1	Maintenance and repair of organizational assets are performed and logged, with approved and controlled tools.	[46, 73, 93, 144]
	PR.MA-2	Remote maintenance of organizational assets is approved, logged, and performed in a manner that prevents unauthorized access.	
	PR.PT-1	Audit/log records are determined, documented, implemented, and reviewed in accordance with policy.	[46, 55, 59, 60, 63, 66, 73, 114, 144, 172]
	PR.PT-2	Removable media is protected and its use restricted according to policy.	
	PR.PT-3	The principle of least functionality is incorporated by configuring systems to provide only essential capabilities.	[33, 73]
	PR.PT-4	Communications and control networks are protected.	[3, 4, 11, 60, 62, 63, 73, 79, 86, 93, 95, 111, 118, 202]
	PR.PT-5	Mechanisms (e.g., failsafe, load balancing, hot swap) are implemented to achieve resilience requirements in normal and adverse situations.	[51, 56, 62, 63, 79, 117, 144, 195, 202, 209]
🟧	DE.AE-1	A baseline of network operations and expected data flows for users and systems is established and managed.	[4, 6, 7, 11, 26, 31, 40, 50, 60, 66, 70, 73, 101, 103, 114, 117, 118, 128, 139, 156, 168, 171, 190, 194, 195, 201, 202, 214, 227, 232, 233]
	DE.AE-2	Detected events are analyzed to understand attack targets and methods.	[7, 11, 13, 21, 40, 44–46, 50, 56, 60, 63, 66, 70, 73, 114, 128, 139, 156, 168, 171, 189, 194, 195, 201, 202, 209, 214, 219]
	DE.AE-3	Event data are collected and correlated from multiple sources and sensors.	[4, 6, 7, 11, 13, 15, 16, 21, 26, 27, 31, 40, 44–46, 50, 56, 59, 60, 62, 63, 66, 70, 73, 85, 86, 95, 101, 114, 117, 118, 121, 128, 133, 135, 137, 139, 144, 156, 164, 168, 171, 189, 190, 195, 199, 201, 202, 209, 214, 219, 227–230, 233]
	DE.AE-4	Impact of events is determined.	[6, 7, 13, 26, 44–46, 50, 56, 60, 70, 73, 94, 103, 128, 139, 144, 156, 168, 202, 219]
	DE.AE-5	Incident alert thresholds are established.	[6, 7, 31, 44, 45, 50, 56, 60, 62, 63, 70, 73, 86, 128, 144, 156, 168, 195, 201, 202, 214, 219]
	DE.CM-1	The network is monitored to detect potential cybersecurity events.	[3, 4, 7, 11, 13, 16, 21, 31, 35, 40, 48–50, 54, 59, 60, 62, 63, 66, 70, 72, 73, 79, 86, 93, 95, 99, 111, 118, 121, 128, 129, 133, 135, 144, 156, 164, 168, 171, 187, 189, 190, 195, 199, 201, 202, 214, 219, 228, 232]
	DE.CM-2	The physical environment is monitored to detect potential cybersecurity events.	[3, 4, 6, 13, 16, 35, 44, 45, 48–51, 59, 62, 63, 66, 72, 73, 79, 85, 93, 99, 103, 114, 117, 129, 133, 135, 137, 168, 180, 187, 195, 199, 201, 202, 228]
	DE.CM-3	Personnel activity is monitored to detect potential cybersecurity events.	[27, 70, 72, 73, 99, 180, 202]
	DE.CM-4	Malicious code is detected.	[46, 70, 72, 73, 99, 202, 219]
	DE.CM-5	Unauthorized mobile code is detected.	[46, 70, 72, 73, 99, 202, 219]
	DE.CM-6	External service provider activity is monitored to detect potential cybersecurity events.	[48, 49, 72, 73, 78, 180]
	DE.CM-7	Monitoring for unauthorized personnel, connections, devices, and software is performed.	[48, 49, 62, 70, 72, 73, 114, 180, 202]
	DE.CM-8	Vulnerability scans are performed.	[13, 46, 53, 56, 73, 78, 129, 213]
	DE.DP-1	Roles and responsibilities for detection are well defined to ensure accountability.	[48, 49]
	DE.DP-2	Detection activities comply with all applicable requirements.	[46, 48, 49, 60, 128]
	DE.DP-3	Detection processes are tested.	[1, 2, 46, 60, 73, 129, 229]
	DE.DP-4	Event detection information is communicated.	[1, 2, 7, 50, 60, 72, 73, 93, 128, 168, 171, 195, 202, 214]
	DE.DP-5	Detection processes are continuously improved.	[60, 72, 73, 99, 101, 128, 171, 187, 202]
🟥	RS.RP-1	Response plan is executed during or after an incident.	[2, 7, 72, 73, 101, 187, 216]
	RS.CO-1	Personnel know their roles and order of operations when a response is needed.	[34, 73, 204, 216]
	RS.CO-2	Incidents are reported consistent with established criteria.	[1, 2, 6, 7, 13, 16, 48–50, 54, 59, 60, 62, 63, 72, 73, 93, 99, 101, 128, 168, 171, 187, 202, 214, 232]
	RS.CO-3	Information is shared consistent with response plans.	[48, 49, 59, 73, 99, 101, 128, 202]
	RS.CO-4	Coordination with stakeholders occurs consistent with response plans.	
	RS.CO-5	Voluntary information sharing occurs with external stakeholders to achieve broader cybersecurity situational awareness.	[59, 73]
	RS.AN-1	Notifications from detection systems are investigated.	[7, 72, 73, 128, 144, 202, 232]
	RS.AN-2	The impact of the incident is understood.	[1, 2, 7, 35, 46, 72, 73, 128, 144, 194, 202, 216]
	RS.AN-3	Forensics are performed.	[7, 46, 55, 73, 144, 202, 216]
	RS.AN-4	Incidents are categorized consistent with response plans.	
	RS.AN-5	Processes are established to receive, analyze and respond to vulnerabilities disclosed to the organization from internal and external sources (e.g. internal testing, security bulletins, or security researchers).	[72, 73, 202]
	RS.MI-1	Incidents are contained.	[1, 2, 70, 72, 73, 85, 144, 151, 187, 202, 232]
	RS.MI-2	Incidents are mitigated.	[2, 3, 70, 72, 73, 85, 144, 151, 187, 202, 209, 216, 230, 232]
	RS.MI-3	Newly identified vulnerabilities are mitigated or documented as accepted risks.	[56, 73, 144, 202]
	RS.IM-1	Response plans incorporate lessons learned.	
	RS.IM-2	Response strategies are updated.	[72, 73, 144, 151]
🟩	RC.RP-1	Recovery plan is executed during or after a cybersecurity incident.	[7, 72, 73, 144, 151, 187, 216]
	RC.IM-1	Recovery plans incorporate lessons learned.	[73, 144, 216]
	RC.IM-2	Recovery strategies are updated.	[72, 73, 144, 216]
	RC.CO-1	Public relations are managed.	
	RC.CO-2	Reputation is repaired after an incident.	
	RC.CO-3	Recovery activities are communicated to internal and external stakeholders as well as executive and management teams.	[7, 72, 73]

REFERENCES

[1] Fatemeh Akbarian, Emma Fitzgerald, and Maria Kihl. 2020. Intrusion Detection in Digital Twins for Industrial Control Systems. In *2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. 1–6.

[2] Fatemeh Akbarian, William Tärneberg, Emma Fitzgerald, and Maria Kihl. 2021. A Security Framework in Digital Twins for Cloud-based Industrial Control Systems: Intrusion Detection and Mitigation. In *IEEE International Conference on Emerging Technologies and Factory Automation*. 01–08.

[3] Fatemeh Akbarian, William Tärneberg, Emma Fitzgerald, and Maria Kihl. 2023. Detecting and Mitigating Actuator Attacks on Cloud Control Systems through Digital Twins. In *2023 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. 1–6.

[4] Abimbola Akerele, William Leppert, Shionta Somerville, and Guy-Alain Amoussou. 2023. The Digital Twins Incident Response To Improve the Security of Power System Critical Infrastructure. *J. Comput. Sci. Coll.* 39, 3 (2023), 86–99.

[5] Tejasvi Alladi, Vinay Chamola, and Sherali Zeadally. 2020. Industrial Control Systems: Cyberattack trends and countermeasures. *Comput. Commun.* 155 (2020), 1–8.

[6] David Allison, Paul Smith, and Kieran McLaughlin. 2022. Digital Twin-Enhanced Methodology for Training Edge-Based Models for Cyber Security Applications. In *2022 IEEE 20th International Conference on Industrial Informatics (INDIN)*. 226–232.

[7] David Allison, Paul Smith, and Kieran Mclaughlin. 2023. Digital Twin-Enhanced Incident Response for Cyber-Physical Systems. In *Proceedings of the 18th International Conference on Availability, Reliability and Security (ARES '23)*.

[8] Sadeq Almeaibed, Saba Al-Rubaye, Antonios Tsourdos, and Nicolas P. Avdelidis. 2021. Digital Twin Analysis to Promote Safety and Security in Autonomous Vehicles. *IEEE Commun. Stand. Mag.* 5, 1 (2021), 40–46.

[9] Daniel Anthony Howard, Zheng Ma, Jesper Mazanti Aaslyng, and Bo Norregaard Jorgensen. 2020. Data Architecture for Digital Twin of Commercial Greenhouse Production. In *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*. 1–7.

[10] Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J. Alex Halderman, Luca Invernizzi, Michalis Kallitsis, Deepak Kumar, Chaz Lever, Zane Ma, Joshua Mason, Damian Menscher, Chad Seaman, Nick Sullivan, Kurt Thomas, and Yi Zhou. 2017. Understanding the Mirai Botnet. In *26th USENIX Security Symposium (USENIX Security 17)*. USENIX Association, Vancouver, BC, 1093–1110.

[11] Varsha Arya, Akshat Gaurav, Brij B. Gupta, Ching-Hsien Hsu, and Hojjat Baghban. 2023. Detection of Malicious Node in VANETs Using Digital Twin. In *Big Data Intelligence and Computing*, Ching-Hsien Hsu, Mengwei Xu, Hung Cao, Hojjat Baghban, and A. B. M. Shawkat Ali (Eds.). Vol. 13864. 204–212.

[12] Giacomo Assenza, Valerio Cozzani, Francesco Flammini, Nadezhda Gotcheva, Tommy Gustafsson, Anders Hansson, Jouko Heikkila, Matteo Iaiani, Sokratis Katsikas, Minna Nissilä, Gabriele Oliva, Eleni Richter, Maaike Roelofs, Mehdi Saman Azari, Roberto Setola, Wouter Stejin, Alessandro Tugnoli, Dolf Vanderbeek, Lars Westerdahl, Marja Ylönen, and Heather Young. 2020. White Paper on Industry Experiences in Critical Information Infrastructure Security. In *Critical Information Infrastructures Security*, Simin Nadjm-Tehrani (Ed.). Vol. 11777. 197–207.

[13] Manolya Atalay and Pelin Angin. 2020. A Digital Twins Approach to Smart Grid Security Testing and Standardization, In International Workshop on Metrology for Industry 4.0 & IoT. *2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT*, 435–440.

[14] Abiodun Ayodeji, Yong-kuo Liu, Nan Chao, and Li-qun Yang. 2020. A New Perspective Towards the Development of Robust Data-driven Intrusion Detection for Industrial Control Systems. *Nucl. Eng. Technol.* 52, 12 (2020), 2687–2698.

[15] V. Ayyalusamy, B. Sivaneasan, NK Kandasamy, J.F. Xiao, Abidi K, and A. Chandra. 2022. Hybrid Digital Twin Architecture for Power System Cyber Security Analysis. In *IEEE PES Innovative Smart Grid Technologies*. 270–274.

[16] Efe C. Balta, Michael Pease, James Moyne, Kira Barton, and Dawn M. Tilbury. 2023. Digital Twin-Based Cyber-Attack Detection Framework for Cyber-Physical Manufacturing Systems. *IEEE Transactions on Automation Science and Engineering* (2023), 1–18.

[17] Xinbo Ban, Ming Ding, Shigang Liu, Chao Chen, Jun Zhang, and Yang Xiang. 2022. TAESim: A Testbed for IoT Security Analysis of Trigger-Action Environment. In *Computer Security. ESORICS 2021 International Workshops*, Sokratis Katsikas, Costas Lambrinoudakis, Nora Cuppens, John Mylopoulos, Christos Kalloniatis, Weizhi Meng, Steven Furnell, Frank Pallas, Jörg Pohle, M. Angela Sasse, Habtamu Abie, Silvio Ranise, Luca Verderame, Enrico Cambiaso, Jorge Maestre Vidal, and Marco Antonio Sotelo Monge (Eds.). Vol. 13106. 218–237.

[18] Matthew Barrett. 2018. Framework for Improving Critical Infrastructure Cybersecurity Version 1.1.

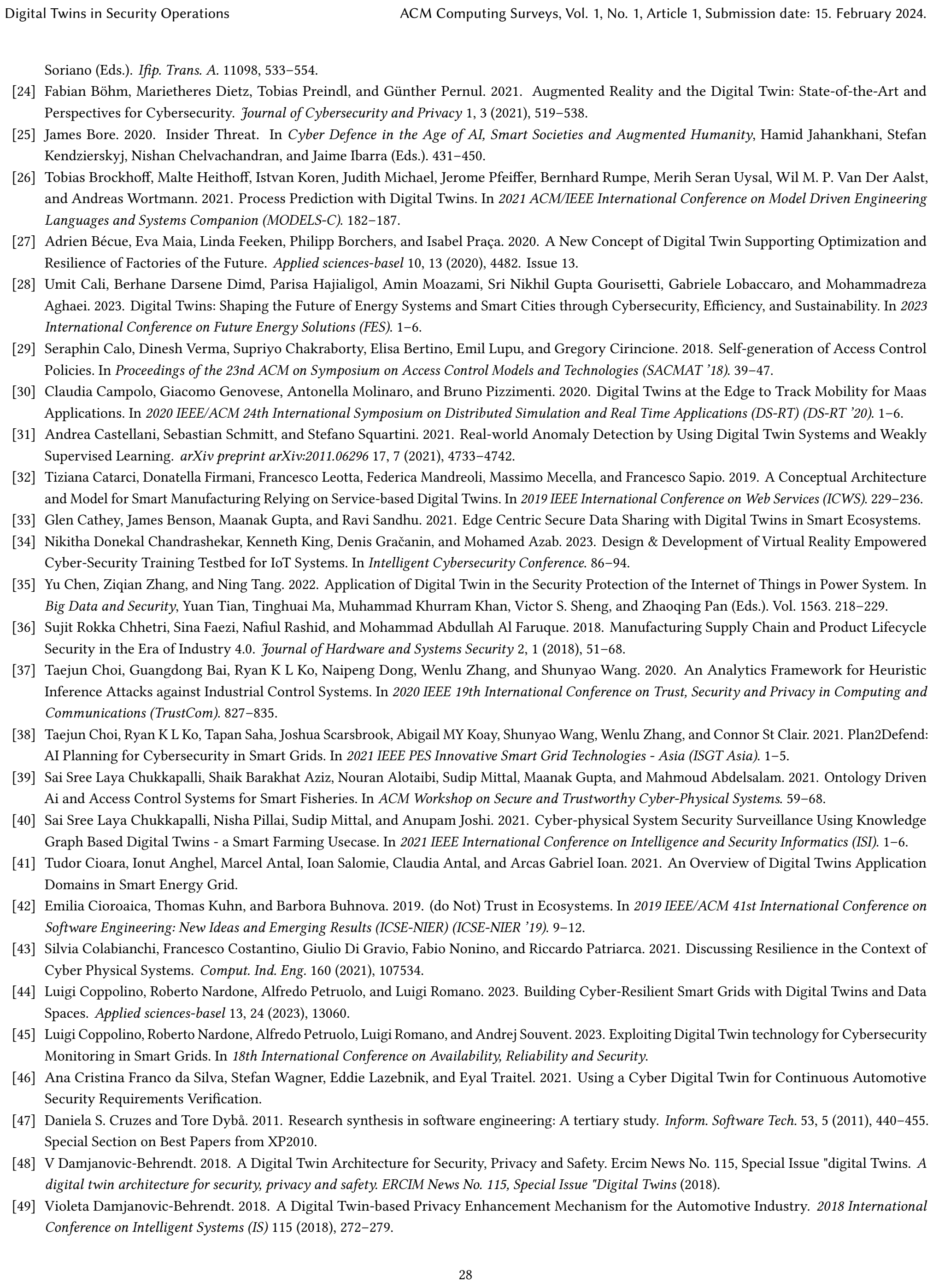
[19] Barbara Rita Barricelli, Elena Casiraghi, and Daniela Fogli. 2019. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access* 7 (2019), 167653–167671.

[20] Adrien Becue, Yannick Fourastier, Isabel Praca, Alexandre Savarit, Claude Baron, Baptiste Gradussofs, Etienne Pouille, and Carsten Thomas. 2018. Cyberfactory#1 — Securing the Industry 4.0 with Cyber-ranges and Digital Twins. *International Workshop on Factory Communication Systems* (2018), 1–4.

[21] Adrien Bécue, Martin Praddaude, Eva Maia, Nicolas Hogrel, Isabel Praça, and Reda Yaich. 2022. Digital Twins for Enhanced Resilience: Aerospace Manufacturing Scenario. In *Advanced Information Systems Engineering Workshops*, Jennifer Horkoff, Estefania Serral, and Jelena Zdravkovic (Eds.). Springer International Publishing, Cham, 107–118.

[22] Stefan Biffl, Matthias Eckhart, Arndt Lüder, and Edgar Weippl. 2019. Conclusion and Outlook on Security and Quality of Complex Cyber-physical Systems Engineering. In *Security and Quality in Cyber-Physical Systems Engineering*, Stefan Biffl, Matthias Eckhart, Arndt Lüder, and Edgar Weippl (Eds.). 497–507.

[23] Ron Bitton, Tomer Gluck, Orly Stan, Masaki Inokuchi, Yoshinobu Ohta, Yoshiyuki Yamada, Tomohiko Yagyu, Yuval Elovici, and Asaf Shabtai. 2018. Deriving a Cost-effective Digital Twin of an Ics to Facilitate Security Evaluation, In Computer Security, Javier López, Jianying Zhou, and Miguel



Soriano (Eds.). *Ifip. Trans. A.* 11098, 533–554.

[24] Fabian Böhm, Marietheres Dietz, Tobias Preindl, and Günther Pernul. 2021. Augmented Reality and the Digital Twin: State-of-the-Art and Perspectives for Cybersecurity. *Journal of Cybersecurity and Privacy* 1, 3 (2021), 519–538.

[25] James Bore. 2020. Insider Threat. In *Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity*, Hamid Jahankhani, Stefan Kendzierskyj, Nishan Chelvachandran, and Jaime Ibarra (Eds.). 431–450.

[26] Tobias Brockhoff, Malte Heithoff, Istvan Koren, Judith Michael, Jerome Pfeiffer, Bernhard Rumpe, Merih Seran Uysal, Wil M. P. Van Der Aalst, and Andreas Wortmann. 2021. Process Prediction with Digital Twins. In *2021 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*. 182–187.

[27] Adrien Bécue, Eva Maia, Linda Feeken, Philipp Borchers, and Isabel Praça. 2020. A New Concept of Digital Twin Supporting Optimization and Resilience of Factories of the Future. *Applied sciences-basel* 10, 13 (2020), 4482. Issue 13.

[28] Umit Cali, Berhane Darsene Dimd, Parisa Hajialigol, Amin Moazami, Sri Nikhil Gupta Gourisetti, Gabriele Lobaccaro, and Mohammadreza Aghaei. 2023. Digital Twins: Shaping the Future of Energy Systems and Smart Cities through Cybersecurity, Efficiency, and Sustainability. In *2023 International Conference on Future Energy Solutions (FES)*. 1–6.

[29] Seraphin Calo, Dinesh Verma, Supriyo Chakraborty, Elisa Bertino, Emil Lupu, and Gregory Cirincione. 2018. Self-generation of Access Control Policies. In *Proceedings of the 23nd ACM on Symposium on Access Control Models and Technologies (SACMAT '18)*. 39–47.

[30] Claudia Campolo, Giacomo Genovese, Antonella Molinaro, and Bruno Pizzimenti. 2020. Digital Twins at the Edge to Track Mobility for Maas Applications. In *2020 IEEE/ACM 24th International Symposium on Distributed Simulation and Real Time Applications (DS-RT) (DS-RT '20)*. 1–6.

[31] Andrea Castellani, Sebastian Schmitt, and Stefano Squartini. 2021. Real-world Anomaly Detection by Using Digital Twin Systems and Weakly Supervised Learning. *arXiv preprint arXiv:2011.06296* 17, 7 (2021), 4733–4742.

[32] Tiziana Catarci, Donatella Firmani, Francesco Leotta, Federica Mandreoli, Massimo Mecella, and Francesco Sapio. 2019. A Conceptual Architecture and Model for Smart Manufacturing Relying on Service-based Digital Twins. In *2019 IEEE International Conference on Web Services (ICWS)*. 229–236.

[33] Glen Cathey, James Benson, Maanak Gupta, and Ravi Sandhu. 2021. Edge Centric Secure Data Sharing with Digital Twins in Smart Ecosystems.

[34] Nikitha Donekal Chandrashekar, Kenneth King, Denis Gračanin, and Mohamed Azab. 2023. Design & Development of Virtual Reality Empowered Cyber-Security Training Testbed for IoT Systems. In *Intelligent Cybersecurity Conference*. 86–94.

[35] Yu Chen, Ziqian Zhang, and Ning Tang. 2022. Application of Digital Twin in the Security Protection of the Internet of Things in Power System. In *Big Data and Security*, Yuan Tian, Tinghuai Ma, Muhammad Khurram Khan, Victor S. Sheng, and Zhaoqing Pan (Eds.). Vol. 1563. 218–229.

[36] Sujit Rokka Chhetri, Sina Faezi, Nafiul Rashid, and Mohammad Abdullah Al Faruque. 2018. Manufacturing Supply Chain and Product Lifecycle Security in the Era of Industry 4.0. *Journal of Hardware and Systems Security* 2, 1 (2018), 51–68.

[37] Taejun Choi, Guangdong Bai, Ryan K L Ko, Naipeng Dong, Wenlu Zhang, and Shunyao Wang. 2020. An Analytics Framework for Heuristic Inference Attacks against Industrial Control Systems. In *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. 827–835.

[38] Taejun Choi, Ryan K L Ko, Tapan Saha, Joshua Scarsbrook, Abigail MY Koay, Shunyao Wang, Wenlu Zhang, and Connor St Clair. 2021. Plan2Defend: AI Planning for Cybersecurity in Smart Grids. In *2021 IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia)*. 1–5.

[39] Sai Sree Laya Chukkapalli, Shaik Barakhat Aziz, Nouran Alotaibi, Sudip Mittal, Maanak Gupta, and Mahmoud Abdelsalam. 2021. Ontology Driven Ai and Access Control Systems for Smart Fisheries. In *ACM Workshop on Secure and Trustworthy Cyber-Physical Systems*. 59–68.

[40] Sai Sree Laya Chukkapalli, Nisha Pillai, Sudip Mittal, and Anupam Joshi. 2021. Cyber-physical System Security Surveillance Using Knowledge Graph Based Digital Twins - a Smart Farming Usecase. In *2021 IEEE International Conference on Intelligence and Security Informatics (ISI)*. 1–6.

[41] Tudor Cioara, Ionut Anghel, Marcel Antal, Ioan Salomie, Claudia Antal, and Arcas Gabriel Ioan. 2021. An Overview of Digital Twins Application Domains in Smart Energy Grid.

[42] Emilia Cioroaica, Thomas Kuhn, and Barbora Buhnova. 2019. (do Not) Trust in Ecosystems. In *2019 IEEE/ACM 41st International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER) (ICSE-NIER '19)*. 9–12.

[43] Silvia Colabianchi, Francesco Costantino, Giulio Di Gravio, Fabio Nonino, and Riccardo Patriarca. 2021. Discussing Resilience in the Context of Cyber Physical Systems. *Comput. Ind. Eng.* 160 (2021), 107534.

[44] Luigi Coppolino, Roberto Nardone, Alfredo Petruolo, and Luigi Romano. 2023. Building Cyber-Resilient Smart Grids with Digital Twins and Data Spaces. *Applied sciences-basel* 13, 24 (2023), 13060.

[45] Luigi Coppolino, Roberto Nardone, Alfredo Petruolo, Luigi Romano, and Andrej Souvent. 2023. Exploiting Digital Twin technology for Cybersecurity Monitoring in Smart Grids. In *18th International Conference on Availability, Reliability and Security*.

[46] Ana Cristina Franco da Silva, Stefan Wagner, Eddie Lazebnik, and Eyal Traitel. 2021. Using a Cyber Digital Twin for Continuous Automotive Security Requirements Verification.

[47] Daniela S. Cruzes and Tore Dybå. 2011. Research synthesis in software engineering: A tertiary study. *Inform. Software Tech.* 53, 5 (2011), 440–455. Special Section on Best Papers from XP2010.

[48] V Damjanovic-Behrendt. 2018. A Digital Twin Architecture for Security, Privacy and Safety. Ercim News No. 115, Special Issue "digital Twins. *A digital twin architecture for security, privacy and safety. ERCIM News No. 115, Special Issue "Digital Twins* (2018).

[49] Violeta Damjanovic-Behrendt. 2018. A Digital Twin-based Privacy Enhancement Mechanism for the Automotive Industry. *2018 International Conference on Intelligent Systems (IS)* 115 (2018), 272–279.

[50] William Danilczyk, Yan Sun, and Haibo He. 2019. Angel: An Intelligent Digital Twin Framework for Microgrid Security, In 2019 North American Power Symposium (NAPS). *2019 North American Power Symposium (NAPS)*, 1–6.

[51] William Danilczyk, Yan Lindsay Sun, and Haibo He. 2021. Smart Grid Anomaly Detection Using a Deep Learning Digital Twin. In *2020 52nd North American Power Symposium (NAPS).* 1–6.

[52] Debak Das. 2019. *An Indian nuclear power plant suffered a cyberattack. Here's what you need to know.* Retrieved June 28, 2021 from https://www.washingtonpost.com/politics/2019/11/04/an-indian-nuclear-power-plant-suffered-cyberattack-heres-what-you-need-know/

[53] Danielle Dauphinais, Michael Zylka, Harris Spahic, Farhan Shaik, Jingda Yang, Isabella Cruz, Jakob Gibson, and Ying Wang. 2023. Automated Vulnerability Testing and Detection Digital Twin Framework for 5G Systems. In *2023 IEEE 9th International Conference on Network Softwarization (NetSoft).* 308–310.

[54] Jorge David de Hoz Diego, Anastasios Temperekidis, Panagiotis Katsaros, and Charalambos Konstantinou. 2022. An IoT Digital Twin for Cyber-Security Defence Based on Runtime Verification. In *Leveraging Applications of Formal Methods, Verification and Validation. Verification Principles (Lecture Notes in Computer Science, Vol. 13701)*, Tiziana Margaria and Bernhard Steffen (Eds.). 556–574.

[55] Marietheres Dietz, Ludwig Englbrecht, and Günther Pernul. 2021. Enhancing Industrial Control System Forensics Using Replication-based Digital Twins. In *Advances in Digital Forensics XVII*, Gilbert Peterson and Sujeet Shenoi (Eds.). Vol. 612. 21–38.

[56] Marietheres Dietz, Leon Hageman, Constantin von Hornung, and Günther Pernul. 2022. Employing Digital Twins for Security-by-Design System Testing. In *Proceedings of the 2022 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems (Sat-CPS '22).* 97–106.

[57] Marietheres Dietz and Günther Pernul. 2020. Digital Twin: Empowering Enterprises Towards a System-of-Systems Approach. *Bus. Inform. Syst. Eng+.+* 62, 2 (2020), 179–184. Issue 2.

[58] Marietheres Dietz and Gunther Pernul. 2020. Unleashing the Digital Twin's Potential for Ics Security. *IEEE Security & Privacy* 18, 4 (2020), 20–27.

[59] Marietheres Dietz, Daniel Schlette, and Gunther Pernul. 2022. Harnessing Digital Twin Security Simulations for systematic Cyber Threat Intelligence. In *46th IEEE Annual Computers, Software, and Applications Conference*, Hong Va Leong, Sahra Sedigh Sarvestani, Yuuichi Teranishi, Alfredo Cuzzocrea, Hiroki Kashiwazaki, Dave Towey, Ji-Jiang Yang, and Hossain Shahriar (Eds.). 789–797.

[60] Marietheres Dietz, Manfred Vielberth, and Günther Pernul. 2020. Integrating Digital Twin Security Simulations in the Security Operations Center. In *15th International Conference on Availability, Reliability and Security*, Melanie Volkamer and Christian Wressnegger (Eds.). 18:1–18:9.

[61] Matthias Eckhart, Bernhard Brenner, Andreas Ekelhart, and Edgar R Weippl. 2019. Quantitative Security Risk Assessment for Industrial Control Systems: Research Opportunities and Challenges. *J. Internet Serv. Inf. Secur.* 9, 3 (2019), 52–73.

[62] Matthias Eckhart and Andreas Ekelhart. 2018. A Specification-based State Replication Approach for Digital Twins. In *Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and Privacy (CPS-SPC '18, Vol. 18).* 36–47.

[63] Matthias Eckhart and Andreas Ekelhart. 2018. Towards Security-aware Virtual Environments for Digital Twins, In Proceedings of the 4th ACM Workshop on Cyber-Physical System Security, CPSS@AsiaCCS 2018, Incheon, Republic of Korea, June 04-08, 2018, Dieter Gollmann and Jianying Zhou (Eds.). *Proceedings of the 4th ACM Workshop on Cyber-Physical System Security -CPSS '18*, 61–72.

[64] Matthias Eckhart and Andreas Ekelhart. 2019. Digital Twins for Cyber-Physical Systems Security: State of the Art and Outlook. In *Security and Quality in Cyber-Physical Systems Engineering*, Stefan Biffl, Matthias Eckhart, Arndt Lüder, and Edgar Weippl (Eds.). 383–412.

[65] Matthias Eckhart, Andreas Ekelhart, David Allison, Magnus Almgren, Katharina Ceesay-Seitz, Helge Janicke, Simin Nadjm-Tehrani, Awais Rashid, and Mark Yampolskiy. 2023. Security-Enhancing Digital Twins: Characteristics, Indicators, and Future Perspectives. *IEEE Security & Privacy* 21, 6 (2023), 64–75.

[66] Matthias Eckhart, Andreas Ekelhart, and Edgar Weippl. 2019. Enhancing Cyber Situational Awareness for Cyber-Physical Systems through Digital Twins. In *Proceedings of the 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA).* 1222–1225.

[67] Ian Elsby. 2019. *Digital Twin Does More Than Designing, Analysing and Processing; It's the Cyber Attack Combatant!* Retrieved June 28, 2021 from https://news.siemens.co.uk/news/digital-twin-does-more-than-designing-analysing-and-processing-its-the-cyber-attack-combatant

[68] Philip Empl, Henric Hager, and Günther Pernul. 2023. Digital Twins for IoT Security Management. In *Data and Applications Security and Privacy XXXVII: 37th Annual IFIP WG 11.3 Conference, DBSec 2023, Sophia-Antipolis, France, July 19–21, 2023, Proceedings* (Sophia-Antipolis, France). Springer-Verlag, Berlin, Heidelberg, 141–149.

[69] Philip Empl and Günther Pernul. 2021. A Flexible Security Analytics Service for the Industrial Iot. In *Proceedings of the 2021 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems (SAT-CPS '21).* 23–32.

[70] Philip Empl and Günther Pernul. 2023. Digital-Twin-Based Security Analytics for the Internet of Things. *Information* 14, 2 (2023).

[71] Philip Empl, Daniel Schlette, Lukas Stöger, and Günther Pernul. 2023. Generating ICS vulnerability playbooks with open standards. *International Journal of Information Security* (2023).

[72] Philip Empl, Daniel Schlette, Daniel Zupfer, and Günther Pernul. 2022. SOAR4IoT: Securing IoT Assets with Digital Twins. In *Proceedings of the 17th International Conference on Availability, Reliability and Security (ARES '22).*

[73] Gregory Epiphaniou, Mohammad Hammoudeh, Hu Yuan, Carsten Maple, and Uchenna Ani. 2023. Digital twins in cyber effects modelling of IoT/CPS points of low resilience. *Simulation Modelling Practice and Theory* 125 (2023), 102744.

[74] Rajiv Faleiro, Lei Pan, Shiva Raj Pokhrel, and Robin Doss. 2022. Digital Twin for Cybersecurity: Towards Enhancing Cyber Resilience. In *Broadband Communications, Networks, and Systems*, Wei Xiang, Fengling Han, and Tran Khoa Phan (Eds.). Springer International Publishing, Cham, 57–76.

[75] Chris Farnell, Estefano Soria, Justin Jackson, and H. Alan Mantooth. 2021. Cyber Protection of Grid-connected Devices through Embedded Online Security. In *2021 IEEE Design Methodologies Conference (DMC).* 1–6.

[76] Apostolos P. Fournaris, Andreas Komninos, Aris S. Lalos, Athanasios P. Kalogeras, Christos Koulamas, and Dimitrios Serpanos. 2019. Design and Run-time Aspects of Secure Cyber-physical Systems. In *Security and Quality in Cyber-Physical Systems Engineering*, Stefan Biffl, Matthias Eckhart, Arndt Lüder, and Edgar Weippl (Eds.). 357–382.

[77] Guillermo Francia and Gregory Hall. 2021. Digital Twins for Industrial Control Systems Security. In *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*. 801–805.

[78] Ana Cristina Franco da Silva, Stefan Wagner, Eddie Lazebnik, and Eyal Traitel. 2022. Using a Cyber Digital Twin for Continuous Automotive Security Requirements Verification. *IEEE Softw.* (2022), 0–0.

[79] Benjamin Fraser, Saba Al-Rubaye, Sohaib Aslam, and Antonios Tsourdos. 2021. Enhancing the Security of Unmanned Aerial Systems Using Digital-twin Technology and Intrusion Detection. In *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*. 1–10.

[80] Aidan Fuller, Zhong Fan, Charles Day, and Chris Barlow. 2020. Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access* 8 (2020), 108952–108971.

[81] Brian Fung and Geneva Sands. 2021. *Ransomware attackers used compromised password to access Colonial Pipeline network.* Retrieved June 28, 2021 from https://edition.cnn.com/2021/06/04/politics/colonial-pipeline-ransomware-attack-password/index.html

[82] Zach Furness. 2019. *Perspectives on Securing Cyber Physical Systems.* 287–299.

[83] Sebastian Gajek, Michael Lees, and Christoph Jansen. 2020. Iiot and Cyber-resilience. *AI & SOCIETY* 36, 3 (2020), 725–735.

[84] Rogelio GAmez Díaz, Qingtian Yu, Yezhe Ding, Fedwa Laamarti, and Abdulmotaleb El Saddik. 2020. Digital Twin Coaching for Physical Activities: A Survey. *Sensors-basel.* 20, 20 (2020), 5936.

[85] Humberto E. Garcia, Steven E. Aumeier, Ahmad Y. Al-Rashdan, and Bri L. Rolston. 2020. Secure Embedded Intelligence in Nuclear Systems: Framework and Methods. *Ann. Nucl. Energy* 140 (2020), 107261.

[86] Akshat Gaurav, Brij B. Gupta, Kwok Tai Chui, Varsha Arya, and Elhadj Benkhelifa. 2023. A DDoS Attack Detection System for Industry 5.0 using Digital Twins and Machine Learning. In *Global Conference on Consumer Electronics.* 1019–1022.

[87] General Electric Research. 2017. *Digital Ghost: Real-time, Active Cyber Defense.* Retrieved June 28, 2021 from https://www.ge.com/research/offering/digital-ghost-real-time-active-cyber-defense

[88] Edward Glaessgen and David Stargel. 2012. The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. In *Proceedings of the 53rd. AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference.* American Institute of Aeronautics and Astronautics.

[89] M. Glawe, L. Feeken, B. Wudka, C.-Y. Kao, E. Mirzaei, T. Weinhold, and A. Szanto. 2020. Towards Resilient Factories of Future – Defining Required Capabilities for a Resilient Factory of Future. (2020), 755–768.

[90] Gerald Glocker. 2019. *A primer on digital twins in the IoT.* Retrieved June 28, 2021 from https://blog.bosch-si.com/bosch-iot-suite/a-primer-on-digital-twins-in-the-iot/

[91] Janis Grabis, Janis Stirna, and Jelena Zdravkovic. 2021. A Capability Based Method for Development of Resilient Digital Services. In *Enterprise Information Systems*, Joaquim Filipe, Michał Śmiałek, Alexander Brodsky, and Slimane Hammoudi (Eds.). Vol. 417. 498–516.

[92] Chiara Grasselli, Andrea Melis, Roberto Girau, and Franco Callegati. 2023. A Digital Twin for Enhanced Cybersecurity in Connected Vehicles. In *2023 23rd International Conference on Transparent Optical Networks (ICTON).* 1–4.

[93] Chiara Grasselli, Andrea Melis, Lorenzo Rinieri, Davide Berardi, Giacomo Gori, and Amir Al Sadi. 2022. An Industrial Network Digital Twin for enhanced security of Cyber-Physical Systems. In *2022 International Symposium on Networks, Computers and Communications (ISNCC).* 1–7.

[94] Amal Guittoum, Francois Aïssaoui, Sébastien Bolle, Fabienne Boyer, and Noel De Palma. 2023. Inferring Threatening IoT Dependencies using Semantic Digital Twins Toward Collaborative IoT Device Management. In *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing (SAC '23).* 1732–1741.

[95] Jiajie Guo, Muhammad Bilal, Yuying Qiu, Cheng Qian, Xiaolong Xu, and Kim-Kwang Raymond Choo. 2022. Survey on digital twins for Internet of Vehicles: Fundamentals, challenges, and opportunities. *Digital Communications and Networks* (2022).

[96] Yun Guo, Xinxin Lou, Edita Bajramovic, and Karl Waedt. 2020. Cybersecurity Risk Analysis and Technical Defense Architecture: Research of ICS in Nuclear Power Plant Construction Stage. In *3rd IAEA International Conference on Nuclear Security: Sustaining and Strengthening Efforts.*

[97] Yun Guo, Aijun Yan, and Junjie Wang. 2021. Cyber Security Risk Analysis of Physical Protection Systems of Nuclear Power Plants and Research on the Cyber Security Test Platform Using Digital Twin Technology. In *2021 International Conference on Power System Technology.* 1889–1892.

[98] Deepti Gupta, Olumide Kayode, Smriti Bhatt, Maanak Gupta, and Ali Saman Tosun. 2021. Hierarchical Federated Learning Based Anomaly Detection Using Digital Twins for Smart Healthcare.

[99] Deepti Gupta, Shafika Showkat Moni, and Ali Saman Tosun. 2023. Integration of Digital Twin and Federated Learning for Securing Vehicular Internet of Things. In *International Conference on Research in Adaptive and Convergent Systems.*

[100] Ethan Hadar, Dmitry Kravchenko, and Alexander Basovskiy. 2020. Cyber Digital Twin Simulator for Automatic Gathering and Prioritization of Security Controls' Requirements. In *2020 IEEE 28th International Requirements Engineering Conference (RE)*, Travis D. Breaux, Andrea Zisman, Samuel Fricker, and Martin Glinz (Eds.). 250–259.

[101] Kim Hammar and Rolf Stadler. 2023. Digital Twins for Security Automation. In *NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium.* 1–6.

[102] Mohammad Hammoudeh. 2020. Blockchain, Internet of Things and Digital Twins in Trustless Security of Critical National Infrastructure. In *The 4th International Conference on Future Networks and Distributed Systems (ICFNDS) (ICFNDS '20).* 43:1.

[103] Jiaxuan Han, Qiteng Hong, Zhiwang Feng, Mazheruddin Syed, Graeme Burt, and Campbell Booth. 2022. Design and Implementation of a Real-Time Hardware-in-the-Loop Platform for Prototyping and Testing Digital Twins of Distributed Energy Resources. *Energies* 15, 18 (2022), 6629.

[104] George Hatzivasilis, Sotiris Ioannidis, Michail Smyrlis, George Spanoudakis, Fulvio Frati, Chiara Braghin, Ernesto Damiani, Hristo Koshutanski, George Tsakirakis, Torsten Hildebrandt, Ludger Goeke, Sebastian Pape, Oleg Blinder, Michael Vinov, George Leftheriotis, Martin Kunc, Fotis Oikonomou, Giovanni Magilo, Vito Petrarolo, Antonio Chieti, and Robert Bordianu. 2021. The Threat-arrest Cyber Range Platform. In *2021 IEEE International Conference on Cyber Security and Resilience (CSR)*. 422–427.

[105] Zhongyuan Hau, John Henry Castellanos, and Jianying Zhou. 2020. Evaluating Cascading Impact of Attacks on Resilience of Industrial Control Systems: A Design-centric Modeling Approach. In *Proceedings of the 6th ACM on Cyber-Physical System Security Workshop (CPSS '20)*. 42–53.

[106] Rui He, Guoming Chen, Che Dong, Shufeng Sun, and Xiaoyu Shen. 2019. Data-driven Digital Twin Technology for Optimized Control in Process Systems. *Isa transactions* 95 (2019), 221–234.

[107] Siemens Healthineers. 2019. *The Value of Digital Twin Technology.* Technical Report 7144 0819. Siemens Healthcare GmbH, Erlangen, Germany.

[108] David Holmes, Maria Papathanasaki, Leandros Maglaras, Mohamed Amine Ferrag, Surya Nepal, and Helge Janicke. 2021. Digital Twins and Cyber Security – Solution or Challenge?. In *2021 South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference*. 1–8.

[109] Tareq Hossen, Mehmetcan Gursoy, and Behrooz Mirafzal. 2021. Digital Twin for Self-security of Smart Inverters. In *2021 IEEE Energy Conversion Congress and Exposition (ECCE)*. 713–718.

[110] Tareq Hossen, Dushyant Sharma, and Behrooz Mirafzal. 2021. Smart Inverter Twin Model for Anomaly Detection. In *2021 IEEE 22nd Workshop on Control and Modelling of Power Electronics (COMPEL)*. 1–6.

[111] Zhe Hou, Qinyi Li, Ernest Foo, Jin Song Dong, and Paulo de Souza. 2022. A Digital Twin Runtime Verification Framework for Protecting Satellites Systems from Cyber Attacks. In *2022 26th International Conference on Engineering of Complex Computer Systems (ICECCS)*. 117–122.

[112] Sihan Huang, Guoxin Wang, Yan Yan, and Xiongbing Fang. 2020. Blockchain-based Data Management for Digital Twin of Product. *J. Manuf. Syst.* 54 (2020), 361–371.

[113] Jithin Jagannath, Keyvan Ramezanpour, and Anu Jagannath. 2022. Digital Twin Virtualization with Machine Learning for IoT and Beyond 5G Networks: Research Directions for Security and Optimal Control. *CoRR* abs/2204.01950 (2022).

[114] Sergej Jakovlev, Tomas Eglynas, Miroslav Voznak, Pavol Partila, Jaromir Tovarek, Mindaugas Jusis, Edvinas Pocevicius, and Zane Purlaura. 2021. Development of an Intelligent Digital Twins Framework for Secure Container Terminal Operations. In *2021 62nd International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)*. 1–4.

[115] David Jones, Chris Snider, Aydin Nassehi, Jason Yon, and Ben Hicks. 2020. Characterising the Digital Twin: A Systematic Literature Review. *CIRP Journal of Manufacturing Science and Technology* 29 (2020), 36–52.

[116] Siavash H. Khajavi, Müge Tetik, Zixuan Liu, Pasi Korhonen, and Jan Holmström. 2023. Digital Twin for Safety and Security: Perspectives on Building Lifecycle. *IEEE Access* 11 (2023), 52339–52356.

[117] Mohammed Masum Siraj Khan, Jairo Giraldo, and Masood Parvania. 2023. Real-Time Cyber Attack Localization in Distribution Systems Using Digital Twin Reference Model. *Ieee transactions on power delivery* 38, 5 (2023), 3238–3249.

[118] Mehdi Kherbache, Moufida Maimour, and Eric Rondeau. 2022. Network Digital Twin for the Industrial Internet of Things. In *2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. 573–578.

[119] Andreas Kind. 2021. The Power of Industrial Cybersecurity. *Digitale Welt* 5, 1 (2021), 18–21.

[120] Dan Klein and Gal Engelberg. 2021. *Get Ahead of Cyberattacks with Digital Twins.* Retrieved June 28, 2021 from https://www.accenture.com/us-en/blogs/technology-innovation/klein-engelberg-get-ahead-of-cyberattacks-with-digital-twins

[121] Nickolaos Koroniotis, Nour Moustafa, Francesco Schiliro, Praveen Gauravaram, and Helge Janicke. 2021. The SAir-IIoT Cyber Testbed as a Service: A Novel Cybertwins Architecture in IIoT-Based Smart Airports. *IEEE Transactions on Intelligent Transportation Systems* 24, 2 (2021), 1–14.

[122] Prabhakar Krishnan, Kurunandan Jain, Rajkumar Buyya, Pandi Vijayakumar, Anand Nayyar, Muhammad Bilal, and Houbing Song. 2021. Mud-based Behavioral Profiling Security Framework for Software-defined Iot Networks. *IEEE Internet Things J.* (2021), 1–1.

[123] Priya R Krishnan and Josephkutti Jacob. 2021. Asset Management And Finite Element Analysis In Smart grid. In *2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA)*. 497–503.

[124] Werner Kritzinger, Matthias Karner, Georg Traar, Jan Henjes, and Wilfried Sihn. 2018. Digital Twin in Manufacturing: A Categorical Literature Review and Classification. *IFAC-PapersOnLine* 51, 11 (2018), 1016–1022. 16th IFAC Symposium on Information Control Problems in Manufacturing.

[125] Jens Krüger. 2020. *Digital Twin for Maximum Cyber Security.* Technical Report. NTT DATA Deutschland GmbH, Munich, Germany.

[126] K. Kruger, A. J. H. Redelinghuys, A. H. Basson, and O. Cardin. 2021. Past and Future Perspectives on Digital Twin Research at Sohoma. In *Service Oriented, Holonic and Multi-Agent Manufacturing Systems for Industry of the Future*, Theodor Borangiu, Damien Trentesaux, Paulo Leitão, Olivier Cardin, and Samir Lamouri (Eds.). Vol. 952. 81–98.

[127] V. M. Krundyshev. 2020. Ensuring Cybersecurity of Digital Production Using Modern Neural Network Methods. *Autom. Control Comput. Sci.* 54, 8 (2020), 786–792.

[128] Prabhat Kumar, Randhir Kumar, Abhinav Kumar, A. Antony Franklin, Sahil Garg, and Satinder Singh. 2022. Blockchain and Deep Learning for Secure Communication in Digital Twin Empowered Industrial IoT Network. *IEEE Trans. Netw. Sci. Eng.* (2022), 1–13.

[129] Nandha Kumar Kandasamy, Sarad Venugopalan, Tin Kit Wong, and Leu Junming Nicholas. 2021. Epictwin: An Electric Power Digital Twin for Cyber Security Testing, Research and Education. *CoRR* abs/2105.04260.

[130] Andre Kummerow, Cristian Monsalve, Dennis Rosch, Kevin Schafer, and Steffen Nicolai. 2020. Cyber-physical Data Stream Assessment Incorporating Digital Twins in Future Power Systems. In *2020 International Conference on Smart Energy Systems and Technologies (SEST)*. 1–6.

[131] Andre Kummerow, Steffen Nicolai, Christoph Brosinsky, Dirk Westermann, Andre Naumann, and Marc Richter. 2020. Digital-twin Based Services for Advanced Monitoring and Control of Future Power Systems. In *2020 IEEE Power Energy Society General Meeting (PESGM)*. 1–5.

[132] Heikki Laaki, Yoan Miche, and Kari Tammi. 2019. Prototyping a Digital Twin for Real Time Remote Control Over Mobile Networks: Application of Remote Surgery. *IEEE Access* 7 (2019), 20325–20336.

[133] Zeqi Lai, Yangtao Deng, Hewu Li, Qian Wu, and Qi Zhang. 2023. Space Digital Twin for Secure Satellite Internet: Vulnerabilities, Methodologies and Future Directions. *IEEE Network* (2023), 1–1.

[134] Zhongcheng Lei, Hong Zhou, Wenshan Hu, Guo-Ping Liu, Shiqi Guan, and Xingle Feng. 2021. Towards a Web-based Digital Twin Thermal Power Plant. *IEEE Trans. Ind. Inf.* (2021).

[135] Qian Li, Dongdong Huo, and Lizhong Jiang. 2022. A Digital Twin System for Monitoring the Security of Theatrical Stages. In *2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta)*. 2224–2230.

[136] Kendrik Yan Hong Lim, Pai Zheng, and Chun-Hsien Chen. 2019. A State-of-the-Art Survey of Digital Twin: Techniques, Engineering Product Lifecycle Management and Business Innovation Perspectives. *J. Intell. Manuf.* 31, 6 (2019), 1313–1337.

[137] Yu-Zheng Lin, Sicong Shao, Md Habibor Rahman, Mohammed Shafae, and Pratik Satam. 2023. DT4I4-Secure: Digital Twin Framework for Industry 4.0 Systems Security. In *2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. 0200–0209.

[138] Mengnan Liu, Shuiliang Fang, Huiyue Dong, and Cunzhi Xu. 2021. Review of Digital Twin about Concepts, Technologies, and Industrial Applications. *J. Manuf. Syst.* 58 (2021), 346–361. Digital Twin towards Smart Manufacturing and Industry 4.0.

[139] Zhen Liu, QingLei Zhang, Jianguo Duan, and Dong Liu. 2022. Digital twin–based testing process management for large and complex equipment components. *International journal of advanced manufacturing technology* 121, 5-6 (2022), 3143–3161.

[140] Salvador Llopis Sanchez, Robert Mazzolin, Ioannis Kechaoglou, Douglas Wiemer, Wim Mees, and Jean Muylaert. 2020. Cybersecurity Space Operation Center: Countering Cyber Threats in the Space Domain. In *Handbook of Space Security*, Kai-Uwe Schrogl (Ed.). 921–939.

[141] Andreas Locklin, Manuel Muller, Tobias Jung, Nasser Jazdi, Dustin White, and Michael Weyrich. 2020. Digital Twin for Verification and Validation of Industrial Automation Systems – a Survey. In *25th International Conference on Emerging Technologies and Factory Automation*, Vol. 1. 851–858.

[142] Javier Lopez, Juan E. Rubio, and Cristina Alcaraz. 2021. Digital Twins for Intelligent Authorization in the B5g-enabled Smart Grid. *IEEE Wirel. Commun.* 28, 2 (2021), 48–55.

[143] Tom H Luan, Ruhan Liu, Longxiang Gao, Rui Li, and Haibo Zhou. 2021. The Paradigm of Digital Twin Communications. *arXiv preprint arXiv:2105.07182* (2021).

[144] Rivera Luis F., Villegas Norha M., Tamura Gabriel, Müller Hausi A., Watts Ian, Erpenbach Eric, Shwartz Laura, and Liu Xiaotong. 2023. Using Digital Twins for Software Change Risk Assessment Toward Proactive AIOps. In *Proceedings of the 33rd Annual International Conference on Computer Science and Software Engineering (CASCON '23)*. 211–216.

[145] Azad M. Madni. 2021. Mbse Testbed for Rapid, Cost-effective Prototyping and Evaluation of System Modeling Approaches. *Applied sciences-basel* 11, 5 (2021), 2321.

[146] David Maher. 2018. *On Software Standards and Solutions for a Trusted Internet of Things.*

[147] Laurent Maillet-Contoz, Emmanuel Michel, Mario Diaz Nava, Paul-Emmanuel Brun, Kevin Lepretre, and Guillemette Massot. 2020. End-to-end Security Validation of Iot Systems Based on Digital Twins of End-devices. In *2020 Global Internet of Things Summit (GIoTS)*. 1–6.

[148] Stefan Marksteiner, Slava Bronfman, Markus Wolf, and Eddie Lazebnik. 2021. Using Cyber Digital Twins for Automated Automotive Cybersecurity Testing. In *2021 IEEE European Symposium on Security and Privacy Workshops (EuroS PW)*. 123–128.

[149] Beatriz F. Martins, Lenin Serrano, José F. Reyes, José Ignacio Panach, Oscar Pastor, and Benny Rochwerger. 2020. Conceptual Characterization of Cybersecurity Ontologies. In *The Practice of Enterprise Modeling*, Janis Grabis and Dominik Bork (Eds.). Vol. 400. 323–338.

[150] Massimiliano Masi, Giovanni Paolo Sellitto, Helder Aranha, and Tanja Pavleska. 2023. Securing critical infrastructures with a cybersecurity digital twin. *Software and Systems Modeling* 22, 2 (2023), 689–707.

[151] Aditya P. Mathur. 2023. Reconfigurable Digital Twin to Support Research, Education, and Training in the Defense of Critical Infrastructure. *IEEE SECURITY & PRIVACY* 21, 4 (2023), 51–60.

[152] Roberto Minerva, Gyu Myoung Lee, and Noel Crespi. 2020. Digital Twin in the Iot Context: A Survey on Technical Features, Scenarios, and Architectural Models. In *Proceedings of the IEEE*, Vol. 108. 1785–1824.

[153] Carlos Miskinis. 2018. *Incorporating Digital Twin into Internet Cyber Security – Creating a Safer Future.* Retrieved June 28, 2021 from https://www.challenge.org/insights/digital-twin-cyber-security/

[154] Saurabh Mittal, Andreas Tolk, Andrew Pyles, Nicolas Van Balen, and Kevin Bergollo. 2019. Digital Twin Modeling, Co-simulation and Cyber Use-case Inclusion Methodology for Iot Systems. In *2019 Winter Simulation Conference (WSC) (WSC '19)*. 2653–2664.

[155] Alvaro Cárdenas Mora, Simin Nadjm-Tehrani, Edgar Weippl, and Matthias Eckhart. 2022. Digital Twins for Cyber-Physical Systems Security (Dagstuhl Seminar 22171). *Dagstuhl Reports* 12, 4 (2022), 54–71.

[156] Fereidoun Moradi, Bahman Pourvatan, Sara Abbaspour Asadollah, and Marjan Sirjani. 2024. Tiny Twins for detecting cyber-attacks at runtime using concise Rebeca time transition system. *J. Parallel and Distrib. Comput.* 184 (2024), 104780.

[157] Jose Andre Morales, Thomas P. Scanlon, Aaron Volkmann, Joseph Yankel, and Hasan Yasar. 2020. Security Impacts of Sub-optimal Devsecops Implementations in a Highly Regulated Environment. In *Proceedings of the 15th International Conference on Availability, Reliability and Security*.

[158] Evandro Pioli Moro. 2021. Semantic Digital Twins: Security and Scale for Constrained Iot Devices. (2021).

[159] Valentin Mullet, Patrick Sondi, and Eric Ramat. 2021. A Review of Cybersecurity Guidelines for Manufacturing Factories in Industry 4.0. *IEEE Access* 9 (2021), 23235–23263.

[160] Andres Murillo, Riccardo Taormina, Nils Tippenhauer, and Stefano Galelli. 2020. Co-simulating Physical Processes and Network Data for High-fidelity Cyber-security Experiments. In *Sixth Annual Industrial Control System Security (ICSS) Workshop (ICSS 2020)*. 13–20.

[161] S Neethirajan and B Kemp. 2021. Digital Twins in Livestock Farming. Animals 2021, 11, 1008. , 1008 pages.

[162] Elisa Negri, Luca Fumagalli, and Marco Macchi. 2017. A Review of the Roles of Digital Twin in CPS-based Production Systems. *Procedia Manufacturing* 11 (2017), 939–948. 27th International Conference on Flexible Automation and Intelligent Manufacturing.

[163] Luong Nguyen, Mariana Segovia, Wissam Mallouli, Edgardo Montes de Oca, and Ana R. Cavalli. 2022. Digital Twin for IoT Environments: A Testing and Simulation Tool. In *Quality of Information and Communications Technology*, Antonio Vallecillo, Joost Visser, and Ricardo Pérez-Castillo (Eds.). Vol. 1621. 205–219.

[164] Maria Nintsiou, Elisavet Grigoriou, Paris Alexandros Karypidis, Theocharis Saoulidis, Eleftherios Fountoukidis, and Panagiotis Sarigiannidis. 2023. Threat intelligence using Digital Twin honeypots in Cybersecurity. In *2023 IEEE International Conference on Cyber Security and Resilience (CSR)*. 530–537.

[165] Juan C. Olivares-Rojas, Enrique Reyes-Archundia, Jose A. Gutierrez-Gnecchi, Ismael Molina-Moreno, Jaime Cerda-Jacobo, and Arturo Mendez-Patino. 2021. Towards Cybersecurity of the Smart Grid Using Digital Twins. *IEEE Internet Computing* (2021), 1–1.

[166] Vittorio Orbinato. 2021. A next-generation platform for Cyber Range-as-a-Service. In *2021 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*. 314–318.

[167] Guy Paré, Marie-Claude Trudel, Mirou Jaana, and Spyros Kitsiou. 2015. Synthesizing Information Systems Knowledge: A Typology of Literature Reviews. *Information & Management* 52, 2 (2015), 183–199.

[168] Abhi Patel, Tim Schenk, Steffi Knorn, Heiko Patzlaff, Dragan Obradovic, and Andrés Botero Halblaub. 2021. Real-time, Simulation-based Identification of Cyber-security Attacks of Industrial Plants. In *2021 IEEE International Conference on Cyber Security and Resilience (CSR)*. 267–272.

[169] Riccardo Patriarca, Francesco Simone, and Giulio Di Gravio. 2022. Modelling cyber resilience in a water treatment and distribution system. *Reliab. Eng. Syst. Safe.* 226 (2022), 108653.

[170] Sandeep Pirbhulal, Habtamu Abie, and Ankur Shukla. 2022. Towards a Novel Framework for Reinforcing Cybersecurity using Digital Twins in IoT-based Healthcare Applications. In *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*. 1–5.

[171] Sandeep Pirbhulal, Habtamu Abie, Ankur Shukla, and Basel Katt. 2023. A Cognitive Digital Twin Architecture for Cybersecurity in IoT-Based Smart Homes. In *Sensing Technology*, Nagender Kumar Suryadevara, Boby George, Krishanthi P. Jayasundera, and Subhas Chandra Mukhopadhyay (Eds.). Vol. 1035. 63–70.

[172] Iakovos Pittaras and George C. Polyzos. 2022. (POSTER) SmartTwins: Secure and Auditable DLT-based Digital Twins for the WoT. In *2022 18th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. 82–84.

[173] Abhishek Pokhrel, Vikash Katta, and Ricardo Colomo-Palacios. 2020. Digital Twin for Cybersecurity Incident Prediction: A Multivocal Literature Review. In *Proceedings of the IEEE/ACM 42nd. International Conference on Software Engineering Workshops (ICSEW'20)*. 671–678.

[174] J. Pushpa and S.A. Kalyani. 2020. Using fog computing/edge computing to leverage Digital Twin. In *The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases*, Pethuru Raj and Preetha Evangeline (Eds.). Advances in Computers, Vol. 117. 51–77.

[175] Benedikt Putz, Marietheres Dietz, Philip Empl, and Günther Pernul. 2021. EtherTwin: Blockchain-based Secure Digital Twin Information Management. *Inform. Process. Manag.* 58, 1 (2021), 102425.

[176] Pethuru Raj. 2021. Empowering Digital Twins with Blockchain. *Adv. Comput.* 121 (2021), 267–283.

[177] Roza Ranjbar, Eric Duviella, Lucien Etienne, and Jose-Maria Maestre. 2020. Framework for a Digital Twin of the Canal of Calais. *Procedia Comput. Sci.* 178 (2020), 27–37.

[178] Adil Rasheed, Omer San, and Trond Kvamsdal. 2020. Digital Twin: Values, Challenges and Enablers from a Modeling Perspective. *IEEE Access* 8 (2020), 21980–22012.

[179] M. Mazhar Rathore, Syed Attique Shah, Dhirendra Shukla, Elmahdi Bentafat, and Spiridon Bakiras. 2021. The Role of Ai, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities. *IEEE Access* 9 (2021), 32030–32052.

[180] Elaine M. Raybourn and Ray Trechter. 2018. Applying Model-based Situational Awareness and Augmented Reality to Next-generation Physical Security Systems. In *Cyber-Physical Systems Security*, Çetin Kaya Koç (Ed.). 331–344.

[181] Lúcio Henrik A. Reis, Andrés Murillo Piedrahita, Sandra Rueda, NatAlia C. Fernandes, Dianne S. V. Medeiros, Marcelo Dias Amorim, and Diogo M. F. Mattos. 2020. Unsupervised and Incremental Learning Orchestration for Cyber Physical Security. *Transactions on Emerging Telecommunications Technologies* 31, 7 (2020), e4011.

[182] Michael Riegler and Johannes Sametinger. 2021. Multi-mode Systems for Resilient Security in Industry 4.0. *Procedia Comput. Sci.* 180 (2021), 301–307.

[183] Roland Rosen, Jan Fischer, and Stefan Boschert. 2019. Next Generation Digital Twin: An Ecosystem for Mechatronic Systems?, In Proc. tmce. *IFAC-PapersOnLine* 52, 15, 265–270.

[184] Juan Enrique Rubio, Cristina Alcaraz, Rodrigo Roman, and Javier Lopez. 2019. Current Cyber-defense Trends in Industrial Control Systems. *Comput. Secur.* 87 (2019), 101561.

[185] Johnson S. A. Osho and David A. Umphress. 2022. IoT Security: Modeling, Development and Validation of IoT.

[186] Ahmed Saad, Samy Faddel, and Osama Mohammed. 2020. Iot-based Digital Twin for Energy Cyber-physical Systems: Design and Implementation. *Energies* 13, 18 (2020), 4762.

[187] Ahmed Saad, Samy Faddel, Tarek Youssef, and Osama A. Mohammed. 2020. On the Implementation of Iot-based Digital Twin for Networked Microgrids Resiliency against Cyber Attacks. *IEEE Trans. Smart Grid* 11, 6 (2020), 5138–5150. Issue 6.

[188] Zoheir Sabeur, Constantinos Marios Angelopoulos, Liam Collick, Natalia Chechina, Deniz Cetinkaya, and Alessandro Bruno. 2021. Advanced Cyber and Physical Situation Awareness in Urban Smart Spaces. In *Advances in Neuroergonomics and Cognitive Engineering*, Hasan Ayaz, Umer Asgher, and Lucas Paletta (Eds.). Vol. 259. 428–441.

[189] Radhya Sahal, Saeed H. Alsamhi, John G. Breslin, Kenneth N. Brown, and Muhammad Intizar Ali. 2021. Digital Twins Collaboration for Automatic Erratic Operational Data Detection in Industry 4.0. *Applied Sciences* 11, 7 (2021), 3186.

[190] Najm Us Sama, Kartinah Zen, Aziz Ud Din, Nazia Azim, and Atiq Ur Rahman. 2023. Security of Cloud-Assisted BANs Using Digital Twin. In *2023 13th International Conference on Information Technology in Asia (CITA)*. 37–42.

[191] Guido Schryen. 2015. Writing Qualitative IS Literature Reviews - Guidelines for Synthesis, Interpretation, and Guidance of Research. *Communication of the Association for Information Systems* 37 (2015), 12.

[192] Kirill Semenkov, Vitaly Promyslov, and Alexey Poletykin. 2020. Verification of Large Scale Control Systems with Hybrid Digital Models and Digital Twins. In *2020 International Russian Automation Conference (RusAutoCon)*. 325–329.

[193] Kirill Semenkov, Vitaly Promyslov, Alexey Poletykin, and Nadir Mengazetdinov. 2021. Validation of Complex Control Systems with Heterogeneous Digital Models in Industry 4.0 Framework. *Machines* 9, 3 (2021), 62.

[194] Omer Sen, Florian Schmidtke, Federico Carere, Francesca Santori, Andreas Ulbig, and Antonello Monti. 2022. Investigating the Cybersecurity of Smart Grids Based on Cyber-Physical Twin Approach. In *2022 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. 439–445.

[195] Ömer Sen, Nathalie Bleser, and Andreas Ulbig. 2023. Digital Twin for Evaluating Detective Countermeasures in Smart Grid Cybersecurity. In *2023 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. 1–6.

[196] Karan Shah, T. V. Prabhakar, Sarweshkumar C. R., Abhishek S. V., and Vasanth Kumar T. 2021. Construction of a Digital Twin Framework Using Free and Open-source Software Programs. *IEEE Internet Comput.* (2021), 1–1.

[197] Andrii Shalaginov, Igor Kotsiuba, and Asif Iqbal. 2019. Cybercrime Investigations in the Era of Smart Applications: Way Forward through Big Data. In *2019 IEEE International Conference on Big Data (Big Data)*. 4309–4314.

[198] Lei Shi, Shanti Krishnan, and Sheng Wen. 2022. Study Cybersecurity of Cyber Physical System in the Virtual Environment: A Survey and New Direction. In *Australasian Computer Science Week 2022 (ACSW 2022)*. 46–55.

[199] Amardeep B. Shitole, Nandha Kumar Kandasamy, Lin Shen Liew, Lewin Sim, and Anh Khoa Bui. 2021. Real-Time Digital Twin of Residential Energy Storage System for Cyber-Security Study. In *2021 International Conference on Smart Technologies for Power, Energy and Control*. 1–6.

[200] Datta Shoumen Palit Austin. 2017. Emergence of Digital Twins. *Emergence of Digital Twins* 5 (2017), 14–33.

[201] Bruno Sousa, Miguel Arieiro, Vasco Pereira, João Correia, Nuno Lourenço, and Tiago Cruz. 2021. Elegant: Security of Critical Infrastructures with Digital Twins. *IEEE Access* 9 (2021), 107574–107588.

[202] Anurag Srivastava, Chen-Ching Liu, Alexandru Stefanov, Sagnik Basumallik, Mohammed M. Hussain, Baza Somda, and Vetrivel S. Rajkumar. 2024. Digital Twins Serving Cybersecurity: More Than a Model: Cybersecurity as a Future Benefit of Digital Twins 2. *IEEE Power and Energy Magazine* 22, 1 (2024), 61–71.

[203] Gayathri Sugumar and Aditya Mathur. 2019. Assessment of a Method for Detecting Process Anomalies Using Digital-twinning. In *2019 15th European Dependable Computing Conference (EDCC)*. 119–126.

[204] Sabah Suhail, Mubashar Iqbal, Rasheed Hussain, and Raja Jurdak. 2023. ENIGMA: An explainable digital twin security solution for cyber–physical systems. *Computers in Industry* 151 (2023), 103961.

[205] Sabah Suhail, Mubashar Iqbal, and Raja Jurdak. 2024. The Perils of Leveraging Evil Digital Twins as Security-Enhancing Enablers. *Commun. ACM* 67, 1 (2024), 39–42.

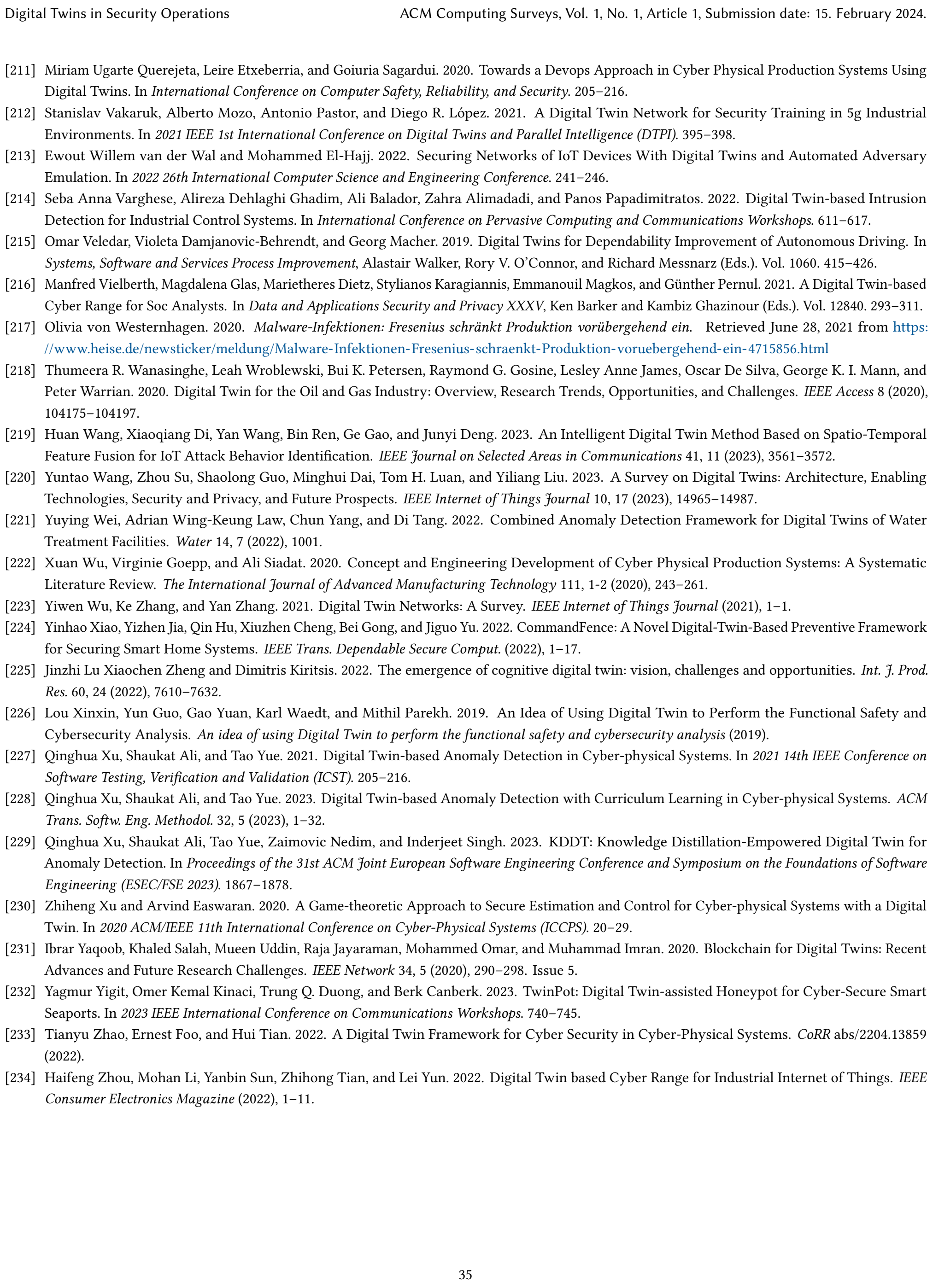
[206] Koen Tange, Michele De Donno, Xenofon Fafoutis, and Nicola Dragoni. 2020. A Systematic Survey of Industrial Internet of Things Security: Requirements and Fog Computing Opportunities. *IEEE Communications Surveys & Tutorials* 22, 4 (2020), 2489–2520.

[207] Fei Tao, Meng Zhang, and A.Y.C. Nee. 2019. Applications of Digital Twin. In *Digital Twin Driven Smart Manufacturing*, Fei Tao, Meng Zhang, and A.Y.C. Nee (Eds.). 29–62.

[208] Nipuna Sankalpa Thalpage and Thebona Arachchige Dushyanthi Nisansala. 2023. Exploring the Opportunities of Applying Digital Twins for Intrusion Detection in Industrial Control Systems of Production and Manufacturing – A Systematic Review. In *Data Protection in a Post-Pandemic Society*, Chaminda Hewage, Yogachandran Rahulamathavan, and Deepthi Ratnayake (Eds.). 113–143.

[209] William Tärneberg, Per Skarin, Christian Gehrmann, and Maria Kihl. 2021. Prototyping Intrusion Detection in an Industrial Cloud-native Digital Twin. In *International Conference on Industrial Technology.*

[210] Latif U. Khan, Walid Saad, Dusit Niyato, Zhu Han, and Choong Seon Hong. 2021. Digital-twin-enabled 6g: Vision, Architectural Trends, and Future Directions.



[211] Miriam Ugarte Querejeta, Leire Etxeberria, and Goiuria Sagardui. 2020. Towards a Devops Approach in Cyber Physical Production Systems Using Digital Twins. In *International Conference on Computer Safety, Reliability, and Security*. 205–216.

[212] Stanislav Vakaruk, Alberto Mozo, Antonio Pastor, and Diego R. López. 2021. A Digital Twin Network for Security Training in 5g Industrial Environments. In *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*. 395–398.

[213] Ewout Willem van der Wal and Mohammed El-Hajj. 2022. Securing Networks of IoT Devices With Digital Twins and Automated Adversary Emulation. In *2022 26th International Computer Science and Engineering Conference*. 241–246.

[214] Seba Anna Varghese, Alireza Dehlaghi Ghadim, Ali Balador, Zahra Alimadadi, and Panos Papadimitratos. 2022. Digital Twin-based Intrusion Detection for Industrial Control Systems. In *International Conference on Pervasive Computing and Communications Workshops*. 611–617.

[215] Omar Veledar, Violeta Damjanovic-Behrendt, and Georg Macher. 2019. Digital Twins for Dependability Improvement of Autonomous Driving. In *Systems, Software and Services Process Improvement*, Alastair Walker, Rory V. O'Connor, and Richard Messnarz (Eds.). Vol. 1060. 415–426.

[216] Manfred Vielberth, Magdalena Glas, Marietheres Dietz, Stylianos Karagiannis, Emmanouil Magkos, and Günther Pernul. 2021. A Digital Twin-based Cyber Range for Soc Analysts. In *Data and Applications Security and Privacy XXXV*, Ken Barker and Kambiz Ghazinour (Eds.). Vol. 12840. 293–311.

[217] Olivia von Westernhagen. 2020. *Malware-Infektionen: Fresenius schränkt Produktion vorübergehend ein.* Retrieved June 28, 2021 from https://www.heise.de/newsticker/meldung/Malware-Infektionen-Fresenius-schraenkt-Produktion-voruebergehend-ein-4715856.html

[218] Thumeera R. Wanasinghe, Leah Wroblewski, Bui K. Petersen, Raymond G. Gosine, Lesley Anne James, Oscar De Silva, George K. I. Mann, and Peter Warrian. 2020. Digital Twin for the Oil and Gas Industry: Overview, Research Trends, Opportunities, and Challenges. *IEEE Access* 8 (2020), 104175–104197.

[219] Huan Wang, Xiaoqiang Di, Yan Wang, Bin Ren, Ge Gao, and Junyi Deng. 2023. An Intelligent Digital Twin Method Based on Spatio-Temporal Feature Fusion for IoT Attack Behavior Identification. *IEEE Journal on Selected Areas in Communications* 41, 11 (2023), 3561–3572.

[220] Yuntao Wang, Zhou Su, Shaolong Guo, Minghui Dai, Tom H. Luan, and Yiliang Liu. 2023. A Survey on Digital Twins: Architecture, Enabling Technologies, Security and Privacy, and Future Prospects. *IEEE Internet of Things Journal* 10, 17 (2023), 14965–14987.

[221] Yuying Wei, Adrian Wing-Keung Law, Chun Yang, and Di Tang. 2022. Combined Anomaly Detection Framework for Digital Twins of Water Treatment Facilities. *Water* 14, 7 (2022), 1001.

[222] Xuan Wu, Virginie Goepp, and Ali Siadat. 2020. Concept and Engineering Development of Cyber Physical Production Systems: A Systematic Literature Review. *The International Journal of Advanced Manufacturing Technology* 111, 1-2 (2020), 243–261.

[223] Yiwen Wu, Ke Zhang, and Yan Zhang. 2021. Digital Twin Networks: A Survey. *IEEE Internet of Things Journal* (2021), 1–1.

[224] Yinhao Xiao, Yizhen Jia, Qin Hu, Xiuzhen Cheng, Bei Gong, and Jiguo Yu. 2022. CommandFence: A Novel Digital-Twin-Based Preventive Framework for Securing Smart Home Systems. *IEEE Trans. Dependable Secure Comput.* (2022), 1–17.

[225] Jinzhi Lu Xiaochen Zheng and Dimitris Kiritsis. 2022. The emergence of cognitive digital twin: vision, challenges and opportunities. *Int. J. Prod. Res.* 60, 24 (2022), 7610–7632.

[226] Lou Xinxin, Yun Guo, Gao Yuan, Karl Waedt, and Mithil Parekh. 2019. An Idea of Using Digital Twin to Perform the Functional Safety and Cybersecurity Analysis. *An idea of using Digital Twin to perform the functional safety and cybersecurity analysis* (2019).

[227] Qinghua Xu, Shaukat Ali, and Tao Yue. 2021. Digital Twin-based Anomaly Detection in Cyber-physical Systems. In *2021 14th IEEE Conference on Software Testing, Verification and Validation (ICST)*. 205–216.

[228] Qinghua Xu, Shaukat Ali, and Tao Yue. 2023. Digital Twin-based Anomaly Detection with Curriculum Learning in Cyber-physical Systems. *ACM Trans. Softw. Eng. Methodol.* 32, 5 (2023), 1–32.

[229] Qinghua Xu, Shaukat Ali, Tao Yue, Zaimovic Nedim, and Inderjeet Singh. 2023. KDDT: Knowledge Distillation-Empowered Digital Twin for Anomaly Detection. In *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2023)*. 1867–1878.

[230] Zhiheng Xu and Arvind Easwaran. 2020. A Game-theoretic Approach to Secure Estimation and Control for Cyber-physical Systems with a Digital Twin. In *2020 ACM/IEEE 11th International Conference on Cyber-Physical Systems (ICCPS)*. 20–29.

[231] Ibrar Yaqoob, Khaled Salah, Mueen Uddin, Raja Jayaraman, Mohammed Omar, and Muhammad Imran. 2020. Blockchain for Digital Twins: Recent Advances and Future Research Challenges. *IEEE Network* 34, 5 (2020), 290–298. Issue 5.

[232] Yagmur Yigit, Omer Kemal Kinaci, Trung Q. Duong, and Berk Canberk. 2023. TwinPot: Digital Twin-assisted Honeypot for Cyber-Secure Smart Seaports. In *2023 IEEE International Conference on Communications Workshops*. 740–745.

[233] Tianyu Zhao, Ernest Foo, and Hui Tian. 2022. A Digital Twin Framework for Cyber Security in Cyber-Physical Systems. *CoRR* abs/2204.13859 (2022).

[234] Haifeng Zhou, Mohan Li, Yanbin Sun, Zhihong Tian, and Lei Yun. 2022. Digital Twin based Cyber Range for Industrial Internet of Things. *IEEE Consumer Electronics Magazine* (2022), 1–11.

P2 Digital Twins for IoT Security Management

Status:	Published
Date of Submission:	05 April 2023
Date of Acceptance:	24 May 2023
Date of Publication:	12 July 2023
Conference:	37th Annual IFIP WG 11.3 Conference on Data and Applications Security and Privacy (DBSec 2023)
Location:	SAP Labs France, Sophia Antipolis, France
Period:	19.07.2023 – 21.07.2023
Authors' Contributions:	Philip Empl 45% Henric Hager 45% Günther Pernul 10%
Full Citation:	Empl, P., Hager, H., & Pernul, G. (2024). Digital Twins for IoT Security Management. In Atluri, V., Ferrara, A.L. (Eds.), *Proceedings of the 37th Annual IFIP WG 11.3 Conference on Data and Applications Security and Privacy* (pp. 141—149). Springer Nature.
DOI:	10.1007/9783031375866_9
Artifact:	github.com/Ric1234567/DigitalTwinsForIoTSecurityManagement

Conference Description: The 37th edition of the Annual IFIP WG 11.3 Conference on Data and Applications Security and Privacy (DBSec 2023) will take place in Sophia Antipolis, France. The conference brings together researchers, practitioners, and experts from academia, industry, and government to share their cutting-edge findings and insights in all theoretical and practical aspects of data protection, privacy, and applications security.

Digital Twins for IoT Security Management

Philip Empl(✉), Henric Hager, and Günther Pernul

Department of Information Systems, University of Regensburg, Universitätsstraße 31,
93053 Regensburg, Germany
{philip.empl,henric.hager,guenther.pernul}@ur.de

Abstract. The proliferation of Internet of Things (IoT) devices has increased the risk of cyber threats to the confidentiality, integrity, and availability of data processed. In this context, proactive security management has emerged as a critical strategy for protecting assets and networks. However, managing the complex nature of IoT devices poses a significant challenge to effective IoT security management. Digital twins - virtual replicas of physical assets - offer a promising solution to address these challenges. By creating a digital twin of an IoT network, security analysts could continuously monitor and regulate the IoT network, detect potential problems before they escalate, and assess the impact of new configurations and updates without risking the physical ones. This paper proposes a concept that uses digital twins for proactive IoT security management. To this end, we implement a proof of concept to demonstrate the practical applicability of this approach for four different security use cases. Our results provide a starting point for further research to leverage digital twins for IoT security management.

Keywords: Internet of Things · Network · Cybersecurity · Security management · Digital twin · Digital twin network

1 Introduction

The pervasive use of Internet of Things (IoT) devices in our daily lives, ranging from smart home appliances to industrial automation systems, has led to significant threats to the availability, integrity, and confidentiality of data processed within IoT networks. The escalating growth of cybercrime is estimated to result in a staggering worldwide cost of $10.5 trillion by 2025, underscoring the importance of addressing security issues (i.e., insecure configurations or communication) in IoT networks [4,16]. In response, proactive security and automation have emerged as crucial strategies to safeguard assets and their associated networks [14,17]. By proactively identifying and addressing security risks, organizations can mitigate the likelihood of potential cyber-attacks, protecting sensitive data and ensuring the privacy of their networks [7]. In light of these

V. Atluri and A. L. Ferrara (Eds.): DBSec 2023, LNCS 13942, pp. 141–149, 2023.
https://doi.org/10.1007/978-3-031-37586-6_9

developments, proactive network security management has become a relevant and necessary area of focus for protecting the security of organizational networks.

However, effective network security management in IoT has become an increasingly daunting task due to the intricate nature of IoT devices [1]. These devices exhibit diverse hardware, operating systems, and communication protocols, rendering them challenging to secure using traditional network security management. Furthermore, IoT devices are often tailored to specific use cases and are constrained by their processing power and memory, compounding network management's difficulty. To address these challenges, digital twins offer a promising solution for leveraging the security of IoT networks [7]. They are often considered as they provide a precise and comprehensive virtual representation of the physical IoT network, making IoT devices visible, which are usually hidden when performing IP network scans. By creating a digital twin of an IoT network, security analysts could continuously monitor the IoT network, detect potential issues before they escalate, and evaluate the effects of new configurations and updates without jeopardizing the physical network. To our knowledge, digital twins have not yet been considered for IoT security management.

This paper leverages digital twins to address the pressing need for proactive security management in IoT networks. Given the increasing threat landscape and the inherent complexity of IoT, traditional network management techniques have proven insufficient, necessitating alternative and more sophisticated solutions. Digital twins are an ideal solution as they allow for proactive network monitoring, analysis, and optimization to enhance security. Thus, this paper's research question is formulated as follows: "*How can digital twins enable IoT security management?*". The contribution of this paper is twofold: firstly, we design a concept for leveraging digital twins in the context of proactive IoT security management. Secondly, we implement a proof of concept, demonstrating this concept's practical applicability through four different security use cases.

The remainder of this paper is structured as follows. Section 2 provides a detailed exposition of the background on IoT network management and digital twins. Section 3 presents our concept showing the use of digital twins to manage IoT networks. In Section 4, we validate our proposed concept by implementing a proof of concept considering four security use cases and measuring the performance outcomes of each. Finally, in Section 5, we conclude the paper by summarizing our research.

2 Background

2.1 IoT Network Management

Network management refers to *monitoring* and *optimizing* a computer network using methods like network verification [15], or testing [8]. Following the network management cycle is crucial to automate tasks, starting from measurement, moving towards decision-making, action strategy, verification, and execution [3].

According to the International Standard Organization (ISO), traditional network management encompasses five fundamental aspects, commonly referred to as fault, configuration, accounting, performance, and security management (FCAPS) [13]. *Fault management (F)* deals with detecting, isolating, and correcting faults or errors in the network. *Configuration management (C)* deals with maintaining accurate and up-to-date information about the network's configuration, e.g., hardware, software, topology, or policies. *Accounting management (A)* deals with tracking network resource usage, including bandwidth, storage, and CPU cycles. *Performance management (P)* monitors and manages the network's performance. *Security management (S)* aims to protect the network from unauthorized access, data theft, and other security threats.

Although these aspects have been specifically designed for computer networks, they apply to IoT as well [1]. The idea behind IoT is based on many heterogeneous devices that are identifiable and interconnected through dedicated communication networks [2]. As defined by the IEEE 802.15.4 standard, such devices are termed reduced functional devices, with a narrower range of functionalities than full functional devices. The German Federal Office for Information Security (BSI) categorizes IoT devices as controllable when operating within a sensor network or addressable when communicating within a TCP/IP network [11]. In this paper, we adopt the definition provided by the BSI.

2.2 Digital Twin Network

A digital twin is a virtual representation of a physical asset used for simulation, replication, or analytics [5]. It maintains bidirectional communication with its physical counterpart. Digital twins are used for many security operations, e.g., cyber ranges [18] or security simulations [6]. A digital twin network addresses the complexities of real-world networks by establishing mappings between multiple digital twins [19]. It incorporates various network metrics, such as topology, routing, and traffic, enabling comprehensive modeling [9,10]. Digital twin networks facilitate "what-if" analyses based on periodic configuration and real-time data collection [12].

3 Digital Twin-Based IoT Security Management

We propose a concept that combines the network management cycle methodology [3] with digital twins to enable security management in IoT networks, as depicted in Fig. 1. This concept addresses the challenges posed by IoT devices' heterogeneity and their networks' dynamic nature. Digital twins play a crucial role in modeling IoT devices, making them visible, providing security recommendations, and facilitating verification and optimization. Our proposed concept comprises three interconnected components: the real world, virtual representation, and network management.

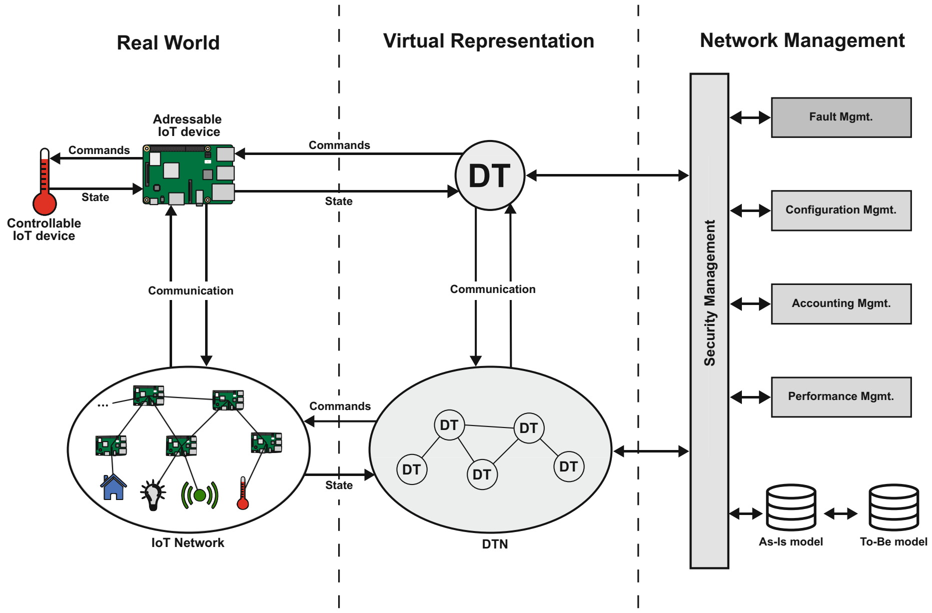


Fig. 1. Digital twin-based IoT security management.

The real world encompasses IoT devices such as laptops, sensors, actuators, controllers, and wearables. Addressable IoT devices serve as coordinators and define access control policies, while controllable IoT devices are limited nodes communicating with coordinators. The IP networks define the primary networks in the network topology, and controllable IoT networks become subnets if an addressable IoT device functions as a coordinator. The network represents the communication relationships between addressable and controllable IoT devices.

The virtual representation models the dependencies of the real world and employs digital twins. Digital twins realistically represent IoT devices and networks, including information about addressable and controllable IoT devices. Digital twins enable bidirectional communication with physical assets and orchestrate commands in the real world. A digital twin network captures the relationships between digital twins and replicates the communication links in the real world. This combination of real-world information and digital twins facilitates subsequent IoT security management.

Network management utilizes digital twins and their networks to monitor and optimize IoT devices in the real world. It encompasses all aspects of FCAPS, focusing on security management, and is coordinated by a central manager. The As-Is model represents the current state of the virtual representation, while the To-Be model outlines the desired objectives regarding FCAPS. Security issues and possible optimizations can be identified by comparing these two models. Recommendations for optimizations are based on the To-Be model, but the final decision lies with the security analyst on whether to deploy them.

4 Proof of Concept

This section will validate the presented concept using four security use cases. These use cases are designed to improve the overall security posture of the IoT network using digital twins in analytics mode. We begin by detailing the security use cases that influence the experimental setting.

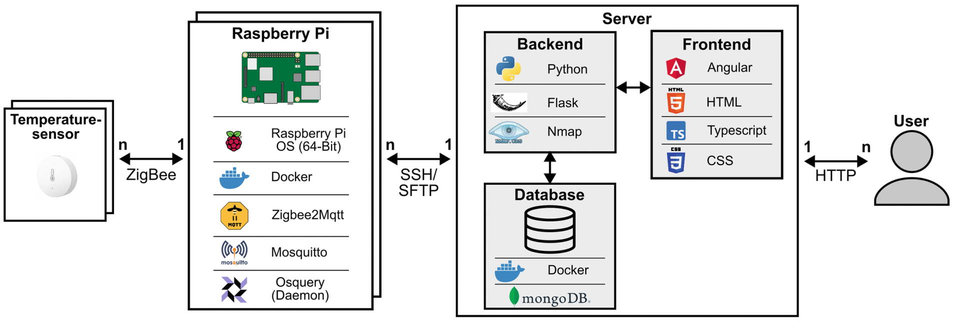


Fig. 2. Experimental setting detailing on the proof of concept.

4.1 Security Use Cases

In the realm of IoT, various security threats can compromise the confidentiality, integrity, and availability of data. We focus on addressable and controllable IoT devices, evaluating our concept against four security threats. *Open ports* (1) on addressable IoT devices can serve as gateways for attackers. To preserve device and network security, closing insecure network ports and leaving essential ones open is crucial. *Malicious USB dongles* (2) present another attack vector. To maintain network security, monitoring USB ports and implementing appropriate measures closely is essential. *Weak access controls* (3) pose significant risks to network confidentiality. Strict access control lists must be enforced to mitigate these risks. *Insecure configurations* (4) allow unauthorized devices access to the network, which should be optimized.

4.2 Experimental Setting

Our experimental setting, as shown in Fig. 2, addresses the security use cases in IoT networks. A digital twin is represented by the network's data stored in the database, while bidirectional communication aims at collecting data and optimizing IoT devices. The experiments were conducted using a Raspberry Pi 3B+ model running Raspberry Pi OS, and a CC2531 flashed ZigBee dongle. A MacOS machine with six cores, 16 GB RAM, and a 256 GB SSD acts as the server. Bidirectional communication between the server and the Raspberry Pi occurs via SSH and SFTP protocols. The Raspberry Pi hosts two critical containerized microservices: Zigbee2MQTT and Mosquitto. Zigbee2MQTT manages the ZigBee network, which includes controllable devices like Xiaomi temperature sensors. The controllable IoT network data is routed from the Zigbee network to the MQTT broker, creating an event-driven architecture. We use osquery, an open-source tool that provides a unified interface to different operating systems, analyzing and optimizing the Zigbee network. In addition to controllable IoT devices, we integrate multiple IP addressable devices, including routers and smartphones, expanding the scope of addressable IoT devices beyond the Raspberry Pi. Our proof of concept is available on GitHub[1].

Table 1. Security use cases, problem statements, and recommendations.

Use case	Problem statement	Recommendation
Open ports	Number of ports exceeds max	Terminate open ports
USB dongle	Unknown USB dongle	Unplug USB dongle
Weak access control	ACL inconsistency	Replace the ACL
Insecure config.	permit_join flag differs	Set permit_join flag

4.3 Results

We present our experimental results based on the optimization process by Arzo et al. [3]. We have findings for the different phases of the process and results related to the performance. In the *security assessment* phase, we extract data from IoT networks to create digital twins and a network representing the As-Is model. We develop an algorithm to strategically identify controllable IoT devices in IP networks, as they cannot be discovered solely through IP network scans. We successfully identify addressable and controllable IoT devices, storing the data in a MongoDB database. This data reflects the respective digital twins. Additional information, such as device logs and configuration files, is gathered for our pre-defined security use cases. All the data is automatically collected.

Optimization Process. During the *To-Be-model comparison* phase, the digital twins compare the respective network model (As-Is model) with the To-Be model. Analysts manually define the To-Be model, so we set the maximum number of open ports and establish secure configurations. The digital twin network emphasizes deviations between those two models as potential security issues. Last, we run the *optimization*, *test*, and *deployment* phases. Table 1 outlines the security use cases, their problem statements, and recommended optimizations. The digital twin network provides recommendations to analysts, who can verify and deploy optimizations to the real network.

[1] https://github.com/Ric1234567/DigitalTwinsForIoTSecurityManagement.

Performance Evaluation. We conduct a performance evaluation to assess the time required for monitoring and optimizing the IoT network (see Fig. 3). The evaluation reveals varying times for different security use cases, with USB dongle monitoring being the shortest and port monitoring being the longest. We successfully resolved all security use cases within the expected time frame, utilizing automation and the guidance of the digital twin network.

Limitations. We must consider the limitations of our research results in the context of IoT security management. Our proof of concept concentrates only on the digital twin in analytics mode and IoT networks' periodic monitoring and optimization. Therefore, we do not employ simulation, e.g., "what-if" analysis, and replication modes that could have aided in a more detailed optimization and verification throughout the process. Our proof of concept may also account for software-defined networks, which could be valuable in managing addressable and controllable IoT devices. However, this should not call into question the concept.

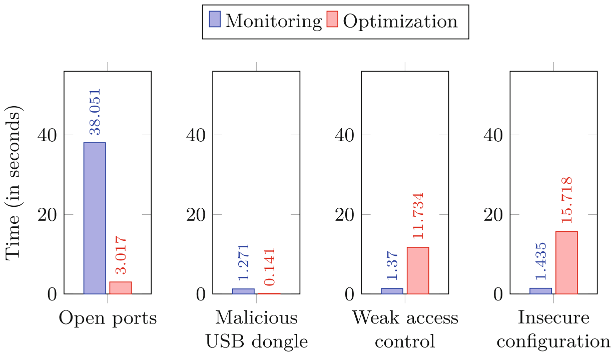


Fig. 3. Performance evaluation based on the four security use cases.

5 Conclusion

The heterogeneous nature of IoT networks poses significant challenges to traditional network management approaches, rendering them inadequate. To address this issue, we propose using digital twin networks to uncover controllable IoT devices that remain undetected through IP network scans. We illustrate the potential of digital twins and networks to enhance IoT security management. Our proof of concept showcases initial steps towards digital twins for IoT security management. Open challenges remain in developing more sophisticated digital twins, ensuring scalability, and enhancing data availability. Nevertheless, we maintain that digital twins possess immense potential to enable IoT security management, which should be further researched in future.

References

1. AboubDakar, M., Kellil, M., Roux, P.: A review of IoT network management: current status and perspectives. J. King Saud Univ. Comput. Inf. Sci. **34**(7), 4163–4176 (2022). https://doi.org/10.1016/j.jksuci.2021.03.006
2. Al-Fuqaha, A.I., Guizani, M., Mohammadi, M., Aledhari, M., Ayyash, M.: Internet of things: A survey on enabling technologies, protocols, and applications. IEEE Commun. Surv. Tutor. **17**(4), 2347–2376 (2015). https://doi.org/10.1109/COMST.2015.2444095
3. Arzo, S.T., Naiga, C., Granelli, F., Bassoli, R., Devetsikiotis, M., Fitzek, F.H.P.: A theoretical discussion and survey of network automation for IoT: challenges and opportunity. IEEE Internet Things J. **8**(15), 12021–12045 (2021). https://doi.org/10.1109/jiot.2021.3075901
4. Ventures, C.: 2021 report: Cyberwarfare in the C-suite. Tech. rep, Cybersecurity Ventures (2021)
5. Dietz, M., Pernul, G.: Digital twin: empowering enterprises towards a system-of-systems approach. Bus. Inf. Syst. Eng. **62**(2), 179–184 (2020). https://doi.org/10.1007/s12599-019-00624-0
6. Dietz, M., Vielberth, M., Pernul, G.: Integrating digital twin security simulations in the security operations center. In: Proceedings of the 15th. International Conference on Availability, Reliability and Security (ARES 2020). Association for Computing Machinery (2020). https://doi.org/10.1145/3407023.3407039
7. Empl, P., Schlette, D., Zupfer, D., Pernul, G.: SOAR4IoT: securing IoT assets with digital twins. In: Proceedings of the 17th. International Conference on Availability, Reliability and Security (ARES 2022), Vienna, Austria, 23–26 August 2022, pp. 1–10. ACM (2022). https://doi.org/10.1145/3538969.3538975
8. Fayaz, S.K., Yu, T., Tobioka, Y., Chaki, S., Sekar, V.: BUZZ: testing context-dependent policies in stateful networks. In: Proceedings of the 13th. USENIX Symposium on Networked Systems Design and Implementation (NSDI 2016), pp. 275–289 (2016)
9. Galmés, M.F., Cheng, X., Shi, X., Xiao, S., Barlet-Ros, P., Cabellos-Aparicio, A.: FlowDT: a flow-aware digital twin for computer networks. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP 2022), pp. 8907–8911. IEEE (2022). https://doi.org/10.1109/ICASSP43922.2022.9746953
10. Galmés, M.F., et al.: Building a digital twin for network optimization using graph neural networks. Comput. Networks **217**, 109329 (2022). https://doi.org/10.1016/j.comnet.2022.109329
11. German Federal Office for Information Security (BSI): Allgemeines zum Einsatz von IoT-Geräten. BSI Grundschutz-Kompendium, Module SYS 4.4, German Federal Office for Information Security (2021)
12. Hui, L., Wang, M., Zhang, L., Lu, L., Cui, Y.: Digital twin for networking: a data-driven performance modeling perspective. IEEE Network, pp. 1–8 (2022). https://doi.org/10.1109/MNET.119.2200080
13. ISO/IEC: Information technology - open systems interconnection - basic reference model: management framework (1989). https://www.iso.org/standard/14258.html. ISO/IEC 7498-4
14. Juniper Networks: the 2020 state of network automation report. Tech. rep. (2020)
15. Li, Y., et al.: A survey on network verification and testing with formal methods: approaches and challenges. IEEE Commun. Surv. Tutor. **21**(1), 940–969 (2019). https://doi.org/10.1109/comst.2018.2868050

16. Networks, N.: OT/IoT security report: cyber war insights, threats and trends, recommendations. Tech. rep, Nozomi Networks (2022)
17. NTT Ltd.: 2022–23 global network report (2022). https://services.global.ntt/zh-cn/insights/2022-23-global-network-report
18. Vielberth, M., Glas, M., Dietz, M., Karagiannis, S., Magkos, E., Pernul, G.: A digital twin-based cyber range for SOC analysts. In: Barker, K., Ghazinour, K. (eds.) DBSec 2021. LNCS, vol. 12840, pp. 293–311. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-81242-3_17
19. Wu, Y., Zhang, K., Zhang, Y.: Digital twin networks: a survey. IEEE Internet Things J. **8**(18), 13789–13804 (2021). https://doi.org/10.1109/jiot.2021.3079510

P3 Digital-Twin-Based Security Analytics for the Internet of Things

Status:	Published	
Date of Submission:	30 November 2022	
Date of Acceptance:	30 January 2023	
Date of Publication:	04 February 2023	
Journal:	Information	
Authors' Contributions:		
	Philip Empl	90%
	Günther Pernul	10%
Full Citation:	EMPL, P., & PERNUL, G. (2023). Digital-Twin-Based Security Analytics for the Internet of Things. *Information 14*(2), Article 95. MDPI.	
DOI:	10.3390/info14020095	
Artifact:	github.com/philipempl/TwinSIGHT	

Journal Description: Information (ISSN 2078-2489) is an international, scientific open access journal of information science and technology, data, knowledge and communication. It publishes reviews, regular research papers and short communications. Our aim is to encourage scientists to publish their experimental and theoretical results in as much detail as possible.

information

MDPI

Article

Digital-Twin-Based Security Analytics for the Internet of Things

Philip Empl * **and Günther Pernul**

Chair of Information Systems, University of Regensburg, 93053 Regensburg, Germany
* Correspondence: philip.empl@ur.de

Abstract: Although there are numerous advantages of the IoT in industrial use, there are also some security problems, such as insecure supply chains or vulnerabilities. These lead to a threatening security posture in organizations. Security analytics is a collection of capabilities and technologies systematically processing and analyzing data to detect or predict threats and imminent incidents. As digital twins improve knowledge generation and sharing, they are an ideal foundation for security analytics in the IoT. Digital twins map physical assets to their respective virtual counterparts along the lifecycle. They leverage the connection between the physical and virtual environments and manage semantics, i.e., ontologies, functional relationships, and behavioral models. This paper presents the DT2SA model that aligns security analytics with digital twins to generate shareable cybersecurity knowledge. The model relies on a formal model resulting from previously defined requirements. We validated the DT2SA model with a microservice architecture called TWINSIGHT, which is publicly available, open-source, and based on a real industry project. The results highlight challenges and strategies for leveraging cybersecurity knowledge in IoT using digital twins.

Keywords: digital twin; security analytics; internet of things; cybersecurity

check for updates

Citation: Empl, P.; Pernul, G. Digital-Twin-Based Security Analytics for the Internet of Things. *Information* **2023**, *14*, 95. https://doi.org/10.3390/info14020095

Academic Editors: Libing Wu and Krzysztof Szczypiorski

Received: 30 November 2022
Revised: 24 January 2023
Accepted: 30 January 2023
Published: 4 February 2023

1. Introduction

Organizations utilize emerging technologies around the Internet of Things (IoT) to stay competitive and build knowledge. In the Industrial Internet of Things (Industrial IoT), sensors complement existing systems to enact more data leading to more accurate predictive maintenance. Thereby, intertwining two complementary worlds is crucial in enabling IoT use cases. Most commonly, information technology (IT) systems are connected to operational technology (OT), processing volumes of data to gain new knowledge about machines. This intertwining is a tremendous cybersecurity challenge for organizations, as the OT remains reachable via the internet. Additionally, the myriad of lifecycle participants, devices, and systems creates an opaque and complex scenario. Knowing the IoT systems well is necessary for efficient cybersecurity management.

Lifecycle-centric cybersecurity management is crucial. Most recently, supply-chain attacks have been emerging, compromising the chain's weakest link. The ENISA states that half of the attacks are attributed to advanced persistent threats (APT) groups, mainly accessing data [1]. Consider a manufacturer of machines, i.e., cyber-physical systems, vending IoT assets built out of physical components from suppliers to business customers. Business customers own IoT assets and operate, maintain, and recycle IoT assets through external service providers. From this example, different lifecycle participants interact with an IoT system, each opening up new attack vectors. Nevertheless, attacks are not limited to supply-chain attacks. Many well-known attacks exist, e.g., Sybil attacks, wormhole attacks, ransomware attacks, and distributed denial service attacks [2]. These attacks can take place throughout the lifecycle of an IoT system, and they are getting more advanced and harder to detect.

Security analytics is a paradigm to enhance cybersecurity in the Industrial IoT. It applies big data analytics techniques aggregating and internalizing external cybersecurity knowledge [3]. Security analytics collects data from various sources and allows the correlation of data, e.g., whether an incident happened or is even likely to happen. Security analytics is essential to organizations, but the less efficient the models, the less efficient the security analytics. Nevertheless, APTs urge organizations to rely on internal and external knowledge covering the unforeseeable as well as possible. Thus, generating knowledge needs to come along with sharing. Sharing knowledge between lifecycle participants about threats and incidents improves cybersecurity [4,5]. For instance, over 90% organizations state that they rely on actionable cybersecurity knowledge from third parties and partners [6]. As generating valuable knowledge to be shared is challenging, organizations deploying complex and lifecycle-centric IoT systems require different and novel approaches.

The Industrial IoT requires more cooperation and collaboration between lifecycle participants (see ISO 27036) and more advanced techniques to cope with sophisticated cybersecurity attacks. The industry has long used the digital twin paradigm to map a physical asset to its virtual representation through the lifecycle by replicating or simulating the IoT [7]. Its use is not limited to operational scenarios, e.g., digital twins assist security testing, enhance cyber situational awareness, and improve intrusion detection systems [8,9]. Estimates say that 80% of organizations instrument some forms of digital twins for cybersecurity scenarios, whereby 85% of security officers agree that digital twins unleash even more efficient detection and mitigation [10]. By abstracting physical assets, digital twins reduce the complexity of problem solving and provide a semantic layer to the virtual representations. Thereby, they ensure a more detailed analysis of the physical asset's state and historical data, leveraging security analytics on top of the data structure of digital twins.

Digital twins show potential in knowledge generation (coupled with security analytics) and knowledge sharing. By bringing security analytics and digital twins together, we aim to enhance cybersecurity in the Industrial IoT along the lifecycle and push cybersecurity knowledge-sharing research. We address the research gap that no unified framework for digital-twin-based security exists and add a replication-based intrusion detection approach to it. This will help future research in the development of digital-twin-based security analytics. The following research question guides this paper: *"How can one align security analytics and digital twins?"* Recent research has already examined digital twins and security analytics specifically but not as a whole. By aligning security analytics with digital twins, this paper contributes to a more secure Industrial IoT by generating cybersecurity knowledge and demonstrating how to share this knowledge throughout the lifecycle. Our contributions are summarized in the following:

1. We comprehensively align security analytics with digital twins and illustrate how to generate and share cybersecurity knowledge between lifecycle participants.
2. We provide a novel formal model for digital twins and security analytics. This formal model assists in implementing digital-twin-based security analytics use cases.
3. We envision the DT2SA model for digital twins and security analytics. This model integrates the Industrial IoT and mediates a global understanding for further research and practical adoption.
4. We instantiate the DT2SA model by implementing a microservice architecture leveraging digital twin-based security analytics based on a real-world research project. TWINSIGHT enables digital-twin-based threat and incident detection using open-source software.

This paper is structured as follows. Section 2 provides the fundamental background knowledge on digital twins and security analytics. We further discuss related research. In Section 3, we conceptually elaborate on knowledge generation and sharing in the cybersecurity domain resulting in requirements. Section 4 takes up these requirements and puts entities and their relationships shaping a formal model. While Section 4.3 outlines the DT2SA model based on the formal model, Section 6 validates the DT2SA model concerning the requirements and introduces TWINSIGHT, a digital-twin-based security analytics microservice architecture. Section 7 concludes the paper and highlights future research.

2. Background and Related Work

In the following sections, we present relevant background on digital twins for security operations in Section 2.1 and for security analytics in Section 2.2. Section 2.3 highlights the related work and existing research focusing on digital twins, security analytics, and knowledge generation and sharing.

2.1. Digital Twins for Security Operations

The digital twin paradigm is deeply grounded in the Industrial IoT, representing a physical asset (e.g., an entity, system, process, or person) that is mapped throughout its lifecycle to a virtual counterpart built on semantics [7]. The digital twin relies on descriptive and dynamic asset data, dynamic environment data, historical asset and dynamic data, and semantics [11]. The application scenarios are diverse and are not limited to operational use. The application of digital twins for cybersecurity scenarios is a trend in research, e.g., Lopez et al. are developing an authorization mechanism through digital twins in 5G environments. Technically, digital twins are executed via so-called operational modes, i.e., simulation, analysis, and replication, which differ as follows:

- *Analytics*—using state data with statistical analysis.
- *Simulation*—using specification data with emulation or simulation techniques.
- *Replication*—using specification and state data with emulation or stimuli techniques.

These operation modes serve different use cases and enable various application scenarios, e.g., intrusion detection, security testing, security training, or penetration testing [8]. Digital twins are also considered an additional layer of security for the IoT that manages incident response [12]. Of course, digital twins present new cybersecurity challenges, but this has already been addressed by research [13]. However, we treat digital twins as an additional layer for more efficient security analytics.

2.2. Security Analytics

In the era of big data, new technologies are emerging to support efficient real-time data processing and analysis, which has led to the term big data analytics [14]. Of course, the need for big data processing in IoT is obvious and key to dealing with the vast amount of heterogeneous IoT data. Big data processing technologies provide a widely used foundation for further analysis. Siow et al. [15] provided an excellent summary of big data analytics in IoT and defined the following five analytical operations:

- Descriptive analytics: *What has happened?*
- Diagnostic analytics: *Why did it happen?*
- Discovery analytics: *What is happening?*
- Predictive analytics: *What will happen?*
- Prescriptive analytics: *What should one do?*

Descriptive or diagnostic analytics provides hindsight, discovery analytics provides insight, and predictive or prescriptive analytics provides foresight. From a cybersecurity perspective, big data analytics is crucial.

Due to the large amount of (un)structured security-relevant data, traditional security information and event management (SIEM) systems are reaching their performance limits [16]. Security analytics is concerned with applying big data processing technologies to cybersecurity and describes the aggregation and analysis of security-relevant data [3]. However, there is no clear definition of security analytics. We define security analytics as follows:

> Security Analytics is a repertoire of capabilities and technologies for the systematic processing and analysis of data to identify threats and imminent incidents.

We see security analytics as an evolution of SIEM with additional operations, e.g., intrusion detection systems, behavioral or network analysis [17], and knowledge sharing.

2.3. Related Work

Big analytics and security analytics.is trending. Ackoff [18] defined the DIKW and presented how to generate wisdom from data. Siow et al. [15] summarized five analytic operations for big data analytics and their alignment with the DIKW. These operations could be anchored in the security analytics domain [3] and were already aligned with security analytics [19]. We go beyond these analytic operations to illustrate how cybersecurity knowledge can be generated by linking operations, expertise, and digital twins.

Generating and sharing cybersecurity knowledge has already been addressed. The incident response process shows data, observables, indicators of compromise, and incidents [20]. Böhm et al. [4] elaborated on knowledge transformation in security analytics and formalized different types of knowledge, i.e., explicit and implicit knowledge. We aim to improve knowledge generation and share research by linking security analytics with digital twins to promote knowledge generation in cybersecurity.

Eckhart et al. have developed CPS Twinning and CPS Replication, both frameworks for creating and deploying digital twins for cybersecurity scenarios [21]. Digital twins for security operations can be equipped with various capabilities, such as analytics, penetration testing, and intrusion detection. For example, Dietz et al. [22] integrated digital twin security simulations into a security operations center to assist analysts with security testing and monitoring IoT assets. Damjanovic-Behrendt [23] defined a microservice architecture of the digital twin that refers to security analytics as data analytics. Other service management tasks, such as incident detection and responding, complement security analytics. Since some research already presented digital twins for security analytics, we comprehensively go beyond this understanding and formalize security analytics and related knowledge sharing with digital twins. In doing so, we select an application scenario to demonstrate our overall model.

In summary, the existing literature does not comprehensively address security analytics in IoT. While there are approaches to align particular analytical operations with digital twins, a comprehensive view still needs to be provided. Our goal is to establish a comprehensive model and bring the previously practice-oriented security analytics into the realm of science. In this way, we will create a balance and improve the use of cybersecurity knowledge in IoT.

3. Managing Cybersecurity Knowledge

The following sections make steps towards the formal model. We first describe the generation of knowledge in Section 3.1. After, we illustrate knowledge sharing between lifecycle participants interacting with digital twins in Section 3.2. We then describe requirements for the formal model in Section 3.3, resulting in the intertwining of environments in Section 3.4.

3.1. Cybersecurity Knowledge Generation

Without knowledge, there is nothing to share. As already mentioned, knowledge is a product of information [18], whereby knowledge generation requires human interaction. In doing so, humans explore data, find insights, formulate hypotheses, and generate knowledge repetitively (sensemaking loop) [24]. However, generating knowledge is challenging and requires interaction with data and models. Research and industry cope with analytical operations by describing the appropriate mix of technologies, techniques, and cognitive abilities. The literature summarizes these analytical operations as descriptive, diagnostic, detective, predictive, and prescriptive ones [15]. We analyzed blog posts from several large organizations, i.e., IBM and Microsoft, confirming this notion but not all covering detective/discovery analytics. However, these five operations ensure knowledge generation all throughout the process.

Cybersecurity is also concerned with analytics. For example, there is an overlap between big data analytics and security analytics for deriving patterns for incident detection. There are similarities between the incident response process [20] and the DIKW [18]. Aside from the naming, the data, observables, indicators of compromise, and incidents share the same relationships as the DIKW. We summarize and transfer these concepts into the cybersecurity context. Figure 1 shows our approach to generating cybersecurity knowledge. Descriptive operations contextualize data into observables that describe specific events within an attack. Diagnostic operations involve analysis and correlation of historical observables, i.e., forensic investigation. Detective operations are suitable for detecting incidents based on indicators of compromise in real-time, e.g., intrusion detection systems. Predictive and prescriptive analytic operations are used to predict incidents and derive actions.

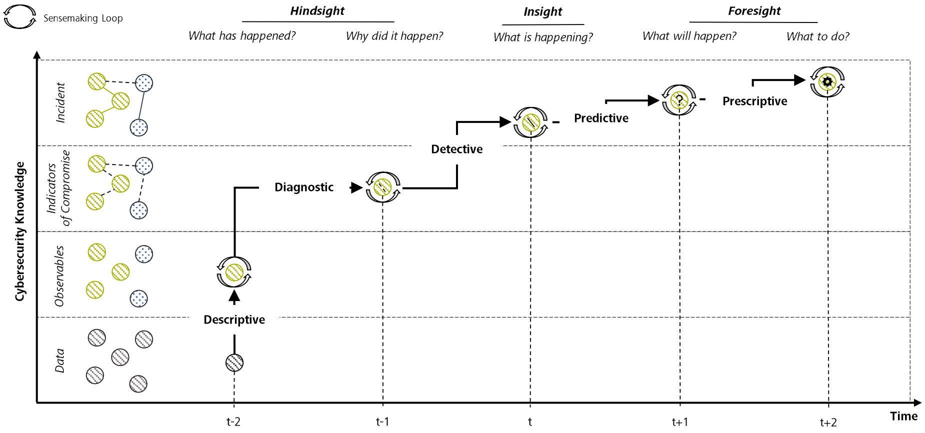


Figure 1. Cybersecurity knowledge generation.

In practice, analysts elaborate observables from many unstructured data sources (i.e., IP addresses). Malicious IP addresses represent so-called indicators of compromise. When they occur in a particular combination with other indicators of compromise, we refer to them as incidents. Nevertheless, these operations are part of the sensemaking loop, since knowledge is shaped by human interaction. Now that we know how knowledge is created in cybersecurity, we examine knowledge sharing.

3.2. Cybersecurity Knowledge Sharing

Knowledge sharing is critical in the Industrial IoT because it is a playground for different players and technologies. It includes many heterogeneous devices, a variety of communication protocols, and standards under development. From a manufacturing perspective, the collaboration between supply chain participants is broader than services and systems. Data must also be available across organizational boundaries. Digital twins have emerged as a paradigm to meet management within the circular economy [25]. They have proven their strength in sharing knowledge with lifecycle participants as long as communication channels are kept secure [26].

Collaboration between lifecycle participants must be strengthened, as this will improve observables and indicators of compromise through external knowledge [4]. For example, the greater the knowledge about indicators of compromise, the more efficient security analytics. To promote collaboration among lifecycle stakeholders, we must first understand their roles in the Industrial IoT. We summarize the leading roles in Industrial IoT: manufacturer, supplier, distributor, maintainer, and owner (see Figure 2). Since digital twins promote knowledge sharing in the Industrial IoT, we illustrate their bidirectional communication with their physical counterparts. A physical asset consists of several components from multiple suppliers assembled by a manufacturer. A distributor brokers the asset to an owner, who contracts a maintainer to provide services and maintenance. Digital twins communicate bidirectionally and serve as a single point of truth, eliminating information asymmetry. For example, they notify an owner if an incident occurs or report vulnerabilities to a supplier or manufacturer. Given this knowledge generation and sharing notion, we can define the key requirements.

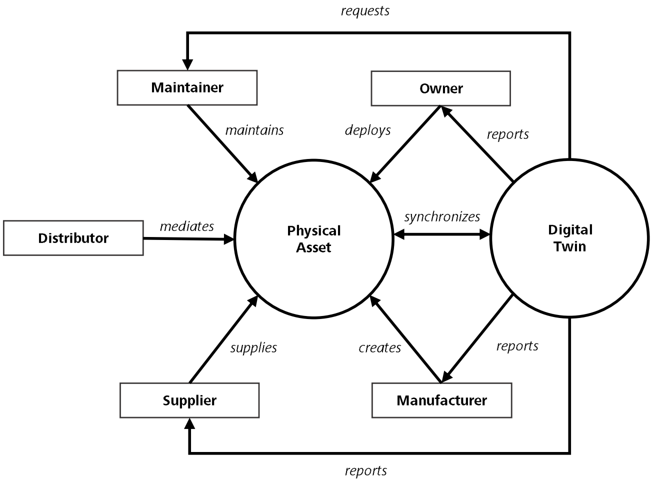


Figure 2. Cybersecurity knowledge sharing.

3.3. Requirements

Aligning security analytics and digital twins requires determining requirements concerning digital twins, security analytics, and knowledge sharing. We define the requirements as follows:

REQUIREMENT 1 (DIGITAL TWIN). The digital twin comprises a physical, a virtual, and a communication component [7,27]. It describes descriptive, dynamic, environmental, historical, and semantical data [11]. For cybersecurity operations, the digital twin operates in simulation, analytics, and replication modes [21,28].

REQUIREMENT 2 (SECURITY ANALYTICS). Security analytics is characterized by heterogeneous data, data warehouses, technologies, system monitoring, and dashboards [3]. It demands descriptive, diagnostic, detective, predictive, and prescriptive operations [15].

REQUIREMENT 3 (KNOWLEDGE). Security analytics enables knowledge generation, which is key for sharing [4,24], whereby digital twins enrich security analytics.

3.4. Intertwining Environments

We denote the primary entities and their relationships using an entity-relationship model (cf. Figure 3) in preparation for the formal model. Gray-colored blocks and dashed lines represent the logical components. Here, we represent the physical environment and the virtual environment as the main components in Industrial IoT. It should be noted that the lifecycle participants and the digital twin data can be mapped to both the physical and virtual environments. However, this plays only a minor role. Furthermore, we define security analytics as an internal component of the virtual environment [23]. The digital twin can be either in one of the three modes of operation: simulation, replication, or analytics.

The physical environment contains physical assets, i.e., IoT devices. A physical asset goes through its lifecycle with different lifecycle participants intertwined. A physical asset is assigned for a lifecycle phase at a given time, and other assets may pass through the same phase. Since digital twins are physical assets, they follow their physical counterparts through the lifecycle. They interact with the same and possibly other lifecycle participants that fit into the current lifecycle phase of the physical asset. In addition, digital twins have specific data, especially metadata, state, and historical data. Security analytics is placed above the digital twin, benefits from the semantics layer, and processes a set of digital twin data.

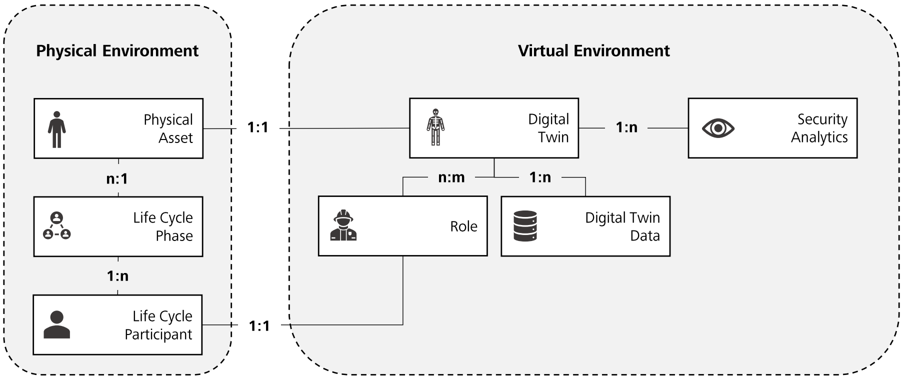


Figure 3. Security analytics through the intertwining of physical assets and digital twins.

4. Formal Model

We develop the formal model for digital-twin-based security analytics in the Industrial IoT in the following. This formal model is key to approaching the DT2SA model that supports digital-twin-based knowledge generation and sharing, enabling collaboration between lifecycle participants. Note that this model only abstractly illustrates which data can be coupled to which operation mode of the digital twins and security analytics operation. In Section 4.1, we define the formal model for the physical environment, and in Section 4.2, that for the virtual environment also containing security analytics described in Section 4.3.

4.1. Physical Environment

The physical environment comprises the physical asset and its associated lifecycle phases. We view physical assets as tangible and intangible artifacts of an organization. The term physical is used to distinguish between real-world and virtual-environment assets. It includes processes, software, and hardware (i.e., IoT). Thus, we define physical IoT assets as $I := \{i_1, ..., i_n\}$.

LIFECYCLE. These physical assets progress through different lifecycle phases—e.g., from design to operation to recycling. We define the lifecycle phases as $L := \{l_1, ..., l_m\}$ and assign exactly one lifecycle phase at a given time e to a physical asset: $f : i^e \mapsto l \mid i \in I \wedge l \in L$. Within a lifecycle phase, participants interact with physical assets, which we define as follows: $P := \{p_1, ..., p_k\}$. Due to this formal definition, a physical asset is always situated in exactly one phase of the lifecycle at a given time and interacts with the respective lifecycle participants.

4.2. Virtual Environment

The virtual environment describes the digital twin and the relation to several lifecycle participants and respective data. This paper considers the digital twin as a mapping to the physical asset. As the digital twin operates one of the three operation modes, namely, analytics, replication, and simulation [28], we define the operation mode analytics as T_{ana}, the replication as T_{rep}, and the simulation as T_{sim}. All digital twins are summarized through $T := \{t_1, ..., t_l\}$ and mapped to the physical assets as follows: $g : t \mapsto i \mid t \in T \wedge i \in I$. As digital twins T are mapped to physical assets A, and these to lifecycle phases L, we conclude $(f \circ g) \Leftrightarrow T \Rightarrow L$. Thus, the digital twin is in the same lifecycle phase as the physical asset.

DIGITAL TWIN DATA. We further define the digital twin data as $D := \{d_1, ..., d_p\}$. As already depicted, digital twins have several types of data, namely, descriptive and dynamic asset data, dynamic environment data, historical asset data, and semantics. Descriptive data D_{desc} includes specifications and static data of a physical asset, such as year of manufacture, model number, or identifiers. Dynamic asset data (we refer to as state data D_{state}) bundles all operational and asset-specific data that virtually represent assets given data structures dynamically. Dynamic environment data D_{env} are external stimuli that surround physical assets, e.g., data captured through external interfaces (i.e., sensors). Dynamic environment data do not only refer to data captured by the asset, e.g., network data. Historical asset data D_{hist} are a collection of persisted data. Semantics D_{sem} represent models that yield relevant data relations (this is also knowledge and wisdom, according to the DIKW). Note that it is not possible to have machines produce wisdom. Wisdom is also perceived differently by each lifecycle participant. Last, we define security-related data as D_{sec}. The latter holds all data for security operations, e.g., policies, SIEM rules, signatures, or threat intelligence data. These data types relate to each other. We define D_{hist} as the historical data storage that aggregates various types of data, whereby $\exists d \in D(d \in D_{state} \vee d \in D_{env} \vee d \in D_{sec})$. D_{desc} is considered not to change over time. The data types D_{state}, D_{env}, and D_{sec} are historically stored and a subset of data type D_{hist} is exceeded if a certain threshold s is breached (i.e., the arrival time of data d^t). We specify this exceeding or persistence as follows, where s is a timely restricted threshold:

$$persist(d_i) = \begin{cases} d_i \in (D_{state} \vee D_{env} \vee D_{sec}) & if\ d^t > s \\ d_i \in D_{hist} & otherwise. \end{cases}$$

4.3. Security Analytics

As described above, we consider security analytics as big data analytics from a cybersecurity perspective, distinguishing between analytical operations, namely, descriptive, diagnostic, detective, predictive, and prescriptive [3,15]. We define all analytical operations as $A := \{a_1, ..., a_q\}$, whereby $a \in A$ holds the respective results of one analytical operation (generated knowledge). Additionally, analytical operations O yield subsets, i.e., analytical operations categories: $(A_{desc} \cup A_{diag} \cup A_{det} \cup A_{pred} \cup A_{pres}) \subseteq A$. We consider all analytical categories, as we provide an approach to holistically integrating security analytics. Digital-twin-based security analytics involves knowledge and bidirectional communication with the physical environment.

Knowledge. We assume data, information, knowledge, and wisdom as the foundation of analytical operations but do not distinguish between these knowledge types as described by Böhm et al. [4]. We consider the sensemaking process a black box, as it is hard to define rationales, but we include cybersecurity knowledge. We refer to specific knowledge about observables as K_{obs}, knowledge about indicators of compromise as K_{ind}, and learning about incidents as K_{inc}.

Incident Response. One of the core properties of digital twins is the bidirectional communication to real-world assets. Therefore, we also specify incident response. Incident response is part of security analytics and involves security orchestration and response. Thereby, playbooks are vital defining actions $C := \{c_1, ..., c_s\}$ given certain knowledge about incidents K_{inc} and threats. Integrating digital twins T into the incident response process is called orchestration, and a reaction to an event using digital twins is a response. A response to an incident is triggered by an event and solved through the orchestration of digital twins. Digital twins enable the orchestration, so we define the orchestration O concerning digital twins as $O \circ T$, and the response is as follows:

$$response := ((k \mapsto ot), c \mid c \in C, k \in K_{inc}, ot \in (O \circ T))$$

Descriptive Analytics. Descriptive analytics descriptively summarizes the data in context, using visualizations and statistical methods. This analytical operation relies on historical data to provide hindsight and identifies relevant observables. Descriptive analytics requires the operation mode T_{ana}. We define a descriptive operation as follows:

$$A_{desc} := (d, t \mid d \in D_{hist}, t \in T_{ana}) \mapsto K_{obs}$$

Diagnostic Analytics. Diagnostic analytics goes beyond descriptive analytics and investigates the causes of phenomena (i.e., incidents). Due to this retro perspective, diagnostic operations incorporate historical data and identify indicators of compromise in observables. Diagnostic operations also require the operation mode T_{ana} and integrate knowledge about observables K_{obs}. We define a diagnostic operation as follows:

$$A_{diag} := (d, k, t \mid d \in D_{hist}, k \in K_{obs}, t \in T_{ana}) \mapsto K_{ind}$$

Detective Analytics. Detective analytics generates new insights and relies on methods such as signatures or rules from a cybersecurity perspective. This analytical operation is often termed discovery analytics in big data analytics and links insights from historical data. We evolved this idea by appending insights from real-time data, which is highly relevant in the case of real-time event correlation (i.e., SIEM). Detective analytics can be in either operation mode, T_{rep} or T_{sim}. A detective operation links several indicators of compromise to incidents. It involves knowledge about observables to generate knowledge about incidents K_{inc}. This knowledge defines analytical and SIEM typical measures, e.g., signatures or rules. Detective operations rely on all kinds of data. We define detective operations as follows:

$$A_{det} := (k, t \mid k \in K_{ind}, t \in T_{rep} \vee t \in T_{sim}) \mapsto K_{inc}$$

Predictive Analytics. Predictive analytics utilizes data and knowledge to predict the future. Thereby, predictive operations deploy semantics (i.e., mathematical models or simulations) and involve knowledge about incidents to predict if an incident is likely to happen. Predictive operations rely on the results of detective operations A_{det} and create new incident knowledge K'_{inc}. Predictive operations can also be in operation mode T_{rep} or T_{sim}. Predictive operations also rely on all kinds of data. We define a predictive operation as follows:

$$A_{pred} := (a, k, t \mid a \in A_{det}, k \in K_{inc}, t \in (T_{rep} \oplus T_{sim})) \mapsto K'_{inc}$$

Prescriptive Analytics. Prescriptive analytics identifies, evaluates, and suggests appropriate security orchestration to mitigate an incident. Thereby, simulations play a decisive role in deriving decisions from the different scenarios, e.g., through what–if simulations. Prescriptive operations can also be either in operation mode T_{rep} or T_{sim} and require results of the analytical operations A_{det} or A_{pred}. Further, prescriptive operations identify appropriate actions for incident response activities out of existing knowledge about incidents K_{inc}. Prescriptive operations may encompass all kinds of data. We define a prescriptive operation as follows:

$$A_{pres} := (a, k, t \mid a \in (A_{det} \oplus A_{pred}), k \in K_{inc}, t \in (T_{rep} \vee \in T_{sim})) \mapsto C$$

Knowledge Sharing. We adopt parts of the data-sharing concept from Dietz et al. [29] in our sharing principles for digital twin data. We add the sharing of cybersecurity knowledge K and security-related data D_{sec}. We assume that roles have access to digital twins, dependent on permissions. If permission is granted, roles are invited to access digital twin data, including existing knowledge and models. Further, lifecycle participants can contribute expertise and write relevant descriptive and security-related data entries or semantics.

$$read := (t, d, k \mid t \in T, d \in D, k \in K)$$

$$write := (t, d, k \mid t \in T, d \in (D_{desc} \oplus D_{sem} \oplus D_{sec}), k \in K)$$

5. DT2SA Model

Based on the formal model in Section 4, we envision the DT2SA model. Figure 4 shows the respective model. We adopt the functional view of the Industrial Internet Reference Architecture (IIRA) [30], based on the five domains: the control domain, information domain, operations domain, functional domain, and business domain. This reference architecture ensures the DT2SA model is embedded in the IoT. In the following, we explain the respective domains of our DT2SA model.

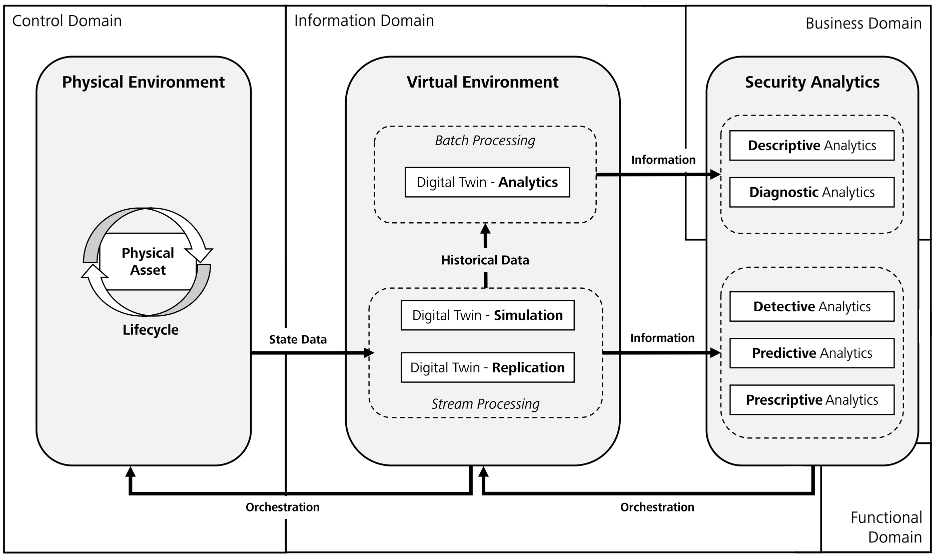


Figure 4. DT2SA model.

Control Domain. The control domain seamlessly captures physical assets and their data. Physical assets refer to artifacts in the Industrial IoT, such as processes, machines, or plants. In this context, physical assets collect data through essential interactions, such as machine-to-machine communication. Asset-specific data of interest includes descriptive, state-related, and security-related data. Physical assets, i.e., machines, are arranged in networks. These networks also generate operational and security-related data, such as network traffic. However, the overarching task of the physical environment is to make all relevant data from the physical assets and their environment available to their respective digital twins. The control domain uses various networks for data transmission. The data pass through so-called proximity, access, and service networks. Proximity networks are responsible for short-range communication and access networks for long-range communication. Service networks are enterprise networks that handle communications between business applications.

Information Domain. The information domain is responsible for data acquisition, processing, and persistence. The physical environment feeds the respective digital twins with data and receives data and commands from the digital twins. Digital twins thus claim bidirectional communication with the physical environment. Technically, bidirectional communication is realized by so-called event-based messaging platforms that rely on hubs or brokers based on publish/subscribe messaging. However, the information domain includes historical, state, and semantic data. The use of different data stores manages the diversity of data. The digital twin in analytics mode processes historical data in a batch-processing pipeline. Digital twins using replication and simulation modes rely heavily on real-time data in a stream processing pipeline. Each pipeline supports different analytical operations. Descriptive and diagnostic operations are coupled with the digital twin's analytics mode to provide hindsight. Detective operations are time-dependent and should be coupled with the replication mode. The digital twin constantly feeds detective operations with real-time state data. Predictive and prescriptive operations build on both simulation and replication modes. Both modes of operation are the foundation for predicting incidents or recommending incident response operations.

FUNCTIONAL DOMAIN. The functional domain operates through rule-based decisions. From a cybersecurity perspective, we map incident response activities to the functional domain because incident response orchestrates assets using predefined logic. For example, incident response is automated using playbooks to secure the IoT [12]. Digital twins provide an additional layer of security and orchestrate physical assets. Incident response is also event-based, receiving information and knowledge from human experts and digital-twin-based security analytics. However, the incident response process is triggered by analytical operations.

BUSINESS DOMAIN. The business domain includes descriptive and diagnostic operations. The business domain promotes the exchange and generation of knowledge. For example, visualized information obtained from descriptive and diagnostic operations helps to identify correlations and gain insights. This knowledge can further optimize digital twin models and analytical operations. We see the potential of a dedicated knowledge-sharing and management platform to explore previously untapped knowledge in cybersecurity visually.

6. Proof of Concept

We validate our DT2SA model in regard to its applicability to an ongoing research project. In Section 6.1, we investigate to what extent all requirements of the DT2SA model have been met. Next, we validate our prototype TWINSIGHT in Section 6.2 and evaluate the applicability of our model to the existing literature on digital-twin-based security analytics in Section 6.3. In Secion 6.4, we define our experimental setting leading to results in Section 6.5, which we discuss thoroughly in Section 6.6.

6.1. DT2SA *Components*

In Section 3.3, we defined the key requirements for the DT2SA model. In the following, we assess and discuss the fulfillment of each requirement.

DIGITAL TWIN (R1). The DT2SA model includes physical assets, virtual representations, communication management, data management, and modes of operation. It seamlessly integrates physical assets into digital twins, making virtual representations indistinguishable from physical assets through synchronization and communication. It also addresses all relevant mechanisms for data management, including how to retain and process data. All requirements for the digital twin are met.

SECURITY ANALYTICS (R2). The DT2SA model incorporates security analytics through the variety of data and semantics used by analytical operations. The model also incorporates different data from different data stores. Although technologies are essential to the implementation of the model, less attention has been paid to them. We also touched on the incident response process that manifests the interaction with the physical asset and the orchestration of cybersecurity operations. We considered all analytical operations to cover security analytics fully. Less attention was paid to interactive dashboards and system monitoring, as these functional components can be effortlessly added and only make the model unnecessarily overcomplex. However, our model covers almost all features of security analytics.

KNOWLEDGE (R3). The DT2SA model enables the generation and sharing of knowledge among lifecycle participants. For cybersecurity knowledge generation and sharing, the model proposes the integration of key lifecycle participants through digital twins and security analytics. The model also embeds the knowledge hierarchy through the formal model. Information sharing is not addressed in detail.

The DT2SA model addresses almost all requirements and provides relevant perspectives on digital twins and security analytics. Although the requirements are met, there is a lack of applicability in practice. Therefore, we refer to an ongoing research project and implement TWINSIGHT to validate our model further.

6.2. Use Case: SISSeC

To further validate the DT2SA model, we implement a microservice architecture called TWINSIGHT. TWINSIGHT relates to SISSeC, an ongoing research project in Germany. SISSeC aims to securely connect a printed circuit board (PCB) manufacturer's machine and sensor data to a cloud via an edge gateway. The PCB manufacturer's overall goal is to collect data and predict likely future operating conditions of machines with digital twins. From a cybersecurity perspective, these digital twins should enable intrusion detection. The project also aims to make security-relevant data available to lifecycle participants via a marketplace. In this way, lifecycle participants and third parties can share their knowledge.

We instantiate the DT2SA model for SISSeC. The PCB manufacturer requires state data, descriptive data, and environmental data to use our model. We focus on a drilling and milling machine and its system-specific operational data. This machine sends data over the proximity network to the edge node, which forwards the data to the cloud. The cloud contains digital twins that enable bidirectional communication, and thus, control of the machines. To instantiate the DT2SA model, we formulate two main objectives of TWINSIGHT: incident detection and threat detection. Then, we choose the analytical operation that satisfies this goal: the detective operation. The next step is to choose an appropriate operation mode for the digital twin.

6.3. Applicability

Before specifying a concrete operation mode for detective operations in SISSeC, we validated the applicability of our DT2SA model to existing literature dealing with digital twins, in particular, intrusion or anomaly detection. Given our use case, we only considered operational modes that enable detective operations, so we focused on simulation and replication modes. In the following, we list relevant literature that presents concrete instantiations of our DT2SA model:

- *Simulation mode* [9,22,23,31–39];
- *Replication mode* [21,36,40–42].

We found that more papers deal with simulation mode and less deal with replication mode. This literature can also serve as references to provide concrete ideas on using simulation- or replication-based digital twins for security analytics. For example, simulation-based digital twins are utilized for incident prediction [35], and state-replication-based digital twins are used for intrusion detection [21]. These approaches fit into the DT2SA model and follow the same scheme: digital twins organize data and models and make them available for subsequent analytic operations. In simulation models for machine learning training, we found, among other things, historical data. Most of the data used are real-time or specification data. Based on the SISSeC use case, the fact that replication-based digital twins for IDS have not been explored, and the lack of sharing capabilities, we develop an experimental environment for replication-based digital twins for detective operations.

6.4. Experimental Setup

We implement TWINSIGHT for digital-twin-based incident and threat detection containing detective operations. TWINSIGHT is publicly available on GitHub. Figure 5 shows the settings in TWINSIGHT. The developments in TWINSIGHT are driven by the SISSeC use case and are intended to illustrate how detective security analytics can be implemented with replication-based digital twins. Note that this experimental setup focuses on only a specific part of the DT2SA framework: replication-based digital twins and detective operations. In this way, we aim to gain new insights into replication-based digital twins and detective security analytics and determine whether the DT2SA model covers these aspects. In the following, we describe these components in more detail.

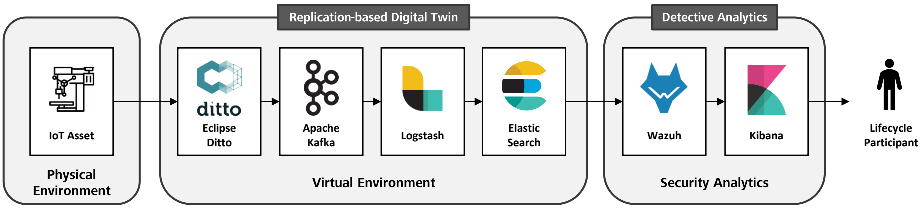


Figure 5. TWINSIGHT setting implemented in the research project SISSeC.

PHYSICAL ENVIRONMENT. The physical environment consists of identical drilling and milling machines relying on Raspberry Pis for data transfer. Depending on the setup, a machine consists of several Raspberry Pis. We use Raspberry Pi 3B+ models running Raspbian GNU/Linux 11 with 1GB RAM. These devices are interconnected using an event-based messaging architecture. More specifically, the devices communicate via MQTT 3.1 with an MQTT broker. In doing so, each client implements the Paho MQTT client library in Python. We can now use MQTT to collect device-related data, such as the state of an asset. Interesting setup details follow in the virtual environment.

VIRTUAL ENVIRONMENT. Eclipse Ditto is the digital twin software built on an MQTT broker that manages messages received from clients. Eclipse Ditto implements an event-based API that allows the definition of device representations and messages. We model each drilling and milling machine in Eclipse Ditto using the built-in JSON schema. In addition, Eclipse Ditto allows us to define connectors, which we use as a bridge to the MQTT broker. Of course, messaging would allow for device orchestration and response, but we focus only on security analytics, not incident response. As Eclipse Ditto only stores the current state of a machine, Apache Kafka collects real-time data; and Logstash subscribes to all topics, transforms messages, and stores them in Elasticsearch. These applications are deployed on a virtual machine running Ubuntu 20.04.3 LTS with 12 GB of RAM, six cores, and 60 GB of storage.

SECURITY ANALYTICS. We implement detective security analytics using Wazuh and its native Kibana integration. Wazuh is open-source software that performs sophisticated security operations and incident response. It relies on agents installed on the assets to be monitored and orchestrated. Kibana is used only for the virtual representation of agent-based information. Both applications also run on the same machine as the virtual environment.

6.5. Results

In implementing replication-based digital twins for detective security analytics, we gained insights and results that we would like to share. Our research provides results related to implementation, security analytics, digital twins, the IoT, and knowledge sharing.

IMPLEMENTATION. We support lifecycle participants by aligning digital twins and their modes of operation with security analytics. The formal model makes implementing software based on digital twins more feasible. We have found that replication-based digital twins fit real-time data processing, and the literature confirms that simulation works decoupled from the live system, e.g., Dietz et al. [43]. Using an event-based microservice architecture ensures flexibility and real-time data processing. We have also found that selecting an appropriate mode of operation, assets, and data is more efficient when starting with the desired goal of security analytics. We consider this to be top-down DT2SA.

SECURITY ANALYTICS. Since most requirements are sufficiently met, the DT2SA model does not lack essential components. In particular, more security monitoring and an interactive dashboard can easily be added. Nevertheless, these features are not considered core features of security analytics. We also found that security analytics technologies, e.g., Wazuh, do not work properly with device representations. They are designed to access a machine's resources using agents. In modern organizations, machines no longer consist of a single component but form complex systems of systems. In such environments, these components should be monitored in relation using security analytics. Appendix A shows the Wazuh user interface and two views that visualize the detection of incidents and threats in real-time and point out the problem at hand. Figure A1 shows all integrated assets and threats with Wazuh agents. Figure A2 shows all attacks against a specific asset. Looking at complex systems that consist of multiple components (and agents) makes traditional security analytics inefficient. Security analytics should integrate asset representations and exploit relationships between assets when considering digital twins. However, correlating events related to complex systems of systems and digital twins is paramount. More efficient analytics should enable the mapping and correlation of agents from assets to digital twins to realize their full potential. Nonetheless, security analytics would benefit from systems of systems approaches for improving the visualization and resolution of security events.

DIGITAL TWINS. Digital twin software allows for easier data processing and analysis, as states are updated dynamically. Eclipse Ditto, a replication-based digital twin focusing on states and functions, provides dynamic user management that defines roles and their privileges.

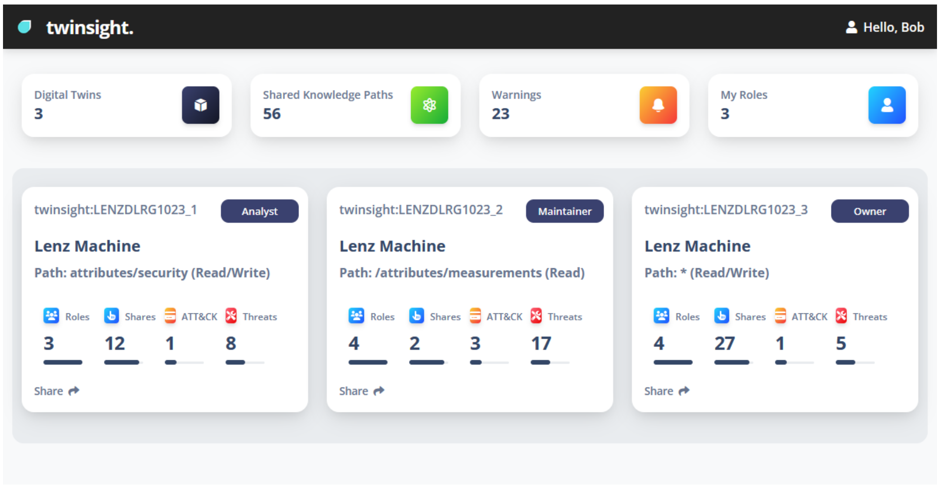
We found that built-in user management enables fine-grained sharing of digital twin data among lifecycle participants. We also found a research gap on digital twins in analytics mode. In addition, we learned in SISSeC that it is possible to implement certain application scenarios with digital twin software. Nevertheless, security analytics is only one of many possible application scenarios for digital twins. Thus, TWINSIGHT can be extended to other analytic operations, and even to more sophisticated security operations.

In addition, digital twins are ideally suited as the ground truth of knowledge. We consider digital twins as hubs for managing device-specific and security-related data. As discussed earlier, they ensure the correct mapping of components, which allows us to model system of systems. Listing 1 shows an Eclipse Ditto device representation, including security-related information and analytics results. In our use case, we included relevant common platform enumerations (CPE) to query vulnerabilities related to the digital twins' components. We can also map IDS agents to a specific digital twin to include the most recent alert in the device representation. The application scenarios and extensions of security analytics are numerous and not limited to component mapping, vulnerability queries, or relationships with IDS agents.

Listing 1. Digital Twin Definition in Eclipse Ditto.

```
{
"thingId": "SISSeC:Lenz_DRB610_1",
"policyId": "SISSeC:policy",
"attributes": {
"manufacturerID": "4302",
"manufacturerName": "Manufacturer",
"dateCode": "20160516",
"model": "Model",
"type": "Drill & Mill Machine",
"image": "/resources/...",
"location": "Hall 1",
"measurements": {...},
"security": {
"cpe": [
{
"name": "Raspberry Pi Model 4",
"usage": "Edge Device",
"enum": "cpe:2.3:h:raspberrypi:raspberry_pi_4_model_b:-:*:*:*:*:*:*:*"
}
],
"alert": {
"sensor": "A002",
"msg": "test alert",
"src_ip_mac": "10.10.10.10",
"dst_ip_mac": "10.10.10.20",
"src_port": "123",
"dst_port": "5065",
"time": "27/08/2022",
"packet_len": "80",
"protocol": "TCP",
"ether_type": "0000"
}
}
}
}
```

Digital twins take security analytics to a new level in knowledge generation and sharing. As shown in Figure 6, our TWINSIGHT UI leverages security-related knowledge from digital twins to inform security professionals and enable knowledge sharing. The user interface integrates device representation (digital twin) and highlights current threats, potential attacks, and granted shares. We believe system-of-systems alerts would lead to more intuitive interaction with Wazuh security analytics software by reducing complexity through digital twins. The alerts would point directly to a system-of-systems component and help understand the big picture and impact of threats and incidents.



Digital Twins 3
Shared Knowledge Paths 56
Warnings 23
My Roles 3

twinsight:LENZDLRG1023_1 Analyst
Lenz Machine
Path: attributes/security (Read/Write)
Roles 3 | Shares 12 | ATT&CK 1 | Threats 8

twinsight:LENZDLRG1023_2 Maintainer
Lenz Machine
Path: /attributes/measurements (Read)
Roles 4 | Shares 2 | ATT&CK 3 | Threats 17

twinsight:LENZDLRG1023_3 Owner
Lenz Machine
Path: * (Read/Write)
Roles 4 | Shares 27 | ATT&CK 1 | Threats 5

Figure 6. TWINSIGHT UI combining security analytics and digital twins.

INTERNET OF THINGS. Indeed, Wazuh relies heavily on an agent-based architecture. Due to its dependence on agents, not all IoT devices are sufficiently addressable. In IoT, a distinction is made between controllable and addressable devices. Addressable devices are reachable via IP addresses and controllable via a controller, i.e., a hub. Wazuh only allows observing addressable devices and leaves out controllable devices, so IoT security analytics still differs from traditional approaches. Controllable devices must be monitored differently to detect and mitigate threats or incidents efficiently.

KNOWLEDGE SHARING. Knowledge sharing in cybersecurity is still in its infancy but will likely increase in the coming years. For example, we found it still difficult to obtain data from machines or their components because the interfaces are not standardized or intentionally hidden. We encountered this problem when we tried to access the interfaces of the machines to get the state data for the digital twin. It took a few months before we could get initial access and start modeling the digital twins. We received excellent feedback from the SISSeC working group, particularly on digital twins as a cybersecurity knowledge generation and sharing facilitator. In this context, data will be shared with the machine maintainer. The PCB manufacturer plans to use parts of TWINSIGHT to generate and share knowledge between the machine owner and IT maintenance. This means that problems can be addressed remotely and do not necessarily have to be solved on-site.

6.6. Discussion

Our research contributes to the scientific community and industry. We also point out the limitations of our research.

SCIENTIFIC CONTRIBUTION. We have summarized research on big data analytics and linked knowledge generation and sharing to cybersecurity. We have formulated a model that envisions digital twins for more efficient security analytics in organizations. We have demonstrated the workings of the digital twin and clarified that simulations are not the only contributors to cybersecurity operations. The overall model helps researchers understand digital twins and aims to draw attention to the use of digital twins for security operations, especially analytics.

PRACTICAL CONTRIBUTION. We have implemented a microservice architecture demonstrating replication-based digital twins for security analytics. We have shown how threat and incident detection can be handled using digital twins, providing a playground for further use cases. Organizations can leverage open-source technologies to deploy their digital twins using the DT2SA model or build sophisticated use cases from scratch using our model. We also highlighted the digital twin paradigm to increase user adoption and make the IoT even more secure. We also drew attention to security analytics technologies, e.g., Wazuh, and their lack of abstraction to form even more complex system of systems. These technologies should take steps toward an asset-specific view that allows users to define complex systems of systems connected to Wazuh agents.

LIMITATIONS. During our research, we encountered several challenges that needed to be solved. Due to the different solution strategies and limited resources, there are limitations to our research. We did not elaborate on access control models. These models provide a clear perspective on inherent roles and grant data access to digital twins. In addition, we validated our model only concerning detective operations. We can only estimate the feasibility of other analytical operations and refer to other literature. However, further research is needed to validate our model in more detail.

7. Conclusions and Future Work

This research aims to leverage cybersecurity knowledge to secure the IoT. We promote cybersecurity knowledge generation and sharing by aligning security analytics with digital twins. Digital twins enable security analytics with high fidelity because they bring semantics and exploit bidirectional communication with their physical counterparts. They take security analytics to a new level, enabling lifecycle centrality and integration among lifecycle participants. This integration promotes the secure sharing of cybersecurity knowledge, such as security states, misconfigurations, or vulnerabilities.

We answered the research question *"How can one align security analytics and digital twins?"* by starting with the foundations of knowledge generation and sharing. We then defined a formal model that elaborates the DT2SA model for adapting security analytics to digital twins. To our knowledge, the DT2SA model is the first to define security analytics comprehensively. We contributed to best practices for research and organizations and bridged the gap between them. Our open-source microservice architecture TWINSIGHT demonstrated practical feasibility and is a starting point for on-building analytical operations. We want to highlight possible future research directions:

- Future research should address decision support for selecting digital twin modes and analytic operations. In particular, whether an analytic operation supports a particular application scenario should be investigated. The goal is to assist analysts in selecting appropriate operation modes for their scenarios. However, the digital twin offers significant cybersecurity opportunities that need to be more fully explored and exploited.
- There is still a considerable need for research, especially in the area of security analytics, since research has focused only on intrusion detection. For example, research should address different analytics implementations based on digital twins. In particular, security monitoring for IoT is urgently needed, as heterogeneous IoT assets form opaque IoT networks. In addition, security analytics research should compare traditional security analytics approaches, such as those implemented in Wazuh, with system-of-systems approaches. It is of the highest interest to evaluate whether analysts using system-of-systems approaches are even more efficient at detecting incidents. Our TWINSIGHT UI highlights opportunities for this evaluation. In addition, there is a significant need for research in implementing a Wazuh plugin for modeling complex system of systems. Finally, future research should work to leverage digital twin recommendations to secure controllable and addressable IoT networks proactively.

While security analytics generates knowledge, digital twins improve overall knowledge generation and enable cybersecurity knowledge sharing. Organizations should leverage cybersecurity knowledge and focus more on digital twins and security analytics. In addition, supply chains should pay more attention to digital twins and their potential for cybersecurity to address sophisticated attacks and APTs. We believe that the digital twin (system of systems) will continue to emerge as a cornerstone of collaboration between lifecycle participants by leveraging cybersecurity knowledge in the Industrial IoT.

Author Contributions: P.E. carried out DT2SA development and the definition of the formal model. P.E. also contributed to knowledge generation and sharing in the cybersecurity domain. G.P. participated in the definition of the DT2SA model. P.E. implemented the microservice architecture TwinSIGHT. P.E. and G.P. also helped to draft the manuscript and revised it critically for important intellectual content. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) and the Bavarian Ministry of Economic Affairs, Regional Development and Energy (BayStMWi). Funding numbers: 16KN085725 and DIK0338/01.

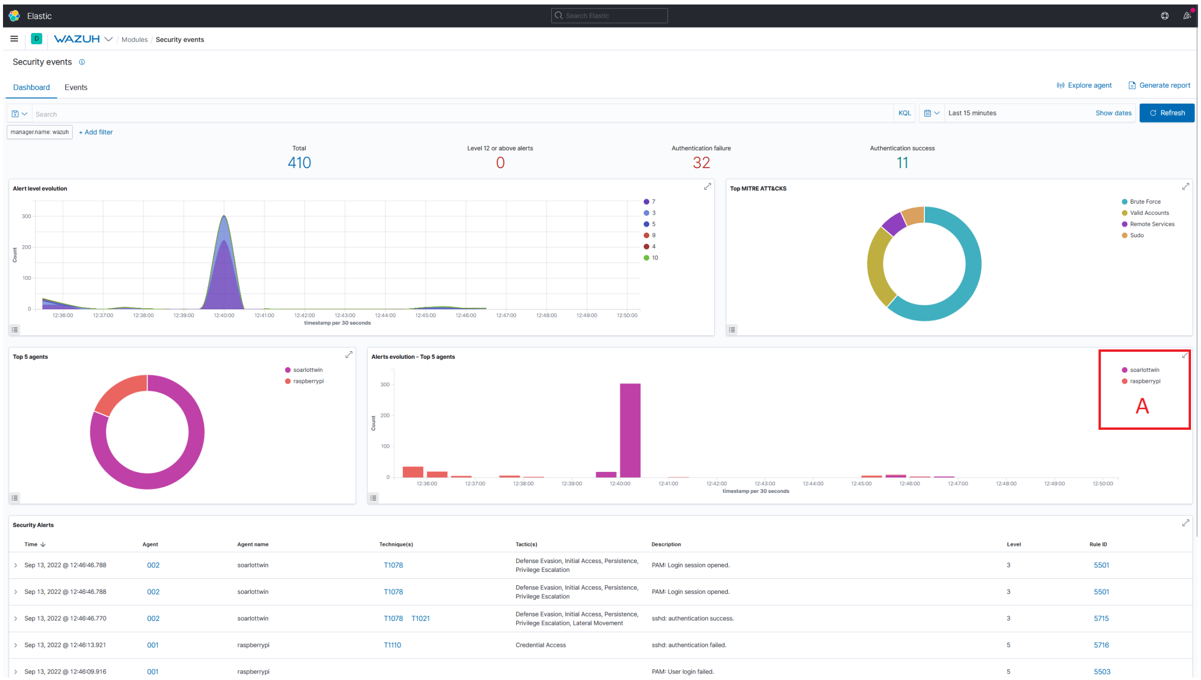
Data Availability Statement: Our data is available on GitHub: https://github.com/philipempl/twinsight.

Acknowledgments: We want to acknowledge all project partners involved in the Secure Industrial Semantic Sensor Cloud (SISSeC) project funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) as part of its central innovation program for small and medium-sized businesses. This research was also partly supported by the Bavarian Ministry of Economic Affairs, Regional Development and Energy (BayStMWi), as part of the INSIST project.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Security Analytics Using Wazuh

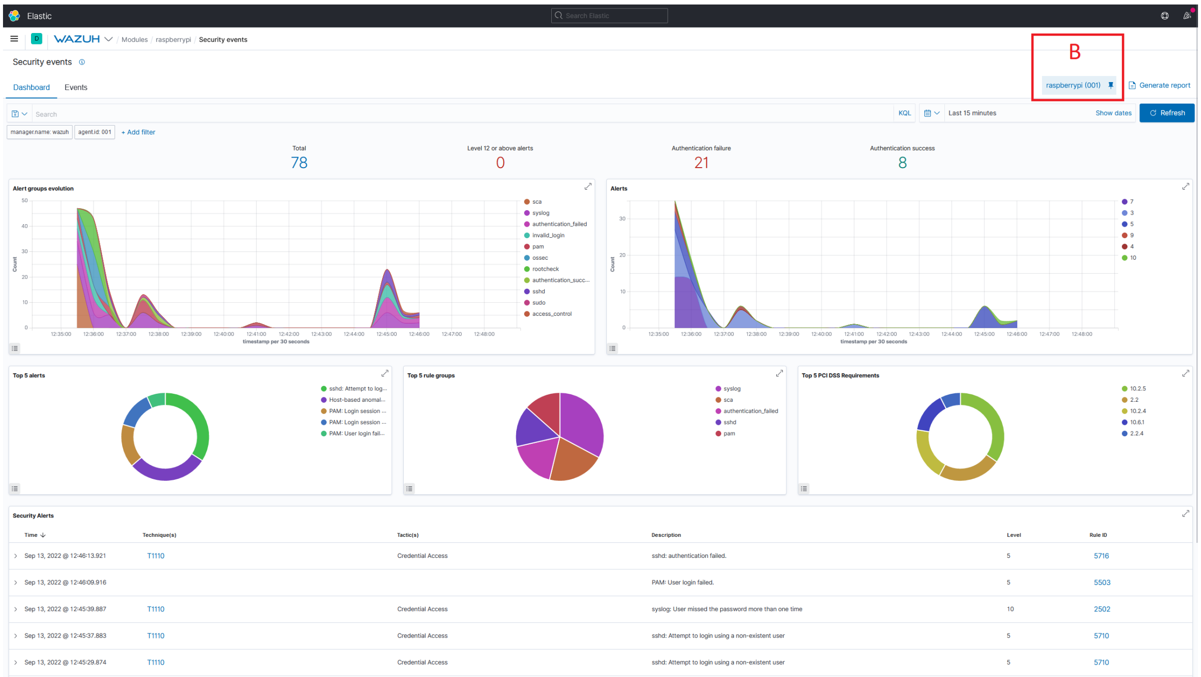
This appendix shows the need for integration and modeling a complex system of systems in Wazuh. Since machines nowadays consist of multiple components, security analytics technology should provide semantic modeling capabilities to analyze systems of systems and their respective components.



Security Alerts

Time ↓	Agent	Agent name	Technique(s)	Tactic(s)	Description	Level	Rule ID
Sep 13, 2022 @ 12:46:46.788	002	soarlottwin	T1078	Defense Evasion, Initial Access, Persistence, Privilege Escalation	PAM: Login session opened.	3	5501
Sep 13, 2022 @ 12:46:46.788	002	soarlottwin	T1078	Defense Evasion, Initial Access, Persistence, Privilege Escalation	PAM: Login session opened.	3	5501
Sep 13, 2022 @ 12:46:46.770	002	soarlottwin	T1078 T1021	Defense Evasion, Initial Access, Persistence, Privilege Escalation, Lateral Movement	sshd: authentication success.	3	5715
Sep 13, 2022 @ 12:46:13.921	001	raspberrypi	T1110	Credential Access	sshd: authentication failed.	5	5716
Sep 13, 2022 @ 12:46:09.916	001	raspberrypi			PAM: User login failed.	5	5503

Figure A1. Analyzing all events of all assets in Wazuh (A shows all agents).



Security Alerts

Time ↓	Technique(s)	Tactic(s)	Description	Level	Rule ID
Sep 13, 2022 @ 12:46:13.921	T1110	Credential Access	sshd: authentication failed.	5	5716
Sep 13, 2022 @ 12:46:09.916			PAM: User login failed.	5	5503
Sep 13, 2022 @ 12:45:39.887	T1110	Credential Access	syslog: User missed the password more than one time	10	2502
Sep 13, 2022 @ 12:45:37.883	T1110	Credential Access	sshd: Attempt to login using a non-existent user	5	5710
Sep 13, 2022 @ 12:45:29.874	T1110	Credential Access	sshd: Attempt to login using a non-existent user	5	5710

Figure A2. Analyzing events of one asset in Wazuh (B showcases the agent filter).

References

1. ENISA. *Threat Landscape for Supply Chain Attacks*; Technical report; ENISA: Athens, Greece, 2021.
2. Ardagna, C.; Corbiaux, S.; Sfakianakis, A.; Douligeris, C. *ENISA Threat Landscape*; Technical report; ENISA: Athens, Greece, 2021.
3. Mahmood, T.; Afzal, U. Security Analytics: Big Data Analytics for Cybersecurity: A Review of Trends, Techniques and Tools. In Proceedings of the 2nd National Conference on Information Assurance (NCIA 2013), Rawalpindi, Pakistan, 11–12 December 2013; pp. 129–134.
4. Böhm, F.; Vielberth, M.; Pernul, G. Bridging Knowledge Gaps in Security Analytics. In Proceedings of the Proceedings of the 7th International Conference on Information Systems Security and Privacy, ICISSP 2021, Online Streaming, 11–13 February 2021; Mori, P., Lenzini, G., Furnell, S., Eds.; SCITEPRESS: Setubal, Portugal, 2021; pp. 98–108. [CrossRef]
5. Skouloudi, C.; Malatras, A.; Naydenov, R.; Dede, G. *Guidelines for Securing the Internet of Things*; Technical report; European Union Agency for Cybersecurity: Athens, Greece, 2020.
6. Pipikaite, A.; Bueermann, G.; Joshi, A.; Jurgen, J.; Bissell, K.; Aguirre, C.; Browder, T.; Pruitt, J. *Global Cybersecurity Outlook 2022: Insight Report*; Technical report; European Union Agency for Cybersecurity: Athens, Greece, 2022.
7. Boschert, S.; Heinrich, C.; Rosen, R. Next Generation Digital Twin. In Proceedings of the 12th International Symposium on Tools and Methods of Competitive Engineering (TMCE), Las Palmas de Gran Canaria, Spain, 7–11 May 2018; Horvath, I., Suarez Riviero, J., Hernandez Castellano, P., Eds.; TMCE 2020 Repository: Dublin, Ireleand, 2018; Volume 2018, pp. 209–218.
8. Eckhart, M.; Ekelhart, A. Digital Twins for Cyber-Physical Systems Security: State of the Art and Outlook. In *Security and Quality in Cyber-Physical Systems Engineering, With Forewords by Robert M. Lee and Tom Gilb*; Biffl, S., Eckhart, M., Lüder, A., Weippl, E.R., Eds.; Springer: Cham, Switzerland, 2019; pp. 383–412. [CrossRef]
9. Pokhrel, A.; Katta, V.; Colomo-Palacios, R. Digital Twin for Cybersecurity Incident Prediction: A Multivocal Literature Review. In Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops, 2020, ICSEW'20, Seoul, Republic of Korea, 27 June–19 July 2020; pp. 671–678. [CrossRef]
10. O'Connor, L. Strengthening Security with Digital Cyber Twins. 2021. Available online: https://www.accenture.com/us-en/blogs/technology-innovation/lisa-oconnor-strengthening-security-with-digital-cyber-twins (accessed on 29 May 2022).
11. Barricelli, B.R.; Casiraghi, E.; Fogli, D. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access* **2019**, *7*, 167653–167671. [CrossRef]
12. Empl, P.; Schlette, D.; Zupfer, D.; Pernul, G. SOAR4IoT: Securing IoT Assets with Digital Twins. In Proceedings of the 17th International Conference on Availability, Reliability and Security (ARES 2022), Vienna, Austria, 23–26 August 2022; Association for Computing Machinery: New York, NY, USA, 2022. [CrossRef]
13. Alcaraz, C.; Lopez, J. Digital Twin: A Comprehensive Survey of Security Threats. *IEEE Commun. Surv. Tutor.* **2022**, *24*, 1475–1503. [CrossRef]

14. Win, T.Y.; Tianfield, H.; Mair, Q. Big Data Based Security Analytics for Protecting Virtualized Infrastructures in Cloud Computing. *IEEE Trans. Big Data* **2018**, *4*, 11–25. [CrossRef]
15. Siow, E.; Tiropanis, T.; Hall, W. Analytics for the Internet of Things: A Survey. *ACM Comput. Surv.* **2018**, *51*, 74:1–74:36. [CrossRef]
16. Cárdenas, A.A.; Manadhata, P.K.; Rajan, S.P. Big Data Analytics for Security. *IEEE Secur. Priv.* **2013**, *11*, 74–76. [CrossRef]
17. Alguliyev, R.; Imamverdiyev, Y. Big Data: Big Promises for Information Security. In Proceedings of the 8th IEEE International Conference on Application of Information and Communication Technologies (AICT), Astana, Kazakhstan, 15–17 October 2014; pp. 1–4. [CrossRef]
18. Ackoff, R.L. From Data to Wisdom. *J. Appl. Syst. Anal.* **1989**, *16*, 3–9.
19. Empl, P.; Pernul, G. A Flexible Security Analytics Service for the Industrial IoT. In Proceedings of the 2021 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems, Virtual Event, Charlotte, NC, USA, 28 April 2021; Gupta, M., Abdelsalam, M., Mittal, S., Eds.; ACM: New York, NY, USA, 2021; pp. 23–32. [CrossRef]
20. Menges, F.; Pernul, G. A comparative analysis of incident reporting formats. *Comput. Secur.* **2018**, *73*, 87–101. [CrossRef]
21. Eckhart, M.; Ekelhart, A. A Specification-based State Replication Approach for Digital Twins. In Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy, CPS-SPC@CCS 2018, Toronto, ON, Canada, 19 October 2018; Lie, D., Mannan, M., Eds.; ACM: New York, NY, USA, 2018; pp. 36–47. [CrossRef]
22. Dietz, M.; Vielberth, M.; Pernul, G. Integrating Digital Twin Security Simulations in the Security Operations Center. In Proceedings of the 15th International Conference on Availability, Reliability and Security, Dublin, Ireland, 25–28 August 2020; Volkamer, M., Wressnegger, C., Eds., 2020, ARES '20, pp. 18:1–18:9. [CrossRef]
23. Damjanovic-Behrendt, V. A Digital Twin Architecture for Security, Privacy and Safety. *ERCIM News* **2018**, *2018*, 25–26.
24. Sacha, D.; Stoffel, A.; Stoffel, F.; Kwon, B.C.; Ellis, G.P.; Keim, D.A. Knowledge Generation Model for Visual Analytics. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 1604–1613. [CrossRef] [PubMed]
25. Preut, A.; Kopka, J.P.; Clausen, U. Digital Twins for the Circular Economy. *Sustainability* **2021**, *13*, 467. [CrossRef]
26. Putz, B.; Dietz, M.; Empl, P.; Pernul, G. EtherTwin: Blockchain-based Secure Digital Twin Information Management. *Inf. Process. Manag.* **2021**, *58*, 102425. [CrossRef]
27. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital Twin in Manufacturing: A Categorical Literature Review and Classification. *IFAC-PapersOnLine* **2018**, *51*, 1016–1022.
28. Dietz, M.; Pernul, G. Unleashing the Digital Twin's Potential for ICS Security. *IEEE Secur. Priv.* **2020**, *18*, 20–27. [CrossRef]
29. Dietz, M.; Putz, B.; Pernul, G. A Distributed Ledger Approach to Digital Twin Secure Data Sharing. In Proceedings of the Data and Applications Security and Privacy XXXIII—33rd Annual IFIP WG 11.3 Conference, DBSec 2019, Charleston, SC, USA, 15–17 July 2019; Lecture Notes in Computer Science; Foley, S.N., Ed.; Springer: Cham, Switzerland, 2019; Volume 11559, pp. 281–300. [CrossRef]
30. Lin, S.W.; Miller, B.; Durand, J.; Joshi, R.; Didier, P.; Chigani, A.; Torenbeek, R.; Duggal, D.; Martin, R.; Bleakley, G. *Industrial Internet Reference Architecture*; Technical report; Industry IoT Consortium: Boston, MA, USA, 2015.
31. Akbarian, F.; Fitzgerald, E.; Kihl, M. Intrusion Detection in Digital Twins for Industrial Control Systems. In Proceedings of the 2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 17–19 September 2020; pp. 1–6. [CrossRef]
32. Atalay, M.; Angin, P. A Digital Twins Approach to Smart Grid Security Testing and Standardization. In Proceedings of the 2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT, Roma, Italy, 3–5 June 2020; pp. 435–440. [CrossRef]
33. Castellani, A.; Schmitt, S.; Squartini, S. Real-world Anomaly Detection by Using Digital Twin Systems and Weakly Supervised Learning. *IEEE Trans. Ind. Inform.* **2021**, *17*, 4733–4742. [CrossRef]
34. Murillo, A.; Taormina, R.; Tippenhauer, N.; Galelli, S. Co-simulating Physical Processes and Network Data for High-fidelity Cyber-security Experiments. In Proceedings of the Sixth Annual Industrial Control System Security (ICSS) Workshop, 2020, ICSS 2020, Austin, TX, USA, 8 December 2020; pp. 13–20. [CrossRef]
35. Saad, A.; Faddel, S.; Mohammed, O. Iot-based Digital Twin for Energy Cyber-physical Systems: Design and Implementation. *Energies* **2020**, *13*, 4762. [CrossRef]
36. Suhail, S.; Jurdak, R.; Matulevicius, R.; Seon Hong, C. Securing Cyber-physical Systems through Blockchain-based Digital Twins and Threat Intelligence. *arXiv* **2021**, arXiv:2105.08886.
37. Chukkapalli, S.S.L.; Pillai, N.; Mittal, S.; Joshi, A. Cyber-physical System Security Surveillance Using Knowledge Graph Based Digital Twins—A Smart Farming Usecase. In Proceedings of the 2021 IEEE International Conference on Intelligence and Security Informatics (ISI), Antonio, TX, USA, 2–3 November 2021; pp. 1–6. [CrossRef]
38. Danilczyk, W.; Sun, Y.L.; He, H. Smart Grid Anomaly Detection Using a Deep Learning Digital Twin. In Proceedings of the 2020 52nd North American Power Symposium (NAPS), Tempe, AZ, USA, 11–13 April 2021; pp. 1–6. [CrossRef]
39. Patel, A.; Schenk, T.; Knorn, S.; Patzlaff, H.; Obradovic, D.; Halblaub, A.B. Real-time, Simulation-based Identification of Cyber-security Attacks of Industrial Plants. In Proceedings of the 2021 IEEE International Conference on Cyber Security and Resilience (CSR), Virtual, 26–28 July 2021; pp. 267–272. [CrossRef]
40. Garcia, H.E.; Aumeier, S.E.; Al-Rashdan, A.Y.; Rolston, B.L. Secure Embedded Intelligence in Nuclear Systems: Framework and Methods. *Ann. Nucl. Energy* **2020**, *140*, 107261. [CrossRef]
41. Tärneberg, W.; Skarin, P.; Gehrmann, C.; Kihl, M. Prototyping Intrusion Detection in an Industrial Cloud-native Digital Twin. In Proceedings of the International Conference on Industrial Technology, Valencia, Spain, 10–12 March 2021.

42. Dietz, M.; Englbrecht, L.; Pernul, G. Enhancing Industrial Control System Forensics Using Replication-based Digital Twins. In *Advances in Digital Forensics XVII*; Peterson, G., Shenoi, S., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; Volume 612, pp. 21–38. [CrossRef]
43. Dietz, M.; Schlette, D.; Pernul, G. Harnessing Digital Twin Security Simulations for systematic Cyber Threat Intelligence. In Proceedings of the 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC), Los Alamitos, CA, USA , 27 June–1 July 2022; pp. 789–797. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

P4 Process-Aware Intrusion Detection in MQTT Networks

Status:	Published
Date of Submission:	11 December 2023
Date of Acceptance:	26 February 2024
Date of Publication:	19 June 2024
Conference:	14th ACM Conference on Data and Application Security and Privacy (CODASPY 2024)
Location:	University of Porto, Porto, Portugal
Period:	19.06.2024 – 21.06.2024
Authors' Contributions:	Philip Empl 60% Fabian Böhm 30% Günther Pernul 10%
Full Citation:	Empl, P., Böhm, F., & Pernul, G. (2024). In *Proceedings of the 14th ACM Conference on Data and Application Security and Privacy* (pp. 91–102). Association for Computing Machinery.
DOI:	10.1145/3626232.3653271
Artifact:	github.com/misssion

Conference Description: Data and applications security and privacy has rapidly expanded as a research field with many important challenges to be addressed. The goal of the ACM Conference on Data and Applications Security (CODASPY) is to discuss novel, exciting research topics in data and application security and privacy, and to lay out directions for further research and development in this area. The conference seeks submissions from diverse communities, including corporate and academic researchers, open-source projects, standardization bodies, governments, system and security administrators, software engineers and application domain experts.

Process-Aware Intrusion Detection in MQTT Networks

Philip Empl
philip.empl@ur.de
University of Regensburg
Regensburg, Bavaria, Germany

Fabian Böhm
fabian.boehm@ur.de
University of Regensburg
Regensburg, Bavaria, Germany

Günther Pernul
guenther.pernul@ur.de
University of Regensburg
Regensburg, Bavaria, Germany

ABSTRACT

Intrusion Detection Systems (IDS) allow for detecting malicious activities in organizational networks and hosts. As the Industrial Internet of Things (Industrial IoT) has gained momentum and attackers become process-aware, it elevates the focus on anomaly-based Network Intrusion Detection Systems (NIDS) in IoT. While previous research has primarily concentrated on fortifying SCADA systems with NIDS, keeping track of the latest advancements in resource-efficient messaging (e.g., MQTT, CoAP, and OPC-UA) is paramount. In our work, we straightforwardly derive IoT processes for NIDS using distributed tracing and process mining. We introduce a pioneering framework called `MISSION` which effectively captures, consolidates, and models MQTT flows, leading to a heightened process awareness in NIDS. Through our prototypical implementation, we demonstrate exceptional performance and high-quality models. Moreover, our experiments provide empirical evidence for rediscovering pre-defined processes and successfully detecting two distinct MQTT attacks in a simulated IoT network.

CCS CONCEPTS

• **Networks** → *Peer-to-peer protocols*; • **Security and privacy** → **Network security**; **Intrusion detection systems**.

KEYWORDS

IDS, MQTT, Internet of Things, Distributed Tracing, Process Mining

ACM Reference Format:
Philip Empl, Fabian Böhm, and Günther Pernul. 2024. Process-Aware Intrusion Detection in MQTT Networks. In *Proceedings of the Fourteenth ACM Conference on Data and Application Security and Privacy (CODASPY '24), June 19–21, 2024, Porto, Portugal.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3626232.3653271

1 INTRODUCTION

Intrusion detection systems (IDS) have long proven themselves as indispensable [68]. They are applicable to many domains [15, 50, 67] and either identify malicious patterns (*signature-based IDS*) or activities deviating from statistically benign behavior (*anomaly-based IDS*) on a host (HIDS) or network (NIDS) [2, 23]. With the proliferation of the Internet of Things (IoT), organizations have increasingly integrated their operational technology (*"physical processes"* [47]) with their IT infrastructure, giving rise to the industrial IoT. Small physical devices and processes shape the IoT but limit performance, communication, and storage capabilities. Keeping the physical processes running is desired to avoid outages [48], necessitating IDS.

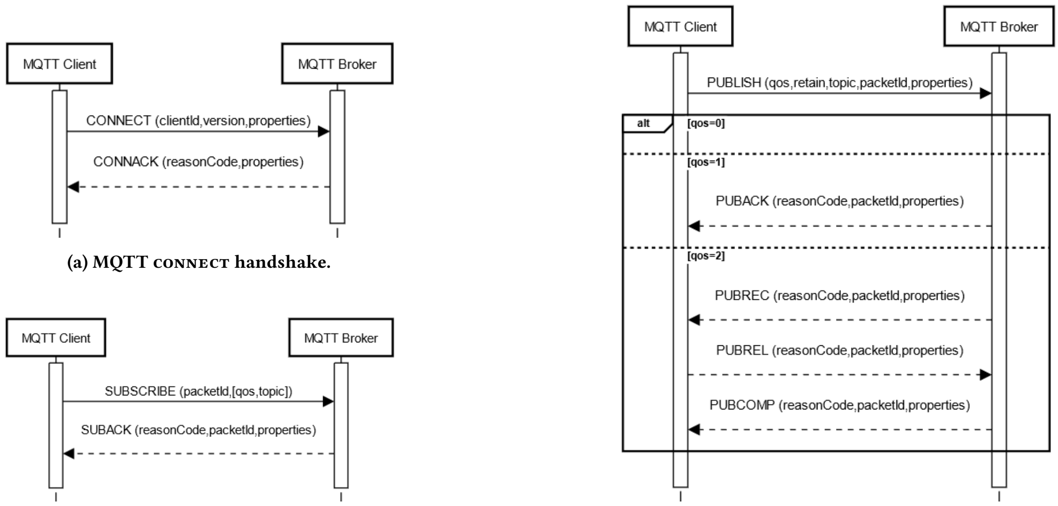
Most notably, research initially focused on systems responsible for managing physical processes. Thereby, SCADA (*Supervisory Control and Data Acquisition*) systems mainly constitute an essential part of nowadays's industrial IoT infrastructure [84]. Different efforts on securing parts of industrial SCADA systems like Modbus or DN3P for communication resulted in plenty of anomaly-based, signature-based and specification-based IDS for both hosts [30, 47, 75], and networks [14, 27, 29, 41, 86]. Signature-based IDS cannot detect unknown attacks, and HIDS may infer physical processes as they require agents running on the host machine [47]. This shifts the focus to anomaly-based NIDS for the Industrial IoT. Additionally, as attacks are more advanced and process-aware [51] like Stuxnet [11] or Industroyer (2) [13, 60], physical processes became a baseline for IDS [7, 9, 16, 17, 26, 59].

SCADA systems are well-researched. Still, these systems are nowadays complemented by resource-efficient messaging protocols like MQTT, CoAP, or OPC-UA, allowing the integration of smaller devices throughout the process and shaping a new era of communication [84]. Despite the benefits of these IoT messaging protocols, their adoption has also increased security threats. Almost half of all organizations cannot detect IoT attacks within their networks [28, 44]. These attacks' increasing frequency and sophistication have raised serious security concerns, but IoT messaging protocols and corresponding communication patterns have only been partially explored for NIDS, yet [10, 18, 45, 55].

We believe that NIDS can benefit from IoT application layer protocols as they carry more contextual information like topic subscriptions than the transport layer, e.g., TCP. As many MQTT clients (such as sensors and actuators) on a single machine result in multiple ports, attacks are unseen by traditional NIDS. The increasing attack surfaces, the potential for additional contextual information, and the lack of research motivate us to investigate the potential of IoT-specific NIDS. However, using process-aware NIDS with IoT messaging protocols presents two primary challenges:

- First, identifying processes, in general, is an extensive manual task [41], whereby different (structured) data like specifications [8, 41], or network traffic packets [26] allow automation, e.g., creating rules or models. Besides, many IoT devices lead to more packets and network complexity, requiring novel, automatic approaches.
- Second, anomaly-based NIDS are mainly based on artificial intelligence lacking explainability [31], failing because of expertise in training and application within organizations [34], and IoT devices communicate via different, heterogeneous (sub-)networks [1]. Organizations require more traceable, distributed, and easy-to-set-up approaches.

CODASPY '24, June 19–21, 2024, Porto, Portugal
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0421-5/24/06.
https://doi.org/10.1145/3626232.3653271


MQTT Client
MQTT Broker
CONNECT (clientId,version,properties)
CONNACK (reasonCode,properties)
MQTT Client
MQTT Broker
SUBSCRIBE (packetId,[qos,topic])
SUBACK (reasonCode,packetId,properties)
MQTT Client
MQTT Broker
PUBLISH (qos,retain,topic,packetId,properties)
alt
[qos=0]
[qos=1]
PUBACK (reasonCode,packetId,properties)
[qos=2]
PUBREC (reasonCode,packetId,properties)
PUBREL (reasonCode,packetId,properties)
PUBCOMP (reasonCode,packetId,properties)

(a) MQTT connect handshake.

(b) MQTT Subscribe handshake.

(c) MQTT Publish handshake.

Figure 1: Relevant MQTT handshakes.

In this paper, we aim to solve these challenges by suggesting network monitoring techniques to automatically and passively record IoT traffic and reduce the number of packets by building flows without risking the physical network's availability. We refer to distributed tracing primarily yet used for auditing [12, 22, 32] and threat investigation and detection [33, 35, 57, 58]. We create explainability and reliability of physical processes as input for NIDS. Consequently, we ask *"How can distributed tracing be utilized to automatically mine IoT processes for NIDS?"* Our main research contributions are the following:

- We envision a distributed and explainable framework `MISSION` to mine IoT processes from network traffic. The process model can be used as an input for NIDS.
- We implement an open-source prototype (GitHub and DockerHub) that allows for real-time probing, collecting, and storing of MQTT flows. Additionally, we provide Jupyter notebooks to showcase the mining of MQTT processes.
- We create an MQTT simulation for research. The simulation consists of multiple sensors and actuators.

This paper is structured as follows. We begin by elaborating on the relevant background for MQTT, IoT network monitoring, and distributed tracing in Section 2. Additionally, related works in this field are discussed in Section 3. In Section 4, we demonstrate the importance of contextual information through two MQTT attack scenarios and the challenges that arise when identifying them. The `MISSION` framework, presented in Section 5, is designed to capture, aggregate, and mine IoT network traffic for NIDS. To evaluate the effectiveness of the `MISSION` framework, we conduct experiments on its performance, process model quality, and its ability to identify two attack scenarios using a prototypical implementation and simulation environment in Sections 6 and 7. Finally, we discuss our research in Section 8 and conclude the paper in Section 9.

2 BACKGROUND

2.1 Message Queuing Telemetry Transport

MQTT stands for Message Queuing Telemetry Transport. It is a lightweight messaging protocol designed for machine-to-machine or IoT communication. MQTT protocol was invented in 1999 by Andy Stanford-Clark of IBM and Arlen Nipper of Arcom Control Systems [24]. MQTT is a publish/subscribe-based protocol that uses a client/server model. The protocol operates over TCP/IP protocol, allowing devices to communicate over unreliable networks with low bandwidth. MQTT is widely used in IoT applications due to its efficiency, low power consumption, and ability to handle intermittent connections. It is an open standard (ISO/IEC 20922 MQTT 3.1.1 [38] and OASIS MQTT 5.0 [21]) and is used in a wide range of industries, including home automation and industrial automation.

The MQTT 5.0 protocol utilizes a packet structure comprising a fixed header, variable header, and payload. The fixed header includes one of 16 control packet types, such as the Connect or Connack, and defines relevant flags for each control type. For instance, the Publish packet includes flags like the quality of service (QoS) level and topic. The variable header contains a packet identifier for request-response mapping and properties, such as the `0x08` property defining the response topic and the `0x09` property pertaining to correlation data. The payload carries data that varies depending on the control packet type. In Figure 1, we summarize three key communication handshakes within an MQTT network, with Figures 1a and 1b relying on the connection to an MQTT broker and topic subscription. A request mostly results in an acknowledgment. Figure 1c illustrates the Publish handshake that varies based on the QoS level. A QoS of 0 is equivalent to fire-and-forget, QoS 1 acknowledges the request, ensuring the packet arrives at least once, and QoS 2 ensures that the packet arrives only once.

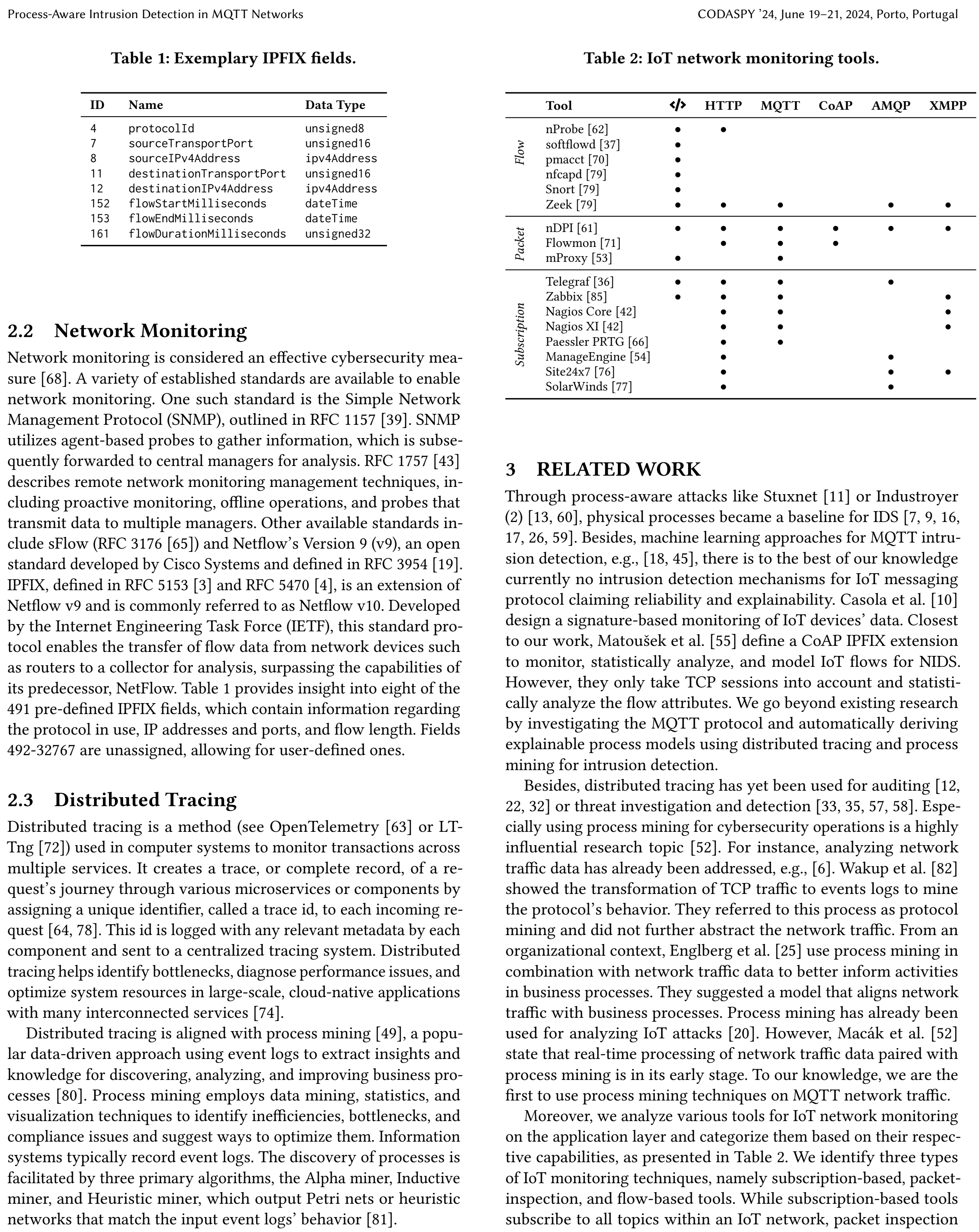


Table 1: Exemplary IPFIX fields.

ID	Name	Data Type
4	`protocolId`	`unsigned8`
7	`sourceTransportPort`	`unsigned16`
8	`sourceIPv4Address`	`ipv4Address`
11	`destinationTransportPort`	`unsigned16`
12	`destinationIPv4Address`	`ipv4Address`
152	`flowStartMilliseconds`	`dateTime`
153	`flowEndMilliseconds`	`dateTime`
161	`flowDurationMilliseconds`	`unsigned32`

2.2 Network Monitoring

Network monitoring is considered an effective cybersecurity measure [68]. A variety of established standards are available to enable network monitoring. One such standard is the Simple Network Management Protocol (SNMP), outlined in RFC 1157 [39]. SNMP utilizes agent-based probes to gather information, which is subsequently forwarded to central managers for analysis. RFC 1757 [43] describes remote network monitoring management techniques, including proactive monitoring, offline operations, and probes that transmit data to multiple managers. Other available standards include sFlow (RFC 3176 [65]) and Netflow's Version 9 (v9), an open standard developed by Cisco Systems and defined in RFC 3954 [19]. IPFIX, defined in RFC 5153 [3] and RFC 5470 [4], is an extension of Netflow v9 and is commonly referred to as Netflow v10. Developed by the Internet Engineering Task Force (IETF), this standard protocol enables the transfer of flow data from network devices such as routers to a collector for analysis, surpassing the capabilities of its predecessor, NetFlow. Table 1 provides insight into eight of the 491 pre-defined IPFIX fields, which contain information regarding the protocol in use, IP addresses and ports, and flow length. Fields 492-32767 are unassigned, allowing for user-defined ones.

2.3 Distributed Tracing

Distributed tracing is a method (see OpenTelemetry [63] or LTTng [72]) used in computer systems to monitor transactions across multiple services. It creates a trace, or complete record, of a request's journey through various microservices or components by assigning a unique identifier, called a trace id, to each incoming request [64, 78]. This id is logged with any relevant metadata by each component and sent to a centralized tracing system. Distributed tracing helps identify bottlenecks, diagnose performance issues, and optimize system resources in large-scale, cloud-native applications with many interconnected services [74].

Distributed tracing is aligned with process mining [49], a popular data-driven approach using event logs to extract insights and knowledge for discovering, analyzing, and improving business processes [80]. Process mining employs data mining, statistics, and visualization techniques to identify inefficiencies, bottlenecks, and compliance issues and suggest ways to optimize them. Information systems typically record event logs. The discovery of processes is facilitated by three primary algorithms, the Alpha miner, Inductive miner, and Heuristic miner, which output Petri nets or heuristic networks that match the input event logs' behavior [81].

Table 2: IoT network monitoring tools.

	Tool	</>	HTTP	MQTT	CoAP	AMQP	XMPP
Flow	nProbe [62]	●	●				
	softflowd [37]	●					
	pmacct [70]	●					
	nfcapd [79]	●					
	Snort [79]	●					
	Zeek [79]	●	●	●		●	●
Packet	nDPI [61]	●	●	●	●	●	●
	Flowmon [71]		●	●	●		
	mProxy [53]	●		●			
Subscription	Telegraf [36]	●	●	●		●	
	Zabbix [85]	●	●	●			●
	Nagios Core [42]		●	●			●
	Nagios XI [42]		●	●			●
	Paessler PRTG [66]		●	●			
	ManageEngine [54]		●			●	
	Site24x7 [76]		●			●	●
	SolarWinds [77]		●			●	

3 RELATED WORK

Through process-aware attacks like Stuxnet [11] or Industroyer (2) [13, 60], physical processes became a baseline for IDS [7, 9, 16, 17, 26, 59]. Besides, machine learning approaches for MQTT intrusion detection, e.g., [18, 45], there is to the best of our knowledge currently no intrusion detection mechanisms for IoT messaging protocol claiming reliability and explainability. Casola et al. [10] design a signature-based monitoring of IoT devices' data. Closest to our work, Matoušek et al. [55] define a CoAP IPFIX extension to monitor, statistically analyze, and model IoT flows for NIDS. However, they only take TCP sessions into account and statistically analyze the flow attributes. We go beyond existing research by investigating the MQTT protocol and automatically deriving explainable process models using distributed tracing and process mining for intrusion detection.

Besides, distributed tracing has yet been used for auditing [12, 22, 32] or threat investigation and detection [33, 35, 57, 58]. Especially using process mining for cybersecurity operations is a highly influential research topic [52]. For instance, analyzing network traffic data has already been addressed, e.g., [6]. Wakup et al. [82] showed the transformation of TCP traffic to events logs to mine the protocol's behavior. They referred to this process as protocol mining and did not further abstract the network traffic. From an organizational context, Englberg et al. [25] use process mining in combination with network traffic data to better inform activities in business processes. They suggested a model that aligns network traffic with business processes. Process mining has already been used for analyzing IoT attacks [20]. However, Macák et al. [52] state that real-time processing of network traffic data paired with process mining is in its early stage. To our knowledge, we are the first to use process mining techniques on MQTT network traffic.

Moreover, we analyze various tools for IoT network monitoring on the application layer and categorize them based on their respective capabilities, as presented in Table 2. We identify three types of IoT monitoring techniques, namely subscription-based, packet-inspection, and flow-based tools. While subscription-based tools subscribe to all topics within an IoT network, packet inspection

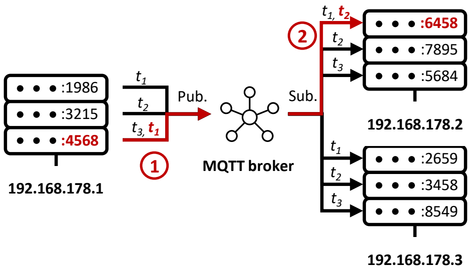


Figure 2: Threat model considering an MQTT network and two possible scenarios.

tools perform deep inspections of specific packet payloads. Flow-based tools observe aggregated packets. However, subscription-based and packet-inspection tools have limitations, especially when dealing with large amounts of data generated by IoT devices, as they may lead to significant performance issues. Moreover, messaging brokers used in subscription-based tools may be vulnerable, compromising the data source's authenticity. Flow-based tools for IoT network monitoring address these challenges. They provide a more comprehensive view of the network traffic and can detect anomalies more effectively. Beside Zeek, we are unaware of tools that probe and aggregate MQTT messages using the flow-based approach. To date, Zeek only targets MQTT 3.1.1 and lacks the correlation field, we concentrate on IPFIX-based MQTT 5.0 monitoring.

4 ADVERSARY MODEL

4.1 Attack Scenarios

This paper considers attacks on IoT networks, especially MQTT. The following discusses the threat model within an MQTT network (see Figure 2). An MQTT network is based on the publish/subscribe architecture of publishers and subscribers. In our proposed network, three devices with different IP addresses run three MQTT clients, with one client publishing data to topics (`192.168.178.1`), while the other two clients (`192.168.178.2` and `192.168.178.3`) subscribe to these topics to receive the data. We define two possible attack scenarios and showcase the associated risks.

The first attack scenario ① involves an attacker gaining unauthorized access to an MQTT client and publishing data to different topics through port 4568. This may occur due to weak access control or insecure configurations. Attackers can conduct denial-of-service attacks by flooding the MQTT network with MQTT Publish messages. In the second attack scenario ②, malicious MQTT clients subscribe to topics they are not authorized to access through port `6458` on `192.168.178.2`. This attack can happen for various reasons (e.g., misconfigurations in the MQTT broker's access control), allowing an attacker to connect successfully to a MQTT client.

Traditional NIDS based on IP address and port information fail to detect such attacks as the attacker behaves normally and sends messages using a valid MQTT client (IP-port combination). Though access control measures such as constraining topic subscriptions can be enforced, many clients publishing on many topics will result in complexity. In this context, we investigate the efficacy and limitations of network-based graphs to detect these attacks.

4.2 Attack Detection

We look closely at an MQTT network's structure and ordinary modeling. We can represent an MQTT network with several communicating devices as a directed graph containing vertices and edges, e.g., by automatically processing a PCAP. The vertices are considered MQTT clients, and the edges as communication links between these clients. We formally define an MQTT network in adoption to Korte et al. [46] as follows:

Theorem 4.1. *An MQTT network structures clients and messages in a directed graph $G = (C, M)$, where*

$$C = \{c_1, c_2, ..., c_n\} \tag{1}$$

is a finite set of MQTT clients and

$$M \subseteq \{(x, y)|(x, y) \in C^2, x \neq y\} \tag{2}$$

defines messages between those clients, containing different attributes a, e.g., $a(m_{type})$, $a(m_{topic})$, or $a(m_{properties})$.

Despite its apparent simplicity, the formal definition of MQTT communication relationships provides a valuable semantic foundation for our research objectives. By leveraging this understanding, we can construct network graphs that reveal fundamental insights into the network's structure. This includes the identification of network centralities, such as brokers, and examining communication patterns between individual devices. As each communication is directed and weighted/labeled, we can deduce the communication's intended purpose. For example, we can discern when a client connects to an MQTT broker and publishes data on different topics.

By utilizing this approach, we can detect attacks occurring in the threat model. For example, if there is no link between 192.168.178.1:4568 and the MQTT broker on topic t_1 in scenario ①, we can deduce an anomaly, such as a denial-of-service, as $t_1 \notin a(m_{topic})$, where $m = (192.168.178.1 : 4568,\ MQTT\ broker) \in C$. Similarly, we can identify the attack in scenario ② as there is typically no connection between 192.168.178.1:6458 and the MQTT broker on topic t_2: $t_2 \notin a(m_{topic})$, where $m = (MQTT\ broker,\ 192.168.178.1 : 6458) \in C$. We can detect attacks in these scenarios by inferring the MQTT broker's communication through Publish and Subscribe messages.

In summary, we can detect attacks using graphs or comparable representations when the attacker is not process-aware. In these scenarios, attack detection is successful when the attacker's behavior differs from the graph's benign behavior definition. As we have motivated our research on attackers becoming increasingly process-aware, we assume that the attacker is aware of the graph structure. An attacker may flood a topic $t_x \in a(m_{topic})$ where m signifies benign communication between two MQTT clients and a is a benign attribute of m, resulting in a denial-of-service attack remaining undetected because there is no deviation from the graph. This motivates our research, as context information is crucial for IoT-specific NIDS. We can concatenate messages or edges in a graph if we know what occurs after client A sends a message. For example, client A sends a message to client B: $t_x \in a_1(m_{topic})|m(A, B) \in M$. Client B processes the information and informs Client C with

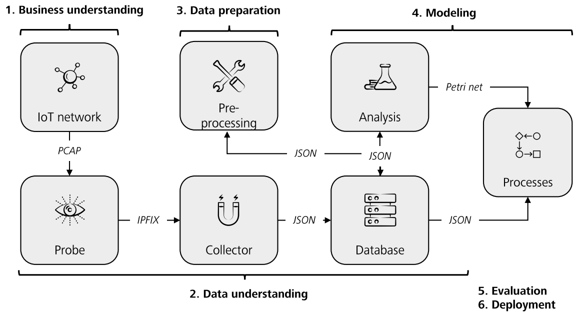


Figure 3: MISSION framework based on CRIPS-DM [83].

$t_y \in a_2(m_{topic}) \mid m(B,C) \in M$, resulting in a sequence, or process P, where $p_1 = a_1 \frown a_2 \mid p_1 \in P$. In the following, we elaborate on a framework detailing how to deduce such processes.

5 MISSION FRAMEWORK

We aim to deduce IoT processes automatically, making them explainable and useful for NIDS without modifying the MQTT 5.0 standard itself. We consider this to be a complex data analysis problem. To ensure rigor, transparency, and traceability, we adopt the *Cross Industry Standard Process for Data Mining* (CRISP-DM) methodology [83], which is highly recognized, aligned with the complexity of the problem and is particularly well-suited for designing data-intensive frameworks in the information systems area [40]. In this paper, we present our approach to mining the processes of IoT networks, especially MQTT, for NIDS, which we call MISSION ("**Mi**ning **S**emantic**s** of **I**oT **N**etworks"). Our framework comprises six phases defined in the CRISP-DM methodology, as illustrated in Figure 3. In the following, we detail the first four phases and provide insights into the phases of evaluation in Section 6.

5.1 Business Understanding

First, we aim to understand the problem at hand. Our business problem is the associated cybersecurity risk through process-aware attacks within IoT networks. To target this problem, we mine IoT-specific processes in network traffic. Specifically, we focus on analyzing MQTT networks, elaborating on their processes, and using them for NIDS. We have the following requirements:

- **Explainability.** Models should be explainable by means humans can reproduce their creation and structure.
- **Distribution.** Data should be reliably collected from different (sub-)networks to integrate the whole network.
- **Standards.** Open standards allow for innovation, interoperability, cost-effectiveness, and security.

We address these requirements by employing established distributed tracing methods in combination with process mining and network monitoring standards.

5.2 Data Understanding

In the second phase, we need to understand the data. MQTT is a publish/subscribe messaging protocol that operates on the application layer using TCP/IP. In line with this definition, an MQTT network consists of clients (C) and messages (M) with weighted attributes, such as $a(m_{topic})$. MQTT handshakes are initiated through different packet control types like Connect or Publish. A handshake mostly comprises of a request (e.g., Connect) and an acknowledgment (e.g., ConnAck). We define such handshakes as flows, a "set of related IP packets". In total, MQTT 5.0 has seven possible flows: Connect, Publish, Subscribe, Unsubscribe, Disconnect, Ping, and Authentication. The length of a flow depends on the control packet type. For example, the Publish flow's vary in length depending on the QoS level. A flow with the "at most once" QoS level is made up of one packet, and with the "at least once" QoS level ($QoS = 1$) out of two. In contrast, a flow with the "exactly once" QoS level ($QoS = 2$) comprises four packets. Each packet comprises one or more header fields, additional characteristics, or computationally derived attributes. A flow is identifiable by a key and is unique. Each flow represents an activity within an IoT process instance. We rely on network monitoring techniques to probe, collect, and match related MQTT flows in IPFIX. In the following, we detail *probing*, *collecting*, and *storing* MQTT flows.

Table 3: Suggested MQTT-specific IPFIX fields.

ID	Name	Data Type
32769	mqttQoS	unsigned16
32770	mqttControlType	unsigned16
32771	mqttPacketId	unsigned16
32772	mqttTopic	string
32773	mqttSrcClientId	string
32774	mqttDstClientId	string
32775	mqttCorrelationData	string

Probing. We propose using an IoT network probe to capture and aggregate MQTT network packets, enriching them with contextual information to match related flows and storing them in a database. To transform these packets into flows, MISSION relies on a well-known approach based on 5-tuples, as defined in RFC 6146, which matches packets according to the source IP address, source port, destination IP address, destination port, and protocol:

$$Tuple(srcIP, dstIP, srcPort, dstPort, prot) \tag{3}$$

The combination of IP addresses and ports is considered unique since the MQTT broker maintains the TCP sessions of the MQTT clients, even when multiple clients connect to the broker using the same IP address. Hash tables ("flow tables") match coherent packets and reverse the flow in case of acknowledgments, e.g., inverting source IP and destination IP. As MQTT flows vary in size, the required flow length depends on the packet control type and QoS level, e.g., a Publish flow in QoS level 2 awaits a length of four packets. To prevent collisions resulting from multiple clients' publish requests, the packet identifier complements the 5-tuple if available (only for $QoS > 0$), ensuring that packets are correctly assigned to the corresponding MQTT flow. As MQTT-specific fields are not included in the 491 pre-defined fields of the IPFIX standard, we complement it with user-defined ones (see Table 3). Please note that these fields are suggestions that can be augmented with additional data points. The definition and number of these fields depends on the users' interest and do not influence our framework. Among these fields, correlation data is of utmost importance as it contains the trace identifier as suggested by distributed tracing. The correlation data should be unique and can be a UUID4 string (an alphanumeric string of 36 characters). MQTT usage has to be modified to transmit the trace identifier. When a client receives correlation data, it must attach it to any potential, following Publish message to keep the trace so that:

$$p = a_1(m_{correlationData})^\frown a_2(m_{correlationData}) \mid p \in P \quad (4)$$

Collecting & storing. It is essential to follow a set of steps to ensure the effective and efficient collection and storage of MQTT flows across multiple distributed networks. RFC 5153 provides guidelines for IPFIX collectors, which must be capable of decoding and encoding information, managing templates, and utilizing a transport protocol like UDP or TCP. When probes are widely dispersed across the network, using SCTP as per RFC 4960 is recommended. The IPFIX records should be stored in a document-oriented database that offers more flexibility than traditional relational databases. This flexibility is crucial because flows can vary depending on the application protocol. The data can be stored using JSON serialization to minimize storage requirements and provide easily readable flow records. The collector should use the same template as the probe.

5.3 Data Preparation

Once the IPFIX flows are stored in a document-oriented database, preprocessing of the MQTT flows becomes possible. However, two quality aspects need to be considered during preprocessing.

- MQTT flows may not contain client identifiers, especially in Publish or Subscribe flows.
- MQTT brokers do not have client identifiers, but friendly names must be assigned to devices without one making them human-readable.

In MQTT networks, clients can decide whether a broker retains their sessions, and this information is transmitted via a Connect flow. Sessions are managed through specific IP addresses and port combinations within an MQTT broker. Therefore, effective preprocessing must be capable of mapping client identifiers to IP addresses and ports. Additionally, preprocessing must resolve and track MQTT flows with unknown client identifiers. Each MQTT flow defines a source and a destination client. MQTT brokers manage client sessions but do not have client identifiers; hence, they only appear in network traffic data with their IP addresses and ports. To account for this, preprocessing must identify occurrences of MQTT brokers, usually on port 1883 or 8883 (SSL), and assign them a friendly name such as "MQTT Broker". It is important to note that preprocessing may vary if an MQTT broker runs on different ports. Finally, with the collection and preprocessing of MQTT flows completed, the next step is to model and correlate the flows semantically.

5.4 Modeling

In the realm of MQTT networks, processes and rules govern their operation. For instance, an industrial process follows business logic and is defined by rules using a control system. These processes may not adhere to strict sequential order and can operate in parallel, resulting in multiple states. Finite-state machines are unsuitable for modeling such networks, as they are designed to model a single client instead of an entire network. On the other hand, Petri nets provide a dynamic model for system behavior and can replicate concurrent processes. These directed networks, comprising vertices and edges, are well-suited to represent MQTT networks. Hence, we adopt the use of Petri nets for modeling MQTT networks, following the model introduced by Petri [69]:

Theorem 5.1. *MQTT processes structured as a Petri net represent a triple $N = (P, T, A)$, where:*

$$P = \{p_1, p_2, ..., p_i\} \quad (5)$$

is a finite set of places holding the MQTT clients' state,

$$T = \{t_1, t_2, ..., t_j\} \quad (6)$$

is a finite set of transitions representing the Publish flows with attributes like $t_1(m_{topic},\ m_{traceId})$, and we can define all arcs of the Petri net as

$$A = (P \times T) \cup (T \times P) \quad (7)$$

whereby inputs direct to places and outputs to transitions given weights, so that:

$$I : t \mapsto p \mid (t, p) \in A \quad (8)$$

$$O : p \mapsto t \mid (p, t) \in A \quad (9)$$

We consider an MQTT client a place so that $c \mapsto p | \ c \in C, p \in P$. We can further specify the MQTT flows gathered in the previous steps as transitions between two MQTT clients, whereby a flow F can be described as $(p, t) \lor (t, p') | (p, t) \lor (t, p') \in A$. A flow may contain different contextual information depending on the packet control type. For instance, Publish flows have fields like $F_{topic,traceId}$, while other packet control types contain different information. Since Publish flows represent the IoT process, the remaining flows, i.e., Connect or Subscribe, are considered prerequisites for publishing messages as there are no Publish flows without prior subscriptions and connections. For instance, due to the publish/subscribe architecture, subscriptions results in Publish flows outgoing from the MQTT broker. Conversely, without authenticating and connecting to the MQTT broker, there are no Subscribe or Publish flows. By considering only the Publish flows, we can deduce and address the remaining ones.

Besides, the Petri net enables the identification of flows that trigger the traversal of different places, creating new markings. Each MQTT client can concurrently hold multiple markings, describing individual process instances in the Petri net. This is represented as $\forall p :\ M(p)$, where $M(p)$ denotes the markings held by an MQTT client. The number of markings held by an MQTT client corresponds to the number of concurrent process instances it can handle. The markings enable the tracking of concurrent processes in the MQTT network as they travel through the Petri net. A transition t is activated by the function $F(p, t)$, which involves MQTT clients and subsequent transitions in the Petri net.

Once the formal model has been created, distributed tracing methods, specifically process mining algorithms, can be used to discover Petri nets automatically. These algorithms rely on event logs containing cases or process instances, activities, and additional details such as resources and timestamps [80]. The mapping of IPFIX fields to the event log structure can be represented as follows:

- **Timestamp:** $flowStartNanoSeconds(156)$
- **Case id:** *mqttCorrelationData*(32775)
- **Resource:** *mqttSrcClientId*(32773)
- **Activity:** *mqttTopic*(32773)

Process mining discovery techniques, e.g., Alpha miner, use this event log structure to discover a Petri net of the MQTT network. Additionally, process mining's conformance-checking algorithms - used to compare the behavior of a process as recorded in an event log to the Petri net's behavior - identify deviations acting as NIDS.

6 EVALUATION

In this section, we conduct experiments on an implementation based on the `MISSION` framework to deduce IoT processes for NIDS. Since there are no tools and research on flow-based MQTT 5.0 NIDS, we cannot make a baseline comparison. So, we rely on the attacker scenarios as defined in Section 6.1, and the experimental setting defining IoT processes in Section 6.2. In Section 6.3, we aim to rediscover the pre-defined rules, assess the quality of the Petri nets, and detail how to detect the attacks.

6.1 Attack scenarios

Our primary objective is to evaluate the vulnerability of the network to two distinct attack scenarios, as outlined in Section 4. The attacker's objective is to manipulate the network's integrity and confidentiality. The first scenario involves an eavesdropping attack, where an infiltrator compromises network confidentiality by subscribing to MQTT topics assessing the data. The second scenario revolves around an attacker capable of deducing network traffic and executing a Denial of Service (DoS) attack by flooding the network with Publish packets. We aim to identify and analyze both of these attacks with the `MISSION` framework within an experimental setting, showcasing the power of conformance checking through process mining. The detection of these attacks relies on previously discovered process models. We assume the attacker can access a single node within the network.

6.2 Experimental Setting

First, we aim to investigate the applicability of distributed tracing and process mining on flows to derive the IoT processes of an MQTT network. To achieve this, we define a set of rules within an experimental MQTT network and simulate the behavior of several industrial IoT assets derived from a real world use case. We then capture and preprocess the resulting network traffic flows using flow tables (hash tables) within our MQTT probe[1], which exports them to a collector for storage. After preprocessing, we apply the remaining phases of the MISSION framework and its prototypical implementation to unravel the underlying process model.

Our MQTT simulation is implemented using Python and the Eclipse Paho library and is publicly available on GitHub[2]. We simulate an industrial environment consisting of various IoT assets, such as sensors and actuators within different systems like a belt circulation system or a log server, interacting with each other over MQTT. Sensors passively sense the environment and publish data to MQTT topics within a pre-defined value range, while actuators

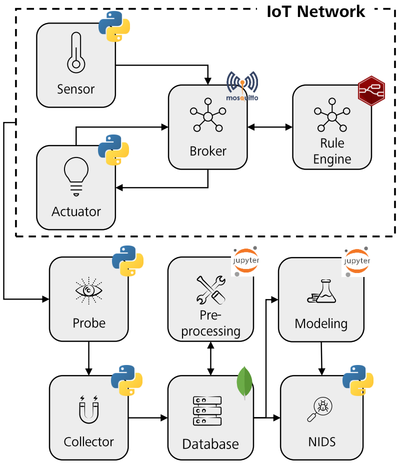


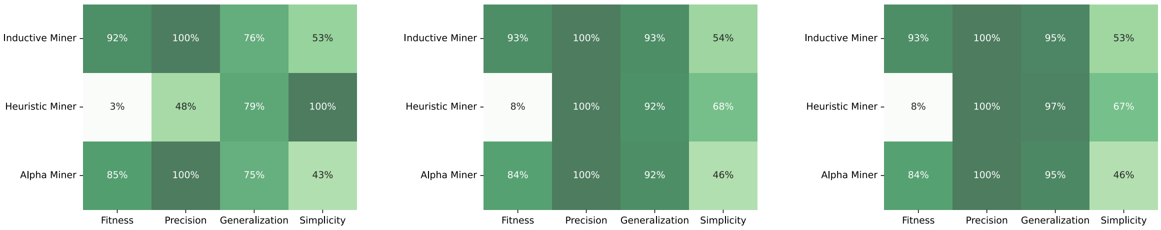
Figure 4: Experimental setting.

actively interact with the environment by subscribing to MQTT topics and reacting to incoming events, e.g., state changes from on to off. Status changes are recognized by a rule engine that reacts differently on the base of predefined threshold values. Note that industrial environments typically are static, so we do not consider scenarios where assets join or leave the network. The simulation is deployed on a virtual machine running Ubuntu 22.04 LTS with six cores, 8GB RAM, and 50GB storage. The MQTT probe runs on a separate virtual machine with Ubuntu 20.04.3 LTS (16GB RAM, eight cores, and 80GB storage).

In our MQTT simulation, we instantiate 57 sensors and actuators, including temperature, motion, and window sensors. Each of these devices regularly sends packets at a fixed frequency that depends on the category of the device. The MQTT Broker is implemented using the open-source software Mosquitto, which is widely used in the industry. To connect to the broker, each IoT device in our simulation uses Eclipse Paho, an open-source MQTT client that supports the latest MQTT specification v5 and user-defined fields. To implement the event-driven architecture of our simulation, we use the low-code programming software Node-RED, which acts as an MQTT client and allows the definition of rules. We pre-define three rules in Node-RED within in an industrial context. Note that these rules are notional to exemplify different devices operating.

The probe is deployed within the same network as the MQTT communication, enabling the capture of all MQTT packets. Subsequently, the captured packets are transformed into flows and exported. The collector then aggregates the MQTT flows and stores them in a document-oriented database, MongoDB[3]. Preprocessing of the data is performed after storing the flows. This involves transforming the flows into an event log by mapping IPFIX fields to event log-specific fields, such as case identifiers or timestamps. Additionally, we filtered by Publish flows.

[1] https://github.com/misssion/probe
[2] https://github.com/misssion/simulation
[3] https://www.mongodb.com



(a) 1,000 events. **(b) 10,000 events.** **(c) 100,000 events.**

Figure 5: Evaluation of process mining discovery techniques using different event log sizes.

6.3 Results

The MQTT simulation was conducted over approximately 12 hours, generating a dataset of 100,000 flows. After preprocessing and storing these flows in a MongoDB database, we perform process mining discovery techniques to unravel the IoT processes. We test all three process mining discovery techniques available in the PM4Py[4] library to ensure a thorough comparison of the results. All of our results are available online through a Jupyter notebook[5].

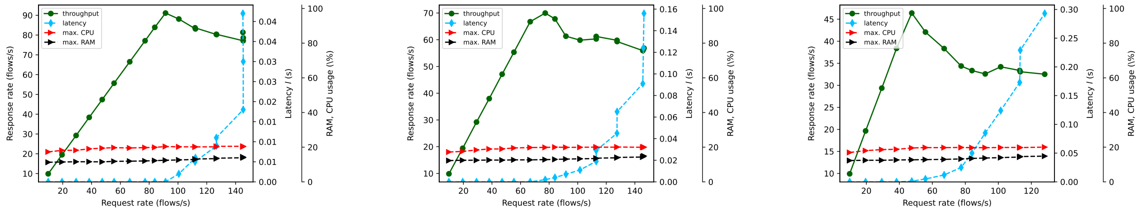
Directly-follows graph. Before assessing the quality of the discovered process models in our event log, we discuss using directly-follows graphs (DFG) to generate a model encompassing all available process instances. DFGs capture all process variants and do not aggregate or simplify process models. They can serve as a baseline for comparison with the process mining discovery output or to provide initial insights.

MQTT processes. After generating DFGs, we apply process mining discovery techniques to our event log, demonstrating the robustness of these techniques in producing accurate process models. Firstly, we use the Inductive miner on our event log, which excels at identifying relationships within event logs. To exclude MQTT communication that may not be pertinent, we set a noise threshold of 0.9. A high noise threshold only considers continuously occurring and thus probable processes. Secondly, we configure the heuristic miner, which is optimal for working with noisy and incomplete data. The heuristic miner relies highly on a relation threshold, which we set at 0.9. A higher relation threshold eliminates less frequent edges in the process model. Lastly, we employ the Alpha miner, adept at identifying parallel activities. However, selecting the process discovery algorithm highly depends on the amount of distinct events in the network. Within this model, we observe that all pre-defined processes have been successfully re-identified. By measuring and averaging the event log data, we determine that our experimental scenario generates a new process instance every 0.31922 seconds (start time) with a case dispersion ratio of 0.31922 seconds (end time) on average. Our experimental setup can sustainably process the received flows as the arrival and dispersion ratio remains constant. The throughput time is 0.00166 seconds, indicating a near real-time operation.

Model quality. To assess the quality of the process models discovered by the Inductive miner, Heuristic miner, and Alpha miner, we evaluate them based on four well-established process mining quality criteria: replay fitness, simplicity, precision, and generalization [5]. Figure 5 presents the evaluation results of the three process mining techniques regarding these four criteria. Additionally, we discuss the soundness of our model as an extra evaluation criterion. It is worth noting that we evaluate the models against three different scenarios with varying amounts of flows and process mining discovery techniques. Note that two of our rules occur less frequently. The first rule appears at approximately 4%, and the less frequent rule at approximately 1.2%.

Replay fitness is a well-established quality criterion in process mining, which quantifies the degree of accuracy with which the discovered model can replicate the event log. It measures the percentage of reproduced process instances and ranges from 0% to 100%. A higher percentage indicates better fitness and a value of 100% indicates perfect fitness. Figures 5a, 5b, and 5c demonstrate that the fitness of the discovered process models remains nearly constant across all three process mining techniques. The Inductive and Alpha miners show the best performance regarding fitness, indicating that reproducing the event log would have been feasible with only 1,000 flows. Precision is another essential criterion that measures the overfitting/underfitting of the discovered model. If a model involves more paths than necessary to represent an event log, the model is overfitted. Precision is also measured in percent and reflects the degree to which the model represents the event log. All three models are precise, with at least 10,000 flows. Generalization measures the ability of a discovered model to represent future process instances reasonably. We can see a slight increase in the generalization of the process model with an increase in the number of flows available for process discovery. The simplicity of a discovered model quantifies the maximum complexity required to represent the event log. A higher percentage indicates less complexity needed to represent the event log. We observe that the simplicity of the discovered model stays almost constant although the number of flows increases, except for the Heuristic miner, which shows a substantial decrease in simplicity. Furthermore, we evaluate the soundness of our model using Woflan algorithm, which indicates a binary decision if the discovered model complies with all modeling rules (e.g., start/end activities or no dead locks). In our experimental

[4] https://pm4py.fit.fraunhofer.de/
[5] https://github.com/misssion/evaluation



(a) MQTT Publish flow with QoS 0.

(b) MQTT Publish flow with QoS 1.

(c) MQTT Publish flow with QoS 2.

Figure 6: Benchmarking MQTT probe running on a virtual machine.

setting, the Inductive miner produces sound models for event logs with at least 10,000 flows. The Heuristic miner has sound models for event logs with more than 100,000 flows. However, the Alpha miner did not produce any sound models.

Attack detection. Conformance checking identifies anomalies concerning the event log utilized for process discovery. Initially, we focused on employing procedural conformance checking, e.g., token-based replay (TBR) and alignment checking to detect the two attacks. However, procedural conformance checking does not account for specific resources' involvement in executing activities. Consequently, we encountered challenges in detecting attacks perpetrated by unauthorized resources following a particular process. We attempted declarative conformance checking using a log skeleton as an alternative approach to examine unusual resource-activity allocations within the event log. However, this approach lacked the incorporation of process awareness within declarative models, but also the resource perspective. We decided to rename activities concerning resource-activity allocation to cope with this limitations. Regarding the first attack scenario, the DoS attack, both procedural methods (TBR and alignment) yielded anomalous results, as the attacker did not behave process-aware. This was observed regardless of whether the attacker used the same trace id for packets or unique ones. The log skeleton approach also identified a mismatch between the event log and the malicious traces. In the case of second attack scenario, we successfully detected anomalous subscriptions using procedural methods. However, the declarative method, log skeleton, failed to identify the anomaly when comparing the benign event log with the malicious ones. This discrepancy arose from the activity name "MQTT Broker ⟶ mission/log_server-sensor-temperature", which, at first glance, appeared benign due to the publish/subscribe architecture obscuring the client receiving the data. In summary, using process models as a baseline for NIDS requires a well-thought definition of input but heightens the respective process awareness.

7 PERFORMANCE EVALUATION

We conduct an assessment of the performance of our MQTT probe. This performance assessment is paramount in determining the probe's ability to process real-time data. To evaluate the probe's performance, we employ the key performance indicators of *maximum sustainable throughput* and *latency*, as proposed by Sedlmeir et al. [73] initially designed for blockchain to measure a node's efficiency. The performance evaluation is documented using Jupyter Notebooks and accessible on GitHub[6] to enable reproducibility.

The maximum sustainable throughput is the highest frequency of network packets that a probe can efficiently transform to IPFIX and export to a collector. To operate the MQTT probe, we deploy a virtual machine running Ubuntu 20.04.3 LTS with 16GB RAM, eight cores, and 80GB disc storage. To determine the maximum throughput, we subject the probe to different fixed frequencies of network packets per second. In a real MQTT network, the probe would receive varying MQTT packets per second, depending on the QoS level. In a network with $QoS = 0$, a probe must export as many flows as the number of packets received. The load increases to two packets in an MQTT network using $QoS = 1$, and four packets are exported for $QoS = 2$. As a result, we create three scenarios where the MQTT probe intends to transform all the packets received. In each scenario, we publish MQTT messages to the MQTT broker at a fixed frequency, beginning with ten packets/s for a minute. The MQTT probe captures the network traffic packets, transforms them into flows, and exports them. We then measure the time t taken to send a message, the export time, the latency ($t_{flow\ exported} - t_{message\ sent}$), and CPU/RAM usage. This process is repeated for different frequencies (step size = 10). Upon successfully executing the experiments, we identify the throughput of each run at a fixed frequency by calculating the linear regressions of responses and requests. The throughput is deemed sustainable and latency-free as long as these linear regressions remain parallel [73]. For instance, throughput is sustainable if the probe receives and exports 40 packets/seconds ($QoS = 0$) with minimum latency. Additionally, we computed the average latency and CPU/RAM usage.

Figure 6 depicts the results obtained from all three scenarios. We transform a total of 92,037 flows for $QoS = 0$, 92,393 flows (184,786 packets) for $QoS = 1$, and 58,675 flows (234,700 packets) for $QoS = 2$. Starting with $QoS = 0$, Figure 6a indicates that our MQTT probe can sustainably export 100 flows/s. After exceeding 100 flows/s, the latency increases, and the response rate decreases. In the $QoS = 1$ scenario (see Figure 6b), the probe can sustainably export 70 flows/s. It is important to note that exporting 70 flows/s with $QoS = 1$ involves a stress factor of 140 packets/s. In the last scenario (see Figure 6c), the maximum sustainable throughput is 50 flows/s (200 packets/s). Interestingly, the RAM and CPU usage remains constant even when the maximum sustainable throughput is achieved. This

[6] https://github.com/misssion/benchmark

phenomenon occurs because our probe is implemented in Python, which has a global interpreter lock to synchronize the execution of threads. Multiprocessing is currently not feasible due to the shared memory of the flow table.

Our MQTT probe performs satisfactorily in all three scenarios, processing between 100 to 200 packets/s and 40 to 100 flows/s, respectively. It is important to note that a collector would aggregate exported flows from multiple probes in a real-world scenario. Scaling to more than one probe would allow for meeting the requirements of more extensive IoT networks. However, in smaller networks, it may be sufficient to operate at least one probe per sub-network to export flows. In summary, our probe, collector and database implementation provides deployable and seamlessly integrated components for existing organizational IoT networks based on MQTT. It focuses on reliability and provides stable and consistent performance for efficient process model discovery and conformance checking on network data enabling security operations.

8 DISCUSSION

This section discusses key learnings and insights during our research, limitations in Section 8.1, and future work in Section 8.2.

OT availability requirements vs. security modifications. The industrial IoT demands high availability to prevent financial loss resulting from outages [48]. Our framework requires modifications to industrial IoT communication to ensure the trace identifier is transmitted whenever a device reacts to a request. However, this may conflict with the high availability requirements of the industrial IoT. Consequently, organizations must weigh the advantages of improved security against the risk of potential outages, constituting a fundamental decision-making process not limited to MISSION.

No need for a sledgehammer cracking a nut? NIDS are shaped by machine learning algorithms that attempt to explain correlations in the data. However, there are cases where it may not be feasible or desirable, particularly in organizations that need more expertise [34] or when models are not explainable [31]. Our more straightforward approach creates reliable and explainable models based on network traffic data and trace identifiers. At present, our framework only generates boolean results for anomalies, whereas machine learning approaches aim to classify attack types directly. Combining both techniques could prove helpful in detecting anomalies and classifying them later, as exemplified by Microsoft Security Copilot's use of large language models [56].

MISSION *as a defensive or offensive measure?* Our research reveals that mining processes in the industrial IoT is feasible with minimal effort, particularly for IoT processes that support business operations. When considering a business process view, manufacturers order numbers to orchestrate their machines, often using plain text communication. These parts of order numbers can also function as trace identifiers resulting in business processes. Our framework, while designed for defensive purposes, could be exploited by attackers to gain insight into operations to launch attacks. It would be interesting to see how attackers have gained process awareness in well-known attacks such as Stuxnet [11] or Industroyer [13, 60]. However, this requires either network traffic in plain text or capabilities for decryption.

8.1 Limitations

Although our framework, MISSION, has shown effective in exploring MQTT processes, several limitations should be considered. Firstly, the scope of our paper is limited to MQTT, and other IoT application protocols may have different messaging patterns that require a distinct data collection and modeling approach. Secondly, we have only estimated the flows ($n = 100,000$)needed to derive sound models. Moreover, we only model Publish flows, as they require corresponding Subscribe or Connect flows in advance. The CPU and RAM utilization of the probe and collector has not reached their limits; more personnel resources could enhance their efficiency, for example, through multi-threading or multi-processing. Additionally, we could have integrated our framework into existing probes, such as softflowd or pmacct. However, to maintain the coherence of the framework, we refrained from doing so. Lastly, we consider network traffic in plain text, which often applies to industrial networks. If encrypted, keys must be available.

8.2 Future Work

We propose several ideas for further improving and developing the MISSION framework. One such idea is to use business process-relevant information (e.g., product identifiers) to reconstruct business processes instead of using correlation data. Future research should also focus on identifying high-level semantic information to further abstract the process, but also to cope with encrypted traffic. Additionally, research should delve into the alignment of resource-based and control flow-based conformance-checking methods to detect intrusions. While we successfully detected both attacks, the absence of a resource perspective in existing conformance-checking methods poses a challenge, necessitating novel aligned approaches for NIDS. We believe the MISSION framework is getting a cornerstone for IoT security and should be expanded to include additional application protocols such as OPC UA, CoAP, and XMPP. Besides, MISSION can provide proactive security monitoring and testing and identify optimizations for IoT networks, such as mandatory access control rights. For example, MQTT brokers could be optimized based on the process model. Finally, the MISSION framework could be implemented in open-source and commercial probes like softflowd to increase its adoption and get production-ready.

9 CONCLUSION

Our paper uses distributed tracing for NIDS to mine IoT processes from network traffic automatically. We found that IoT application protocols contain contextual information relevant to NIDS and can be captured near real-time using network monitoring techniques. We use corresponding flows as input to derive process models using process mining discovery techniques and conformance checking to detect anomalies. Our prototype, including probe, collector, and modeling, is open-source, deployable, and available on GitHub with multi-arch images to be found on DockerHub[7]. We highlight the MISSION framework as a fundamental cornerstone in creating transparency, explainability, and enhancing cybersecurity in IoT networks, enabling more sophisticated security operations beyond network intrusion detection systems.

[7] https://hub.docker.com/u/iotmission

REFERENCES

[1] Ala I. Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, and Moussa Ayyash. 2015. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Communications Surveys & Tutorials* 17, 4 (2015), 2347–2376. https://doi.org/10.1109/COMST.2015.2444095

[2] Stefan Axelsson. 2000. *Intrusion detection systems: A survey and taxonomy.* Technical Report.

[3] B. Claise and B. Trammell and P. Aitken and S. Zseby and J. Quittek. 2008. *IP Flow Information Export (IPFIX) Implementation Guidelines.* Technical Report. https://doi.org/10.17487/rfc5153 RFC 5153.

[4] B. Trammell and E. Boschi and T. Zseby and D. Quittek and M. Stiemerling and M. Claise. 2009. *Architecture for IP Flow Information Export.* Technical Report. https://doi.org/10.17487/rfc5470 RFC 5470.

[5] Joos C. A. M. Buijs, Boudewijn F. van Dongen, and Wil M. P. van der Aalst. 2012. On the Role of Fitness, Precision, Generalization and Simplicity in Process Discovery. In *Proceedings of the On the Move to Meaningful Internet Systems (OTM 2012)* (2012), Robert Meersman, Hervé Panetto, Tharam Dillon, Stefanie Rinderle-Ma, Peter Dadam, Xiaofang Zhou, Siani Pearson, Alois Ferscha, Sonia Bergamaschi, and Isabel F. Cruz (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 305–322. https://doi.org/10.1007/978-3-642-33606-5_19

[6] Javier Bustos-Jiménez, Cecilia Saint-Pierre, and Alvaro Graves. 2014. Applying Process Mining Techniques to DNS Traces Analysis. In *Proceedings of the 33rd International Conference of the Chilean Computer Science Society, (SCCC 2014).* IEEE Computer Society, 12–16. https://doi.org/10.1109/SCCC.2014.9

[7] Alvaro A. Cárdenas, Saurabh Amin, Zong-Syun Lin, Yu-Lun Huang, Chi-Yen Huang, and Shankar Sastry. 2011. Attacks against process control systems: risk assessment, detection, and response. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security (ASIACCS 2011)* (2011-03), Bruce S. N. Cheung, Lucas Chi Kwong Hui, Ravi S. Sandhu, and Duncan S. Wong (Eds.). ACM, 355–366. https://doi.org/10.1145/1966913.1966959

[8] Marco Caselli, Emmanuele Zambon, Johanna Amann, Robin Sommer, and Frank Kargl. 2016. Specification Mining for Intrusion Detection in Networked Control Systems. In *Proceedings of the 25th USENIX Security Symposium (USENIX Security 2016),* Thorsten Holz and Stefan Savage (Eds.). USENIX Association, 791–806.

[9] Marco Caselli, Emmanuele Zambon, and Frank Kargl. 2015. Sequence-aware Intrusion Detection in Industrial Control Systems. In *Proceedings of the 1st ACM Workshop on Cyber-Physical System Security (CPSS 2015),* Jianying Zhou and Douglas Jones (Eds.). ACM, 13–24. https://doi.org/10.1145/2732198.2732200

[10] Valentina Casola, Alessandra De Benedictis, Antonio Riccio, Diego Rivera, Wissam Mallouli, and Edgardo Montes de Oca. 2019. A security monitoring system for internet of things. *Internet of Things* 7 (2019), 100080. https://doi.org/10.1016/j.iot.2019.100080

[11] Thomas M. Chen and Saeed Abu-Nimeh. 2011. Lessons from Stuxnet. *Computer* 44, 4 (2011), 91–93. https://doi.org/10.1109/MC.2011.115

[12] Xutong Chen, Hassaan Irshad, Yan Chen, Ashish Gehani, and Vinod Yegneswaran. 2021. CLARION: Sound and Clear Provenance Tracking for Microservice Deployments. In *Proceedings of the 30th USENIX Security Symposium (USENIX Security 2021),* Michael Bailey and Rachel Greenstadt (Eds.). USENIX Association, 3989–4006.

[13] Anton Cherepanov. 2017. WIN32/INDUSTROYER: A new threat for industrial control systems. *White paper, ESET (June 2017)* (2017).

[14] Steven Cheung, Bruno Dutertre, Martin Fong, Ulf Lindqvist, Keith Skinner, and Alfonso Valdes. 2007. Using model-based intrusion detection for SCADA networks. In *Proceedings of the SCADA security scientific symposium*, Vol. 46. SRI International, 1–12.

[15] Ronny Chevalier, Maugan Villatel, David Plaquin, and Guillaume Hiet. 2017. Coprocessor-based Behavior Monitoring: Application to the Detection of Attacks Against the System Management Mode. In *Proceedings of the 33rd Annual Computer Security Applications Conference (ACSAC 2017)* (2017-12). ACM, 399–411. https://doi.org/10.1145/3134600.3134622

[16] Justyna J. Chromik, Anne Remke, and Boudewijn R. Haverkort. 2016. What's under the hood? Improving SCADA security with process awareness. In *Proceedings of the 2016 Joint Workshop on Cyber- Physical Security and Resilience in Smart Grids (CPSR-SG 2016).* IEEE, 1–6. https://doi.org/10.1109/CPSRSG.2016.7684100

[17] Justyna J Chromik, Anne Remke, and Boudewijn R Haverkort. 2018. Bro in SCADA: Dynamic intrusion detection policies based on a system model. In *Proceedings of the 5th International Symposium for ICS & SCADA Cyber Security Research* (2018-08). BCS Learning & Development, 112–121. https://doi.org/10.14236/ewic/ics2018.13

[18] Ege Ciklabakkal, Ataberk Donmez, Mert Erdemir, Emre Süren, Mert Kaan Yilmaz, and Pelin Angin. 2019. ARTEMIS: An Intrusion Detection System for MQTT Attacks in Internet of Things. In *Proceedings of the 38th Symposium on Reliable Distributed Systems (SRDS 2019)* (2019-10). IEEE, 369–371. https://doi.org/10.1109/SRDS47363.2019.00053

[19] B. Claise, P. Aitken, and N. Ben-Dvora. 2004. *Cisco Systems NetFlow Services Export Version 9.* Technical Report. https://doi.org/10.17487/rfc6759 RFC 3954.

[20] Simone Coltellese, Fabrizio Maria Maggi, Andrea Marrella, Luca Massarelli, and Leonardo Querzoni. 2019. Triage of IoT Attacks Through Process Mining. In *Proceedings of the On the Move to Meaningful Internet Systems (OTM 2019)* (2019) *(Lecture Notes in Computer Science, Vol. 11877).* Springer, 326–344. https://doi.org/10.1007/978-3-030-33246-4_22

[21] Richard Coppen. 2019. *MQTT Version 5.0 Specification.* Technical Specification. Organization for the Advancement of Structured Information Standards (OASIS). https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html

[22] Pubali Datta, Isaac Polinsky, Muhammad Adil Inam, Adam Bates, and William Enck. 2022. ALASTOR: Reconstructing the Provenance of Serverless Intrusions. In *Proceedings of the 31st USENIX Security Symposium (USENIX Security 2022)* (2023), Kevin R. B. Butler and Kurt Thomas (Eds.). USENIX Association, 2443–2460.

[23] Hervé Debar, Marc Dacier, and Andreas Wespi. 1999. Towards a taxonomy of intrusion-detection systems. *Computer Networks* 31, 8 (1999), 805–822. https://doi.org/10.1016/S1389-1286(98)00017-6

[24] Eclipse Foundation. 2021. Eclipse Newsletter - February 2021. https://www.eclipse.org/community/eclipse_newsletter/2021/february/1.php. Accessed: November 5, 2023.

[25] Gal Engelberg, Moshe Hadad, and Pnina Soffer. 2021. From Network Traffic Data to Business Activities: A Process Mining Driven Conceptualization. In *Proceedings of the 22nd International Conference on Business Process Modeling, Development and Support (BPMDS 2021)* (2021) *(Lecture Notes in Business Information Processing, Vol. 421)*, Adriano Augusto, Asif Gill, Selmin Nurcan, Iris Reinhartz-Berger, Rainer Schmidt, and Jelena Zdravkovic (Eds.). Springer, 3–18. https://doi.org/10.1007/978-3-030-79186-5_1

[26] Robert Flosbach, Justyna Joanna Chromik, and Anne Remke. 2019. Architecture and Prototype Implementation for Process-Aware Intrusion Detection in Electrical Grids. In *Proceedings of the 38th Symposium on Reliable Distributed Systems (SRDS)* (2019-10). IEEE, 42–51. https://doi.org/10.1109/SRDS47363.2019.00015

[27] Igor Nai Fovino, Andrea Carcano, Thibault De Lacheze Murel, Alberto Trombetta, and Marcelo Masera. 2010. Modbus/DNP3 State-Based Intrusion Detection System. In *Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications (AINA 2010).* IEEE Computer Society, 729–736. https://doi.org/10.1109/AINA.2010.86

[28] Gemalto. 2019. Gemalto: State of IoT Security. *Network Security* 2019, 2 (2019), 4. https://doi.org/10.1016/S1353-4858(19)30018-2

[29] Niv Goldenberg and Avishai Wool. 2013. Accurate modeling of Modbus/TCP for intrusion detection in SCADA systems. *International Journal of Critical Infrastructure Protection* 6, 2 (2013), 63–75. https://doi.org/10.1016/j.ijcip.2013.05.001

[30] Dina Hadziosmanovic, Damiano Bolzoni, and Pieter H. Hartel. 2012. A log mining approach for process monitoring in SCADA. *International Journal of Information Security* 11, 4 (2012), 231–251. https://doi.org/10.1007/s10207-012-0163-8

[31] Dongqi Han, Zhiliang Wang, Wenqi Chen, Ying Zhong, Su Wang, Han Zhang, Jiahai Yang, Xingang Shi, and Xia Yin. 2021. DeepAID: Interpreting and Improving Deep Learning-based Anomaly Detection in Security Applications. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security (CCS 2021),* Yongdae Kim, Jong Kim, Giovanni Vigna, and Elaine Shi (Eds.). ACM, 3197–3217. https://doi.org/10.1145/3460120.3484589

[32] Wajih Ul Hassan, Mark Lemay, Nuraini Aguse, Adam Bates, and Thomas Moyer. 2018. Towards Scalable Cluster Auditing through Grammatical Inference over Provenance Graphs. In *Proceedings of the 25th Annual Network and Distributed System Security Symposium, (NDSS 2018)* (2018). The Internet Society. https://doi.org/10.14722/ndss.2018.23141

[33] Wajih Ul Hassan, Mohammad A. Noureddine, Pubali Datta, and Adam Bates. 2020. OmegaLog: High-Fidelity Attack Investigation via Transparent Multilayer Log Analysis. In *Proceedings of the 27th Annual Network and Distributed System Security Symposium (NDSS 2020)* (2020). The Internet Society. https://doi.org/10.14722/ndss.2020.24270

[34] Khari Hernandez. 2022. For 1 in 4 companies, half of all AI projects fail. https://venturebeat.com/ai/idc-for-1-in-4-companies-half-of-all-ai-projects-fail/

[35] Tim Hübener, Michel R. V. Chaudron, Yaping Luo, Pieter Vallen, Jonck van der Kogel, and Tom Liefheid. 2022. Automatic Anti-Pattern Detection in Microservice Architectures Based on Distributed Tracing. In *Proceedings of the 44th IEEE/ACM International Conference on Software Engineering: Software Engineering in Practice, (ICSE 2022).* IEEE, 75–76. https://doi.org/10.1109/ICSE-SEIP55303.2022.9794000

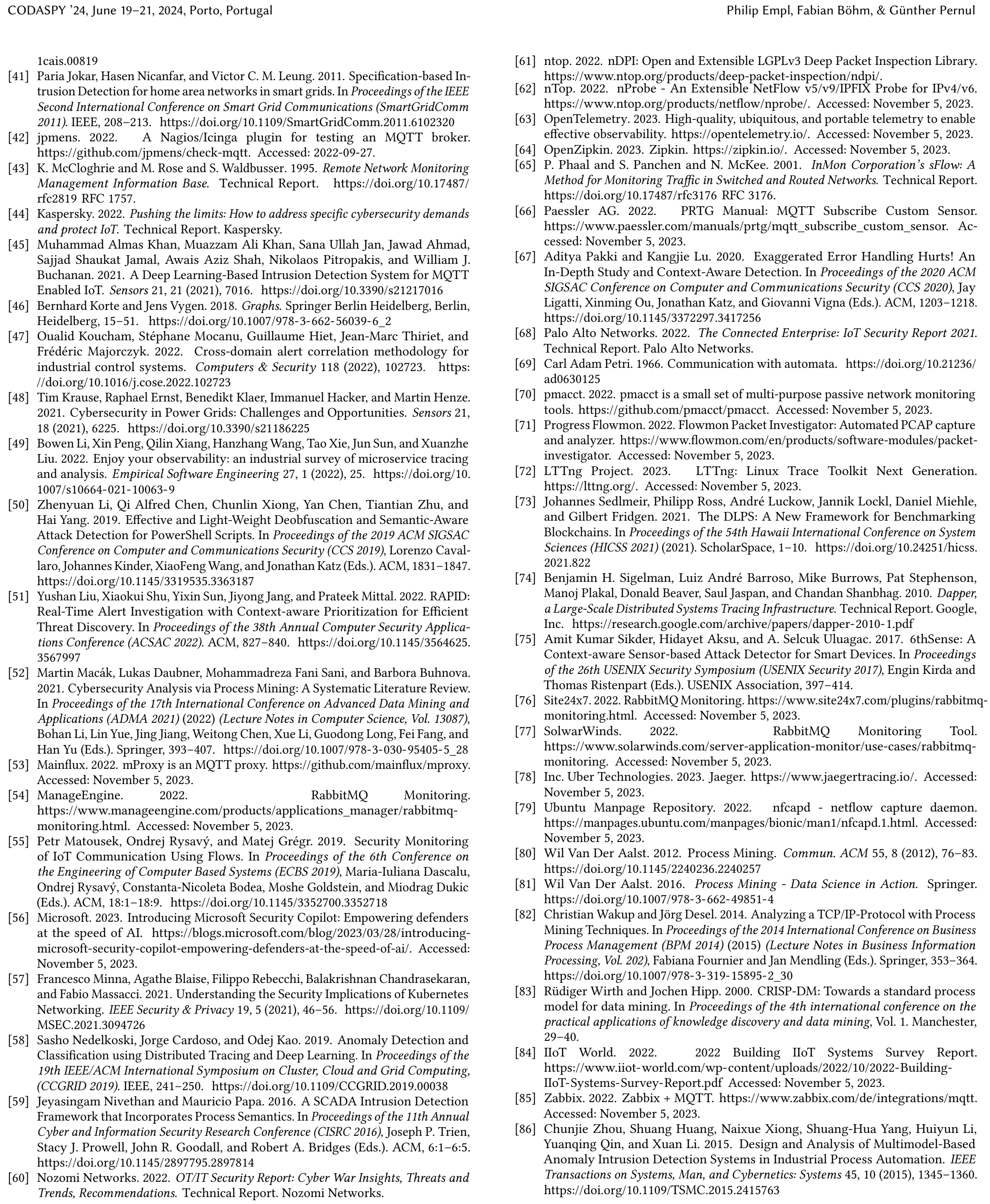
[36] InfluxData. 2022. Telegraf: MQTT Consumer Input Plugin. https://www.flowmon.com/en/products/software-modules/packet-investigator. Accessed: November 5, 2023.

[37] irino. 2022. softflowd: A flow-based network traffic analyser capable of Cisco NetFlow data export software. https://github.com/irino/softflowd. Accessed: November 5, 2023.

[38] ISO/IEC JTC 1/SC 29/WG 11. 2016. *Information technology – Message Queuing Telemetry Transport (MQTT).* International Standard ISO/IEC 20922:2016. Geneva, Switzerland. https://www.iso.org/obp/ui/#iso:std:iso-iec:20922:ed-1:v1:en

[39] J. Case and K. McCloghrie and M. Rose and S. Waldbusser. 1990. *Simple Network Management Protocol (SNMP).* Technical Report. https://doi.org/10.17487/rfc1448 RFC 1157.

[40] Joyce Jackson. 2002. Data Mining; A Conceptual Overview. *Communications of the Association for Information Systems* 8 (2002), 19. https://doi.org/10.17705/



1cais.00819

[41] Paria Jokar, Hasen Nicanfar, and Victor C. M. Leung. 2011. Specification-based Intrusion Detection for home area networks in smart grids. In *Proceedings of the IEEE Second International Conference on Smart Grid Communications (SmartGridComm 2011).* IEEE, 208–213. https://doi.org/10.1109/SmartGridComm.2011.6102320

[42] jpmens. 2022. A Nagios/Icinga plugin for testing an MQTT broker. https://github.com/jpmens/check-mqtt. Accessed: 2022-09-27.

[43] K. McCloghrie and M. Rose and S. Waldbusser. 1995. *Remote Network Monitoring Management Information Base.* Technical Report. https://doi.org/10.17487/rfc2819 RFC 1757.

[44] Kaspersky. 2022. *Pushing the limits: How to address specific cybersecurity demands and protect IoT.* Technical Report. Kaspersky.

[45] Muhammad Almas Khan, Muazzam Ali Khan, Sana Ullah Jan, Jawad Ahmad, Sajjad Shaukat Jamal, Awais Aziz Shah, Nikolaos Pitropakis, and William J. Buchanan. 2021. A Deep Learning-Based Intrusion Detection System for MQTT Enabled IoT. *Sensors* 21, 21 (2021), 7016. https://doi.org/10.3390/s21217016

[46] Bernhard Korte and Jens Vygen. 2018. *Graphs.* Springer Berlin Heidelberg, Berlin, Heidelberg, 15–51. https://doi.org/10.1007/978-3-662-56039-6_2

[47] Oualid Koucham, Stéphane Mocanu, Guillaume Hiet, Jean-Marc Thiriet, and Frédéric Majorczyk. 2022. Cross-domain alert correlation methodology for industrial control systems. *Computers & Security* 118 (2022), 102723. https://doi.org/10.1016/j.cose.2022.102723

[48] Tim Krause, Raphael Ernst, Benedikt Klaer, Immanuel Hacker, and Martin Henze. 2021. Cybersecurity in Power Grids: Challenges and Opportunities. *Sensors* 21, 18 (2021), 6225. https://doi.org/10.3390/s21186225

[49] Bowen Li, Xin Peng, Qilin Xiang, Hanzhang Wang, Tao Xie, Jun Sun, and Xuanzhe Liu. 2022. Enjoy your observability: an industrial survey of microservice tracing and analysis. *Empirical Software Engineering* 27, 1 (2022), 25. https://doi.org/10.1007/s10664-021-10063-9

[50] Zhenyuan Li, Qi Alfred Chen, Chunlin Xiong, Yan Chen, Tiantian Zhu, and Hai Yang. 2019. Effective and Light-Weight Deobfuscation and Semantic-Aware Attack Detection for PowerShell Scripts. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS 2019)*, Lorenzo Cavallaro, Johannes Kinder, XiaoFeng Wang, and Jonathan Katz (Eds.). ACM, 1831–1847. https://doi.org/10.1145/3319535.3363187

[51] Yushan Liu, Xiaokui Shu, Yixin Sun, Jiyong Jang, and Prateek Mittal. 2022. RAPID: Real-Time Alert Investigation with Context-aware Prioritization for Efficient Threat Discovery. In *Proceedings of the 38th Annual Computer Security Applications Conference (ACSAC 2022).* ACM, 827–840. https://doi.org/10.1145/3564625.3567997

[52] Martin Macák, Lukas Daubner, Mohammadreza Fani Sani, and Barbora Buhnova. 2021. Cybersecurity Analysis via Process Mining: A Systematic Literature Review. In *Proceedings of the 17th International Conference on Advanced Data Mining and Applications (ADMA 2021)* (2022) *(Lecture Notes in Computer Science, Vol. 13087)*, Bohan Li, Lin Yue, Jing Jiang, Weitong Chen, Xue Li, Guodong Long, Fei Fang, and Han Yu (Eds.). Springer, 393–407. https://doi.org/10.1007/978-3-030-95405-5_28

[53] Mainflux. 2022. mProxy is an MQTT proxy. https://github.com/mainflux/mproxy. Accessed: November 5, 2023.

[54] ManageEngine. 2022. RabbitMQ Monitoring. https://www.manageengine.com/products/applications_manager/rabbitmq-monitoring.html. Accessed: November 5, 2023.

[55] Petr Matousek, Ondrej Rysavý, and Matej Grégr. 2019. Security Monitoring of IoT Communication Using Flows. In *Proceedings of the 6th Conference on the Engineering of Computer Based Systems (ECBS 2019)*, Maria-Iuliana Dascalu, Ondrej Rysavý, Constanta-Nicoleta Bodea, Moshe Goldstein, and Miodrag Dukic (Eds.). ACM, 18:1–18:9. https://doi.org/10.1145/3352700.3352718

[56] Microsoft. 2023. Introducing Microsoft Security Copilot: Empowering defenders at the speed of AI. https://blogs.microsoft.com/blog/2023/03/28/introducing-microsoft-security-copilot-empowering-defenders-at-the-speed-of-ai/. Accessed: November 5, 2023.

[57] Francesco Minna, Agathe Blaise, Filippo Rebecchi, Balakrishnan Chandrasekaran, and Fabio Massacci. 2021. Understanding the Security Implications of Kubernetes Networking. *IEEE Security & Privacy* 19, 5 (2021), 46–56. https://doi.org/10.1109/MSEC.2021.3094726

[58] Sasho Nedelkoski, Jorge Cardoso, and Odej Kao. 2019. Anomaly Detection and Classification using Distributed Tracing and Deep Learning. In *Proceedings of the 19th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, (CCGRID 2019).* IEEE, 241–250. https://doi.org/10.1109/CCGRID.2019.00038

[59] Jeyasingam Nivethan and Mauricio Papa. 2016. A SCADA Intrusion Detection Framework that Incorporates Process Semantics. In *Proceedings of the 11th Annual Cyber and Information Security Research Conference (CISRC 2016)*, Joseph P. Trien, Stacy J. Prowell, John R. Goodall, and Robert A. Bridges (Eds.). ACM, 6:1–6:5. https://doi.org/10.1145/2897795.2897814

[60] Nozomi Networks. 2022. *OT/IT Security Report: Cyber War Insights, Threats and Trends, Recommendations.* Technical Report. Nozomi Networks.

[61] ntop. 2022. nDPI: Open and Extensible LGPLv3 Deep Packet Inspection Library. https://www.ntop.org/products/deep-packet-inspection/ndpi/.

[62] nTop. 2022. nProbe - An Extensible NetFlow v5/v9/IPFIX Probe for IPv4/v6. https://www.ntop.org/products/netflow/nprobe/. Accessed: November 5, 2023.

[63] OpenTelemetry. 2023. High-quality, ubiquitous, and portable telemetry to enable effective observability. https://opentelemetry.io/. Accessed: November 5, 2023.

[64] OpenZipkin. 2023. Zipkin. https://zipkin.io/. Accessed: November 5, 2023.

[65] P. Phaal and S. Panchen and N. McKee. 2001. *InMon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks.* Technical Report. https://doi.org/10.17487/rfc3176 RFC 3176.

[66] Paessler AG. 2022. PRTG Manual: MQTT Subscribe Custom Sensor. https://www.paessler.com/manuals/prtg/mqtt_subscribe_custom_sensor. Accessed: November 5, 2023.

[67] Aditya Pakki and Kangjie Lu. 2020. Exaggerated Error Handling Hurts! An In-Depth Study and Context-Aware Detection. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS 2020)*, Jay Ligatti, Xinming Ou, Jonathan Katz, and Giovanni Vigna (Eds.). ACM, 1203–1218. https://doi.org/10.1145/3372297.3417256

[68] Palo Alto Networks. 2022. *The Connected Enterprise: IoT Security Report 2021.* Technical Report. Palo Alto Networks.

[69] Carl Adam Petri. 1966. Communication with automata. https://doi.org/10.21236/ad0630125

[70] pmacct. 2022. pmacct is a small set of multi-purpose passive network monitoring tools. https://github.com/pmacct/pmacct. Accessed: November 5, 2023.

[71] Progress Flowmon. 2022. Flowmon Packet Investigator: Automated PCAP capture and analyzer. https://www.flowmon.com/en/products/software-modules/packet-investigator. Accessed: November 5, 2023.

[72] LTTng Project. 2023. LTTng: Linux Trace Toolkit Next Generation. https://lttng.org/. Accessed: November 5, 2023.

[73] Johannes Sedlmeir, Philipp Ross, André Luckow, Jannik Lockl, Daniel Miehle, and Gilbert Fridgen. 2021. The DLPS: A New Framework for Benchmarking Blockchains. In *Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS 2021)* (2021). ScholarSpace, 1–10. https://doi.org/10.24251/hicss.2021.822

[74] Benjamin H. Sigelman, Luiz André Barroso, Mike Burrows, Pat Stephenson, Manoj Plakal, Donald Beaver, Saul Jaspan, and Chandan Shanbhag. 2010. *Dapper, a Large-Scale Distributed Systems Tracing Infrastructure.* Technical Report. Google, Inc. https://research.google.com/archive/papers/dapper-2010-1.pdf

[75] Amit Kumar Sikder, Hidayet Aksu, and A. Selcuk Uluagac. 2017. 6thSense: A Context-aware Sensor-based Attack Detector for Smart Devices. In *Proceedings of the 26th USENIX Security Symposium (USENIX Security 2017)*, Engin Kirda and Thomas Ristenpart (Eds.). USENIX Association, 397–414.

[76] Site24x7. 2022. RabbitMQ Monitoring. https://www.site24x7.com/plugins/rabbitmq-monitoring.html. Accessed: November 5, 2023.

[77] SolwarWinds. 2022. RabbitMQ Monitoring Tool. https://www.solarwinds.com/server-application-monitor/use-cases/rabbitmq-monitoring. Accessed: November 5, 2023.

[78] Inc. Uber Technologies. 2023. Jaeger. https://www.jaegertracing.io/. Accessed: November 5, 2023.

[79] Ubuntu Manpage Repository. 2022. nfcapd - netflow capture daemon. https://manpages.ubuntu.com/manpages/bionic/man1/nfcapd.1.html. Accessed: November 5, 2023.

[80] Wil Van Der Aalst. 2012. Process Mining. *Commun. ACM* 55, 8 (2012), 76–83. https://doi.org/10.1145/2240236.2240257

[81] Wil Van Der Aalst. 2016. *Process Mining - Data Science in Action.* Springer. https://doi.org/10.1007/978-3-662-49851-4

[82] Christian Wakup and Jörg Desel. 2014. Analyzing a TCP/IP-Protocol with Process Mining Techniques. In *Proceedings of the 2014 International Conference on Business Process Management (BPM 2014)* (2015) *(Lecture Notes in Business Information Processing, Vol. 202)*, Fabiana Fournier and Jan Mendling (Eds.). Springer, 353–364. https://doi.org/10.1007/978-3-319-15895-2_30

[83] Rüdiger Wirth and Jochen Hipp. 2000. CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, Vol. 1. Manchester, 29–40.

[84] IIoT World. 2022. 2022 Building IIoT Systems Survey Report. https://www.iiot-world.com/wp-content/uploads/2022/10/2022-Building-IIoT-Systems-Survey-Report.pdf Accessed: November 5, 2023.

[85] Zabbix. 2022. Zabbix + MQTT. https://www.zabbix.com/de/integrations/mqtt. Accessed: November 5, 2023.

[86] Chunjie Zhou, Shuang Huang, Naixue Xiong, Shuang-Hua Yang, Huiyun Li, Yuanqing Qin, and Xuan Li. 2015. Design and Analysis of Multimodel-Based Anomaly Intrusion Detection Systems in Industrial Process Automation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 45, 10 (2015), 1345–1360. https://doi.org/10.1109/TSMC.2015.2415763

P5 Do You Play It by the Books? A Study on Incident Response Playbooks and Influencing Factors

Status:	Published
Date of Submission:	13 April 2023
Date of Acceptance:	10 July 2023
Date of Publication:	20 October 2023
Conference:	45th IEEE Symposium on Security and Privacy (S&P 2024)
Location:	Hilton Hotel, San Francisco, CA, United States of America
Period:	20.05.2024 – 22.05.2024

Authors’ Contributions:

Daniel Schlette	30%
Philip Empl	30%
Marco Caselli	20%
Thomas Schreck	10%
Günther Pernul	10%

Full Citation: Schlette, D., Empl, P., Caselli, M., Schreck, T., & Pernul, G. (2024). Do you Play It by the Books? A Study on Incident Response Playbooks and Influencing Factors. In *Proceedings of the 45th IEEE Symposium on Security and Privacy* (pp. 59:1–59:19). IEEE Computer Society.

DOI: 10.1109/SP54263.2024.00060

Artifact: github.com/luduslibrum/awesome-playbooks

Conference Description: Since 1980, the IEEE Symposium on Security and Privacy has been the premier forum for presenting developments in computer security and electronic privacy, and for bringing together researchers and practitioners in the field. The 2024 Symposium will mark the 45th annual meeting of this flagship conference.

Do You Play It by the Books? A Study on Incident Response Playbooks and Influencing Factors

Daniel Schlette*, Philip Empl*, Marco Caselli†, Thomas Schreck‡, Günther Pernul*
**University of Regensburg, {firstname.lastname}@ur.de*
†*Siemens AG, marco.caselli@siemens.de*
‡*HM Munich University of Applied Sciences, thomas.schreck@hm.de*

Abstract—**Incident response "playbooks" are structured sets of operational procedures organizations use to instruct humans or machines on performing countermeasures against cybersecurity threats. These playbooks generally combine information about a given threat and organizational aspects relevant within the context of an organization. Both types of information are crucial for using, maintaining, and sharing playbooks across organizations as they ensure effectiveness and confidentiality. While practitioners show great interest in playbooks, their characteristics have not yet been thoroughly investigated from a research perspective. For this reason, we explore the topic by analyzing what is inside a playbook. Our approach consists of a comprehensive empirical assessment of available data (1217 playbooks), an online study with 147 participants, and final in-depth interviews with nine security professionals to consolidate and validate our findings. We notably find intrinsic ambiguities in the way practitioners and organizations define their playbooks. Furthermore, we notice that available playbooks cannot be used outright which might currently impair their wide use across different cybersecurity actors. As a result, we can conclude that organizations do "play it by the books" but individually define what is inside their playbooks and which areas of incident response they might address.**

1. Introduction

Organizations facing cybersecurity threats should define and document standard operating procedures to ensure consistent cybersecurity operations and incident response [1], [2], [3], [4]. The shortage of skilled cybersecurity analysts [5], [6] draws additional attention to (automated) processes that can increase efficiency by reducing and eliminating errors. Integrating security tools, data, and teams is another pain point identified by cybersecurity practitioners [7], [8]. Dedicated processes can help organizations use resources effectively, particularly host and network tools, internal and external threat intelligence, and multiple security teams.

Playbooks represent processes and procedures which are part of every organization. In cybersecurity, more specifically in incident response and security operations, playbooks make for a novel field of research. While these playbooks address a specific threat or incident, they complement other well-established areas, such as system hardening [9], [10], vulnerability handling [11], [12], [13], [14], and business process management [15], [16]. System hardening, being mainly proactive, is based on checklists to fulfill generic security requirements (i.e., confidentiality, integrity, availability) during deployment. Vulnerability handling covers proactive scanning and security advisories where process representation (i.e., lists of remediations) is incomplete. Business process management and playbook-based IT operations (e.g., with Ansible [17]) show overlap but intentionally lack a clear cybersecurity focus. However, incident response playbooks address reactive scenarios, specifying what to do when things go wrong [18], [19].

With threats on the rise [20], there is a move toward playbooks as they promise consistency and automation. This development can be seen most clearly in the proliferation of Security Orchestration, Automation, and Response (SOAR) platforms [21], [22]. Playbooks are at the core of SOAR platforms that streamline security operations. Beyond vendors, standardization efforts and industry groups document the interest in playbooks and automation. The OASIS CACAO technical committee and the FIRST Automation special interest group aim to advance the representation of playbooks and explore user perspectives [23], [24]. As a result, playbooks are widely recognized, praised by professionals and playbook data has started to become available.

Nevertheless, open challenges around playbooks remain. Above all, we need to tackle the understanding of playbooks. Only a few previous works have discussed playbooks [18], [25], [26], [27]. Therefore, our work explores the overarching question: *What is inside a playbook?*

Initially, we observe the absence of an established theoretical playbook foundation. From an academic perspective, we suppose that playbooks are based on generic, technical information and contain additional organization-specific information. *Community playbooks*, describing generic operations and addressing existing threats, are stripped of any specific organizational context. They can be shared across different organizations as they do not convey confidential information. We expect to find these community playbooks analyzing repositories of different SOAR vendors.

When used within organizations, playbooks encompass additional organizational aspects defined by context. Making playbooks fit the context is about considering *influencing factors* and complementing community playbooks with organization-specific information. An influencing factor (e.g., EU GDPR) may induce changes and makes a community playbook operational. We expect organizations that define incident response processes and use playbooks to point out influencing factors to some extent. While incident response playbooks are a novel research field, our theoretical considerations are relevant for organizations using, maintaining, and sharing playbooks.

Against this backdrop, we aim to validate our playbook hypothesis by exploring the following research questions:

RQ1: *What are characteristics of community playbooks made available by trusted sources?*

RQ2: *Which influencing factors shape incident response processes and organization-specific playbooks?*

To answer these questions we gather, preprocess, and analyze playbook data from leading SOAR vendors (e.g., Splunk), cybersecurity institutions (e.g., CISA), and open-source platforms (e.g., Shuffle). In our three-step approach, we further perform an online study ($n = 147$) and conduct interviews ($n = 9$) to deepen the discussion on playbook content with specific attention to influencing factors. In particular, this paper makes the following contributions:

- We introduce a theoretical foundation to the playbook concept. Considering community playbooks and influencing factors specific to an organization is crucial when using, maintaining, and sharing playbooks.
- We are the first to empirically analyze 1217 playbooks from 14 sources. Our analysis reveals that community playbooks are modular and contain a significant amount of information about tools and ticketing which are likely handled differently across organizations.
- We find that incident response processes and playbooks are indeed shaped by influencing factors but this is not necessarily recognized by practitioners. Our online study shows that the number of steps and the workflow inside a playbook varies. Interview participants mention specific implications of influencing factors.
- We report that while playbooks are well-known, understanding and use depend on implementation tools and abstraction levels.

The remainder of this paper is structured as follows. In Section 2, we outline playbook foundations. We describe our methodology and introduce a research model for influencing factors in Section 3. We report on the characteristics of community playbooks in Section 4. We present results form our online study in Section 5. Interview findings are reported in Section 6. We discuss the implications in Section 7. Section 8 presents related work before we conclude our work in Section 9.

2. Playbook Foundations

We define the playbook concept and show basic forms of playbooks representation. Influencing factors must be considered when sharing, maintaining, and using playbooks in organizations as they transform community playbooks.

TABLE 1. TEXTUAL, GRAPHICAL, OR CODE-BASED ELEMENTS AND NOTATIONS CAN BE USED FOR PLAYBOOK REPRESENTATION.

Type	Elements	Notation
Textual	Checklist ☰, Table ⊞	Markdown
Graphical	Diagram ⧉	BPMN
Code	File 🗋, Script </>	JSON, Python

2.1. Playbook definition

Despite the growing interest, the notion of playbooks still needs to be clarified as other terms such as script, automation, runbook, checklist, standard operating procedure, workflow, or process are frequently used. Playbooks have been defined in a few other academic works [18], [28], [29], [30], [31], by cloud providers, and cybersecurity vendors. Most notable, according to the US President's Executive Order 14028, playbooks are "a standard set of operational procedures to be used in planning and conducting cybersecurity activity" [2]. Besides, standardization efforts (e.g., OASIS CACAO [23]) provide reliable information on playbooks and their components (see Appendix C for a detailed comparison of terminology and concepts).

In absence of an established incident response playbook definition, three characteristics guide our work: 1) *Security context.* Playbooks are used for cybersecurity purposes. 2) *Process representation.* Playbooks describe instructions to be followed. 3) *Technology integration.* Playbooks assist humans with technology and data. Upon these characteristics, a playbook to us describes a specific cybersecurity process or procedure based on a workflow with individual steps. Identified by [27], workflow steps are essentially actuator-action-artifact triplets describing who (e.g., firewall) performs an action (e.g., block) on an object (e.g., IP address). A playbook addresses a more or less specific threat or incident. We explicitly acknowledge different levels of abstraction found in playbooks.

2.2. Playbook representation

As playbooks represent processes and procedures, they logically structure information. Workflow steps can be structured sequentially, one after the other, or in parallel if the tasks and resources permit. Conditions can define the subsequent workflow steps. A text-based community phishing playbook might consist of the following steps:

- **Start:** Phishing email received
  - □ Create ticket and assign incident handler
  - □ Analyze email header, content and attachments
  - □ Correlate threat intelligence (DNS logs, IPs, hashes)
  - □ Search and remove (unread) phishing emails
  - □ Update protection systems and notify users
- **End:** Incident closed

Who the playbook is for – humans or machines – determines the decision for either textual, graphical, or code-based representation. The phishing playbook above offers a generic human-readable description. Information is presented in text form with a checklist. Other playbook representations include graphical or code-based elements with different notations (see Table 1 and Appendix A, B). In the case of low-code/no-code SOAR platforms, human playbook designers construct graphical playbooks enriched with text and commands. The tools then generate and save a code-based representation for automation.

2.3. Playbook sharing, maintenance, and use

Inspired by collaborative cybersecurity and Cyber Threat Intelligence (CTI) sharing [29], [32], [33], we assume a basic setting as illustrated in Figure 1. Organization A (producer) shares community playbooks with Organization B (consumer). We expect that any organization-specific information is excluded as organizations value confidentiality. If Organization A is a SOAR vendor, playbooks are kept generic to promote widespread adoption of its platform.

On the receiving end, Organization B intends to use the playbooks for automation, onboarding, or other purposes. The playbooks' instructions contain essential elements (e.g., actions) that apply across organizations. However, the playbooks lack contextual elements of incident response specific to Organization B. For instance, Organization B might decide to sinkhole DNS traffic instead of blocking via firewall. Also, Organization B might only have a few incident handlers or outsourced services constraining in-depth malware analyses. Therefore, Organization B and other playbook consumers must transform playbooks to fit their context. The knowledge of influencing factors is a prerequisite for transforming and using community playbooks.

A push toward collaboration and actionable advice amplifies the relevance of playbook sharing. Industry groups (e.g., Financial Services Information Sharing and Analysis Center, FS-ISAC [34]) and (national) coordination teams can provide their constituents with playbooks. Sharing playbooks will help a small organization build incident response processes from the ground up. A large conglomerate can use playbooks to improve existing processes. Sharing can also aid in establishing a language to discuss incident response.

Influencing factors are also crucial for playbook maintenance within an organization. As organizations evolve, the organizational context changes. Consequently, knowing about influencing factors and how they manifest helps to keep organization-specific playbooks effective. For example, an organization might extend its business operations to another country with a different jurisdiction. As a result, changes in influencing factors (e.g., laws and regulations) require the transformation of existing playbooks by adding further reporting steps. The same applies when technology is discarded as threats and playbooks become obsolete.

2.4. Playbook transformation

Using and maintaining playbooks relies on two separate inputs: community playbooks and influencing factors. Via transformation, these inputs lead to organization-specific playbooks. Inputs might also be modified over time:

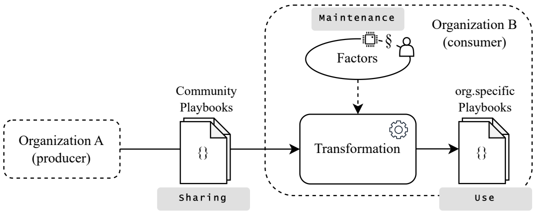


Figure 1. Sharing playbooks across organizations, maintenance, and use of playbooks put focus on influencing factors.

- *Modified community playbooks* reflect a changing threat landscape where new threats emerge. Community playbooks cover the various threats that demand countermeasures.
- *Modified influencing factors* reflect new organizational characteristics and external requirements. Influencing factors have organization-specific values and implications.

Our research, builds on these foundations and the study of "what is inside a playbook" focuses on analyzing community playbooks before examining how these are shaped by influencing factors.

3. Methodology

We follow a mixed-methods approach combining quantitative and qualitative research to thoroughly investigate playbooks. Playbook data is analyzed to answer the first research question on community playbooks. With our online study and qualitative interviews we aim to answer the second research question on influencing factors. Data collection for community playbooks, online study, and interviews was carried out between May 2022 and February 2023.

3.1. Playbook data

From an initial 1347 (18 sources), we empirically analyze 1217 community playbooks (14 sources) for their characteristics. After collecting the data, we opted to automatically preprocess and analyze the playbooks (leading to the exclusion of non-processable ones), making the playbooks and analysis transparent[1].

Data collection. We initially tried to collect playbooks from all major SOAR vendors based on two popular market reports [21], [22]. Thus, we acquired access through GitHub repositories, engaged in a demonstration, or implemented the software through local deployment. However, we faced different obstacles. In particular, playbooks from several vendors were not publicly accessible, our requests remained

[1] https://github.com/luduslibrum/awesome-playbooks

unanswered, or vendors demanded non-disclosure agreements. To prevent any potential negative impact upon these vendors, explicit mention of their identities is deliberately withheld. To the best of our knowledge, we are the first to collect and compile a comprehensive playbook repository.

Data preprocessing. Using Python, we built custom connectors for each vendor to create schemata and extract relevant data from the heterogeneous playbooks. In addition, we used Jupyter notebooks to review and label 8,623 machine-generated actuator-action-artifact triplets. By extracting triplets, we can categorize each playbook step.

Data analysis. Jupyter notebooks also describe our data analysis. We use NLP stemming, tokenization, and dependency parsing to extract the actuator-action-artifacts triplets based on the step name. We tested different clustering methods (e.g., k-Means, based on text vectorization and PCA) to identify step categories and also tried topic modeling methods. However, the methods remained unsuccessful mainly because the clusters or topics were ambiguous. We noticed that step names are not distinctive, e.g., "investigate" or "block IP address" are two different topics, and observed that vendors use different vocabularies to describe their steps. This leads us to manually label 370 unique actions to identify the categories.

3.2. Research model and rationale

Extending the scope of our research to organization-specific playbooks, we set up a research model (see Figure 2). The model divides influencing factors in two groups – external and internal factors – and is used for our online study and interviews. Influenced by factors, incident response processes and playbooks are characterized by workflow logic and the content of individual steps. For groups and factors, we borrow from information security policy research [35], [36] and business process literature [37], [38].

We presented our idea of influencing factors at a major industry conference for incident response teams. When asked via interactive poll, all participants agreed that factors do influence incident response. However, answers started to vary when pressed, which is the most important factor. The feedback we received points to legal aspects, people, and technology. While organizations must determine if they can perform incident response processes as envisioned, they are also influenced by staff and available systems.

External factors. We assume attacker characteristics, including motivation, behavior, location, and others, to affect processes. Industry standards and guidelines, such as frameworks, can constitute an external factor as they serve as process references. In addition, laws and regulations require organizations to conduct incident response and implement compliant processes. These legal aspects capture, for example, business structure, location, and sector. We further assume that business partners and other organizations within the supply chain formulate expectations influencing organizational processes. Factor → impact examples are:

- Sophisticated attacker → long-term monitoring
- ISO/IEC 2700x → evidence collection
- EU GDPR → informing supervisory authorities
- Outsourced services → contact external team

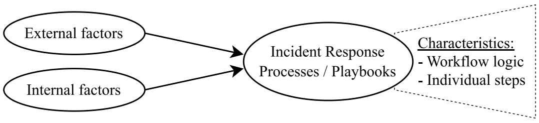


Figure 2. Theoretical research model with two factor groups influencing incident response processes and playbook use.

Internal factors. Organizations can choose to abide by arbitrary internal rules or directives. Such internal factors can center on incident response data operations and attack targets. Besides, people form another internal factor which encompasses security culture and various security team-related elements (e.g., team size, chain of command, and responsibilities). The technology, general and cybersecurity specific, within an organization provides a setting in which incident response processes are placed. Thus, IT infrastructure and security tools might also explain variations in incident response processes. Factor → impact examples are:

- CRUD constraints → no data copying
- CTI team → CTI sharing
- Security automation tools → querying endpoint agents

Observable differences. Discussing influencing factors directs attention to their observable influence on processes. We identify three possible changes caused by influencing factors – the first two center on workflow logic and the last on step semantics. We anticipate that incident response processes and playbooks can differ across organizations in the *number of steps*. Influencing factors might lead to more or fewer process steps changing workflow logic. Besides, we account for *sequential or parallel workflow*. Factors can introduce conditions to guide process flow, and process steps might be aligned sequentially or in parallel. Last, influencing factors can cause the actuator-action-artifact triplet to change, while the number of steps remains the same (*altered steps*).

3.3. Online study

The primary objective of the online study is to assess the relevance of influencing factors within organizations. Therefore, we collected data from 147 participants reaching a broad audience within incident response. One of our co-authors with experience in online research methodology, provided expert guidance throughout our iterative process. To assess the validity and reliability of our questionnaire, we administered it to a group of colleagues in a controlled setting. In our questionnaire (available at our GitHub repository[2]) we first provide contextual cues pertaining to the authors' affiliations and the scientific nature of the study, while also emphasizing the voluntarily participation. Furthermore, participants are instructed to consider their responses within

[2] https://github.com/luduslibrum/awesome-playbooks/tree/main/factors

the context of their respective organizations. Among other background questions, we directly ask for influencing factors (in three incident scenarios) and beliefs thereof. We have extra questions exclusively for participants using playbooks which $\approx 85\%$ do.

Participant recruitment and data. We recruit participants using our professional networks. We ask for participation via email, social media (i.e., LinkedIn), and industry groups (i.e., FIRST). Participants currently employed in cybersecurity are requested seven minutes of their time to partake. As it is difficult to get participants' time, we provide a monetary lottery incentive (100€) with a 1:147 chance if they participate within the next 15 days. Facing unsolicited responses ($\approx 90\%$), we completely exclude any observations which contain invalid (i.e., non-English), redundant, and implausible answers. Exclusion is mainly based on free text fields (e.g., What is a playbook to you?) indicating missing security context (e.g., stage play) and numerical values. Further, redundant observations within a few-minute threshold containing only partial variations are excluded. In order to prevent the participation of bots, we included a math CAPTCHA as a security measure. We used the statistical software Stata to analyze the data.

3.4. Interviews

We interviewed security professionals working on incident response topics to gain deeper insights into influencing factors. While data saturation is arguable, we ceased the interview process after nine interviews as the results showed only marginal additions. The interviews were semi-structured to allow for flexibility and open-ended questions. Participants were not made aware of our research questions and instead asked about the effect of influencing factors in specific "what-if" scenarios. We opted for questions on incident response processes and procedures to remain unbiased of playbook notions. However, we actively asked additional questions whether and how participants use playbooks.

Interview guide. Following a pretest (with two colleagues working in incident response 5+ years), the interview questions were adapted and reviewed by all authors. Our interview guide consists of four parts: 1) *Demographics.* We ask about a participant's role and organizational characteristics. 2) *Incident response basics.* We ask about incident response within a participant's organization. 3) *Incident response scenarios.* We outline three incidents (Ransomware, DDoS, APT) and ask about a participant's organizational processes and implications of specific influencing factors. 4) *Playbooks.* We ask about playbook understanding and use.

Interview procedure. We conducted the interviews using cognitive interviewing methodology [39]. Interviews were held in German or English virtually via Zoom (invites send by email), except for one in-person meeting. At the outset of the interviews, we introduced the topic, emphasizing the significance and relevance of the interview, and highlighting the research methods to collect and evaluate the data. Participants consented that notes were taken during the interviews. We further asked the participants if we could use their answers in anonymous or aggregated form for scientific publication, to which participants agreed. The interviews lasted from one hour to one hour and thirty minutes.

Participant recruitment. Participants working on incident response topics were selected based on personal contacts and the authors' professional networks reflecting industry experience. Most interview participants work for large (> 5000 employees), multi-national corporations headquartered in Germany. Asked about how they would rate the maturity level for cybersecurity in their organization, most participants mentioned a high maturity compared to peers.

Coding and analysis. Interview questions and transcribed interview data are associated with influencing factors during data coding by two coders. We use thematic analysis [40], [41], [42] and corresponding deductive/inductive coding to relate questions and answers to themes (see Appendix E). Our themes are based on the research model with its influencing factors (Section 3.2). Conflicts (e.g., supply chain and laws) are resolved with re-reading and team discussions.

3.5. Ethics

Cybersecurity incidents and organizational processes are delicate topics. We validated our research approach with our institutions' ethics committees via email, explaining our research and data handling. We got confirmed that there is no legal obligation to obtain an ethics vote as very limited personal data is handled. While we were met with great openness, we aim to assure the anonymity of our study participants. Our questionnaire consisted of 43 questions, of which we left 37 open-ended. Participants gave their consent and had the option to opt-out at any point during the process. We deliberately refrained from participant monitoring (e.g., drop-outs, completion time) for privacy reasons. Notification was only provided to the winner of the lottery incentive, via email, while the rest remained unnoticed. We contacted potential interviewees twice, via email, after an initial contact attempt. We collected only a limited amount of demographic information, report mostly aggregated data, and in order to minimize errors, we allowed each interview participant to check and correct their quotes. We did not require non-disclosure agreements from any of our participants.

3.6. Limitations

While our study provides valuable insights, it is important to acknowledge its limitations, particularly with regard to the *playbook data*: The playbook data we analyze is mostly code-based, includes only a selection of vendors mentioned in two SOAR market reports, and is not representative of all platforms. This limits the scope of our analysis. Given the rapid pace of technological advancements and market consolidation, the community playbooks collected could become outdated fast. Another challenge we faced was the poor data quality in the community playbooks. The actuator-action-artifact triplets often contain

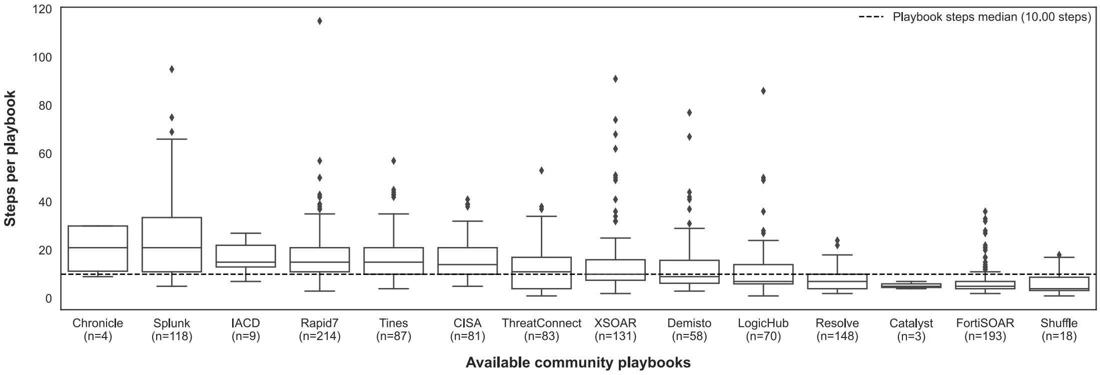


Figure 3. Exploring the characteristics of 1217 playbooks based on the number of steps.

typos, programming syntax, or incomplete data, making it almost impossible to analyze the data using clustering and topic modeling techniques. Although the step categories are highly specific, their contents can overlap due to blurry task boundaries (e.g., append or update data), The situation requires manual labeling of triplets, which can be prone to errors. We encountered ambiguity in labeling steps, often requiring us to assign them to multiple categories, e.g, the action "check" could be part of an investigation but also a logical gateway. In the *user study*, there could be potential biases and undetected duplicates related to the financial motivations of the participants, which may have influenced their responses. This may limit the reliability of our findings. Lastly, for the *interviews*, we faced challenges in identifying and accessing incident response professionals, making it difficult to obtain a representative sample.

4. Community Playbook Analysis

This section targets RQ1 and investigates the structure and content of community playbooks (see Figure 3 and Table 2), thereby partly confirming existing theories [18], [27]. We present our findings of community playbook characteristics at playbook and workflow step level. Moreover, we categorize steps and show multi-category playbooks.

4.1. Playbook level

Playbooks exhibit high structural homogeneity across vendors, with variances predominantly found in their implementation-specific nuances. Below we summarize the components of an incident response playbook:

Meta information Descriptive elements, including an identifier, name, description, and categorization tags, serve to identify, retrieve, and understand a playbook. Playbooks also comprise specific details on attributes such as visibility, ownership, inheritance, or versioning.

Workflow At the core of a playbook lies the workflow. The workflow denotes a security process's systematic and logical sequencing, characterized by a discrete set of individual steps discussed in more detail in Section 4.2.

Parameter Playbooks receive input and produce output. Inputs, outputs and other environmental parameters (i.e., playbook variables) can be defined globally at the playbook or workflow step levels.

Features As features, we define all playbook components that were used only occasionally. Such features encompass deployment specifications, playbook validation, testing procedures, process visualization, sharing policies, return on investment, or priority handling.

Community playbooks are mostly code-based and predominantly represented by JSON, YAML, and XML, with some sources opting for alternative tabular or graphical approaches in Business Process Modeling Notation (BPMN) [43]. There is a discernible inclination towards modularity to support community playbooks' reusability. For example, some vendors rely on playbook hierarchies to allow "playbooks calling playbooks" and their unrestricted reusability across individual use cases. These use cases leverage the logical linking of community playbooks. We find malware playbooks confirming this modularity. While one malware playbook is solely responsible for email notifications, the other handles malware hunting, and containment. We also find playbooks that include and test vendors' platform integrations, e.g., testing the connection to Microsoft Teams.

4.2. Workflow step level

The workflow establishes a coherent sequence of steps. Each step comprises meta-information, including an identifier, a name, and a description. Moreover, a workflow step encapsulates the underlying logic, which guides the orchestration of the workflow. Notably, every workflow step contains an actuator-action-artifact triplet.

TABLE 2. Workflow step categories, their frequencies and illustrative actuators, actions, and artifacts.

Category	Frequency (abs./rel.)	Actuator	Action	Artifact
Logic	3,770 (*22.5%*)	Platform, human, API	trigger, loop, decide	Playbook, while, condition
Utility	3,834 (*22.8%*)	Platform, JMESPath, JoinArray	extract, filter, format	Status, data, report
Ticketing	2,344 (*13.9%*)	Microsoft Teams, Slack, Jira	send, document, close	Email, ticket, case
Investigation	5,943 (*35.3%*)	Human, VirusTotal, HashIt	link, search, lookup	Indicator (IP, domain/URL, file/hash)
Countermeasure	930 (*5.5%*)	AD, firewall, endpoint protection	block, quarantine, reset	Indicator, user, endpoint

Workflows differ in their complexity and scope. Figure 3 compares community playbooks, indicating a median of 10 steps and a mean of 13.78 steps per playbook. We calculated the step size based on the respective array length of the workflow. However, the number of steps varies depending on the source, with 25% of the playbooks possessing less than six or more than 17 steps. Notably, a crucial aspect of SOAR is the automation of workflows, which is reflected in the degree of automation of the workflow steps. On average, 96.93% of the steps in the workflows are automated, whereas the remaining 3.07% of the steps necessitate human interaction. The rationale for human interaction stems from the workflow step type involved.

Workflow steps vary in their type. Based on 16,821 workflow steps, we identify seven types: start/end, single, decisions, loop, trigger, information, and playbook, confirming the CACAO [23] notion except for a parallel workflow step type. The start and end step types demarcate the scope of the playbook. More than half of the workflow steps are single, encompassing security operations, including HTTP requests, filtering, and transformation tasks. Workflow steps orient on well-known process gateways, such as decisions or loops. Decisions are made automatically via conditions or manually via ChatOps, prompts, and workflow interactions. Loops either wait for conditions to finish or conduct actions for each element in a list or array. Another type of workflow step comprises triggers that await a specific event. Informational workflow steps are also pertinent for informing humans, as they provide details about the current status or display relevant data. Lastly, certain workflow step types call other playbooks.

Accurately distinguishing between step types is challenging as some sources have integrated the process flow directly into each step. This means each step knows the predecessors and subsequent steps, including loops and conditions. Contrarily, other sources define the process flow globally in the community playbook, related to a node-link diagram. These varying approaches have far-reaching implications for the playbook's readability, length, number of steps, and step types. Another interesting aspect concerns how vendors integrate applications into their SOAR platform, including native integrations with their custom adapters or direct API calls to the application. Last but not least, from the total of 16,821 workflow steps, we identified 8,623 (51.3%) unique workflow steps indicating the reuse of steps within playbooks.

4.3. Categorization

We manually label unique workflow steps using the actuator, action, and artifact information and assign a specific category to a workflow step. We sometimes assign multiple step categories to a particular workflow step due to ambiguities. In these cases, a workflow step is related to multiple categories. Our findings on step categories are summarized in Table 2 and explained below.

Logic includes all the elements that play a crucial role in organizing and maintaining the process flow. The platform and the human are the primary actuators responsible for this category. For example, the security analyst decides the subsequent process path(s), calls additional playbooks, or initiates further workflow steps.

Utility encompasses all operations that are ancillary to the workflow's logic but crucial for supporting the other step categories. For instance, the platform utilizes regular expressions to extract data from a report.

Ticketing or alerting refers to the incident or case management to initiate a case and inform relevant stakeholders. A workflow step within this category may subsequently be reapplied to update a ticket.

Investigation deals with the search and analysis of relevant threat information, i.e., tactics, techniques, and procedures (TTP). A case in point would be a security analyst scrutinizing a phishing email, identifying an unfamiliar domain, examining its geographic location, and classifying it as an indicator of compromise.

Countermeasures go beyond analysis and take measures to counteract a threat through remediation, containment, and recovery. For instance, the security analyst mentioned in the previous example would block the IP address or domain of the phishing email.

Multi-category playbooks. As playbooks often comprise multiple step categories (e.g., logic steps are commonly integrated into workflows), it is imperative to investigate the composition of playbooks based on their step categories.

We present the composition of multi-category playbooks and show the likelihood of encountering a particular step category within a playbook of n categories (see Appendix D Figure 4). To arrive at these findings, we analyzed playbooks containing at least two categories, accounting for 91.4% of the total playbooks. On average, playbooks mainly comprise logic and utility steps, with varying degrees of alerting/ticketing, investigation, and countermeasure steps. Besides, 7.2% of the playbooks feature all step categories, whereby the likelihood of encountering investigation steps

TABLE 3. Online study participants and their beliefs on influencing factors

Continent (n = 144)			
EMEA	84 (58.3%)	Asia	13 (9.1%)
America	47 (32.6%)		
Sector (n = 141)			
IT	74 (52.5%)	Industrials	13 (9.2%)
Public	21 (14.9%)	Consumer	7 (5%)
Financials	16 (11.3%)	Other	10 (7.1%)
Role (n = 147)			
Sen. sec. manager	58 (39.5%)	IT operations	11 (7.5%)
Sen. sec. expert	20 (13.4%)	Sec. researcher	11 (7.5%)
Sec. operations	15 (10.2%)	Sec. consultant	11 (7.5%)
Incident handler	15 (10.2%)	Other	6 (4.1%)
Influencing factors (n = 142)			
Technology	112 (78.9%)	Laws & regul.	74 (52.1%)
IR directives	98 (69%)	Attacker charact.	71 (50%)
People	94 (66.2%)	Supply chain	46 (32.4%)
Industry stand.	79 (55.6%)		

within these playbooks is 34%. However, most playbooks have three different categories and the majority of these does not include countermeasure steps instead focusing on investigation (e.g., sighting inactive users or reviewing indicator reputations).

Additionally, we investigate whether playbooks consisting of five categories are intentionally designed for a specific incident type (e.g., malware or phishing). Upon analysis of 82 playbooks, we often find varying purposes, but also playbooks aligned to incident types. We derive that if a countermeasure step is included, there are likely ticketing and investigation steps in advance. This observation may account for the increased likelihood of countermeasures when more playbook categories are involved. Across all 1217 playbooks, we found 50 phishing (4.1%) and 27 malware (2.2%) playbooks.

4.4. Implications for playbook adoption

Our results show that community playbooks are predominantly generic but contain organization-specific aspects. Workflow steps cover handling alerts and ticketing, which likely vary by organization. Meeting compliance requirements by adhering to service level agreements (SLAs) is present in community playbooks but must be determined in an organizational context. Further, community playbooks include people, technology, and best practices, but their adoption and use can vary significantly from organization to organization. For example, an organization may rely on a single analyst managing tickets or use divergent technologies, such as firewalls or communication tools, not integrated into vendor offerings. Ultimately, community playbooks are designed to match the vendors' product portfolios and focus on the vendors' customers as best as possible, not necessarily addressing the broad incident response community. Therefore, tailoring community playbooks to the organization's context is crucial to cater to the diverse needs (e.g., beyond automation) of different organizations. In the following sections, we will delve into the organizational context to better understand the factors that influence the transformation of community playbooks into organization-specific playbooks.

TABLE 4. Online study participants and their organizational playbooks.

	Mean (Std.)	Median	n
Employees	35,756 (120,871)	600	145
Experience [years]	9.8 (6.5)	8	143
Security teams	2.9 (1.79)	3	142
IR team size	20.5 (36.2)	7	136
Team maturity [0-4]	2.7 (1.1)	3	145
Process maturity [0-4]	2.7 (0.9)	3	144
Tech. maturity [0-4]	2.7 (1.0)	3	144
Playbook contribution [0-4]	2.5 (1.25)	3	124
Org. playbooks	24.5 (46.8)	9	125
Malware steps	10.2 (7.9)	7	108
Phishing steps	8.4 (7.1)	6	102
Account comp. steps	8.6 (6.7)	9	100

5. Online Study Results

In this section, we address RQ2 and employ an online study to first investigate the relevance and impact of the influencing factors responsible for the transformation of community playbooks. We report the findings of our online study, which included a sample of $n = 147$ participants. The data is presented in Table 3 and Table 4 for better clarity. First, we provide an overview of the participants' demographics and their respective organizations, including incident response capabilities. Subsequently, we investigate the prevalence of playbook sharing, use, and maintenance in the surveyed organizations. Additionally, we explore the differences between malware, phishing, and account compromise playbooks within these organizations, with participants indicating the number of steps required for each playbook. Finally, we detail how the participants perceive the significance of influencing factors.

Playbook understanding. The study primarily consists of participants from Europe, the Middle East, and Africa (EMEA). Most participants are associated with the information technology (IT) sector, occupying senior positions like security managers, including SOC and CTI managers. On average, the surveyed organizations have three security teams in place, namely, incident response (CERT/CSIRT), threat intelligence, and security operations center (SOC). An incident response team, on average, comprises seven employees. All participants report that their organizations possess mature people, processes, and technologies for incident response. The term "playbook" is predominantly used by most participants, while some mention related terms like "runbook/workflow" (an automated playbook), "stories" (use cases), "standard operating procedure", or "incident management framework". This indicates a partial variance in the participants' perception of playbooks. The majority of respondents mention an active involvement in the playbook design. On average, an organization has 24.5 playbooks, with a high standard deviation indicating stark differences.

Playbook use. According to the participants, the primary reasons for using playbooks are documentation (79.7%), automation (52.8%), compliance (51.2%), and onboarding (44.7%). The majority of playbooks are text-based (76.7%), followed by graphical (46.8%) or code-based (38.7%), which highlights the significance of documentation. In terms of operational use cases, playbooks are primarily utilized for investigation/analysis (83.6%), countermeasures/mitigation (68.9%), and alerting/ticketing (60.7%). About 52.5% of respondents use playbook hierarchies also seen in community playbooks. Regarding the relevance of the aforementioned use cases, 40% of respondents indicate using external playbooks, all of which have been modified for organizational context. Moreover, 71.1% of respondents share playbooks internally, while 13.2% also share them externally.

Playbook differences. We analyze the differences between organization-specific incident-type processes or playbooks. Therefore, we asked the participants about their malware, phishing, and account compromise workflow characteristics: step count, parallel steps, and involved influencing factors. On average, 67% of respondents report that their incident-type playbooks include parallel steps. Conversely, this indicates that these playbooks can be logically distinguished from other playbooks. Regarding the number of steps, we observe a high standard deviation (± seven steps) in all kinds of playbooks. Malware playbooks contain the most steps on average compared to the other two, while the account compromise playbook has the highest median. All in all, we see a varying number of steps in the individual playbooks, which we attribute to the influence of factors.

Influencing factors vs. beliefs. Specifically, we asked participants to identify factors affecting incident response processes and whether their incident-type playbooks include steps that reflect these factors. The participants believe that technology is the most crucial factor, followed by incident response directives and people. This implies that technological advancements within an organization are seen to have an impact on incident response processes and playbooks. Incident response directives and the security team(s) also play a crucial role. In the case of malware incidents, technology remains the most significant factor, followed by incident response directives and attacker characteristics. Similarly, phishing playbooks are influenced most by technology, incident response directives, and the security team. Last, in the case of account compromise, attacker characteristics are stated as the most relevant factor, followed by technology and team. In all three scenarios, it is noticeable that the characteristics of the attacker are important, although the participants believe that they generally play a relatively minor role. This means that organizations strive to identify the attacker's motivation, particularly regarding account compromise. In general, we observe that different influencing factors shape incident response processes and playbooks.

To sum up, our observations indicate that organizations have specific playbooks indicated by step count, parallel workflow and factors. They commonly use them to document organizational security processes. Playbooks are also frequently shared and modified. We can not derive which factor influences playbooks the most as it strongly depends on the use case. However, technology, incident response directives, people, and attacker characteristics seem to be prevalent. It remains unanswered which influence a factor exerts on community playbooks. Therefore, our interviews with incident response professionals go into detail.

TABLE 5. Interview Participants are characterized by industry sector, position, and ID.

Sector	Position	ID
Information Technology	Tech / Mgmt	P3, P6, P7
Industrials / Materials / Energy	Tech / Mgmt	P2, P4, P8, P9
Public Institution	Tech	P1, P5

$^{\text{Tech}}$ technical positions $^{\text{Mgmt}}$ management positions

6. Interview Findings

This section validates and complements Section 5 with in-depth discussions of influencing factors through interviews. We describe for each external and internal factor the (missing) influence on incident response processes and playbooks. Table 5 presents details about the participants and their ID. Influencing factors and selected statements made by interview participants are listed in Table 6. The findings below depict aggregated data from the interviews.

Processes and playbooks are organization-specific. We had interview participants elaborate on their organization's incident response processes and playbooks. They offered detailed insights into what they do and what others might not be doing (e.g., investigations, CTI sharing, or coordination). Mostly, processes and playbooks are generic and aligned to incident types (e.g., phishing, malware, or account/system compromise) or recurring events. We can report that all interview participants were aware of playbooks or used playbooks as *"they can reduce time spent on tasks, improve task quality, and allow rookie teams to have stable operations." (P8)* Whether processes and playbooks are organization-specific, participant P3 sums up a common stance: *"The incident response team creates playbooks [...]. [Community playbooks] never fit the organizational context and are either too generic or too specific for certain products." (P3)* Combined with other participants' statements this leads us to conclude that organizational context and factors matter. In community playbooks, we find domain (e.g., ICS), vulnerability (e.g., 8×Log4J and 5×WannaCry), and generic playbooks.

Attacker behavior and motivation beat location. Behind every attack or incident, there is a threat actor. Organizations fending off attacks and coping with incidents take attacker characteristics into account. However, interview participants paint a nuanced picture when asked about the influence of more specific attacker characteristics, such as presumed location, behavior, and motivation, on their processes. Unintuitively, precise attribution, whether attacks originate from a specific group and country, is considered

TABLE 6. INFLUENCING FACTORS AND STATEMENTS MADE BY INTERVIEWED PROFESSIONALS.

	Influencing Factor	Participant statement
External	Attacker characteristics	“We investigate for longer when we think there is more to know [in the case of an APT].” (P4)
	Industry standards	“We started with ISO 27035 and developed a [high-level] process according to our organizational needs.” (P8)
	Laws and regulations	“Playbooks and actions are pre-approved by legal. We keep lawyers in the loop [during incident response].” (P7)
	Business structure	“We have a general [workflow] step to check if authorities need to be informed.” (P4)
	Location	“We perform EU GDPR assessment and inform authorities.” (P2)
	Sector	“Only incident handlers in the US are allowed to handle incidents affecting US military contracts.” (P3)
	Supply chain	“We collaborate with an ISP [for DoS attacks].” (P1)
Internal	Incident response directives	“Playbooks go into detail. A step refers to ’investigate headers’ and which specific tools to be used.” (P7)
	Data operations	“We manually extend data retention adapting the log settings [in the case of an incident].” (P1)
	Targets	“We have a crisis mode, a dedicated crisis team, and retainers with [...] forensic companies.” (P3)
	People	“We do shift operations, on-call duty, and follow the sun for incident response.” (P9)
	Security culture	“Permissions and pre-authorizations are important. At times you have to act without [explicit] permission.” (P2)
	Security team	“We diligently share information, insert data in MISP, and correlate incidents.” (P4)
	Technology	“Business units or third parties operate our infrastructure. We advise and coordinate measures.” (P6)
	IT infrastructure	“If necessary, we integrate third parties (AWS, Microsoft) into our process. They are notified by email.” (P9)
	Security tools	“Without a [SIEM/SOAR] dashboard, we rely on [administration] tools to detect and respond to incidents.” (P5)

less relevant as *“clarifying attack origin is clearly not the duty of private corporations but should be delegated to law enforcement.” (P2)* This is contrasted by 16 “geolocate IP address” steps in community playbooks.

As explained by participants’ narrow understanding of attribution as solely identifying the attacker’s location, *“it is all about technical matters rather than where the attack originated from.” (P6)* Consequently, the focus is shifted to the influence of attacker behavior and motivations. Beyond prioritization, security teams escalate to management and conduct additional threat hunting in an Advanced Persistent Threat (APT) scenario. Community playbooks also prioritize alerts, incidents, and users, but do not address APTs. Different motivations (e.g., espionage, financial gain, or reputation) are reflected in decisions to observe or eliminate threats: *“We investigate for longer when we think there is more to know [in the case of an APT].” (P4)*

Industry standards provide broad guidance only. Organizations build on industry standards and guidelines for their high-level incident response process. Overarching frameworks (e.g., NIST Cybersecurity Framework, Incident Response Life Cycle, or FIRST CSIRT framework) are often adapted: *“We started with ISO 27035 and developed a [high-level] process according to our organizational needs.” (P8)* This finding is replicated by community playbooks containing 45 steps aligned to the MITRE ATT&CK framework and CISA’s playbooks being primarily based on NIST SP 800-61. On a more detailed level addressed by playbooks and automation scripts, we cannot infer any influence of industry standards and guidelines from the interviews.

Laws and regulations apply throughout. A central finding in our validation was the relevance of legal aspects for incident response. In our interviews, participant P7 summarized the influence of laws and regulations on processes at different levels: *“Playbooks and actions are pre-approved by [our] legal [department]. We keep lawyers in the loop [during incident response].” (P7)* We conclude that any initial process design and playbook transformation must consider legal influences. Defined roles and communication between security and legal teams can help streamline security processes, especially in unfamiliar situations.

When examining the influence of more specific legal aspects, we make the following observations:

- Business structure requires compliance, but the influence is vague: *“We have a general [workflow] step to check if authorities need to be informed.” (P4)*
- Location-based influence is mainly about privacy and data protection: *“We perform EU GDPR assessment and inform authorities.” (P2)*
- Sector-based influence, most notable in the defense industry, has implications on security operations: *“Only incident handlers in the US are allowed to handle incidents affecting US military contracts. We do reassign responsibilities accordingly.” (P3)*

Overall, there is a strong emphasis among participants that technical incident response is separate from legal processes. For instance, while we expected to hear about SEC 8-K filings for NYSE-listed corporations or regional equivalent regulations, our participants from large organizations instead delegated responsibilities. Nevertheless, participants indicated the relevance of organizational aspects in their processes: *“Our legal department gets [corresponding] tickets in their ticketing system when [a GDPR-related] incident happens.” (P7)* Beyond Europe, a legal influence was mentioned for the USA but not for other countries: *“If the US is concerned, we have separate procedures.” (P4)* In community playbooks, we do not find specific laws and regulations.

When in doubt, collaborate with business partners. Recent attacks and proposed legislation (e.g., EU Cyber Resilience Act) put focus on cybersecurity within the supply chain. However, we do not see a clear indication of the influence of business partners’ expectations on specific processes within the interview results. We note that most participants work for larger organizations which, whenever necessary, collaborate with other organizations to resolve incidents: *"We collaborate with an [Internet Service Provider,] ISP [for DoS attacks]." (P1)* Another explanation for missing influence is the overlap with other factors, such as laws and regulations or internal technology sourcing strategies. Community playbooks feature sharing indicators with partners via dedicated platforms (e.g., MISP), reporting malware samples, and notifying security vendors.

Besides external influencing factors, we evaluate internal factors shaping incident response processes and playbooks.

Directives address the finer points of incident response. Organizations have flexibility in how to handle incidents. Thus, individual workflow steps and actions can contain precise instructions permitting or constraining specific data operations (e.g., data copying) and addressing specific attack targets (e.g., devices of board members). Our interview participants mention SLAs that define incident response. For instance, when to start working on a case or report to management, but there is no definite time to close a case. On a technical level, tool selection is covered: *"Playbooks go into detail. A [workflow] step [within a phishing playbook] refers to 'investigate headers' and which specific tools to be used." (P7)* Beyond these observations, we find:

- When organizations have to deal with storage constraints they perform additional data operations: *"We manually extend data retention adapting the log settings [in the case of an incident]." (P1)* This is partially contrasted by large organizations that have nearly "unlimited" storage capacity: *"We aim for storing logs for a year. The more data, the better." (P4)* Nevertheless, costs are considered.
- Organizations adapt incident response when attack targets are business critical (e.g., systems or accounts). We note that these situations require a special process, are handled with priority, or are taken "off-process". As a result, precautions are taken: *"We have a crisis mode, a dedicated crisis team, and retainers with incident response and forensic companies." (P3)* Additionally, access to information might be restricted. Some participants object that incident response is always on a need-to-know basis, but they agree that different people (e.g., CISO) must be informed.

Directives found in community playbooks include, e.g., violating SLAs (188 steps) and notifying public relations.

More people imply more tasks and coordination. Depending on the organization, one or more security teams are involved in incident response. Most participants state that they have incident response and other teams (e.g., IT operations, red team, threat intelligence, human resources) work together. Communication and the chain-of-command are relevant prerequisites as organizations *"do shift operations, on-call duty, and follow the sun for incident response." (P9)* In favor of fast operations, organizations opt for specialization so tasks can be shared and executed in parallel. Security culture and teams influence incident response processes:

- Security culture is about management support. Influence on communication and permissions can be observed: *"Permissions and pre-authorizations are important. At times you have to act without [explicit] permission." (P2)* In general, participants mention good management support within their organization but those without face consequences (e.g., missing resources), which affect tasks. As a result, tasks cannot be performed, and security becomes a topic of personal motivation: *"Within a public institution, processes take time, and focus is on essential tasks." (P5)*
- In multi-national organizations when specialists and different teams work on incidents, additional tasks cover forensic investigations, dark web searches and CTI sharing: *"We diligently share information, insert data in MISP, and correlate incidents." (P4)* Among the interview participants we notice different operation modes such as double incident assignments, task rotation, or feedback loops and different team structures. While in one organization certain tasks are performed by the incident response team itself, in other organizations these are handled by CTI, SOC/monitoring, or other dedicated teams: *"We need to involve the email team [in the case of account compromise]." (P3)*

Inside community playbooks, we find analysts-in-the-loop (e.g., ChatOps) deciding about critical process branches, reviewing machine-generated output, and conducting complex tasks (e.g., investigate and remediate).

On the shoulders of technology. Technology is a broad category influencing incident response processes and playbooks in various ways. Consequences range from general applicability (i.e., a process or playbook is relevant to the organization) to which specific actions can be taken (e.g., query EDR agents on every endpoint). Overall, influence is dependent on existing technology and its characteristics (e.g., centralized vs. decentralized, homogeneous vs. heterogeneous). Some participants point out that they only advise and coordinate actions because they do not (directly) operate the IT infrastructure: *"Business units or third parties operate our infrastructure. We advise and coordinate measures." (P6)* Technology being used to fulfill business and security requirements has the following implications:

- IT infrastructure and sourcing strategies play an essential role. Nowadays, web services are typically hosted externally: *"If necessary, we will integrate third parties (AWS, Microsoft) into our process. They will be notified by email but not via incident tracking." (P9)* Participants further mention that organizations can realize more streamlined processes if they have all their services in the cloud using one cloud provider. In contrast, OT networks and production systems are more diverse, demanding different measures.
- Organizations strive for redundant and independent incident response infrastructure: *"Semi-self-sufficient operation is possible. We cover cases when no domain controller is available. The CERT should still be functioning and able to access their workstations." (P8)*

> This implies that processes cover switching to alternative infrastructure which includes communication channels (e.g., email, Mattermost, or war rooms). In addition, security tools shape what is done. Although most participants use commercial SOAR platforms, some have different solutions: *"Without a [SIEM/-SOAR] dashboard, we rely on [administration] tools to detect and respond to incidents." (P5)*

Technology drives community playbooks in communications (e.g., MS Teams or Slack), collaborations (e.g., war rooms), and security operations. SOAR platforms are primarily responsible for logic, utility, and ticketing while investigation and countermeasures rely on different tools.

7. Discussion

We discuss our research implications along the dimensions of community playbook *data quality*, influencing factors forming a *common language* to discuss different playbook perspectives, and the question of *technology assistance*. Our actionable insights emphasize how to cope with terminology, playbook content and handling playbooks.

7.1. How to talk about playbooks

Although organizations use playbooks, *what is a playbook to you?* must remain a key question due to ambiguous definitions. Toward a detailed understanding of playbooks offered by a given community or used within a given organizational context, our analysis recommends the following:

- Clarify playbook representation and implementation.
- Clarify abstraction levels and management instruments.

When it comes to community playbooks, it is best to pay attention whether or not they are code-based and linked to a SOAR platform. Emphasized by our community playbook analysis, these playbooks typically rely on additional plugins (e.g., Docker containers, APIs, or other software artifacts) offered by the SOAR platform integrating (external) services and functions. Therefore, community playbooks abstract technical measures to some extent and largely do not contain specific CLI commands or code snippets. Community playbooks need to be checked for technical content and overarching concepts (e.g., use cases) to grasp the meaning of the term playbook.

When it comes to organization-specific playbooks, it is necessary to account for other organizational management instruments. Thereby, policies, plans, checklists, and security tools can define different abstraction levels and guide processes. Broadly speaking, playbooks can tilt toward text-based instructions within a knowledge base (e.g., wiki) or toward dedicated automation scripts both being possibly referred to as playbooks. Thus, it is necessary to inquire management instruments and tools. Additionally, a separation in response and detection playbooks should be checked.

7.2. How to define playbook information

Playbooks should be precise and contain information to fulfill their intended purpose. Analyzing community playbooks and manually labeling workflow steps, we discover that basic information on who is performing an action on an object (i.e., actuator, action, artifact), as mentioned in [27], is difficult to extract, obstructing comparison and perhaps use. Thus, we see potential for standardization, including naming conventions. While we initially agreed with research findings in [18] that (executable) commands are required inside a playbook, this is opposite to how organizations see their playbooks and community playbooks look like. Instead, most organizations aim for abstract playbooks and introduce another tool-based implementation layer below.

Nevertheless, organizations show intrinsic ambiguities in the way they define their playbooks. Even without clear statistical significance, we believe that there is merit in discussing influencing factors on incident response processes and playbooks as they can offer a common language. As community playbooks cannot be used outright due to missing or obsolete information (e.g., logic, tools, ticketing), transformation according to organizational context and consideration of influencing factors is needed. Providing an initially stepping stone for theory building, we want to emphasize three insights.

First, there is no clear picture on influencing factors. In our online study, most participants believe technology exerts influence, and they have steps based on (security) technology. In our interviews, participants show a deep knowledge but are not fully aware of what shapes their playbooks, possibly explaining the absence of some factors. Despite multiple participants claiming that incident response is solely about technical matters, their responses indicate otherwise. In line with other works on security management [44], [45], we argue that interfaces to other teams and technology are the most important factors. Acknowledging challenges on boundaries, organizations should define the scope of incident response and its playbooks.

Second, as community playbooks partially contain organization-specific information, we hypothesize that organizations prefer adapting available playbooks to building them from scratch. In [18], playbook frameworks guide playbook design. Intending to make intuitive influencing factors visible, we showcase an additional path to playbook design, assisting organizations in adapting community playbooks. Consequently, this might reduce the time and resources required to build organization-specific playbooks while including all essential elements.

Third, what should be inside an organizational playbook must be squared with organizations' intend. Organizations might opt for creativity and critical thinking of their incident handlers, avoiding specifics. In contrast to behavioral research emphasizing bias and unwanted variability (noise) [46], [47], [48], we reason that some threats (APTs) demand flexibility and that more specific playbooks could introduce further challenges (e.g., prioritization, reusability, hierarchies). Nevertheless, playbooks can detail when and which decision to make, thus building a cornerstone to combat bias, ensure consistency, and speed up repetitive tasks.

7.3. How to use, maintain, and share playbooks

For organizations using playbooks, the next step involves automation, tool selection, and deployment. However, it is crucial to determine if automation does help or harm. Interviewees mention that security professionals must be kept in the loop when making decisions. Besides automation and seen in the online study, playbooks can also aid documentation, reporting, and onboarding thereby improving existing processes. Ideally, building organization-specific playbooks is based on systematically merging community playbooks and influencing factors. Here, it is on future research to address recent developments of generative artificial intelligence for incident response (e.g., Microsoft Security Copilot's promptbook), investigating options to build highly contextual playbooks automatically. For playbook maintenance, organizations need to recognize and keep track of changes. Influencing factors could be kept in a dedicated repository and frequently checked. With separated repositories, organizations can build their organization-specific playbooks on-the-fly coping with a growing number of threats and playbooks over time. Playbook sharing is the foundation toward establishing a common language for incident response. In few years, with even more playbook data available, causal effects and the meaning of data (i.e., a high/low number of steps) can be explored. We see challenges in diverse or new data formats and different sharing modes (e.g., internal recipients wish for commands) to be addressed in the future.

8. Related Work

The idea of structured guidance achieved with playbooks is present in adjacent research areas. We first mention related work on incident response playbooks before discussing system hardening, vulnerability handling, IT operations, and business processes on which playbooks are based.

8.1. Incident response playbooks

Incident response playbooks caught researchers' interest. Closest to our work, Stevens et al. [18] investigated playbook design pointing to organizational constraints. Schlette et al. [27] compared structured playbook representations building a prerequisite to playbook sharing. In [26] the removal of confidential information is briefly discussed. Similar to how playbooks provide strategic guidance in sports, their use in fighting DDoS attacks and establishing incident response programs centers on outlining different options [19], [25]. In cybersecurity, playbooks can be seen as a continuation of threat intelligence research aiming to present and disseminate actionable security information [32], [49], [50]. We find security standardization efforts encompassing playbooks and courses of action [29], [31], [51]. Aside from academic research, two GitHub repositories aggregate information on incident response [52], [53].

8.2. Structured guidance and systematic processes

Organizations perform system hardening to reduce their attack surface with securely configured systems [9], [10]. System hardening and security configuration rely on best practices and their technical implementation relates to playbooks. What must be done to conform with a given security baseline is defined by vendors (e.g., Microsoft security baselines [54]), community organizations (e.g., Center for Internet Security, CIS Benchmarks [55]), and governmental institutions (e.g., US Department of Defense Security Technical Implementation Guides, DoD STIGs [56]). Typically, as part of the Security Content Automation Protocol (SCAP), XCCDF-structured security checklists are used. They contain rules, fixes and can link to OVAL-based vulnerability checks. Implementation is based on tools (e.g., Ansible, Chef, Terraform), dedicated scripts (e.g., PowerShell, bash), and other configuration options (e.g., GPO backups). We notice similarities (e.g., abstraction, tools) and differences (e.g., proactive) compared to incident response playbooks.

Once vulnerabilities have been discovered and disclosed [14], organizational vulnerability handling involves preventing exploitation by following security advisories and patching systems [11], [12], [13], [57]. Typically, organizations are guided by CVRF/CSAF-structured security advisories describing how to fix a vulnerability [58]. We note that incident response playbooks focus on organizational processes and thus go beyond the rich stream of vulnerability research.

Extending the scope to non-security areas, IT operations, system administration, and configuration management cover systematic deployment, maintenance, and monitoring [59], [60]. Consequently, regular tasks are structured and automated with scripts or runbooks to ensure consistency [61]. Automation relies heavily on tools, with Ansible and its playbooks being a prominent example [17], [62]. In cloud environments, playbook-like concepts are part of *Infrastructure as Code* [63] but also remain security agnostic.

8.3. Business process orchestration

Process orchestration and automation are tightly coupled with SOAR platforms but are rooted in business process research [15], [64]. Business processes and process orchestration are necessary when there are multiple activities and diversity in people, processes, and technology [65], [66]. Business Process Management Systems (BPMS) define execution engines and perform fundamental tasks, such as modeling, simulating, instantiating, and monitoring workflows (i.e., automated business processes) [67], [68]. Besides, business process re-engineering relates to influencing factors and targets "order-of-magnitude improvements" [37], [38], [69]. Dedicated software solutions (e.g., Camunda, ServiceNow) implement execution engines and automation platforms for business processes. Comparing architectures and functionalities, we note that SOAR platforms and security orchestration largely borrow from business process research but instantiate concepts within the security context. A notable difference is the deep integration of security tools.

9. Conclusion

In brief, we investigated incident response playbooks and different factors that influence and shape their sharing, maintenance, and use. Actionable insights on data quality, common language, and technology assistance document necessary considerations for playbooks. Our initial belief that playbooks have a clear and concise definition was challenged as we discovered that individuals have different understandings of what constitutes a playbook. Besides, organizations intentionally keep playbook information abstract to support flexibility and reuse, which is partially reflected in community playbooks. Although our research does not statistically clarify the factors' relevance, it makes factors visible, guiding playbook design and use with a common language. We observe that organizations do play it by the books but adjust playbooks to their context. Our findings suggest that technology and the security team(s) are critical drivers in shaping incident response playbooks. For future work, we see opportunities to explore the systematic merging of community playbooks and influencing factors to build organization-specific playbooks. The emergence of generative artificial intelligence has already begun to reshape incident response processes, and we are keen to observe how playbooks and their use will continue to evolve in the future.

Acknowledgment

The authors would like to sincerely thank all interviewees and study participants for their invaluable insights and time spent, without which this research would not have been possible. This work was supported by the German Federal Ministry of Education and Research (BMBF) as part of the DEVISE project and by the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 830927.

References

[1] ISO/IEC, "ISO/IEC 27001:2022 - Information security, cybersecurity and privacy protection — Information security management systems — Requirements," International Organization for Standardization, Tech. Rep., 2022.

[2] Executive Office of the President, "Executive Order 14028 of May 12, 2021 – Improving the Nation's Cybersecurity," 2021, last accessed 2023-03-01. [Online]. Available: https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/

[3] P. Cichonski, T. Millar, T. Grance, K. Scarfone *et al.*, "Computer Security Incident Handling Guide," National Institute of Standards and Technology (NIST), Tech. Rep. SP 800-61, Rev2, 2012.

[4] M. J. West-Brown, D. Stikvoort, K.-P. Kossakowski, G. Killcrece, and R. Ruefle, "Handbook for Computer Security Incident Response Teams (CSIRTs)," Carnegie-Mellon Software Engineering Institute, Tech. Rep., 2003.

[5] (ISC)^2, "(ISC)^2 Cybersecurity Workforce Study 2022," (ISC)^2, Tech. Rep., 2022.

[6] J. Dykstra and C. L. Paul, "Cyber Operations Stress Survey (COSS): Studying fatigue, frustration, and cognitive workload in cybersecurity operations," in *11th USENIX Workshop on Cyber Security Experimentation and Test (CSET 18)*, 2018.

[7] R. Brown and P. Stirparo, "SANS 2022 Cyber Threat Intelligence Survey," SANS, Tech. Rep., 2022.

[8] C. Crowley and B. Filkins, "SANS 2022 SOC Survey," SANS, Tech. Rep., 2022.

[9] K. Scarfone, W. Jansen, and T. Miles, "Guide to General Server Security," National Institute of Standards and Technology (NIST), Tech. Rep. SP 800-123, 2008.

[10] J. Christensen, I. M. Anghel, R. Taglang, M. Chiroiu, and R. Sion, "DECAF: Automatic, Adaptive De-bloating and Hardening of COTS Firmware," in *29th USENIX Security Symposium, USENIX Security 2020, August 12-14, 2020*, 2020, pp. 1713–1730.

[11] S. de Smale, R. van Dijk, X. Bouwman, J. van der Ham, and M. van Eeten, "No One Drinks From the Firehose: How Organizations Filter and Prioritize Vulnerability Information," in *2023 IEEE Symposium on Security and Privacy, SP 2023, Proceedings, 22-25 May, 2023, San Francisco, California, USA*, 2023.

[12] N. Alomar, P. Wijesekera, E. Qiu, and S. Egelman, ""You've Got Your Nice List of Bugs, Now What?" Vulnerability Discovery and Management Processes in the Wild," in *Sixteenth Symposium on Usable Privacy and Security, SOUPS 2020, August 7-11, 2020*, 2020, pp. 319–339.

[13] K. A. Farris, A. Shah, G. Cybenko, R. Ganesan, and S. Jajodia, "VULCON: A System for Vulnerability Prioritization, Mitigation, and Management," *ACM Transactions on Privacy and Security (TOPS)*, vol. 21, no. 4, pp. 16:1–16:28, 2018.

[14] D. Votipka, R. Stevens, E. M. Redmiles, J. Hu, and M. L. Mazurek, "Hackers vs. Testers: A Comparison of Software Vulnerability Discovery Processes," in *2018 IEEE Symposium on Security and Privacy, SP 2018, Proceedings, 21-23 May 2018, San Francisco, California, USA*, 2018, pp. 374–391.

[15] M. Dumas, M. L. Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of Business Process Management*. Springer, 2013.

[16] K. Salimifard and M. Wright, "Petri net-based modelling of workflow systems: An overview," *European Journal of Operational Research*, vol. 134, no. 3, pp. 664–676, 2001.

[17] Red Hat, "Using Ansible playbooks," 2023, last accessed 2023-03-01. [Online]. Available: https://docs.ansible.com/ansible/latest/playbook_guide/index.html

[18] R. Stevens, D. Votipka, J. Dykstra, F. Tomlinson, E. Quartararo, C. Ahern, and M. L. Mazurek, "How Ready is Your Ready? Assessing the Usability of Incident Response Playbook Frameworks," in *CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–18.

[19] J. Bollinger, B. Enright, and M. Valites, *Crafting the InfoSec playbook: security monitoring and incident response master plan*. O'Reilly Media, Inc., 2015.

[20] I. Lella, E. Tsekmezoglou, R. S. Naydenov, C. Ciobanu, A. Malatras, and M. Theocharidou, "ENISA Threat Landscape 2022," European Union Agency for Network and Information Security (ENISA), Tech. Rep., 2022.

[21] C. Lawson and A. Price, "2022 Market Guide for Security Orchestration, Automation and Response Solutions," Gartner, Tech. Rep., 2022.

[22] MarketsandMarkets Research, "Security Orchestration Automation and Response (SOAR) Market Size, Share and Global Market Forecast to 2027," MarketsandMarkets, Tech. Rep., 2022.

[23] OASIS, "CACAO Security Playbooks Version 1.0 - Committee Specification 02," OASIS, Tech. Rep., 2021, last accessed 2023-03-01. [Online]. Available: https://docs.oasis-open.org/cacao/security-playbooks/v1.0/security-playbooks-v1.0.html

[24] FIRST Forum of Incident Response and Security Teams, "Automation SIG," 2023, last accessed 2023-03-01. [Online]. Available: https://www.first.org/global/sigs/automation/

[25] A. S. M. Rizvi, L. M. Bertholdo, J. M. Ceron, and J. S. Heidemann, “Anycast agility: Network playbooks to fight ddos,” in *31st USENIX Security Symposium, USENIX Security 2022, Boston, MA, USA, August 10-12, 2022*, 2022, pp. 4201–4218.

[26] M. A. Gurabi, A. Mandal, J. Popanda, R. Rapp, and S. Decker, “SASP: a Semantic web-based Approach for management of Sharable cybersecurity Playbooks,” in *ARES 2022: The 17th International Conference on Availability, Reliability and Security, Vienna, Austria, August 23 - 26, 2022*, 2022, pp. 109:1–109:8.

[27] D. Schlette, M. Caselli, and G. Pernul, “A Comparative Study on Cyber Threat Intelligence: The Security Incident Response Perspective,” *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2525–2556, 2021.

[28] A. Shaked, Y. Cherdantseva, and P. Burnap, “Model-Based Incident Response Playbooks,” in *ARES 2022: The 17th International Conference on Availability, Reliability and Security, Vienna, Austria, August 23 - 26, 2022*, 2022, pp. 26:1–26:7.

[29] V. Mavroeidis, P. Eis, M. Zádník, M. Caselli, and B. Jordan, “On the Integration of Course of Action Playbooks into Shareable Cyber Threat Intelligence,” in *2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, December 15-18, 2021*, 2021, pp. 2104–2108.

[30] C. Onwubiko and K. Ouazzane, “SOTER: A playbook for cybersecurity incident management,” *IEEE Transactions on Engineering Management*, vol. 69, no. 6, pp. 3771–3791, 2020.

[31] A. Applebaum, S. Johnson, M. Limiero, and M. Smith, “Playbook Oriented Cyber Response,” in *2018 National Cyber Summit (NCS)*. IEEE, 2018, pp. 8–15.

[32] X. Bouwman, V. L. Pochat, P. Foremski, T. van Goethem, C. H. Gañán, G. C. M. Moura, S. Tajalizadehkhoob, W. Joosen, and M. van Eeten, “Helping hands: Measuring the impact of a large threat intelligence sharing community,” in *31st USENIX Security Symposium, USENIX Security 2022, Boston, MA, USA, August 10-12, 2022*, 2022, pp. 1149–1165.

[33] B. Stojkovski, G. Lenzini, V. Koenig, and S. Rivas, “What’s in a Cyber Threat Intelligence sharing platform?: A mixed-methods user experience investigation of MISP,” in *ACSAC ’21: Annual Computer Security Applications Conference, Virtual Event, USA, December 6 - 10, 2021*, 2021, pp. 385–398.

[34] Financial Services Information Sharing and Analysis Center, “Reducing Cyber Risk Through Intelligence Sharing,” 2023, last accessed 2023-03-01. [Online]. Available: https://www.fsisac.com/who-we-are

[35] H. Paananen, M. Lapke, and M. Siponen, “State of the Art in Information Security Policy Development,” *Computers & Security*, vol. 88, 2020.

[36] K. J. Knapp, R. Franklin Morris, T. E. Marshall, and T. A. Byrd, “Information Security Policy: An Organizational-Level Process Model,” *Computers & Security*, vol. 28, no. 7, pp. 493—508, 2009.

[37] J. vom Brocke, S. Zelt, and T. Schmiedel, “On the role of context in business process management,” *International Journal of Information Management*, vol. 36, no. 3, pp. 486–495, 2016.

[38] M. Rosemann, J. Recker, and C. Flender, “Contextualisation of Business Processes,” *International Journal of Business Process Integration and Management*, vol. 3, no. 1, pp. 47–60, 2008.

[39] G. B. Willis and P. Royston, *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. SAGE Publications, Inc., 2005.

[40] V. Braun and V. Clarke, “Using thematic analysis in psychology,” *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[41] D. S. Cruzes and T. Dybå, “Research synthesis in software engineering: A tertiary study,” *Information and Software Technology*, vol. 53, no. 5, pp. 440–455, 2011.

[42] V. Braun and V. Clarke, *Thematic Analysis: A Practical Guide*. London, UK: SAGE Publications, Ltd., 2019.

[43] Object Management Group (OMG), “Business Process Model and Notation (BPMN) Specification Version 2.0.2,” OMG, Tech. Rep., 2013, last accessed 2023-03-01. [Online]. Available: https://www.omg.org/spec/BPMN/2.0.2/PDF

[44] D. Ashenden, “Information Security management: A human challenge?” *Information security technical report*, vol. 13, no. 4, pp. 195–201, 2008.

[45] G. D. Bhatt, “Knowledge management in organizations: examining the interaction between technologies, techniques, and people,” *Journal of knowledge management*, 2001.

[46] B. Flyvbjerg, “Top ten behavioral biases in project management: An overview,” *Project Management Journal*, vol. 52, no. 6, pp. 531–546, 2021.

[47] T. D. Wilson and N. Brekke, “Mental contamination and mental correction: unwanted influences on judgments and evaluations,” *Psychological bulletin*, vol. 116, no. 1, pp. 117–142, 1994.

[48] D. Kahneman and A. Tversky, “Intuitive prediction: Biases and corrective procedures,” Decisions and Designs Inc, Tech. Rep., 1977.

[49] W. Tounsi and H. Rais, “A survey on technical threat intelligence in the age of sophisticated cyber attacks,” *Computers & Security*, vol. 72, pp. 212–233, 2018.

[50] F. Skopik, G. Settanni, and R. Fiedler, “A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing,” *Computers & Security*, vol. 60, pp. 154–176, 2016.

[51] S. Barnum, “Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX),” MITRE Corporation, Tech. Rep., 2012.

[52] M. Wahnon, “Awesome Incident Response,” 2023, last accessed 2023-03-01. [Online]. Available: https://github.com/meirwah/awesome-incident-response

[53] Correlated Security, “Awesome SOAR,” 2023, last accessed 2023-03-01. [Online]. Available: https://github.com/correlatedsecurity/Awesome-SOAR

[54] Microsoft Community, “Windows security baselines,” 2023, last accessed 2023-03-01. [Online]. Available: https://learn.microsoft.com/en-us/windows/security/threat-protection/windows-security-configuration-framework/windows-security-baselines

[55] Center for Internet Security, “CIS Benchmarks,” 2023, last accessed 2023-03-01. [Online]. Available: https://www.cisecurity.org/cis-benchmarks

[56] US Department of Defense, “Security Technical Implementation Guides (STIGs),” 2023, last accessed 2023-03-01. [Online]. Available: https://public.cyber.mil/stigs/

[57] F. Li and V. Paxson, “A large-scale empirical study of security patches,” in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 2201–2215.

[58] OASIS, “Common Security Advisory Framework Version 2.0 OASIS Standard,” OASIS, Tech. Rep., 2022, last accessed 2022-03-01. [Online]. Available: https://docs.oasis-open.org/csaf/csaf/v2.0/csaf-v2.0.html

[59] A. Frisch, *Essential system administration: Tools and techniques for linux and unix administration*. O’Reilly Media, Inc., 2002.

[60] E. Nemeth, G. Snyder, T. R. Hein, B. Whaley, and D. Mackin, *UNIX and Linux system administration handbook*. Pearson Education, 2018.

[61] M. C. Langston, *Short Topics in Systems Administration - Documentation Writing for System Administrators*. USENIX/SAGE, 2003.

[62] L. Hochstein and R. Moser, *Ansible: Up and Running: Automating configuration management and deployment the easy way*. O’Reilly Media, Inc., 2017.

[63] K. Morris, *Infrastructure as code*. O'Reilly Media, Inc., 2020.

[64] W. M. P. van der Aalst, "The Application of Petri Nets to Workflow Management," *Journal of Circuits, Systems, and Computers*, vol. 8, no. 1, pp. 21–66, 1998.

[65] M. Weske, "Process Orchestrations," in *Business Process Management: Concepts, Languages, Architectures*. Springer, 2019, pp. 123–240.

[66] Camunda, "The Process Orchestration Handbook," 2021, last accessed 2023-03-01. [Online]. Available: https://camunda.com/process-orchestration/

[67] A. Oberweis, "An integrated approach for the specification of processes and related complex structured objects in business applications," *Decision Support Systems*, vol. 17, no. 1, pp. 31–53, 1996.

[68] D. Georgakopoulos, M. F. Hornick, and A. P. Sheth, "An Overview of Workflow Management: From Process Modeling to Workflow Automation Infrastructure," *Distributed Parallel Databases*, vol. 3, no. 2, pp. 119–153, 1995.

[69] P. O'Neill and A. S. Sohal, "Business Process Reengineering A review of recent literature," *Technovation*, vol. 19, no. 9, pp. 571–581, 1999.

[70] OASIS, "Open Command and Control (OpenC2) Language Specification Version 1.0 - Committee Specification 02," OASIS, Tech. Rep., 2019, last accessed 2023-03-01. [Online]. Available: https://docs.oasis-open.org/openc2/oc2ls/v1.0/oc2ls-v1.0.html

Appendix A. Graphical phishing playbook

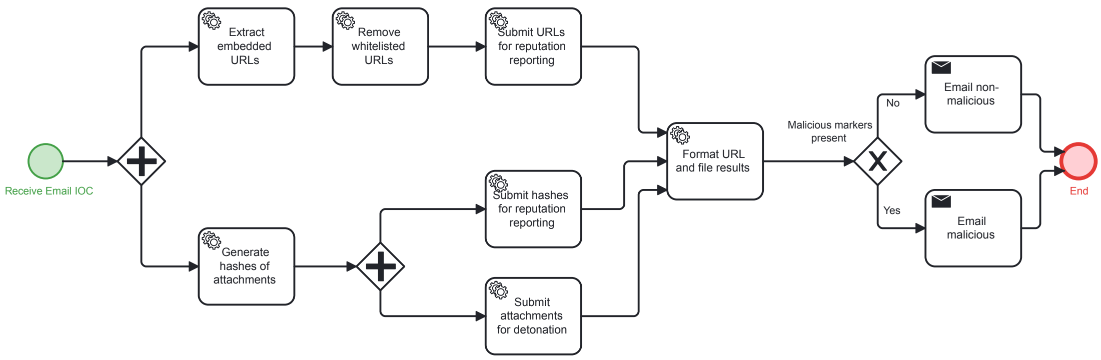

Receive Email IOC
Extract embedded URLs
Remove whitelisted URLs
Submit URLs for reputation reporting
Generate hashes of attachments
Submit hashes for reputation reporting
Submit attachments for detonation
Format URL and file results
Malicious markers present
No
Yes
Email non-malicious
Email malicious
End

Figure 5. Graphical phishing playbook with BPMN (source: IACD).

Appendix B. Code-based phishing playbook workflow step (excerpt)

```
{
  "action": "URL reputation",
  "action_type": "investigate",
  "assets": [
    {
      "app_name": "VirusTotal",
      "app_version": "1.2.40",
      "output": [
        {
          "data_path": "action_result.status"
        },
        {
          "data_path": "action_result.parameter.url"
        },
        {
          "data_path": "action_result.data.*.resource"
        },
        {
          "data_path": "action_result.data.*.scan_date"
        },
        {
          "data_path": "action_result.data.*.scan_id"
        },
        {
          "data_path": "action_result.data.*.url"
        },
        {
          "data_path": "action_result.message"
        }
      ],
      "parameters": {
        "url": {
          "data_type": "string",
          "description": "URL to query",
          "key": "url",
          "required": true
        }
      }
    }
  ]
}
```

Figure 6. Excerpt of a workflow step investigating URL reputation within a code-based phishing playbook (source: Splunk SOAR).

Appendix C. Playbook terminology

TABLE 7. Security standards and process notations help to understand playbook terminology.

Term	Description	CACAO [23]	OpenC2 [70]	CSAF [58]	BPMN [43]
Playbook	A playbook describes a specific cybersecurity process or procedure based on a workflow with individual steps or actions. Playbooks also include metadata. (Example: Phishing playbook)	**Playbook:** “[...] a playbook consisting of one or more security actions combined into a sequence or algorithmically-defined use.” “A **template playbook** will not be immediately executable by a receiving organization but may inform their own executable playbook for their specific environment or organization.”	N/A	**CSAF document:** “security advisory text document [...].” **Advisory:** “reporting item that describes a condition present in an artifact and that requires action by the consumers.”	**BPMN Model / BPMN Diagram:** “[...] a BPMN diagram is a particular snapshot of a BPMN model at a certain point in time.”
Workflow and Workflow Step	A workflow or course of action captures multiple workflow steps and procedural logic. A workflow step is defined by its position and its structural components. (Example: Start - 1) Step X, 2) Step Y, 3) Step Z - End)	“**Workflows** contain a series of **steps** [...]. Workflows process steps either sequentially, in parallel, or both depending on the type of steps required by the playbook.”	**OpenC2 Command:** “The Command describes an Action to be performed on a Target and may include information identifying the Actuator or Actuators that are to execute the Command.” **Command** ≈ workflow step	**Remediations:** “Every Remediation item [...] specifies details on how to handle (and presumably, fix) a vulnerability.” **Remediations - Details:** “[...] contains a thorough human-readable discussion of the remediation.”	**Business Process:** “A defined set of business activities that represent the steps required to achieve a business objective. It includes the flow and use of information and resources.” **Activity:** “Work that a company or organization performs using business processes. An activity can be atomic or non-atomic (compound).”
Actuator	An actuator represents an entity performing an action. Information systems, applications, or humans are actuators and provide specific capabilities. (Example: Incident handler / linux server)	**Targets:** “The CACAO target object contains detailed information about the entities or devices that accept, receive, process, or execute one or more commands as defined in a workflow step. Targets contain the information needed to send commands as defined in steps to devices or humans” (e.g., individual, ssh, http-api, net-address).	**Actuator:** “The Actuator executes the Command. The Actuator will be defined within the context of an Actuator Profile” (e.g., Stateless Paket Filtering Profile/-function).	N/A	**Pool:** “A Pool represents a Participant in a Collaboration.” **Lane:** “A partition that is used to organize and categorize activities within a Pool. [...] Lanes are often used for such things as internal roles (e.g., Manager, Associate), systems (e.g., an enterprise application), or an internal department (e.g., shipping, finance).”
Action	An action is an executable instruction or task representing a precise cybersecurity measure. Actions are manifold but center on operations towards systems, networks, or humans. (Example: Investigate / block)	**Action:** “[...] security activity in an organization [...]. Those actions may represent an activity to investigate, prevent, mitigate or remediate a specific security state that has either occurred or the organization is taking action to ensure the security state never occurs.” **Commands:** “The CACAO command object [...] contains detailed information about the commands that are to be executed or processed automatically or manually as part of a workflow step” (e.g., last; netstat -n; ls -l -a /root or Disconnect the machine from the network and call the SOC on-call person).	**Action:** “The task or activity to be performed” (e.g., scan, deny).	N/A	**Task:** “An atomic activity that is included within a Process. A Task is used when the work in the Process is not broken down to a finer level of Process Model detail. Generally, an end-user, an application, or both will perform the Task.”
Artifact	An artifact serves as the object or input of an action. Artifacts allow identification and refer to threat intelligence. (Example: IP address)	“**Variables** can be defined and used as the playbook is executed” (e.g., $$ipv4-addr, $$sha256-hash). **in_args:** “The optional list of arguments passed to the target(s) as input to the step.” **Variables and in_args** ≈ artifact	**Target:** “The object of the action. The Action is performed on the Target” (e.g., ipv4_net, device, file, mac_addr).	N/A (related to **Vulnerabilities Property - Remediations - Group Ids** and **Vulnerabilities Property - Remediations - Product Ids** “[...] the current remediation item applies to.”)	**Data Objects:** “The primary construct for modeling data within the Process flow is the DataObject element. A DataObject has a well-defined lifecycle, with resulting access constraints.” **Data Inputs:** “Data requirements are captured as Data Inputs [...].”
Output	An output is the result of a workflow step. Outputs are optional and can contain status codes or other data. (Example: Success)	**out_args:** “The optional list of arguments that are returned from this step after execution of the commands by the targets.”	**OpenC2 Response:** “The Response is a Message sent from the recipient of a Command. Response messages provide acknowledgment, status, results from a query, or other information.”	N/A	**Data Objects:** “The primary construct for modeling data within the Process flow is the DataObject element. A DataObject has a well-defined lifecycle, with resulting access constraints.” **Data Outputs:** “Data that is produced is captured using Data Outputs [...].”

Appendix D. Multi-category playbooks

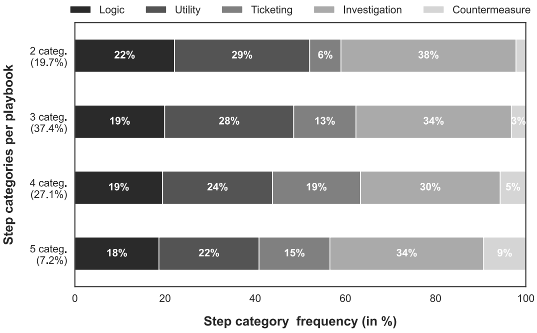


Figure 4. Multi-category playbooks and their workflow step composition.

Appendix E. Interview codebook

The codebook matches interview questions to external and internal influencing factors. Most factors relate to the three incident scenarios (i.e., Ransomware, DDoS, APT) and are addressed by "what-if" questions. In addition, we used the overarching questions for each scenario. The overarching questions provided in-depth results emphasizing the relevance of individual factors.

Overarching Questions:

- How does your organization handle a [incident type] scenario?
- How does your organization handle this scenario differently than other organizations?

External Factor 1: *Attacker characteristics (motivation, behavior, location)*
What would your organization do differently if ...

- it is operating in Israel/Ukraine?
- it assumes a state-sponsored attacker group located in Iran/Russia behind the attack?
- it has specified incident response time limits for APT attacks to 1 month?

External Factor 2: *Industry standards and guidelines (including frameworks)*
Does your organization use ...

- generic security process descriptions (e.g., BPMN, NIST Incident Response Life Cycle, NIST Cybersecurity Framework, FIRST CSIRT Framework, etc.)?
- incident response maturity models (e.g., SIM3, etc.)?
- incident response standards (e.g., CACAO, OpenC2, RE&CT, MITRE D3FEND, etc.)?
- an incident response policy to define guiding principles?

External Factor 3: *Laws and regulations (business structure, location/privacy, sector)*
What would your organization do differently if ...

- it is a publicly traded company in the United States?
- it has customers in the European Union?
- it is a sub-contractor for the defense industry?

External Factor 4: *Supply chain and business partners expectations*
What would your organization do differently if ...

- it has sourced server hosting to a third party?
- it is a sub-contractor for the defense industry?
- it observed compromised email accounts?

Internal Factor 1: *Incident response directives (data operations, attack targets/assets)*
What would your organization do differently if ...

- it has proxy log retention set to 1 week?
- it has business-critical applications running on the targeted web server?
- it has defined formal reporting requirements?
- its CFO's laptop is affected?

Internal Factor 2: *People (security culture/mandate, security team)*
What would your organization do differently if ...

- it has a dedicated Cyber Threat Intelligence team?
- it has defined email for incident response communication?
- it has a security team with 100 security experts?
- it has specified to contact the CISO, but the CISO is unavailable?

Internal Factor 3: *Technology (infrastructure/tech stack, security tools)*
What would your organization do differently if ...

- it defined an incident budget of $100k?
- it has sourced server hosting to a third party?
- it has a Web Application Firewall and additional server capacity?

Does your organization use ...

- Security Orchestration, Automation and Response (SOAR) tools (e.g., Splunk Phantom, Cortex XSOAR, Tines, etc.)?

Appendix F. Meta-Review

F.1. Summary

This paper explores the content of publicly-shared community incident response playbooks and the factors influencing playbook customization and deployment. Through a multistep experiment involving a measurement study, online survey, and interviews with security professionals, the authors find that organizations customize their playbooks based on their own definitions and areas of incident response.

F.2. Scientific Contribution

- Provides a New Data Set For Public Use
- Provides a Valuable Step Forward in an Established Field

F.3. Reasons for Acceptance

1) The paper curates a set of playbooks and conducts a structured analysis, making an important step in the study of playbooks. The inclusion of interviews with various actors adds value to the research. While some findings were considered expected, the mention of counter-intuitive results, such as "Attacker behavior and motivation beat location", was appreciated by reviewers.
2) The paper highlights failed approaches, explains why they likely failed, and describes the restrictions of the selected alternative approach. This practice was appreciated as it promotes future replication or related research by providing insights into what works and what doesn't.
3) The authors provide a GitHub repository containing the playbook data, research artifacts, and the questionnaire and interview guide. Reviews mentioned transparency and their increased confidence in the technical accuracy of the work.

F.4. Noteworthy Concerns

Insufficient findings and actionable insight: Reviews expressed dissatisfaction with the lack of concrete insights and absence of specific evidence supporting key claims. Reviewers emphasized the need for clearer insights and practical implications for researchers and practitioners based on the study's findings. It was also suggested that future studies may benefit from building upon the data presented in the paper to generate more meaningful insights and practical implications.

Appendix G. Response to the Meta-Review

We thank the reviewers and our shepherd for the time and effort spent reviewing our manuscript and for the fair and positive feedback. By sharing our compiled playbook dataset, committing to keep it public on GitHub, and writing this paper on playbook characteristics and influencing factors, we wanted to contribute to the academic understanding of a topic comprehensively known only by practitioners and further promote research in the cybersecurity incident response field. Our aim was to provide a solid scientific foundation on the use of playbooks and to position our paper as a strong groundwork on top of which future research can build on (or even challenge) our findings. What follows provides some details on how we addressed the main concern from the meta-review:

G.1. Main concern

We acknowledge that our findings may not be definitive in all their aspects. However, we believe they offer valuable insights that can guide future research on playbooks. We understand the reviewers' concerns regarding the lack of actionable insights and specific evidence supporting our key claims. To rectify this, we consolidated and expanded upon the insights derived from our research, providing clearer and more tangible implications for both researchers and practitioners. More specifically:

- The actionable insights under review were: addressing data quality issues, identifying technological assistance opportunities, and establishing a common language for discussing incident response playbooks. To ensure greater clarity, we have expounded on such insights and findings, which are now highlighted in the discussion and conclusion sections of the paper.
- To further substantiate our claims, we included additional information where deemed necessary by the reviewers. For instance, we have supplemented Section 4 (RQ1) with incident-specific playbook information. Moreover, our research connects three distinct data sets, facilitating in-depth discussions on influencing factors and playbook content. While RQ1 focuses on community playbooks, RQ2 delves into the influencing factors that determine organization-specific playbooks. We recognize that the connection between the research questions might not have been perfect, but by incorporating examples (e.g., location look-up, SLAs, communication tools) from RQ1, we better corroborate or challenge interviewees' perceptions of influencing factors.

P6 SOAR4IoT: Securing IoT Assets with Digital Twins

Status:	Published	
Date of Submission:	02 March 2022	
Date of Acceptance:	16 May 2022	
Date of Publication:	23 August 2022	
Conference:	17th International Conference on Availability, Reliability and Security (ARES 2022)	
Location:	University of Vienna, Vienna, Austria	
Period:	23.08.2022 – 26.08.2022	
Authors' Contributions:		
	Philip Empl	40%
	Daniel Schlette	40%
	Daniel Zupfer	10%
	Günther Pernul	10%
Full Citation:	Empl, P., Schlette, D., Zupfer, D., & Pernul, G. (2022). SOAR4IoT: Securing IoT Assets with Digital Twins. In *Proceedings of the 17th International Conference on Availability, Reliability and Security* (pp. 4:1–4:10). Association for Computing Machinery.	
DOI:	10.1145/3538969.3538975	
Artifact:	git.ur.de/soar4iot	

Conference Description: The International Conference on Availability, Reliability and Security (ARES) brings together researchers and practitioners in the field of IT security & privacy. Since 2005, ARES serves as an important platform to exchange, discuss and transfer knowledge and is hosted every year in another European city.

SOAR4IoT: Securing IoT Assets with Digital Twins

Philip Empl*
philip.empl@ur.de
University of Regensburg
Germany

Daniel Schlette
daniel.schlette@ur.de
University of Regensburg
Germany

Daniel Zupfer
daniel.zupfer@ur.de
University of Regensburg
Germany

Günther Pernul
guenther.pernul@ur.de
University of Regensburg
Germany

ABSTRACT

As more and more security tools provide organizations with cybersecurity capabilities, security analysts are overwhelmed by security events. Resolving these events is challenging due to extensive manual processes, limited financial resources, and human errors. Security Orchestration, Automation, and Response (SOAR) is an established approach to manage security tools and assets. However, SOAR platforms typically integrate traditional IT systems only. Additional considerations are required to deal with the Internet of Things (IoT), its multiple devices and complex networks. Therefore, we adapt SOAR to IoT. We first aggregate existing research and information on SOAR and SOAR platforms. We envision the `SOAR4IoT` framework, making IoT assets manageable for SOAR via middleware. We implement a prototypical digital twin-based SOAR application integrating IoT assets and security tools to validate our framework. The experimental setup includes two playbooks coping with Mirai and Sybil attacks. Results show feasibility as our SOAR application enables securing IoT assets with digital twins.

CCS CONCEPTS

• **Security and privacy** → *Network security*; *Systems security*; *Security services*; • **Computer systems organization**; • **Information systems**;

KEYWORDS

Internet of Things, Security Orchestration, Incident Response, SOAR, Digital Twin

ACM Reference Format:
Philip Empl, Daniel Schlette, Daniel Zupfer, and Günther Pernul. 2022. SOAR4IoT: Securing IoT Assets with Digital Twins. In *The 17th International Conference on Availability, Reliability and Security (ARES 2022), August 23–26, 2022, Vienna, Austria.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3538969.3538975

1 INTRODUCTION

Attackers and defenders shape cybersecurity. Sophisticated attacks on networked information systems are countered by defenders' use of tools for security monitoring and operations. However, there is an ongoing challenge for security analysts. While more and more security tools are being used, analysts can face up to 11,000 security alerts per day (including false positives) [11]. Therefore, organizations use Security Orchestration, Automation and Response (SOAR) platforms promising tool integration, automation, and streamlined workflows for rapid incident response [19, 25].

SOAR platforms are based on security events. Security events concern traditional IT resources but also the Internet of Things (IoT). The new IoT frontier (e.g., smart factories or automated home systems) with its multitude of heterogeneous devices contributes to the ongoing datafication but currently neglects cybersecurity. Inadequate or missing security measures caused by a "set-it-and-forget-it manner" [20] are illustrative for the insecurity of IoT assets. Attackers notice these IoT security issues, as Kaspersky reports 1.5 billion attacks against their IoT honeypots in the first half of 2021 [30]. Eventually, networked IoT devices exposing default username/password authentication will become part of botnets. Estimates see the approximate time to compromise an IoT device at just five minutes [20]. Thus, it is necessary to extend security operations to IoT assets for which digital twins provide promising features [9]. Digital twins are used for security to simulate IoT attacks [8] and can assist incident response [7, 10].

Whether IoT-specific or not, security analysts cannot process security events manually. SOAR platforms greatly help analysts perform investigations and initiate adequate incident response actions. Analysts can reduce time and resources spent on low-priority events and manual actions using automated playbooks. Thus, SOAR documents a shift towards more effective security operations within organizations. As SOAR attracts attention in research and provides the dynamics to abstract complex environments, we investigate its potential for the IoT. Consequently, we ask *"how to use Security Orchestration, Automation and Response for the Internet of Things?"* We expect the general applicability of SOAR for IoT as it is a flexible construct. Still, it is crucial to showcase adaptation rigorously.

In this paper, we aim to answer the following questions: (1) What defines SOAR? (2) How to secure the IoT? (3) How to implement SOAR for IoT with digital twins? These questions lead to our main contributions:

- We enlighten SOAR core activities and platform features by analyzing the few academic works and current SOAR platforms.
- We envision our `SOAR4IoT` framework built on IoT attacks and mitigation strategies. Our framework encompasses IoT assets, middleware, SOAR platform, and security tools.
- We provide a SOAR4IoT implementation leveraging digital twins. The experimental setup documents the straightforward, ground-up implementation of a SOAR platform, including Eclipse Ditto-based digital twins, which researchers and practitioners can easily adapt and extend.

*Corresponding author

ARES 2022, August 23–26, 2022, Vienna, Austria
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9670-7/22/08.
https://doi.org/10.1145/3538969.3538975

- We explore two security issues of IoT assets. We address IoT security operations by designing and implementing two generic playbooks for orchestration and automated response to the Mirai botnet and the Sybil attack.

The paper is organized as follows. Section 2 outlines IoT, digital twins for cybersecurity, SOAR foundations and describes related work. Section 3 elaborates the framework defining the characteristics of SOAR, discussing the objectives of secure IoT assets, and describing technologies abstracting the IoT. Then, formal requirements lead to the overall `SOAR4IoT` framework. We validate our framework in Section 4 through the implementation of a digital twin-based SOAR platform integrating two use cases. We conclude our paper in Section 5.

2 BACKGROUND AND RELATED WORK

This section elaborates the background on IoT (Section 2.1), digital twins for cybersecurity (Section 2.2) and SOAR (Section 2.3), concluding with related work (Section 2.4).

2.1 Internet of Things

The IoT is characterized by identifiable networking objects (sensors or actuators) advertising their services to assemble semantic-rich applications [1]. Beyond scrutinizing particular devices, the IoT involves communication, applications, and processes. Heterogeneous devices and machines of widely ranging specifications and data operate seamlessly and collaboratively to assist business processes. The heterogeneity of IoT devices and networks is mainly caused by various manufacturers and (communication) protocols. As a result, there are plenty of cybersecurity issues demanding 1) automated security operations (detection and mitigation) and 2) orchestration of security functions for the IoT [17]. When it comes to integrating IoT assets, middleware is reliable, and a common choice [27, 32]. Organizations can choose between different types of middleware according to technology preferences and use cases (see Figure 1).

2.2 Digital Twins for Cybersecurity

In general, digital twins can be conceived as middleware. At its core, the digital twin links a virtual representation to a physical asset aiming to mirror the asset along its life cycle with semantic technologies [3]. The digital twin synchronizes system states using bidirectional communication with its physical counterpart. Implementing digital twins is a challenging task. Digital twins (e.g., Eclipse Ditto or Azure Digital Twins) can be used standalone or connected to IoT platforms (e.g., Eclipse Kapua or Azure IoT Hub).

From a security perspective, digital twins concern three primary security-operation modes: replication, simulation, and historical data analytics [8]. *Historical data analytics* deals with the documented behavior of IoT assets in the past and draws conclusions for the future. *Simulations* build on user-specific parameters and model the semantics of the real world. Last, the *replication* integrates real-world data to semantically model and operate a digital twin identical to its real-world counterpart. These operation modes assist security operations. For instance, behavior-based modeling supports more efficient intrusion detection, and the virtual representation of the digital twin is suitable for security training [9]. Moreover, replication-based digital twins indicate security orchestration and incident response features.

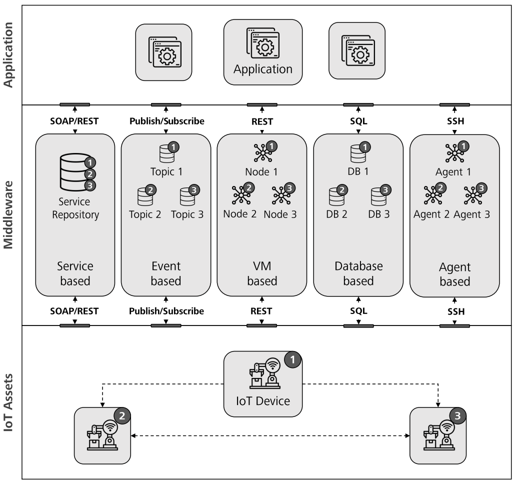


Figure 1: IoT architecture and middleware types

2.3 Security Orchestration, Automation and Response (SOAR)

Platforms promising *Security Orchestration, Automation and Response (SOAR)* capabilities for organizations are the latest solutions proposed by cybersecurity vendors [19]. Like other solutions before, the underlying concept has not received much research attention while products are being pushed to market. SOAR is not a standalone concept but part of continuous development. From related concepts like log management to Security Information and Event Management (SIEM), Cyber Threat Intelligence (CTI), and security orchestration, it can be observed that succeeding concepts build on previous ones. Examining SOAR, it becomes evident that platforms, system architectures, and data are crucial to understanding and implementing the concept.

In the organizational context, SOAR and corresponding platforms are associated with the Security Operations Center (SOC) or Computer Security Incident Response Team (CSIRT) [31]. Intuitively, SOAR aims to assist activities within the three domains of 1) security orchestration, 2) automation, and 3) incident response.

For *security orchestration*, SOAR subsumes the functionality of SIEM and integrates multiple devices, systems, and security tools [13]. Additionally, integration and unification aspects of SOAR relate to threat intelligence as relevant information about threats, attacks, and vulnerabilities is aggregated from internal and external sources. For *automation*, SOAR relies on events and defined courses of action to enable rapid security operations. Thus, automation bridges the gap between security orchestration and incident response. For *incident response*, containment, eradication, and recovery activities demand to derive and perform appropriate measures.

Therefore, SOAR includes the instrumentalization of endpoints and security tools to execute commands.

Related to SOAR is the standardization and representation of incident response [28]. While current systems are often based on ticketing systems for security incidents, incident response playbooks are central. In essence, incident response playbooks define how to conduct a specified defensive procedure. Towards standardization, the incident response community initiated the development of dedicated data formats. These formats specify structural elements and required meta-data for incident response use cases. For instance, the two formats *Open Command and Control (OpenC2)* [23] and *Collaborative Automated Course of Action Operations (CACAO) for Cyber Security* [24] document different focal areas such as executable commands and procedural workflows, respectively.

2.4 Related Work

IoT devices and networks are susceptible to cyberattacks. Providing security measures for IoT is a practical problem and has attracted researchers' attention. As outdated firmware enables attacks on IoT devices, the literature emphasizes security orchestration by using a firmware update scenario (e.g., [2]). RFC 9019 describes updating IoT firmware in detail [18] while others use distributed ledger technologies [5]. As a consequence, we consider IoT firmware updates to validate our work. From a network perspective, the European Telecommunication Standard Institute proposes central security orchestration based on automated configurations and deployments [15]. We build on existing research and unify security orchestration activities. We include network and device layers within a single SOAR framework.

Digital twins for incident response is a trending research topic. Digital twins assist analysts in SOC [8] and are proposed for response measures [12]. Especially for operational systems, digital twins should implement cybersecurity services (e.g., access control, intrusion detection, or incident response) [7]. In a recent publication, Eckhart and Ekelhart [10] emphasize digital twins of real-world IoT systems as a new method for incident response. Existing literature only conceptualizes digital twin-based incident response. We are taking research further and implement digital twins for incident response.

Scoping the topic of SOAR, we identified additional related work. Most notably, Islam et al. [13] provide a survey on security orchestration. In a follow-up work on SOAR architecture, the authors propose the layered integration of security tools and map tools to response activities [14]. For CTI sources in SOAR, security enumerations have been discussed in the context of the IoT [29]. We go beyond security tools and include application aspects and IoT assets in our approach.

Further, SOAR has been examined in the context of incident response. Complementary to incident response formats, Schlette et al. [28] outline the vast SOAR product landscape. As SOAR platforms assist organizations' incident response, research addressed the appropriate selection [22] and quantitative evaluation of features [21]. SOAR platforms evolve and existing works provide a snapshot. Based on these works, we aggregate common features of SOAR platforms and settle on agreed-upon characteristics.

Table 1: SOAR requirements

	Requirement	Description	IoT
Core activities	Security Orchestration	Integration of IT assets, security tools, and threat intelligence	*
	Automation	Use of technologies and logic to perform security operations	✓
	Incident Response	Investigation, mitigation, and remediation of incidents	*
Platform features	User Interface	Dashboard or console for human interaction	✓
	Playbooks	Workflows, courses of action, or scripts	✓
	Ticketing System	Case management for security incidents	✓
	User Management	Access control and communication	✓

✓ is applicable * requires modification

3 SOAR4IOT FRAMEWORK

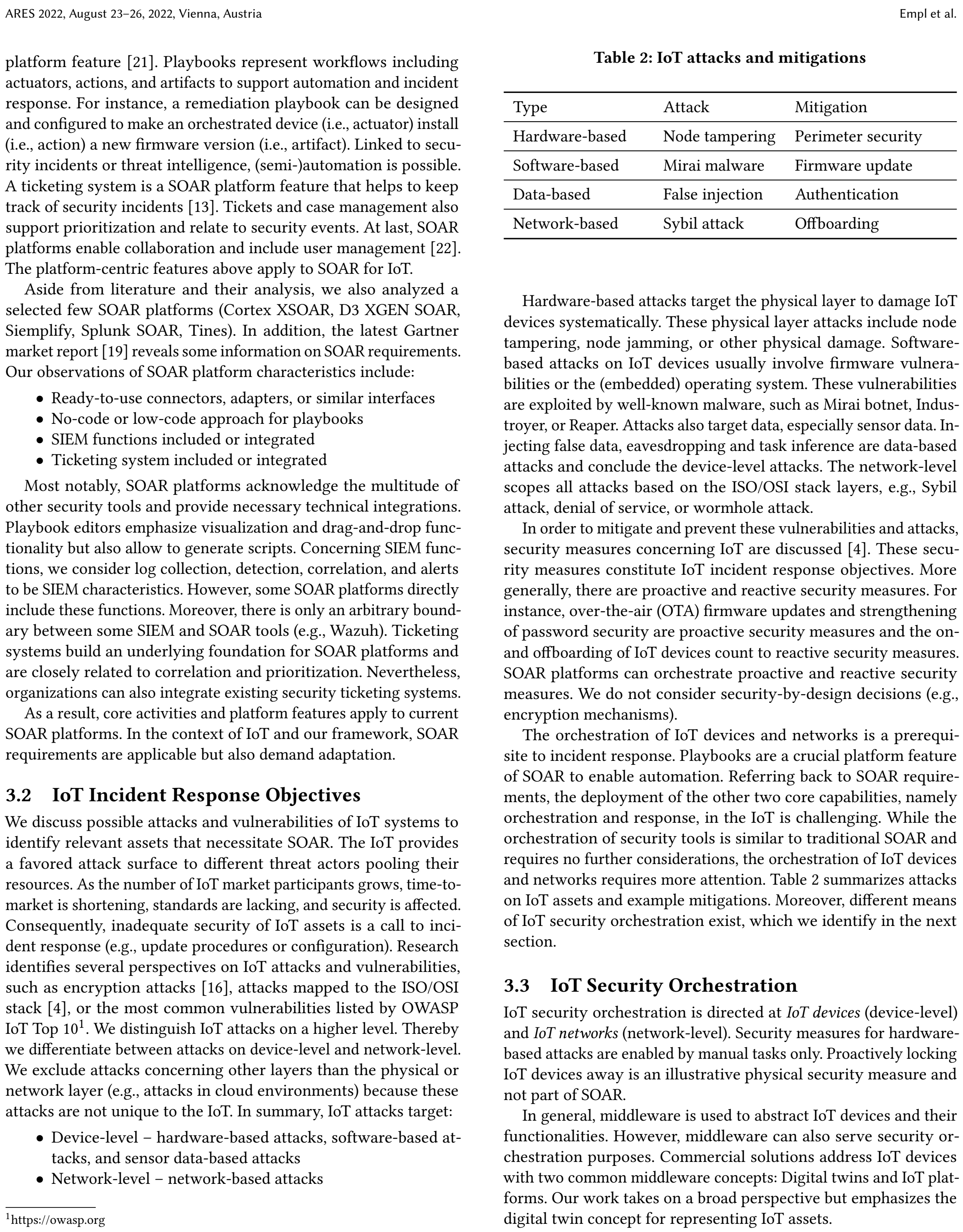
To apply SOAR to IoT, we first identify general SOAR requirements (Section 3.1). Examining attacks on the IoT, we then derive IoT incident response objectives (Section 3.2). These objectives guide us towards required IoT security orchestration (Section 3.3). Based on our formal model (Section 3.4), we conceptualize a `SOAR4IoT` framework (Section 3.5) that integrates IoT systems using digital twins.

3.1 SOAR Requirements

SOAR requirements describe essential characteristics for the implementation of SOAR. Ultimately, SOAR requirements can assist the development of a SOAR platform, the evaluation of existing ones, or the adaptation to IoT devices and networks. In the following, we aggregate SOAR requirements from existing literature and validate the findings by examining current SOAR platforms. Table 1 describes core activities and platform features.

Core activities (i.e., security orchestration, automation, and incident response) constitute one group of requirements. They represent platform capabilities. For IoT, security orchestration demands modification as heterogeneous, dispersed devices form dynamic networks. Task automation remains largely unaffected, is conducted at SOAR platform level, and applies to IoT. Incident response measures directly involve IoT assets and thus demand modification.

Platform features constitute the second group of SOAR requirements. They represent technical aspects of a SOAR platform. Typically, a SOAR platform provides a user interface such as a dashboard or a console to assist orchestration and response activities [14]. More precisely, the user interface allows querying data and triggering courses of action. Playbooks are another dedicated SOAR



platform feature [21]. Playbooks represent workflows including actuators, actions, and artifacts to support automation and incident response. For instance, a remediation playbook can be designed and configured to make an orchestrated device (i.e., actuator) install (i.e., action) a new firmware version (i.e., artifact). Linked to security incidents or threat intelligence, (semi-)automation is possible. A ticketing system is a SOAR platform feature that helps to keep track of security incidents [13]. Tickets and case management also support prioritization and relate to security events. At last, SOAR platforms enable collaboration and include user management [22]. The platform-centric features above apply to SOAR for IoT.

Aside from literature and their analysis, we also analyzed a selected few SOAR platforms (Cortex XSOAR, D3 XGEN SOAR, Siemplify, Splunk SOAR, Tines). In addition, the latest Gartner market report [19] reveals some information on SOAR requirements. Our observations of SOAR platform characteristics include:

- Ready-to-use connectors, adapters, or similar interfaces
- No-code or low-code approach for playbooks
- SIEM functions included or integrated
- Ticketing system included or integrated

Most notably, SOAR platforms acknowledge the multitude of other security tools and provide necessary technical integrations. Playbook editors emphasize visualization and drag-and-drop functionality but also allow to generate scripts. Concerning SIEM functions, we consider log collection, detection, correlation, and alerts to be SIEM characteristics. However, some SOAR platforms directly include these functions. Moreover, there is only an arbitrary boundary between some SIEM and SOAR tools (e.g., Wazuh). Ticketing systems build an underlying foundation for SOAR platforms and are closely related to correlation and prioritization. Nevertheless, organizations can also integrate existing security ticketing systems.

As a result, core activities and platform features apply to current SOAR platforms. In the context of IoT and our framework, SOAR requirements are applicable but also demand adaptation.

3.2 IoT Incident Response Objectives

We discuss possible attacks and vulnerabilities of IoT systems to identify relevant assets that necessitate SOAR. The IoT provides a favored attack surface to different threat actors pooling their resources. As the number of IoT market participants grows, time-to-market is shortening, standards are lacking, and security is affected. Consequently, inadequate security of IoT assets is a call to incident response (e.g., update procedures or configuration). Research identifies several perspectives on IoT attacks and vulnerabilities, such as encryption attacks [16], attacks mapped to the ISO/OSI stack [4], or the most common vulnerabilities listed by OWASP IoT Top 10[1]. We distinguish IoT attacks on a higher level. Thereby we differentiate between attacks on device-level and network-level. We exclude attacks concerning other layers than the physical or network layer (e.g., attacks in cloud environments) because these attacks are not unique to the IoT. In summary, IoT attacks target:

- Device-level – hardware-based attacks, software-based attacks, and sensor data-based attacks
- Network-level – network-based attacks

[1] https://owasp.org

Table 2: IoT attacks and mitigations

Type	Attack	Mitigation
Hardware-based	Node tampering	Perimeter security
Software-based	Mirai malware	Firmware update
Data-based	False injection	Authentication
Network-based	Sybil attack	Offboarding

Hardware-based attacks target the physical layer to damage IoT devices systematically. These physical layer attacks include node tampering, node jamming, or other physical damage. Software-based attacks on IoT devices usually involve firmware vulnerabilities or the (embedded) operating system. These vulnerabilities are exploited by well-known malware, such as Mirai botnet, Industroyer, or Reaper. Attacks also target data, especially sensor data. Injecting false data, eavesdropping and task inference are data-based attacks and conclude the device-level attacks. The network-level scopes all attacks based on the ISO/OSI stack layers, e.g., Sybil attack, denial of service, or wormhole attack.

In order to mitigate and prevent these vulnerabilities and attacks, security measures concerning IoT are discussed [4]. These security measures constitute IoT incident response objectives. More generally, there are proactive and reactive security measures. For instance, over-the-air (OTA) firmware updates and strengthening of password security are proactive security measures and the on- and offboarding of IoT devices count to reactive security measures. SOAR platforms can orchestrate proactive and reactive security measures. We do not consider security-by-design decisions (e.g., encryption mechanisms).

The orchestration of IoT devices and networks is a prerequisite to incident response. Playbooks are a crucial platform feature of SOAR to enable automation. Referring back to SOAR requirements, the deployment of the other two core capabilities, namely orchestration and response, in the IoT is challenging. While the orchestration of security tools is similar to traditional SOAR and requires no further considerations, the orchestration of IoT devices and networks requires more attention. Table 2 summarizes attacks on IoT assets and example mitigations. Moreover, different means of IoT security orchestration exist, which we identify in the next section.

3.3 IoT Security Orchestration

IoT security orchestration is directed at *IoT devices* (device-level) and *IoT networks* (network-level). Security measures for hardware-based attacks are enabled by manual tasks only. Proactively locking IoT devices away is an illustrative physical security measure and not part of SOAR.

In general, middleware is used to abstract IoT devices and their functionalities. However, middleware can also serve security orchestration purposes. Commercial solutions address IoT devices with two common middleware concepts: Digital twins and IoT platforms. Our work takes on a broad perspective but emphasizes the digital twin concept for representing IoT assets.

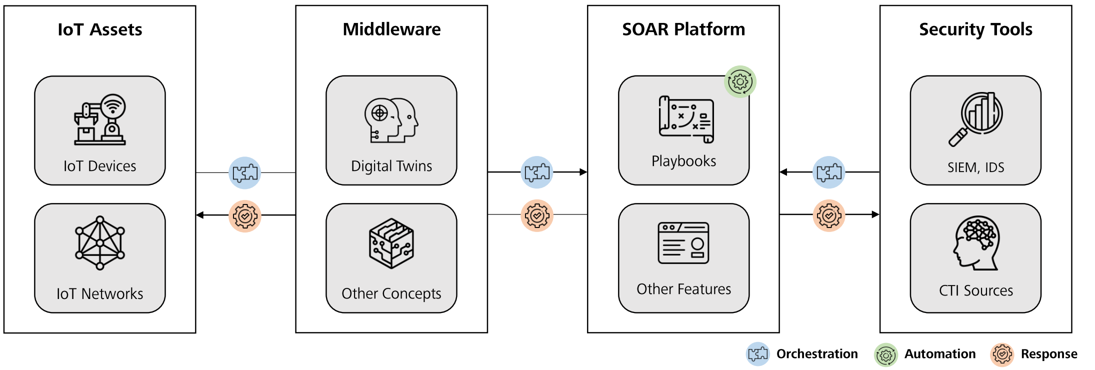


Figure 2: SOAR4IoT framework

Digital twins provide many features that enable security orchestration for IoT devices. They go beyond IoT platforms that are centered on common management tasks (e.g., onboard, monitor, and offboard devices). In particular, digital twins in replication mode provide IoT device modeling and security features. The bidirectional communication between the digital twin and IoT asset is beneficial as synchronizing sensor data and receiving commands can fulfill security orchestration. For instance, digital twins can store threat information acquired from third-party apps and synchronize information about vulnerabilities with their physical counterparts.

Besides IoT devices, digital twins and IoT platforms also extend to IoT networks. In this regard, digital twins allow the representation of dedicated edge nodes. Edge nodes are used in IoT networks as they control device communication. Using digital twins of edge nodes is thus a node-centric approach to communication-related security orchestration.

To sum things up, IoT device orchestration is enabled by digital twins. Further, IoT network orchestration requires the integration of edge nodes. Edge nodes are crucial as they control sub-networks containing several IoT devices. Therefore, we also include some node-centric aspects of IoT networks in our framework. We consider edge nodes represented by digital twins.

3.4 Formal Model

Concerning the security objectives of IoT, we define requirements targeting the three core capabilities of SOAR. These requirements are essential for the implementation of SOAR platforms and the definition of playbooks. The formal model includes:

Requirement 1 (Orchestration of IoT assets). We denote IoT assets as $A = \{a_1, a_2, ..., a_n\}$, whereby an asset is either a device, network or security tool. These assets are integrated into *SOAR*:

$$a \longrightarrow SOAR$$

Requirement 2 (Automation of security measures). Automation depends on security measures strategically executed for a specific event $E = \{e_1, e_2, ..., e_o\}$ mapping an asset to a playbook $P = \{p_1, p_2, ..., p_m\}$. Thereby, a playbook is generic and could be linked to one or more assets, located inside a SOAR platform. An asset does not necessarily require a playbook:

$$\exists e \in E:\ e \longrightarrow SOAR(p \circ a)\ \wedge\ SOAR(p) \longrightarrow a$$

$$\exists a \in A:\ \neg SOAR(p \circ a)$$

Requirement 3 (Deployment of responses to IoT assets). Response of the SOAR platform depends on whether at least one playbook fulfills or characterizes appropriate security measures for an event. Otherwise, no response is automatically deployed:

$$respond(e) = \begin{cases} SOAR(p) \longrightarrow a & if\ \exists p \in P:\ SOAR(p \circ e) \\ notify(e) & otherwise. \end{cases}$$

In the next step, we outline our framework, its components and middleware integration.

3.5 Framework Overview

Middleware integration complements our SOAR4IoT framework. We emphasize using digital twin middleware to extend existing SOAR platforms based on the previously established SOAR requirements and IoT security objectives. Figure 2 depicts the SOAR4IoT framework and the middleware integration.

IoT assets. The SOAR4IoT framework is based on IoT assets. IoT assets are classified as IoT devices (i.e., sensors or actuators) or IoT networks (i.e., edge nodes and communication). Intertwined, IoT devices and networks form complex IoT systems accessible through applications. IoT security orchestration implies that IoT assets are known to the SOAR platform. Consequently, there is an information flow from IoT assets to the SOAR platform. In the opposite direction, incident response measures target IoT assets.

Middleware. The SOAR4IoT framework integrates middleware. Besides digital twins, other middleware concepts (e.g., IoT platforms) exist. The middleware is located between IoT assets and the SOAR platform. We argue that middleware is beneficial for abstracting IoT assets. Also, IoT asset data is aggregated. Digital twins, in particular, provide semantic features (e.g., modeling components), a dedicated interface, and different perspectives (e.g., data views) for orchestration and response. In our case, digital twins offer a comprehensive summary of the asset's (security) state and enable the validation of security measures.

SOAR platform. The SOAR4IoT framework contains a SOAR platform at its core. Most importantly, the SOAR platform emphasizes playbooks and their automation but includes other typical features such as ticketing, user interface, and user management. Data flows from the middleware and connected security tools to the SOAR platform for security orchestration. Then, appropriate incident response measures are disseminated.

Security tools. The SOAR4IoT framework includes security tools. Security tools (e.g., SIEM – Security Information and Event Management systems or IDS – Intrusion Detection Systems) are queried or actively provide security-relevant information. Various Cyber Threat Intelligence sources (e.g., CTI feeds) can also provide input to the SOAR platform and serve as a trigger to response actions. However, incident response actions also address security tools (e.g., updating SIEM rules or disseminating CTI).

4 PROOF OF CONCEPT

We implement the SOAR4IoT framework to validate its feasibility. Defining two use cases, we represent security measures in two playbooks (Section 4.1 and 4.2). More specifically, our experimental setup includes the SOAR platform, replication-based digital twin middleware, and IoT assets (Section 4.3). Further, we demonstrate security orchestration, automation, and incident response and show experimental results (Section 4.4). At last, we conclude our proof of concept by discussing the impact and limitations (Section 4.5).

4.1 Mirai Botnet – Use Case 1

The Mirai malware is scanning IoT devices for vulnerabilities. The attacker's goal is to use the IoT devices for malicious purposes. Consequently, IoT assets need to be secured at the device level. This scenario represents our first use case. The following SOAR playbook describes courses of action to address Mirai-like situations that require firmware updates.

Playbook 1 Mirai Botnet (proactive)

```
procedure MIRAI
    $a \leftarrow$ IoT devices
    for all $d \in a$ do
        $e \leftarrow$ CTI for d
        if $isVulnerable(e, d)$ and $d.checkFirmware()$ then
            $d.updateFirmware()$
    if $checkAuthentication(e, a)$ then
        $changeAuthentication(a)$
        $a.permitJoin(true,\ 30s)$
```

Organizational security operations to cope with Mirai or similar malware include threat intelligence. Organizations monitor their IoT devices and pay attention to vulnerabilities. Either manually or automated, organizations analyze CTI reports. CTI describes severe vulnerabilities and triggers security operations. Such security operations include checking affected IoT device status and whether a new firmware update is available. This procedure is necessary to keep IoT devices secure and ensure continuous operation. Otherwise, IoT devices can easily contribute to malicious activities, such as distributed denial of service (DDoS) attacks.

4.2 Sybil Attack – Use Case 2

A Sybil attack in IoT describes the fake creation of identities (i.e., IoT assets) in IoT networks [26]. Thereby, attackers attempt to forward data selectively, drop data packets or manipulate data. Consequently, IoT assets need to be secured at the network level. This scenario represents our second use case. The following SOAR playbook describes courses of action to address Sybil attack situations that require device removal.

Playbook 2 Sybil Attack (reactive)

```
procedure SYBIL
    $e \leftarrow$ SIEM event
    $a \leftarrow$ IoT network
    for all $n \in a$ do
        if $isSybilNode(e, n)$ then
            $a.removeDevice(n)$
        $a.permitJoin(false)$
```

Organizational security operations to cope with a Sybil attack center on adequate monitoring of additional edge nodes or other IoT network components. Digital twins include detailed information about trusted IoT assets. Thus, they can be leveraged once a trigger (e.g., a SIEM event containing the loss of several data packets) from a security tool is received. Assessing the IoT network components, organizations can identify additional fake nodes or even missing ones and start response measures. This procedure is crucial to avoid malfunctioning IoT applications.

Multiple attacks on IoT assets demand SOAR capabilities. We opted for the two exemplary use cases based on the Mirai botnet and Sybil attack to document our SOAR4IoT framework implementation. Next, we describe our technological setup, including hardware and software.

4.3 Experimental Setup

Our experimental setup implements the SOAR4IoT framework. The source code is available in Gitlab[2]. Figure 3 describes our prototypical implementation and documents technology and data flows. This overview is further specified by categorizing and listing the underlying hardware (see Table 3).

IoT assets. We deploy two Xiaomi Aqara temperature sensors and two IKEA Tradfri LED bulb actuators in our lab environment. The sensors measure temperature and humidity. The actuators control brightness, color temperature, and state of connected LEDs. For communication purposes, sensors and actuators use the Zigbee protocol. Additionally, we deploy a Raspberry Pi 3B+ edge node. Zigbee communication between IoT assets and the edge node is controlled with a CC2531 Zigbee USB-Stick. This Zigbee controller is physically plugged into the edge node, but communication is

[2] https://git.ur.de/soar4iot

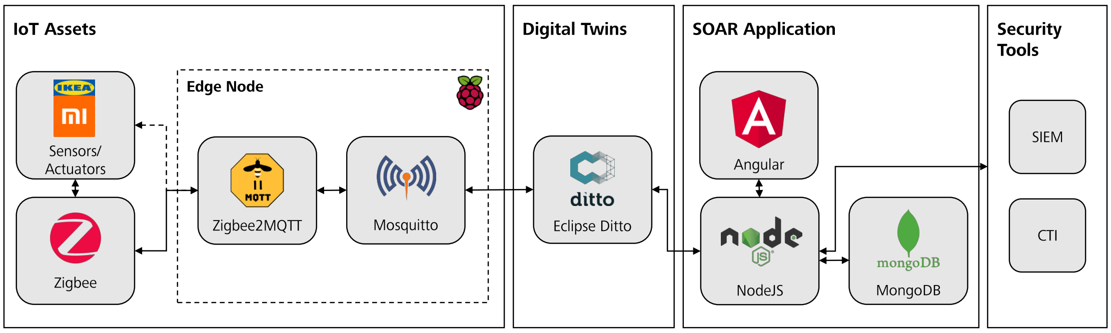


Figure 3: Experimental setting

Table 3: Hardware list

Device	Category	Characteristics
Xiaomi Aqara Temperature	Sensor	WSDCGQ01LM, Zigbee protocol
IKEA Tradfri LED Bulb E14	Actuator	LED1733G7, Zigbee protocol
CC2531 Zigbee USB flash drive	Controller	USB interface, Zigbee protocol
Raspberry Pi 3B+	Edge Node	Raspbian GNU/Linux 11, 1GB RAM, RJ-45 Ethernet
Virtual Machine	Server	Ubuntu 20.04.3 LTS, 16GB RAM, 8 cores, 80GB disc

wireless. At the edge node, the Zigbee data is transformed into MQTT data using the Zigbee2MQTT[3] bridge. Zigbee2MQTT acts as a client sending data from sensors and actuators to the MQTT broker. In our setup, the open-source MQTT broker Mosquitto[4] is installed on the edge node. As MQTT data is structured in topics, Zigbee2MQTT publishes/subscribes to an IoT asset-specific topic (e.g., SOAR4IoT/Lab_Actuator_Bulb1). In the same way, Mosquitto uses MQTT topics for upstream data. Similar IoT assets and edge nodes to our experimental setup might be used as part of an industrial oven or assembly line.

Digital twins. We implement digital twins representing the middleware of our `SOAR4IoT` framework. For each IoT asset there is one digital twin. Using the open-source digital twin software Eclipse Ditto[5] allows us to integrate and replicate heterogeneous IoT assets. Eclipse Ditto enables message-oriented communication with IoT assets through their digital twin. Besides, it supports the definition of policies (i.e., access control) and the integration of specific brokers for several IoT protocols (e.g., MQTT, AMQP, or CoAP). In our experimental setup, Eclipse Ditto runs on a virtual machine (Ubuntu, 16GB RAM, 8 kernels, and 80GB storage) and connects to Mosquitto.

We design and configure our Eclipse Ditto-based digital twins (see Figure 4). First, we define the primary policy. This policy grants an admin user read and write access to the digital twins and restricts a demo user to read access only. We then create five IoT assets, including the edge node. Each IoT asset is structured using JSON data serialization that defines a primary data schema for its digital twin. We further define messages in Eclipse Ditto. These messages allow users to interact directly with the digital twin of an IoT asset. Digital twins process all messages received from users separately and behave according to the message-defined function. However, not all messages are equally feasible for all IoT assets. While sensors and actuators implement firmware and state/effect messages, the edge node (network administrator) can remove or permit devices to join the network. For instance, if a new IoT asset is invited to onboard the network, the edge node temporarily allows new devices to join for 20 seconds by messaging *permitJoin(true,20)*. Last, we connect Eclipse Ditto to the Mosquitto MQTT broker to establish bidirectional communication between the digital twins and the IoT assets. On the one side, data received from the MQTT broker fills the pre-defined data schemata of the digital twins, and on the other side, digital twins can send commands to the IoT assets.

We opted for Eclipse Ditto because event-based middleware is most qualified for real-time data processing [6] and SOAR use cases. Eclipse Ditto implements the publish/subscribe approach with topics and events (see Figure 1). Nevertheless, there are several ways of implementing digital twins (e.g., physics-based modeling vs. data-driven techniques). Eclipse Ditto uses a data-driven technique with messages to represent IoT asset functions. This type of middleware fits SOAR best, as SOAR does not require simulation capabilities and other aspects of physics-based digital twins. Additionally, Eclipse Ditto is established and used by industrial companies (e.g., Bosch or Aloxy).

SOAR application. The SOAR platform application is deployed on the same virtual machine that runs Eclipse Ditto. We implemented the frontend of the SOAR platform using the Angular[6]

[3] https://www.zigbee2mqtt.io
[4] https://mosquitto.org
[5] https://www.eclipse.org/ditto
[6] https://angular.io

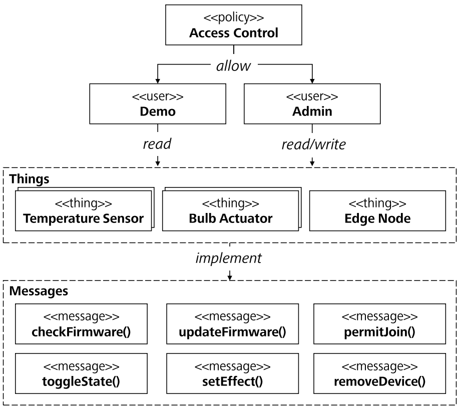


Figure 4: Digital twin setting in Eclipse Ditto

web application framework and Typescript[7]. The backend of our SOAR application is based on NodeJS[8] storing data in a MongoDB[9] database. Developing the SOAR application, we find microservice architecture to fit the purposes of SOAR best. Our SOAR4IoT implementation integrates four main microservices: core app (central microservice), Eclipse Ditto app, CTI app, and a SIEM app. The SIEM app generates pseudo-events used to trigger the execution of playbooks. The CTI app queries vulnerabilities, and the Eclipse Ditto app integrates IoT assets. For ease of deployment, we use Docker Compose and Docker Images. A detailed description of the SOAR platform features is described in Section 4.4.

Security tools. At last, our experimental setup includes the use of security tools. We pursue a twofold approach. First, we integrate existing CTI sources for security-relevant information. Thus, information about vulnerabilities in applications, hardware, or operating systems can be queried from the US National Vulnerability Database (NVD) and is structured by its Common Vulnerabilities and Exposures (CVE) enumeration. CVE descriptions are particularly relevant as attackers widely use available exploits for known vulnerabilities. Also, firmware update information can be queried for our actuators. Second, we directly include a security event feature. This feature is based on predefined security events representing SIEM alarms or incident notifications. Contrary to our experimental setting, organizations will integrate their existing SIEM systems or ticketing systems instead.

4.4 Results

Our research yields results concerning the demonstration of two IoT security use cases. Implementing our digital twins and IoT-centric SOAR application enables security workflows based on user interface (UI) and playbook execution.

We created three playbooks, of which two are addressing the Sybil attack and one the Mirai botnet use case. Therefore, our UI[10] includes an intuitive playbook editor for configuration. In general, the UI of our SOAR application follows a minimalistic approach and provides a single point of contact. Figure 5 documents three main views: (a) security event list, (b) IoT assets (digital twins), and (c) playbooks. Our digital twin and security-focused UI goes beyond the generic Mosquitto UI and the Ansible Semaphore UI[11]. We reason that designing and implementing a customized SOAR application along SOAR requirements is feasible with open-source technologies.

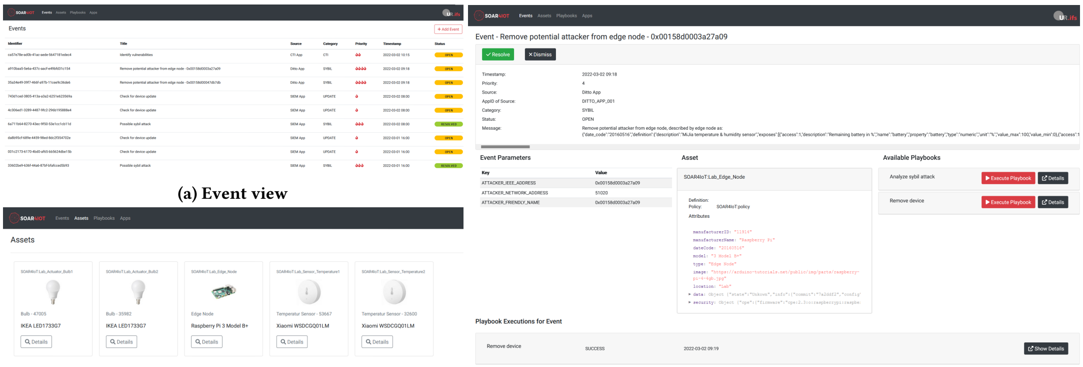
We define a generic `SOAR4IoT` workflow to showcase playbook execution. The workflow involves IoT assets (digital twins), apps, actions, playbooks, and events. Apps (i.e., individual microservices) implement specific actions (e.g., API calls) relevant for security operations. These actions are then structured and instantiated within playbooks. At last, given a specific security event (received by app or created via the UI), playbook execution is triggered. Playbook execution is dependent on event parameters and matching logic. As events are linked to IoT assets, matched playbooks must refer to the same IoT assets. During playbook execution the SOAR core service checks an app's availability, documents action status and starts subsequent actions. The playbook status indicates success, timeout or failure.

The *Mirai playbook* is used for vulnerable IoT assets (e.g., missing updates or default passwords). Its actions include fetching CTI data, updating IoT assets OTA, and requesting analysts to check the IoT assets' authentication manually. Our experimental setup includes no vulnerable IoT assets, so we define a repetitive update event. This event triggers playbook execution regularly. We successfully achieved firmware updates for the IKEA Tradfri LED bulb using digital twin messaging functions and Zigbee2MQTT. Changing authentication and validating playbook execution (e.g., comparing firmware versions) are subsequent manual tasks.

The *Sybil playbooks* address rogue devices. The actions include identifying and removing Sybil nodes from the network. This is followed by preventing new devices to join the network. Leveraging our SIEM app, we create events indicating a possible Sybil attack. In `SOAR4IoT`, the security analyst can then execute a playbook to analyze IoT assets not represented by a digital twin. If so, new removal events are created and listed with the asset's network address (see Figure 5a). A security analyst can also check manually if the network address is linked to a known IoT asset (see Figure 5b). The analyst is assisted in resolving the removal event by executing the corresponding playbook (see Figure 5c). Observing the status of playbook execution, the Sybil node is successfully removed. Validation might include comparing connected IoT assets at the edge node before and after playbook execution. In general, playbook selection depends on analyst's assessment of whether a playbook's actions meet the desired objective.

Lessons Learned. We learned that a logical separation of security orchestration and incident response using microservices benefits the SOAR application. We consider security orchestration a data collection task (e.g., querying device status or available CTI) and

[7] https://www.typescriptlang.org
[8] https://nodejs.org
[9] https://www.mongodb.com
[10] https://soar4iot.ur.de
[11] https://github.com/ansible-semaphore/semaphore



Events

Identifier	Title	Source	Category	Priority	Timestamp	Status
ca57e78e-ad0b-41ac-aede-5647181edec4	Identify vulnerabilities	CTI App	CTI	🔥🔥	2022-03-02 10:15	OPEN
a910baa5-5e6a-437c-aacf-e49bfd31c154	Remove potential attacker from edge node - 0x00158d0003a27a09	Ditto App	SYBIL	🔥🔥🔥🔥	2022-03-02 09:18	OPEN
35a34e49-39f7-466f-a97b-11cee9c36de6	Remove potential attacker from edge node - 0x00158d00047db7db	Ditto App	SYBIL	🔥🔥🔥🔥	2022-03-02 09:18	OPEN
743d1ced-3805-413a-a3a2-6251e623569a	Check for device update	SIEM App	UPDATE	🔥	2022-03-02 08:00	OPEN
4c306ed1-3289-4487-9fc2-296b195888e4	Check for device update	SIEM App	UPDATE	🔥	2022-03-02 08:00	OPEN
6a711b64-8270-43ec-9f50-53e1cc1cb11d	Possible sybil attack	SIEM App	SYBIL	🔥🔥🔥	2022-03-02 08:00	RESOLVED
da8b95cf-689e-4459-98ed-8dc2f354702e	Check for device update	SIEM App	UPDATE	🔥	2022-03-01 16:00	OPEN
001c2173-6170-4bd0-af65-bb5624dbe15b	Check for device update	SIEM App	UPDATE	🔥	2022-03-01 16:00	OPEN
33602be9-636f-44a6-87bf-bfafcced5b93	Possible sybil attack	SIEM App	SYBIL	🔥🔥🔥	2022-03-01 16:00	RESOLVED

(a) Event view

Assets

SOAR4IoT:Lab_Actuator_Bulb1	SOAR4IoT:Lab_Actuator_Bulb2	SOAR4IoT:Lab_Edge_Node	SOAR4IoT:Lab_Sensor_Temperature1	SOAR4IoT:Lab_Sensor_Temperature2
Bulb - 47005	Bulb - 35982	Edge Node	Temperatur Sensor - 53667	Temperatur Sensor - 32600
IKEA LED1733G7	IKEA LED1733G7	Raspberry Pi 3 Model B+	Xiaomi WSDCGQ01LM	Xiaomi WSDCGQ01LM
Details	Details	Details	Details	Details

Event - Remove potential attacker from edge node - 0x00158d0003a27a09

✓ Resolve ✕ Dismiss

Timestamp: 2022-03-02 09:18
Priority: 4
Source: Ditto App
AppID of Source: DITTO_APP_001
Category: SYBIL
Status: OPEN
Message: Remove potential attacker from edge node, described by edge node as:
{"date_code":"20160516","definition":{"description":"MiJia temperature & humidity sensor","exposes":[{"access":1,"description":"Remaining battery in %","name":"battery","property":"battery","type":"numeric","unit":"%","value_max":100,"value_min":0},{"access":1

Event Parameters

Key	Value
ATTACKER_IEEE_ADDRESS	0x00158d0003a27a09
ATTACKER_NETWORK_ADDRESS	51020
ATTACKER_FRIENDLY_NAME	0x00158d0003a27a09

Asset

SOAR4IoT:Lab_Edge_Node

Definition:
Policy: SOAR4IoT:policy
Attributes

```
manufacturerID: "11914"
manufacturerName: "Raspberry Pi"
dateCode: "20160516"
model: "3 Model B+"
type: "Edge Node"
image: "https://arduino-tutorials.net/public/img/parts/raspberry-
pi-4-4gb.jpg"
location: "Lab"
▸ data: Object {"state":"Unkown","info":{"commit":"7a2ddf2","config'
▸ security: Object {"cpe":{"firmware":"cpe:2.3:o:raspberrypi:raspbe:
```

Available Playbooks

Analyze sybil attack ▶ Execute Playbook Details
Remove device ▶ Execute Playbook Details

Playbook Executions for Event

Remove device SUCCESS 2022-03-02 09:19 Show Details

(b) Asset view

(c) Playbook execution view

Figure 5: SOAR4IoT UI

incident response a modification task. Digital twins prove relevant as they provide an additional layer with unified access and control to make the IoT manageable for security purposes. We experienced excellent feedback from the Eclipse Ditto community during development. It leads us to conclude that, in practice, digital twins go beyond the functional scope of IoT platforms, and digital twin research is relatively narrow. Overall, SOAR application development is a challenging task, but complexity can be reduced (e.g., via microservices, virtualization, and deployment tools).

4.5 Discussion

We discuss both the scientific and practical impact of our SOAR4IoT framework before mentioning limitations.

Scientific impact. Only a few academic works have addressed security orchestration and SOAR platforms. Our work is an attempt towards leveling the playing field with the large number of commercial SOAR platforms. This attempt includes a list of SOAR platform features. Eventually, documented by our SOAR application, open-source technologies can be used. We contrast user reluctance with the potential use cases for security and open-source frameworks like Eclipse Ditto digital twins. We direct attention to digital twins for security operations beyond current simulations.

Practical impact. To cope with the current IoT trend, organizations must manage IoT assets and extend existing SOAR platforms. Our work can be seen as an innovative approach using open-source technologies. Pointing at the benefits of small-scale, customized SOAR platforms, we contrast commercial SOAR platforms. Our publicly available source code can serve for future extensions.

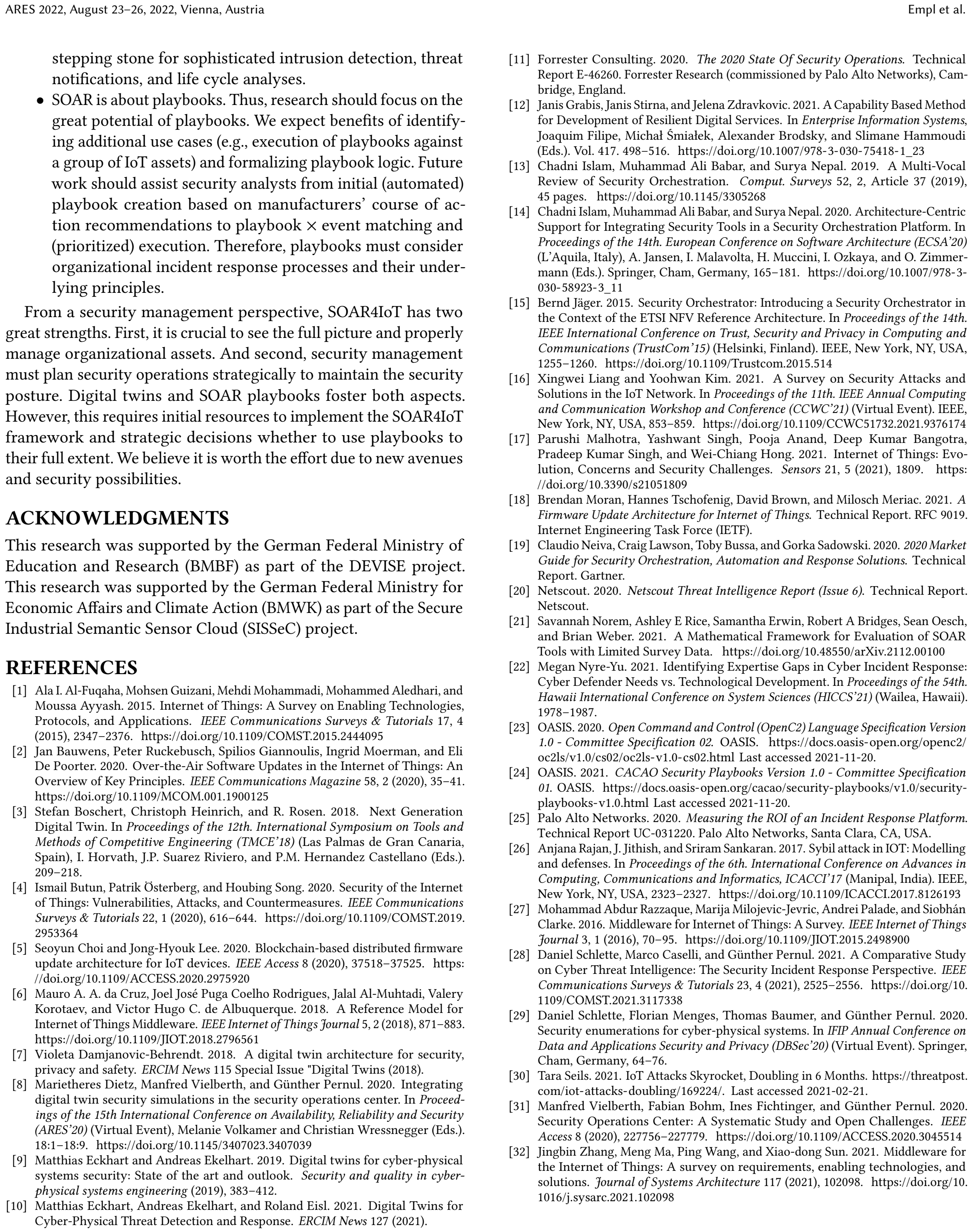
Limitations. There are several aspects that our work does not address. We attempted to select appropriate technologies and justify our decisions, but there are no best practices for digital twins in cybersecurity. CPS Twinning[12] is an alternative digital twin framework worth further investigation. Additionally, we excluded security for IoT cloud applications (e.g., predictive maintenance) typically used with IoT assets. Our SOAR application does not consider communication features (e.g., messaging or task delegation) found in commercial SOAR platforms. Access control, available for digital twins, is missing at SOAR application level but is required in production environments. Due to the small quantity of IoT assets, we can not assess the scalability of our SOAR application. Since many IoT devices will never experience updates, organizations should pay attention when buying them. Also, we did not exploit the full range of possibilities as our SOAR application integrates only a few security tools.

5 CONCLUSION AND FUTURE WORK

The question *"How to use Security Orchestration, Automation and Response for the Internet of Things?"* was the starting point of our work. While investigating the SOAR concept and SOAR platforms, we derived a detailed understanding of SOAR and its requirements. Defined by its orchestration, automation, and incident response capabilities, SOAR is mainly centered on playbooks and security tool integration for security operations. Extending the security operations to the IoT is a necessary step, as IoT attacks and IoT objectives show. Among different options to secure the IoT, digital twins provide a feasible, lightweight solution abstracting heterogeneous assets. Thus, our SOAR4IoT framework integrates a digital twin-based middleware. More precisely, we establish a prototypical implementation using Eclipse Ditto and a microservice SOAR application. Implications of our conceptual design and SOAR4IoT implementation include the following:

- Digital twins provide abstraction and a unified interface for the plethora of IoT assets. The security community should further compare different digital twin frameworks' abilities (e.g., advanced behavior or process modeling). To the best of our knowledge, our Eclipse Ditto implementation is the first, with security use cases built on top. It can serve as a

[12] https://github.com/sbaresearch/cps-twinning



stepping stone for sophisticated intrusion detection, threat notifications, and life cycle analyses.

- SOAR is about playbooks. Thus, research should focus on the great potential of playbooks. We expect benefits of identifying additional use cases (e.g., execution of playbooks against a group of IoT assets) and formalizing playbook logic. Future work should assist security analysts from initial (automated) playbook creation based on manufacturers' course of action recommendations to playbook × event matching and (prioritized) execution. Therefore, playbooks must consider organizational incident response processes and their underlying principles.

From a security management perspective, SOAR4IoT has two great strengths. First, it is crucial to see the full picture and properly manage organizational assets. And second, security management must plan security operations strategically to maintain the security posture. Digital twins and SOAR playbooks foster both aspects. However, this requires initial resources to implement the SOAR4IoT framework and strategic decisions whether to use playbooks to their full extent. We believe it is worth the effort due to new avenues and security possibilities.

ACKNOWLEDGMENTS

This research was supported by the German Federal Ministry of Education and Research (BMBF) as part of the DEVISE project. This research was supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) as part of the Secure Industrial Semantic Sensor Cloud (SISSeC) project.

REFERENCES

[1] Ala I. Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, and Moussa Ayyash. 2015. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Communications Surveys & Tutorials* 17, 4 (2015), 2347–2376. https://doi.org/10.1109/COMST.2015.2444095

[2] Jan Bauwens, Peter Ruckebusch, Spilios Giannoulis, Ingrid Moerman, and Eli De Poorter. 2020. Over-the-Air Software Updates in the Internet of Things: An Overview of Key Principles. *IEEE Communications Magazine* 58, 2 (2020), 35–41. https://doi.org/10.1109/MCOM.001.1900125

[3] Stefan Boschert, Christoph Heinrich, and R. Rosen. 2018. Next Generation Digital Twin. In *Proceedings of the 12th. International Symposium on Tools and Methods of Competitive Engineering (TMCE'18)* (Las Palmas de Gran Canaria, Spain), I. Horvath, J.P. Suarez Riviero, and P.M. Hernandez Castellano (Eds.). 209–218.

[4] Ismail Butun, Patrik Österberg, and Houbing Song. 2020. Security of the Internet of Things: Vulnerabilities, Attacks, and Countermeasures. *IEEE Communications Surveys & Tutorials* 22, 1 (2020), 616–644. https://doi.org/10.1109/COMST.2019.2953364

[5] Seoyun Choi and Jong-Hyouk Lee. 2020. Blockchain-based distributed firmware update architecture for IoT devices. *IEEE Access* 8 (2020), 37518–37525. https://doi.org/10.1109/ACCESS.2020.2975920

[6] Mauro A. A. da Cruz, Joel José Puga Coelho Rodrigues, Jalal Al-Muhtadi, Valery Korotaev, and Victor Hugo C. de Albuquerque. 2018. A Reference Model for Internet of Things Middleware. *IEEE Internet of Things Journal* 5, 2 (2018), 871–883. https://doi.org/10.1109/JIOT.2018.2796561

[7] Violeta Damjanovic-Behrendt. 2018. A digital twin architecture for security, privacy and safety. *ERCIM News* 115 Special Issue "Digital Twins (2018).

[8] Marietheres Dietz, Manfred Vielberth, and Günther Pernul. 2020. Integrating digital twin security simulations in the security operations center. In *Proceedings of the 15th International Conference on Availability, Reliability and Security (ARES'20)* (Virtual Event), Melanie Volkamer and Christian Wressnegger (Eds.). 18:1–18:9. https://doi.org/10.1145/3407023.3407039

[9] Matthias Eckhart and Andreas Ekelhart. 2019. Digital twins for cyber-physical systems security: State of the art and outlook. *Security and quality in cyber-physical systems engineering* (2019), 383–412.

[10] Matthias Eckhart, Andreas Ekelhart, and Roland Eisl. 2021. Digital Twins for Cyber-Physical Threat Detection and Response. *ERCIM News* 127 (2021).

[11] Forrester Consulting. 2020. *The 2020 State Of Security Operations.* Technical Report E-46260. Forrester Research (commissioned by Palo Alto Networks), Cambridge, England.

[12] Janis Grabis, Janis Stirna, and Jelena Zdravkovic. 2021. A Capability Based Method for Development of Resilient Digital Services. In *Enterprise Information Systems*, Joaquim Filipe, Michał Śmiałek, Alexander Brodsky, and Slimane Hammoudi (Eds.). Vol. 417. 498–516. https://doi.org/10.1007/978-3-030-75418-1_23

[13] Chadni Islam, Muhammad Ali Babar, and Surya Nepal. 2019. A Multi-Vocal Review of Security Orchestration. *Comput. Surveys* 52, 2, Article 37 (2019), 45 pages. https://doi.org/10.1145/3305268

[14] Chadni Islam, Muhammad Ali Babar, and Surya Nepal. 2020. Architecture-Centric Support for Integrating Security Tools in a Security Orchestration Platform. In *Proceedings of the 14th. European Conference on Software Architecture (ECSA'20)* (L'Aquila, Italy), A. Jansen, I. Malavolta, H. Muccini, I. Ozkaya, and O. Zimmermann (Eds.). Springer, Cham, Germany, 165–181. https://doi.org/10.1007/978-3-030-58923-3_11

[15] Bernd Jäger. 2015. Security Orchestrator: Introducing a Security Orchestrator in the Context of the ETSI NFV Reference Architecture. In *Proceedings of the 14th. IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom'15)* (Helsinki, Finland). IEEE, New York, NY, USA, 1255–1260. https://doi.org/10.1109/Trustcom.2015.514

[16] Xingwei Liang and Yoohwan Kim. 2021. A Survey on Security Attacks and Solutions in the IoT Network. In *Proceedings of the 11th. IEEE Annual Computing and Communication Workshop and Conference (CCWC'21)* (Virtual Event). IEEE, New York, NY, USA, 853–859. https://doi.org/10.1109/CCWC51732.2021.9376174

[17] Parushi Malhotra, Yashwant Singh, Pooja Anand, Deep Kumar Bangotra, Pradeep Kumar Singh, and Wei-Chiang Hong. 2021. Internet of Things: Evolution, Concerns and Security Challenges. *Sensors* 21, 5 (2021), 1809. https://doi.org/10.3390/s21051809

[18] Brendan Moran, Hannes Tschofenig, David Brown, and Milosch Meriac. 2021. *A Firmware Update Architecture for Internet of Things.* Technical Report. RFC 9019. Internet Engineering Task Force (IETF).

[19] Claudio Neiva, Craig Lawson, Toby Bussa, and Gorka Sadowski. 2020. *2020 Market Guide for Security Orchestration, Automation and Response Solutions.* Technical Report. Gartner.

[20] Netscout. 2020. *Netscout Threat Intelligence Report (Issue 6).* Technical Report. Netscout.

[21] Savannah Norem, Ashley E Rice, Samantha Erwin, Robert A Bridges, Sean Oesch, and Brian Weber. 2021. A Mathematical Framework for Evaluation of SOAR Tools with Limited Survey Data. https://doi.org/10.48550/arXiv.2112.00100

[22] Megan Nyre-Yu. 2021. Identifying Expertise Gaps in Cyber Incident Response: Cyber Defender Needs vs. Technological Development. In *Proceedings of the 54th. Hawaii International Conference on System Sciences (HICCS'21)* (Wailea, Hawaii). 1978–1987.

[23] OASIS. 2020. *Open Command and Control (OpenC2) Language Specification Version 1.0 - Committee Specification 02.* OASIS. https://docs.oasis-open.org/openc2/oc2ls/v1.0/cs02/oc2ls-v1.0-cs02.html Last accessed 2021-11-20.

[24] OASIS. 2021. *CACAO Security Playbooks Version 1.0 - Committee Specification 01.* OASIS. https://docs.oasis-open.org/cacao/security-playbooks/v1.0/security-playbooks-v1.0.html Last accessed 2021-11-20.

[25] Palo Alto Networks. 2020. *Measuring the ROI of an Incident Response Platform.* Technical Report UC-031220. Palo Alto Networks, Santa Clara, CA, USA.

[26] Anjana Rajan, J. Jithish, and Sriram Sankaran. 2017. Sybil attack in IOT: Modelling and defenses. In *Proceedings of the 6th. International Conference on Advances in Computing, Communications and Informatics, ICACCI'17* (Manipal, India). IEEE, New York, NY, USA, 2323–2327. https://doi.org/10.1109/ICACCI.2017.8126193

[27] Mohammad Abdur Razzaque, Marija Milojevic-Jevric, Andrei Palade, and Siobhán Clarke. 2016. Middleware for Internet of Things: A Survey. *IEEE Internet of Things Journal* 3, 1 (2016), 70–95. https://doi.org/10.1109/JIOT.2015.2498900

[28] Daniel Schlette, Marco Caselli, and Günther Pernul. 2021. A Comparative Study on Cyber Threat Intelligence: The Security Incident Response Perspective. *IEEE Communications Surveys & Tutorials* 23, 4 (2021), 2525–2556. https://doi.org/10.1109/COMST.2021.3117338

[29] Daniel Schlette, Florian Menges, Thomas Baumer, and Günther Pernul. 2020. Security enumerations for cyber-physical systems. In *IFIP Annual Conference on Data and Applications Security and Privacy (DBSec'20)* (Virtual Event). Springer, Cham, Germany, 64–76.

[30] Tara Seils. 2021. IoT Attacks Skyrocket, Doubling in 6 Months. https://threatpost.com/iot-attacks-doubling/169224/. Last accessed 2021-02-21.

[31] Manfred Vielberth, Fabian Bohm, Ines Fichtinger, and Günther Pernul. 2020. Security Operations Center: A Systematic Study and Open Challenges. *IEEE Access* 8 (2020), 227756–227779. https://doi.org/10.1109/ACCESS.2020.3045514

[32] Jingbin Zhang, Meng Ma, Ping Wang, and Xiao-dong Sun. 2021. Middleware for the Internet of Things: A survey on requirements, enabling technologies, and solutions. *Journal of Systems Architecture* 117 (2021), 102098. https://doi.org/10.1016/j.sysarc.2021.102098

P7 Generating ICS Vulnerability Playbooks with Open Standards

Status:	Published	
Date of Submission:	07 June 2023	
Date of Acceptance:	11 September 2023	
Date of Publication:	11 December 2023	
Journal:	International Journal of Information Security	
Authors' Contributions:		
	Philip Empl	40%
	Daniel Schlette	40%
	Lukas Stöger	10%
	Günther Pernul	10%
Full Citation:	Empl, P., Schlette, D., Stöger, L., & Pernul, G. (2023). Generating ICS Vulnerability Playbooks with Open Standards. *International Journal of Information Security 23*(2), 1215–1230. Springer.	
DOI:	10.1007/s10207023007605	
Artifact:	github.com/ad2play	

Journal Description: The International Journal of Information Security is an English language international journal on research in information security. Information security builds on computer security and applied cryptography, but also reaches out to other branches of the information sciences. Information security is an important aspect of protecting the information society from a wide variety of threats. In this new century, The International Journal of Information Security will provide prompt publication of important technical work in information security, whether theoretical, applicable, or related to implementation.

International Journal of Information Security (2024) 23:1215–1230
https://doi.org/10.1007/s10207-023-00760-5

REGULAR CONTRIBUTION

Generating ICS vulnerability playbooks with open standards

Philip Empl[1] · Daniel Schlette[1] · Lukas Stöger[2] · Günther Pernul[1]

Published online: 11 December 2023

Abstract

Organizations face attacks on industrial control systems (ICS) as vulnerabilities are pervasive. However, patching vulnerable systems by simply updating to the newest version is often not an option and shifts focus to workarounds. Beyond pure patching, workarounds specify other remediation measures (e.g., firewall or VPN configuration) that must be taken due to system availability requirements, complexity, or heterogeneous devices. In this paper, we introduce vulnerability playbooks based on open standards. Pushing the envelope of cybersecurity playbooks—steps organizations should follow when responding to cybersecurity incidents reactively—for ICS vulnerability management offers organizations a more transparent, repeatable process and faster, possibly automated actions. We have designed a process model to collect and transform security advisories in *Common Security Advisory Framework* (CSAF) format and generate *Collaborative Automated Course of Action Operations* (CACAO) playbooks based on listed remediation advice. With a proof of concept, we demonstrate that structured CSAF documents can be seamlessly transformed into CACAO playbooks. For our industrial use case, we must also use unstructured security advice highlighting quality differences (compared to CSAF). Our generated 79 standard conformant CACAO playbooks with 485 identified actions hint at imbalanced advice toward patching. Preferably, vendors should include detailed technical remediation advice, provide APIs, and go beyond patching recommendations in their security advisories. Subscribers should structure their assets and use machine learning to normalize, generate, and prioritize CACAO playbooks. With CSAF and CACAO, we see two open standards for handling vulnerabilities.

Keywords Vulnerability playbook · Security advisory · Industrial control system · CSAF · CVRF · CACAO

1 Introduction

Cybersecurity playbooks are about knowing what to do when insecurity becomes apparent. The heavily promoted notion of playbooks captures the description of organizational processes, specified workflows, and individual actions. Security Orchestration, Automation and Response (SOAR) tools rely on playbooks [1], and the US government, special interest groups, and researchers are eager to develop playbooks [2–4]. With industry support, the *Collaborative Automated Course of Action Operations (CACAO)* playbook format aims to standardize playbooks upholding the principle of open standards [5].

Existing playbooks often address incident types (e.g., phishing or malware), and research is focused on incident response [4]. However, using playbooks to handle specific vulnerabilities is another promising field that vulnerability management tools have only partially explored [6]. Industrial Control System (ICS) vulnerability playbooks—steps organizations should follow when dealing with vulnerabilities proactively—can fill the gap and provide additional remediation advice to organizations. We make a first approach to answer the question: Is it possible to generate ICS vulnerability playbooks?

Our work focuses on vulnerability playbooks for ICS and the industrial Internet of Things (IoT). These systems are affected by numerous vulnerabilities and countless attacks. For instance, according to the National Vulner-

Philip Empl
philip.empl@ur.de

Daniel Schlette
daniel.schlette@ur.de

Lukas Stöger
lukas.stoeger@dehn.de

Günther Pernul
guenther.pernul@ur.de

[1] University of Regensburg, Universitätsstr. 31, 93053 Regensburg, Germany

[2] Dehn SE, Hans-Dehn-Straße 1, 92318 Neumarkt in der Oberpfalz, Germany

ability Database (NVD), 72 vulnerabilities for SIMATIC S7 were discovered in the last ten years and caused the vendor to issue patches and security advisories. Moreover, ICS high-availability requirements, complexity, and many heterogeneous devices complicate (manual) vulnerability management and demand measures beyond updating [7]. Thus, ICS vendors typically offer security advisories detailing workarounds for remediation when system availability is a must and patching is not a direct option. In addition, the US Cybersecurity & Infrastructure Security Agency (CISA) maintains a collection of ICS advisories [8].

Looking at security advisories, we see different vendors use different data formats. One such format is the *Common Security Advisory Framework (CSAF)*, an open standard foreseen to exchange machine-readable information [9]. It includes a dedicated section on remediation options which builds the basis for our streamlined, automated vulnerability playbook generation. Organizations can benefit from ICS vulnerability playbooks by reducing the manual handling of workarounds in multiple ways. Most notably, organizations can limit error-prone information extraction and structuring. Automating the process further increases process transparency and data provenance. These improvements are typically associated with playbooks, which leads us to create vulnerability playbooks based on security advisories.

In this work, we design and implement a process model on top of open standards for security advisories (i.e., CSAF) and playbooks (i.e., CACAO) to generate ICS vulnerability playbooks. We aim at demonstrating the practical benefits of structured security advisories making both security advisory publishers and consumers aware of this. In particular, we leverage public advisory sources and preprocess their data. Thereby, we model devices representing Siemens and Cisco assets. In our proof of concept implementation, we query security advisories from two leading ICS vendors and CISA relevant to our use case. In total, we generate 79 vulnerability playbooks and identify 485 workflow actions. Matching terms, which can be customized, are used to classify playbook steps containing the workflow actions. Our main contributions are:

- A process model for generating ICS vulnerability playbooks. The process model covers four phases: (1) querying vulnerability information, (2) sourcing security advisories, (3) converting data in CSAF, and (4) leveraging matching terms to create CACAO playbooks with workflow actions.
- An open-source application[1] to generate vulnerability playbooks with open standards and industry use case.
- Recommendations for improvement and use of security advisory and playbook standards.

The paper is structured as follows. In Sect. 2, we present a motivating ICS vulnerability and the associated security advisory. Additional background on open standards for vulnerabilities, incident response playbooks, and related work is part of Sect. 3. Section 4 details our process model automating the creation of vulnerability playbooks for ICS. Then, we evaluate our approach with a use case and open-source tool implementation in Sect. 5. In Sect. 6, we outline recommendations for better vulnerability handling. In Sect. 7, we conclude with future research directions.

2 Motivation

We illustrate the representation of security advisories with a highly critical (CVSS[2] base score of 10) ICS vulnerability affecting Siemens SIMATIC CP devices, communication processors used in digital factories [10]. Identified by CVE-2022-34819, the vulnerability centers on improper input validation and the resulting heap-based buffer overflow. As a consequence, attackers could execute malicious code and cause production to halt. We use this vulnerability to emphasize aspects of ICS security advisories and their representation in CSAF format.

Figure 1 shows the abbreviated CSAF document. Upfront metadata inform about the CSAF format and the security advisory publisher, typically the vendor of the affected product(s). A string-based title is used to refer to the security advisory. However, product users are mostly interested in security advisories to extract relevant information on vulnerability remediation. Therefore, crucial remediation advice in CSAF is listed inside a remediations array. Besides vendor fixes instructing to update to the newest version (omitted for brevity), Fig. 1 details workarounds as alternative remediation steps. These workarounds help to harden SIMATIC CP devices until patches are installed. In the example CSAF, these include blocking access to a specific port by using an external firewall and disabling a VPN feature.

In the following, we elaborate on data quality issues concerning security advisories and possibilities of CACAO playbooks. Security advisories and (if available) their CSAF documents do not always contain detailed and executable information. The CSAF example in Fig. 1 represents a best-case scenario. Subscribers are faced with security advisories in various data formats, which are often not machine-readable.

Generating CACAO playbooks (Un)structured security advisories, e.g., CSAF, cannot be automatically applied because they have to be put into an executable format

[1] https://www.github.com/ad2play/ad2play.

[2] Common Vulnerability Scoring System (CVSS) is a standardized framework used to assess and communicate the severity of software vulnerabilities.

and mapped to ICS assets. Considering also heterogeneous devices, multiple vulnerabilities, and security advisory sources, automated vulnerability playbook generation is evident. In contrast to manual advisory processing, process consistency can be improved. For instance, the manual vulnerability handling is error-prone or takes even more time. We would like to emphasize that the best-case scenario is not always given. Unfortunately, organizations currently use proprietary data formats to represent playbooks, even though OASIS has published an open playbook standard, CACAO. CACAO promotes standardization and interoperability, allowing seamless integration of different cybersecurity tools. By enabling automation, it reduces response times and manual intervention in vulnerability management. Its human-readable format ensures ease of use and customization, while its integration with various tools streamlines orchestration and automation, ultimately enhancing an organization's overall cybersecurity defense. However, when automatically creating CACAO playbooks from security advisories, we must also deal with unstructured remediation advice until the CSAF standard is established across the industry.

3 Background and related work

Open standards for vulnerability management and incident response playbooks represent foundations for our work. We further discuss related work within this section.

3.1 Open security standards

Vulnerability management relies on a shared understanding of concepts. Open security standards provide the means to cope with low information quality by assisting with uniform representation and content structure. The following standards and data formats are widely recognized in cybersecurity and help organizations handle vulnerabilities.

CVE [11]—Common Vulnerability Enumeration, used to identify and describe vulnerabilities.

CPE [12]—Common Platform Enumeration, used to identify IT/OT assets.

CVSS [13]—Common Vulnerability Scoring System, used to define and assign severity scores.

CVRF/CSAF [14]—Common Vulnerability Reporting Framework/Common Security Advisory Framework, used to describe security advisories.

The open standards and data formats are intended to inform others about vulnerabilities, exploits, and remediation advice [15]. They answer the crucial questions: What characterizes a vulnerability? What systems are affected? How severe is the vulnerability? And what do others need to know about vulnerability remediation?

```
{
  "document": {
    "category": "csaf_security_advisory",
    "csaf_version": 2.0,
    "publisher": {
      "category": "vendor",
      "name": "Siemens ProductCERT"

    },
    "title": "SSA-517377: Multiple
        Vulnerabilities in the SRCS VPN
        Feature in SIMATIC CP Devices"
  },
  "vulnerabilities": [
    {
      "cve": "CVE-2022-34819",
      "remediations": [
        {
          "category":
          "workaround",
          "details": "Block access to port
              5243/udp e.g. with an external
               firewall if possible"
        },
        {
          "category": "workaround",
          "details": "Disable the SINEMA
              Remote Connect Server (SRCS)
              VPN feature"
        }
      ] ...
```

Fig. 1 Excerpt of a CSAF document for CVE-2022-34819 with remediation advice that specifies two workarounds

3.2 Incident response playbooks

Organizations need to define processes, procedures, and actions for incident response. Threat intelligence is also necessary to handle security incidents [16]. Thus, incident response representations with playbooks bridge the gap between processes and data containing both [17]. Mainly two major use cases—the automation of incident response and the sharing of playbooks—have resulted in the development of open standards and data formats (e.g., CACAO, OpenC2, MITRE D3FEND, or RE&CT) for playbooks and individual actions [18–20].

CACAO. Collaboratively developed by the Organization for the Advancement of Structured Information Standards (OASIS) and its members, the open CACAO format targets playbooks. In contrast to more action-focused standards, CACAO playbooks can describe information on various granularity levels. As a result, the CACAO format is comprehensive and a promising candidate for the description of vulnerability playbooks. To the best of our knowledge, there are no other maintained and open playbook standards with similar characteristics. Using CACAO playbooks, organizations can follow defined workflows and have the ability to automate repetitive, error-prone tasks.

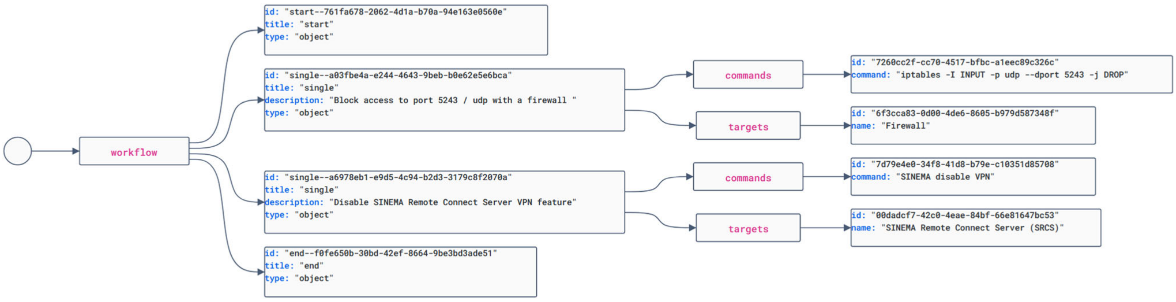

workflow
id: "start--761fa678-2062-4d1a-b70a-94e163e0560e"
title: "start"
type: "object"
id: "single--a03fbe4a-e244-4643-9beb-b0e62e5e6bca"
title: "single"
description: "Block access to port 5243 / udp with a firewall "
type: "object"
id: "single--a6978eb1-e9d5-4c94-b2d3-3179c8f2070a"
title: "single"
description: "Disable SINEMA Remote Connect Server VPN feature"
type: "object"
id: "end--f0fe650b-30bd-42ef-8664-9be3bd3ade51"
title: "end"
type: "object"
commands
targets
commands
targets
id: "7260cc2f-cc70-4517-bfbc-a1eec89c326c"
command: "iptables -I INPUT -p udp --dport 5243 -j DROP"
id: "6f3cca83-0d00-4de6-8605-b979d587348f"
name: "Firewall"
id: "7d79e4e0-34f8-41d8-b79e-c10351d85708"
command: "SINEMA disable VPN"
id: "00dadcf7-42c0-4eae-84bf-66e81647bc53"
name: "SINEMA Remote Connect Server (SRCS)"

Fig. 2 Schematic visualization of a CACAO vulnerability playbook that includes workflow, command, and target objects

The benefits of the CACAO playbook format are best understood by looking at its structure and object types. Figure 2 shows the visualization of a vulnerability playbook for CVE-2022-34819 and the underlying attribute–value pairs in JSON. The playbook is based on real-world vendor remediation advice augmented with commands. CACAO playbooks contain workflows to structure workflow steps. Typically, start and end steps enclose single action steps outlining specific actions. On a more granular level, command and target objects describe executable information and its recipients. In the example, organizations can derive two remediation actions, systems involved (i.e., firewall, server), and commands (i.e., iptables, disable). CACAO is broad in scope, and command and target types also support manual actions for individuals. Adding conditional workflow steps helps to represent sophisticated workflows. We use CACAO as it can capture multiple CSAF-based remediation measures and hand action-based workflows to organizations. In the remainder of this paper, we refer to CACAO workflow steps as workflow actions to differentiate between CSAF and CACAO.

3.3 Related work

Vulnerabilities and vulnerability management are of interest to researchers and organizations alike. Organizations are advised to systematically handle vulnerabilities as they can lead to severe security incidents [21]. A steady stream of research covers general and ICS-specific vulnerabilities [22, 23]. From a management perspective, CISA provides a so-called vulnerability response playbook to assist organizations in deciding about vulnerability handling [6]. From vulnerability discovery [24], to vulnerability assessment [25] and security advisories [26], transparent processes and standards are important. While Fenz et al. [27] introduce automated handling of security advisories, other work has taken on the challenge to provide commit-level patch advice for vulnerabilities [28]. Besides, vulnerability management is of practical interest as vendors of commercial vulnerability management tools address the need to keep track of assets and vulnerabilities.

Academic work on cybersecurity playbooks is sparse. Nevertheless, playbooks are an emerging research topic related to threat intelligence and security standards [17]. In a recent study, Stevens et al. [4] explored human playbook creation with available frameworks indicating playbook variety. As different approaches and sharing use cases exist, integration and semantics of playbooks are investigated [29–31]. Against the backdrop of security orchestration and a plethora of commercial SOAR tools [1], specific use cases (e.g., an IoT context with digital twins) have been discussed [32, 33]. It can be seen that playbook generation is crucial to leverage SOAR tools.

We go beyond related work in the following ways. Our approach is the first to combine the two areas of security advisories and playbooks. Building vulnerability playbooks offers organizations more process-oriented advice on what to do. While some vulnerability management tools incorporate the idea of playbooks, we see benefits in following a similar path with open security standards. Our focus on ICS security advisories capitalizes on the fact that remediation measures are most important when simply patching is not an option. Playbooks can introduce transparent processes and automation toward better vulnerability management for ICS.

4 Vulnerability playbook generation

Driven by the problem of ICS vulnerability handling and inspired by related works, we develop a process model. Our approach follows the design science research methodology [34] starting with a problem and developing an artifact to be evaluated and communicated. Our artifact is a process model that aims for a complete output (playbooks). From *security advisory to vulnerability playbook*, the process captures automated ICS vulnerability playbook generation with four phases shown as a Business Process Model and Notation (BPMN) diagram in Fig. 3. The subsequent sections are dedicated to the illustrated process phases.

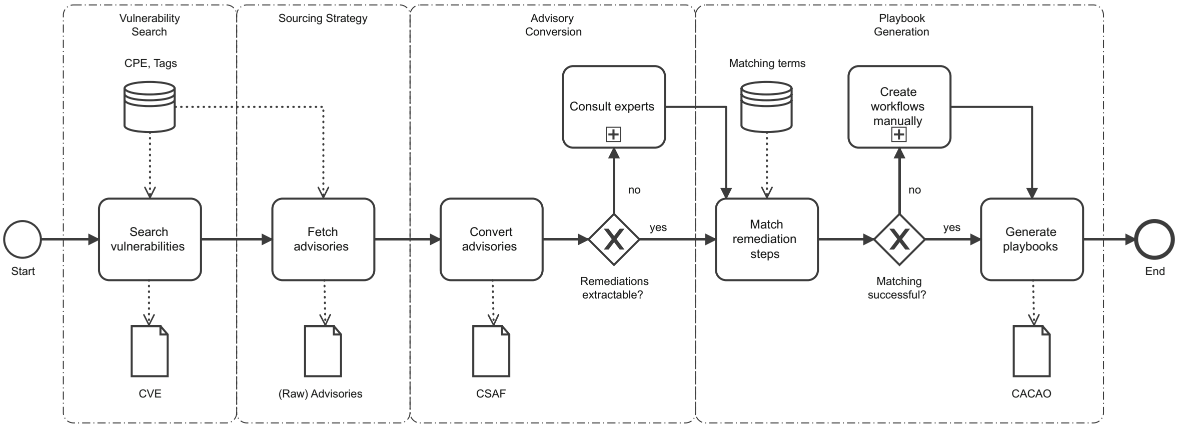


Fig. 3 Process description from security advisory to vulnerability playbook

4.1 Vulnerability search

Vulnerability handling and the playbook generation process start with assets and the question of whether these assets are vulnerable or not. Thus, we define the activity *Search vulnerabilities* to get an overview of relevant vulnerabilities. As a prerequisite, organizations must already carefully document their assets and components (e.g., virtual representations or SBOM—Software Bill of Materials). Using this documentation, assets and respective identifiers (e.g., CPE-ID) are used to find vulnerabilities. However, the specific characteristics of ICS need to be considered. Most notably, ICS assets are built of multiple components forming complex systems-of-systems [35]. Each of the components can run its own software on dedicated hardware. Searching for relevant security vulnerabilities requires identifiers—for the vulnerability and the components. Vulnerabilities are given a CVE-ID. ICS components (i.e., hardware or software) have a CPE-ID or other tags. If a component is described by CPE, querying associated CVEs is straightforward. Without CPE, other information (e.g., model or version) must be used to search vulnerability databases. In our process model intended for automation, we rely on CPE or, when not available, use device-specific tags. Both approaches enable automated vulnerability searches, but using tags might lead to more errors. The first phase yields vulnerabilities regardless of available security advisories.

4.2 Sourcing strategy

The activity *Fetch advisories* is part of the second process phase. Our sourcing strategy involves security advisory acquisition from product vendors. These product vendors often have Product Security Incident Response Teams (PSIRTs/ProductCERTs) that offer vulnerability remediation advice for their products. In addition, other institutions (e.g., national or coordination Computer Emergency Response Teams, CERTs/CSIRTs) and commercial security vendors partially aggregate security advice. In most cases, security advisories can be fetched with CPEs or tags and link to CVEs. Focusing on ICS and fetching security advisories for systems-of-systems involves multiple sources varying in format, structure, and content. We compared security advisory publishers and data formats. Most sources provide access to PDF security advisories or embed these directly on their website. In the best-case scenario, we find dedicated formats such as CSAF or its predecessor CVRF, but they can be retrieved less often. Formats like HTML or PDF require deep traversal and scraping. Since both formats do not provide options to directly map assets to remediation advice, we need to filter whether the remediation advice actually affects devices of interest. Therefore, we recommend the device representations to skim and filter these documents automatically. As a means of communication, many of the listed sources offer RSS feeds, email notifications, or communicate the latest advice via Twitter.

The various communication channels do not solve the problem of data heterogeneity and do not always allow the exchange of remediation advice and feedback. Additionally, many sources do not provide an API to fetch security advisories for specific vulnerabilities. Automating the filtering of RSS feeds or emails to match the advisories of interest is an unnecessarily complex intermediate step. Organizations relying on different sources and advisory formats must convert and standardize these advisories to enable automation.

4.3 Advisory conversion

The activity *Convert advisories* targets standardization and the results are uniform security advisories. Since different

Table 1 Action classification and related terms based on the OpenC2 commands

Class	Terms
Update	[[“patch”, “update”, “upgrade”], [“version”, “v”, “ver”]]
Investigation	[[“investigation”, “investigate”, “scan”, “examine”, “inspect”, “inspection”, “review”, “check”]]
Locating	[[“locate”, “find”, “detect”, “discover”, “uncover”], [“object”, “artifact”, “file”, “directory”, “instance”]]
Data operation	[[“query”, “create”, “alter”, “delete”, “copy”], [“data”, “entity”, “directory”, “file”]]
Isolation	[[“contain”, “containment”, “isolation”, “avoid”], [“file”, “process”, “entity”, “asset”]]
Privileges	[[“access”, “credentials”, “right”], [“allow”, “restrict”, “grant”, “assign”, “give”, “permit”, “reduce”, “regulate”, “block”, “limit”]]
System	[[“start”, “stop”, “restart”, “cancel”, “enable”, “disable”], [“process”, “application”, “system”, “activity”, “action”, “environment”, “function”, “port”]]
Configuration	[[“set”, “change”, “apply”, “put”, “restore”], [“value”, “configuration”, “state”, “property”, “attribute”]]
Network	[[“redirect”, “switch”, “block”, “intercept”], [“traffic”, “destination”, “url”, “ip”, “port”, “address”, “packet”, “network”]]
Observation	[[“detonate”, “execute”, “observe”, “examine”, “monitor”, “discover”], [“behaviour”, “malware”, “target”, “action”, “attack”, “activity”]]

vendors use different formats for representing and sharing security advisories, it is essential to convert these heterogeneous security advisories into a uniform format before generating playbooks. We rely on the open standard CSAF for the structuring and presentation of remediation steps. Thus, it is the objective of this activity to convert security advisories into CSAF documents.

There are three possible cases. First, CVRF is converted to CSAF using semantically identical fields to store remediation steps. Second, when security advisories are provided as CSAF documents, they are not converted and taken as is. However, we discard remediation steps not matching the CPE identifiers or tags. Last, also other source-specific types of security advisory formats are converted. Here, remediation advice needs to be extracted. Dependent on the data format, steps may include HTML/PDF parsing and scraping to identify and extract nested remediation steps. For instance, in the case of CISA security advisories, we suggest to extract the mitigation section, the executive summary, and the technical details besides relevant metadata, i.e., title, date, or URL. Note that these steps require logic to filter the remediation advice as unstructured data do not maintain a reliable mapping between remediation advice and devices of interest. In a scenario where no remediation advice is available, we include user interaction and consult experts. Possible calls to action include the search for internal playbooks targeting similar vulnerabilities. These playbooks might provide remediation steps that can fit the currently investigated vulnerability. Regardless of the scenario, this process phase results in standardized security advisories with remediation steps for vulnerability playbook generation.

4.4 Playbook generation

After unifying security advisories, we move from security advisories to playbooks involving activities to *Match remediation steps* and *Generate playbooks*. We rely on the open standard CACAO for structuring playbook-related information and workflow actions. In CSAF, remediation steps are mostly textual descriptions that are not actionable. We aim at deriving workflow actions for playbook execution. Thereby, we take remediation steps from CSAF, classify them, and put appropriate predefined actions into the CACAO workflow action section.

We introduce the concept of *matching terms* deciding about the class of a specific workflow action. In advance, organizations must define and assign action templates to specific classes. This allows playbooks to be dynamically populated with respective actions matching a class. These matching terms resemble a two-dimensional search on remediation steps. One dimension describes the action and the other dimension the target. Matching both dimensions is essential to meet a stemmed matching term fully. A given string must at least match one word per dimension to reach the next one. With the approach, it is possible to define complex matching terms. We initially define the action classes based on individual actions (e.g., create, update, or delete) proposed by the OASIS OpenC2 standard. Note that organizations are flexible in their choice of classification, mapping, and creation of specific action templates; OpenC2 is only one option to classify. The classification helps to tag the playbook accordingly. While these actions are used in our work to classify workflow steps, CACAO command objects can capture OpenC2 commands supporting agnostic automation. Table 1 shows possible action classes and related matching terms. For full automation, organizations must create and assign action templates to specific classes and matching terms to populate the playbook dynamically. These action templates might be selected based on the matched terms and dynamically fed by variables (e.g., port = 5243 and target = firewall).

Applying matching terms to remediation steps requires disassembling these steps into sentences and understanding their intention. Natural language processing (NLP) is an accepted method to process and understand human-readable language. Breaking a remediation step into sentences and tokenizing each sentence lead to a set of words. Then, these words are brought into the basic form using stemming. Finally, the action class is identified if a stemmed matching term applies to a sentence across its dimensions. The following example demonstrates the two-dimensional mapping matching the terms "block" and "port" and resulting in the "network" action class:

Remediation step: *block access to port 5243/udp*
→ Stemming: *[block, access, to, port, 5243/udp]*
→ Matching: *[[block],[port]]*
→ Tag: *Class* → *network*
Suggested action: *Block port (port: 5243, target: firewall)*

As can be seen, the matching terms identify the class of a workflow action. Playbook-relevant parameters can be passed in this context. Ideally, actions should rely on predefined commands fitting match term combinations to automate the vulnerability handling completely. Toward automated execution of vulnerability playbooks, more granular action classes with more matching term combinations are necessary. Nevertheless, workflow steps are only one part of a CACAO playbook. Besides the workflow, a CACAO playbook also contains metadata and targets. The playbook generation activity places the remediation steps in the workflow section of the CACAO playbook and fills the remaining fields with metadata and additional information.

Our process model tends not to automate the whole process, from identifying a vulnerability to its remediation. We see this process model as a means to assist analysts by identifying and suggesting asset-relevant security advice. The playbook generation phase also involves two manual steps. First, if there is a matching error, e.g., no classification is possible, analysts can manually label workflow actions to continue the process. Second, the process model ends after suggesting a vulnerability playbook to the analyst. It is then up to the analysts whether they would like to execute, adjust, or delete the playbook. Of course, in a best-case scenario, these steps would be automated, although it is questionable whether organizations are willed to apply remediation advice to critical assets without reviewing them.

5 Evaluation

We show that it is feasible to seamlessly generate vulnerability playbooks from structured security advisories with a reasonable amount of effort. Additionally, we compare the quality and completeness of playbooks generated using structured and unstructured security advisories. In doing so, we implement our process model with a proof of concept satisfying a real-world industrial use case. Our use case defines two device representations to model systems-of-systems with vendor-specific components. Our application implements the *security advisory to vulnerability playbook* process aggregating remediation advice from three sources differing in data format, namely *Siemens ProductCERT*—CSAF, *Cisco Security Advisories*—CVRF, and *CISA ICS CERT*—HTML.

5.1 Industrial use case

Our real-world industrial use case describes an enterprise, namely Dehn SE, that is a market leader in plant and building technology, traffic and telecommunications systems, the process industry, and photovoltaic and wind power plants. As a manufacturing enterprise with over 2000 employees, the ICS consists of several assets from Siemens and Cisco. The enterprise already tracks the vulnerabilities of IT assets, such as software packages. The monitoring of vulnerabilities in the ICS is currently still under development. Tracking vulnerabilities and managing remediation advice is perceived as a mammoth task due to the heterogeneity and plethora of assets in use. The enterprise is highly interested in an automated solution gathering vulnerabilities and remediation advice for its assets.

In order to track their ICS assets, we model two virtual representations (i.e., components, CPE identifier, and tags) detailing ICS assets in use. These representations aggregate assets by vendor, thereby forming complex systems-of-systems. To not reveal the actual assets in use, we have augmented them with several other products of the respective vendor. In doing so, we created two obfuscated representations. Of course, other use cases may have other system representations. The first system comprises 22 Siemens field devices, e.g., Siemens SIMATIC S7 (see Appendix 1), typically used in industrial automation and control systems. The second system defines 17 Cisco networking devices, e.g., used as gateways or controllers, found in ICS networks.

5.2 Experimental setting

Our experimental setting consists of adequate hardware and software serving the industrial use case. We have implemented an application with a user interface to efficiently integrate analysts into the vulnerability playbook generation process.

We run all experiments on a single virtual machine with Ubuntu 22.04 LTS operating system, 8GB RAM, and 80GB storage. The device representations are structured using JSON, similar to the widely used Eclipse Ditto[3] representation. The application is based on a front-end/back-end architecture and fully conforms to the CSAF and CACAO standards. The front end is based on Vue.js and the back end on Node.js. The front end is the entry point for the user to verify the correct processing of the security advisories. It provides several functions: CSAF and CACAO visualization, task overview and execution, matching term management, a CSAF converter, and a playbook configurator. A task[4] is considered open if no workflow actions can be derived. A task is done when the workflow actions have been successfully processed, but the final human assessment and approval are pending. The back end relies on the model–view–controller principle and stores CACAO and CSAF documents in a MongoDB. We provide a dashboard for all tasks and their states. Additionally, an analyst can manage the device representations and integrated sources. The pattern section is dedicated to the definition of matching terms.

Our evaluation is threefold. We first run the application, gathering the security advisories (input) to generate playbooks (output). Afterward, we manually assess the input and compare it with the output to assess the overall playbook quality and completeness. As input, we rely on security advisories from different sources for the respective devices. Therefore, we have integrated security advisories from three sources: Siemens ProductCERT, Cisco PSIRT, and CISA ICS CERT. Our application automatically fetches remediation advice from these sources and prevents us from fetching the same advisories multiple times. We selected these sources as they offer vendor-specific or aggregated security advice. Second, these sources ultimately use different data formats to evaluate whether structured security advisories lead to more qualitative and complete playbooks. We collected security advisories over the last 150 days for the playbook quality evaluation. As we also had to assess the security advisories manually, we considered only a collection period of 150 days, although our application could fetch and process even more advisories. After these advisories passed the whole process, we compare the following key indicators to evaluate the playbook's quality and completeness:

- Quantity of workflows actions
- Mistaken acceptance of workflow actions (*type I error*)
- Mistaken rejection of workflow actions (*type II error*)
- Classification of workflow actions

Third, we evaluate the performance of our automated process model showing that automation changes the game in managing vulnerabilities for ICS assets. For the performance measurement, we collect security advisories targeting our assets from the last five years. Through the manual labeling process, the human assessment, and performance measurements, our experimental setting led to several results.

5.3 Experimental results

We have grouped our results according to the process phases from security advisory to vulnerability playbook. Additionally, we show results concerning playbook quality, completeness, and performance. The results are documented using a Jupyter notebook to create transparency, which is available on GitHub.[5]

Vulnerability search. In the industrial use case, device representations hold asset information, including CPE-IDs. We noticed that we could not assign a CPE-ID to each component. This problem has also been pointed out by previous research [36]. We found 13 CPE-IDs for the 17 Cisco assets and 22 CPE-IDs for the 20 Siemens assets. At first glance, these numbers sound reasonable, but considering that CPE can address assets' firmware and hardware, we expected 34 and 40 CPE-IDs, respectively. In addition to the CPE-IDs, we added device-specific tags (i.e., model number). We found 35 vulnerabilities for our devices. Grasping the insecurity of ICS with these asset-specific vulnerabilities, we follow up with the search for security advisories.

Sourcing strategy. Integrating the security advisory sources was a significant challenge due to their heterogeneity. The Siemens ProductCERT does not provide an API. Instead, they offer an Atom feed to query CSAF security advisories using the SSA ID[6], CVE, title, product, sector, or tags. We use the advisory identifier within the Atom feed to manipulate the Siemens website URL and request the advisory in CSAF format. The Cisco PSIRT provides an API based on open security standards (e.g., CVE, CVSS, and CVRF).[7] Since the API does only respond with XML-based CVRF, we still need to convert it. Finally, the CISA ICS CERT does not provide an API or feed to retrieve security advisories. Using the device tags of the device representations, we search within the HTML document and scrape information from its remediation section. As can be seen, searching for remediation advice without any interface and filtering options is a fundamental problem. Therefore, we had to use the device tags and CPEs to automate filtering and verify

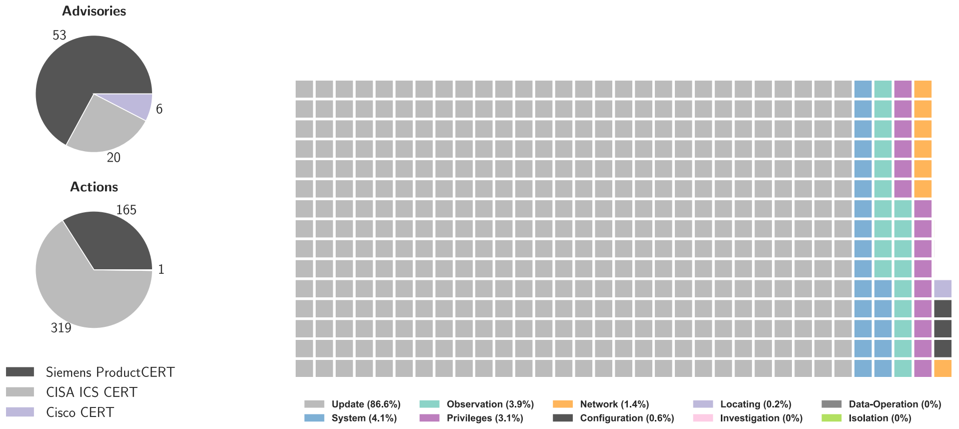
[3] https://www.eclipse.org/ditto/.

[4] A task manages the generation of one playbook.

[5] https://github.com/ad2play/evaluation.

[6] Siemens Security Advisory (SSA) is a Siemens global security advisory identifier.

[7] We noticed that the Cisco OpenVuln API recently added support for CSAF documents.



(a) 79 Advisories and 485 actions by CERT.

(b) Classification of 485 identified workflow actions.

Fig. 4 Analyzing the workflow actions in the generated CACAO playbooks

whether the security advisory is associated with the asset and the respective vulnerability. Since they categorize vulnerabilities, products, and remediation steps, filtering is only a minor problem within CSAF/CVRF documents.

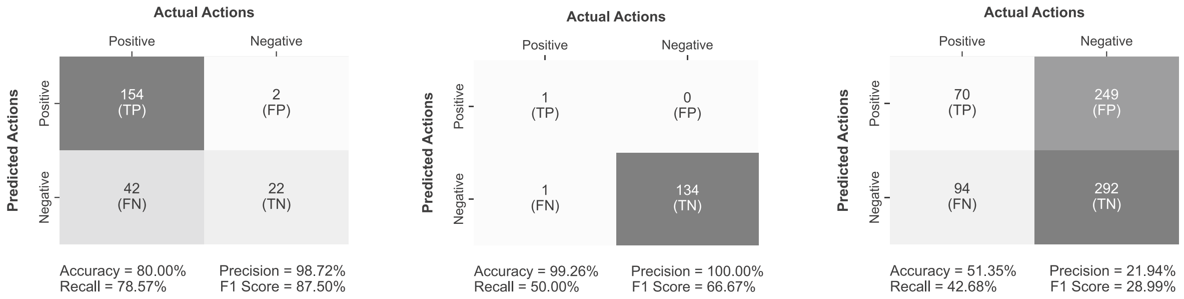
We identified 79 security advisories (see Fig. 4a). Siemens offers 53 advisories, and Cisco offers six. CISA usually lists security advisories for both Siemens and Cisco devices, but the CISA advisories that have been fetched do not contain remediation advice for the Cisco device. However, Cisco has generally listed fewer advisories in the period in question. Also, CISA ICS CERT advisories primarily focus on ICS and do not cover Cisco products for IT enterprise networks. CISA provides a total of 20 security advisories for Siemens assets. It is also noticeable that Siemens offers several versions of advisories, but most overlap considerably in content. Therefore, the total number of Siemens advisories is significantly higher than those from CISA. In addition to the three sources mentioned above, we skimmed IBM X-Force Exchange and NVD [37][8] for security advisories. There, we could find remediation advice only in linked external vendor documents creating complexity for our use case. At the end, we notice that different sources imply different obstacles in obtaining security advisories for specific assets, making sourcing inconvenient.

Advisory conversion. After successfully acquiring security advisories, they are automatically converted into the CSAF data format. For Siemens advisories, already available in CSAF format, no further steps are necessary. The security advisories from Cisco and CISA are converted into CSAF using CVRF and HTML adapters, respectively. When the security advisories from all sources have been converted to CSAF, we analyze these documents.

A closer look at the remediation steps leading to workflow actions (see Fig. 4a) also shows that the amount varies by vendor. While CISA has 319 remediation steps in 20 advisories (16 steps per advisory), Siemens captures 165 remediation steps (3 steps per advisory), and Cisco provides only one remediation step. The identified number of workflow steps in CISA might indicate a high type I error, but it is noticeable that CISA offers additional remediation advice compared to vendor-specific ones. Most interestingly, vendors even advertise remediation advice to inform customers that there is currently no fix available. None of the vendors directly offers technical commands (e.g., in OpenC2 or else) in the remediation steps, whereby dealing with textual descriptions of remediation advice is crucial. In conclusion, advisory conversion is strongly action-centric identifying individual remediation steps.

Playbook generation. The standardized security advisories in CSAF enable the generation of CACAO playbooks. CACAO is an extensive standard, and its implementation is challenging. Emblematic for this fact, the generated CACAO playbooks have a total length of 29,100 lines of code, which leads to 410 lines per playbook. Appendix 1 shows an excerpt of a generated CACAO playbook. However, generated CACAO playbooks are shorter than the initial CSAF documents. One reason is that CSAF also lists remediation advice for other assets, which were not required for our industrial use case. We successfully generated 71 CACAO playbooks out of 79 CSAF documents. Eight advisories require manual post-processing as actions could not be classified correctly. These eight reworks can be traced back to two issues. Seven errors are due to the NLP procedure, which has problems processing placeholders in version numbers, such

[8] NVD by NIST is a comprehensive repository of information related to publicly known cybersecurity vulnerabilities.



(a) Siemens ProductCERT ($n = 220$). (b) Cisco PSIRT ($n = 136$) (c) CISA ICS CERT ($n = 705$).

Fig. 5 Measuring the CACAO playbook quality along workflow action classes (n equals the number of workflow actions)

as “update to version 3.X.” The other error occurred because one remediation step could potentially be assigned to two different classes. Still, we can reduce the manual effort by roughly 90% and automating remediation advice can be seen as a success. Of course, final human assessment is crucial to performing the correct workflow actions to the right target at the right time.

Another elementary part of the CACAO playbook generation is the classification of the individual workflow actions (see Fig. 4b). It is striking that 86.6% of the workflow actions force an update, whereas system (4.1%), observation (3.9%), and access (3.1%) play a rather subordinate role. We also found that the class observation is only mentioned in CISA security advisories. They have a dedicated section advising to observe malicious activity and to report security incidents. The lack of contextual understanding is also a problem while using NLP. These above numbers are the output we yield within the automated process. Ensuring that the generated playbooks match the security advisories’ content requires determining the overall CACAO playbook quality and completeness.

Playbook quality and completeness. We have already seen that the automated creation of CACAO playbooks is feasible and promising. We evaluate the extent to which these results are actually correct in the following. We measure the playbook quality and completeness by referring to confusion matrices (see Fig. 5). These confusion matrices shows the three sources and an overall estimation of the playbooks’ quality and completeness. We thereby include the correct amount of actions and their classification. We calculate the type I error as falsely identified workflow actions. The type II error represents the incorrectly rejected workflow actions. The total number of potential actions is given by the total number of sentences in the remediation steps ($= n$) because, except for one remediation step, all workflow actions were assigned unambiguously to a specific class. We assume that each workflow action is targeted by one sentence. Figure 5a shows Siemens security advisories’ precision, accuracy, recall, and $F1$ score. The high precision (98.72%) shows a high quality of the generated CACAO playbooks. This indicates that the playbook quality is kept high when vendors provide security advisories in CSAF format. Only relevant remediation advice is included in the playbook generation process, while insignificant workflow actions are disregarded. The recall of 78.57% shows acceptable completeness of workflow actions indicating that only a small proportion of workflow actions is actually missing within the playbook. The CACAO playbook generation ($F1$ score = 87.5%) using structured and machine-readable security advisories is outstanding. The type I error is 1%, and the type II error is 19%, which signifies that the matching terms may be too soft. For example, the locating and isolation classes have been matched several times on the first dimension, but did not succeed on dimension two. Figure 5b portrays the results for the Cisco PSIRT using the structured CSAF predecessor CVRF. This leads to an averaged result with an $F1$ score of 66.67%, a type I error of 0%, and a type II error of 0.007%. These results are insignificant, but we decided to include them for completeness. In contrast, unstructured security advisories from CISA deliver different stats (see Fig. 5c). The generated playbooks for CISA are qualitatively inferior compared to Siemens, which is reflected by a low precision of 21.94%. The identified workflow actions show higher incompleteness (precision = 42.68%), leading to an $F1$ score of 28.99%. The type I error is 35.3%, and the type II error is 13.3%. The direct comparison reveals that clear structured, machine-readable security advisories lead to more qualitative and complete playbooks, which in turn results in fewer manual corrections. The matching terms are an adjustment screw to balance the type I and type II errors, but the quality of the fetched security advisories is decisive.

Performance. We have found that fully automating the process, starting with vulnerability search and ending with playbook generation, saves time and reduces effort. For measuring the performance, we use the experimental setting mentioned above. We have collected vulnerabilities and secu-

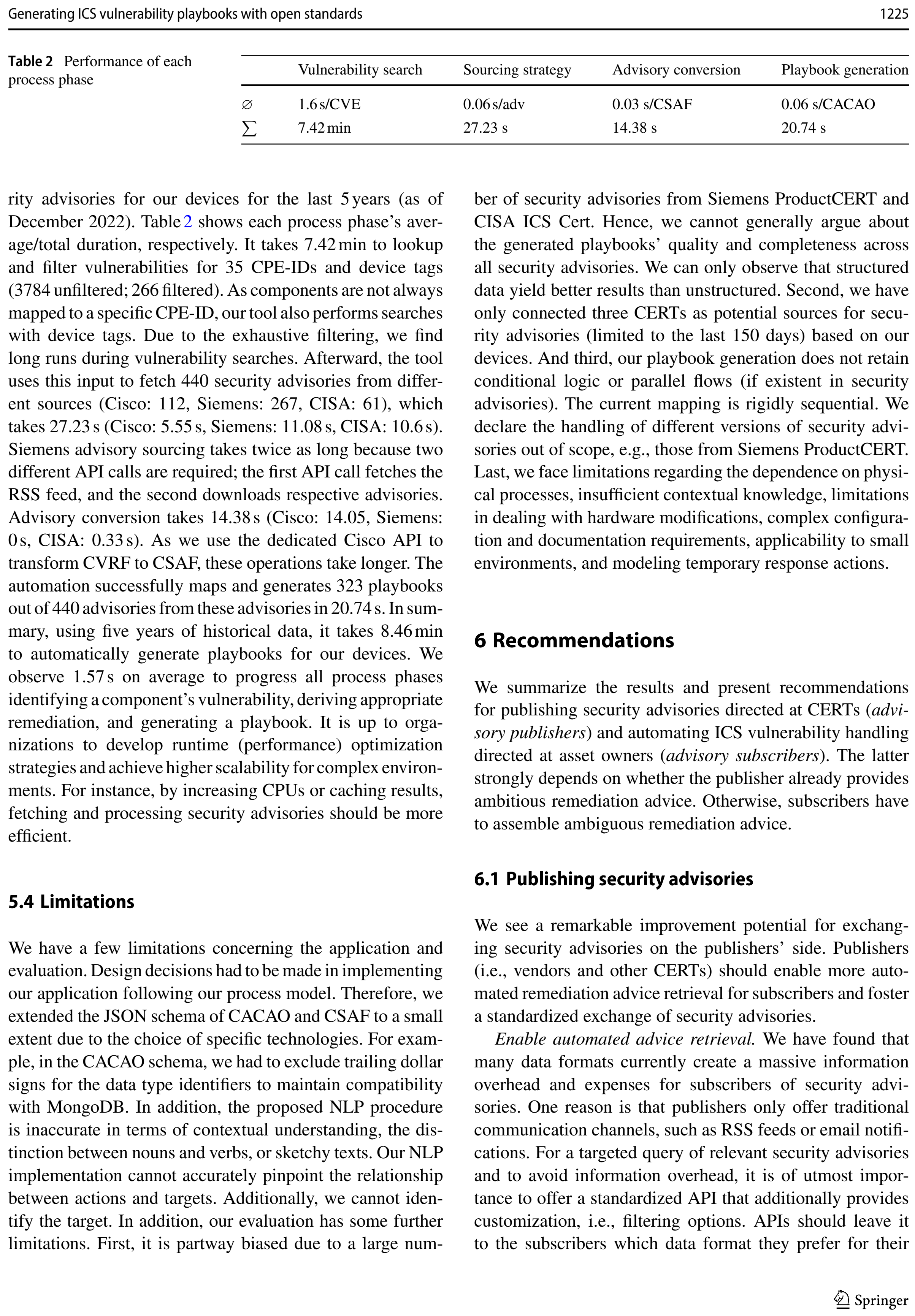


Table 2 Performance of each process phase

	Vulnerability search	Sourcing strategy	Advisory conversion	Playbook generation
∅	1.6 s/CVE	0.06 s/adv	0.03 s/CSAF	0.06 s/CACAO
∑	7.42 min	27.23 s	14.38 s	20.74 s

rity advisories for our devices for the last 5 years (as of December 2022). Table 2 shows each process phase's average/total duration, respectively. It takes 7.42 min to lookup and filter vulnerabilities for 35 CPE-IDs and device tags (3784 unfiltered; 266 filtered). As components are not always mapped to a specific CPE-ID, our tool also performs searches with device tags. Due to the exhaustive filtering, we find long runs during vulnerability searches. Afterward, the tool uses this input to fetch 440 security advisories from different sources (Cisco: 112, Siemens: 267, CISA: 61), which takes 27.23 s (Cisco: 5.55 s, Siemens: 11.08 s, CISA: 10.6 s). Siemens advisory sourcing takes twice as long because two different API calls are required; the first API call fetches the RSS feed, and the second downloads respective advisories. Advisory conversion takes 14.38 s (Cisco: 14.05, Siemens: 0 s, CISA: 0.33 s). As we use the dedicated Cisco API to transform CVRF to CSAF, these operations take longer. The automation successfully maps and generates 323 playbooks out of 440 advisories from these advisories in 20.74 s. In summary, using five years of historical data, it takes 8.46 min to automatically generate playbooks for our devices. We observe 1.57 s on average to progress all process phases identifying a component's vulnerability, deriving appropriate remediation, and generating a playbook. It is up to organizations to develop runtime (performance) optimization strategies and achieve higher scalability for complex environments. For instance, by increasing CPUs or caching results, fetching and processing security advisories should be more efficient.

5.4 Limitations

We have a few limitations concerning the application and evaluation. Design decisions had to be made in implementing our application following our process model. Therefore, we extended the JSON schema of CACAO and CSAF to a small extent due to the choice of specific technologies. For example, in the CACAO schema, we had to exclude trailing dollar signs for the data type identifiers to maintain compatibility with MongoDB. In addition, the proposed NLP procedure is inaccurate in terms of contextual understanding, the distinction between nouns and verbs, or sketchy texts. Our NLP implementation cannot accurately pinpoint the relationship between actions and targets. Additionally, we cannot identify the target. In addition, our evaluation has some further limitations. First, it is partway biased due to a large number of security advisories from Siemens ProductCERT and CISA ICS Cert. Hence, we cannot generally argue about the generated playbooks' quality and completeness across all security advisories. We can only observe that structured data yield better results than unstructured. Second, we have only connected three CERTs as potential sources for security advisories (limited to the last 150 days) based on our devices. And third, our playbook generation does not retain conditional logic or parallel flows (if existent in security advisories). The current mapping is rigidly sequential. We declare the handling of different versions of security advisories out of scope, e.g., those from Siemens ProductCERT. Last, we face limitations regarding the dependence on physical processes, insufficient contextual knowledge, limitations in dealing with hardware modifications, complex configuration and documentation requirements, applicability to small environments, and modeling temporary response actions.

6 Recommendations

We summarize the results and present recommendations for publishing security advisories directed at CERTs (*advisory publishers*) and automating ICS vulnerability handling directed at asset owners (*advisory subscribers*). The latter strongly depends on whether the publisher already provides ambitious remediation advice. Otherwise, subscribers have to assemble ambiguous remediation advice.

6.1 Publishing security advisories

We see a remarkable improvement potential for exchanging security advisories on the publishers' side. Publishers (i.e., vendors and other CERTs) should enable more automated remediation advice retrieval for subscribers and foster a standardized exchange of security advisories.

Enable automated advice retrieval. We have found that many data formats currently create a massive information overhead and expenses for subscribers of security advisories. One reason is that publishers only offer traditional communication channels, such as RSS feeds or email notifications. For a targeted query of relevant security advisories and to avoid information overhead, it is of utmost importance to offer a standardized API that additionally provides customization, i.e., filtering options. APIs should leave it to the subscribers which data format they prefer for their remediation advice. This would make searching for security advisories less painful and more efficient. Additionally, API access allows security advisories to be retrieved in real time or near real time, ensuring that subscribers receive the most current information about vulnerabilities. This is crucial for promptly addressing potential security risks. Beside playbook generation, APIs enable seamless integration of security advisory data into various systems, applications, and tools (e.g., SOAR) used by security professionals. This integration facilitates automated processes for vulnerability scanning, patch management, and incident response, reducing manual effort and potential human errors. Last, APIs are designed to handle high volumes of requests, making them suitable for distributing security advisories to a large number of subscribers and systems efficiently. Overall, providing CSAF-structured security advisories via API fosters a more efficient, interconnected, and responsive cybersecurity ecosystem, enabling organizations to stay proactive and better defend against emerging threats.

Use structured security advisories. Publishers should offer structured security advisories making the content easily machine-readable. Most data formats (i.e., HTML or PDF) for exchanging security advisories differ in structure and content. We have found that structured data formats (i.e., CSAF) better support automation than unstructured data by providing dedicated sections for actions and targets and tend to be more machine-readable. Translating unstructured data into machine-readable advisories requires sophisticated techniques coined by errors. In addition, structured data simplify uniform handling without striving for different conversions of the security advisories. We also came across some best practices for the security advisories' content. First, publishers should only include relevant information in security advisories to keep the remediation advice clean and to prevent information overhead. We propose to structure advisories using the actuator–action–artifact triplet [17]. This triplet helps organize information about the actuator (e.g., firewall), action (e.g., blocking), and artifact (e.g., IP address), normalizing the content of advisories. Second, publishers should be aware of streamlining, maintaining, and optimizing remediation advice. We believe versioning of security advisories to be helpful, as additional remediation advice extends to newly affected assets while keeping the total quantity of security advisories the same. Next, publishers should dedicate a sentence to each remediation step to foster automation. Additionally, as updates are not always feasible, publishers should include more "real" workarounds. Ideally, publishers should keep the CVE and product identifiers within security advisories. It can be observed that some security advisories do not list CVE-IDs. However, there is different remediation advice for different products and publishers should continue mapping product identifiers to individual remediation steps. If this mapping is missing, subscribers cannot ensure that the remediation advice is meant for their assets. Last, we recommend publishing asset-specific commands, needed for automated playbook execution. For that purpose, CACAO defines command types that can capture OpenC2 commands.

6.2 Automating vulnerability handling

Automating vulnerability handling is crucial to cope with the increased number of threats. We summarize our key learning and provide recommendations for security advisory subscribers. Structured device representations, a clear prioritization and sourcing strategy, the integration of machine learning, and the adoption of CACAO playbooks are enablers for automation.

Use structured device representation. Subscribers must know their devices, components (including versions), and vulnerabilities. Organizations need to keep track of hardware modifications and require configuration and documentation management. Comprehensive, well-structured, integrated device representations are the cornerstone for identifying and automating relevant remediation advice. We recommend using a structured format (e.g., JSON or SBOM) and device representations to model complex systems-of-systems. Enriching and maintaining these representations with security-relevant information (e.g., CPE) is crucial to identify vulnerabilities, exploits, and remediation advice.

Integrate machine learning. Subscribers should pay particular attention when selecting appropriate security advisory sources. As these sources differ in many aspects, subscribers have to decide whether the added value of a potential source outweighs the effort involved. The effort usually results from the additional development for security advisories' conversion. For high quality, subscribers should directly integrate vendor-specific advisory sources if they plan automated processing. Free-to-use sources that aggregate remediation advice (e.g., CISA ICS CERT) list advisories from several vendors but are less suitable for automation. Alternatively, subscribers can obtain aggregated security advisories from security vendors without worrying about integrating different vendors.

Integrate machine learning. The integration of machine learning for the automated identification of actions and targets is promising. As long as some CERTs advertise remediation steps in plain text, subscribers should consider whether the application of machine learning can lead to a general improvement in automation. Sophisticated machine learning techniques could lead to sounder contextual understanding and, thus, better automation, quality, and completeness of workflow actions. In particular, large language models hold tremendous potential. With their advanced natural language processing capabilities, large language models can efficiently analyze and interpret textual information to identify appropriate matches and classifications of actuators, actions, and artifacts. In detail, large language models can better capture the subject, verb, and object of a sentence in order to address the current heterogeneity of different data formats for security advisories. In addition, these models can identify commands and modify them to fit within an organization's landscape. When they incorporate organizational knowledge through embedding, they may predict the relationship between assets (actuators) and security advisories.

Adopt CACAO playbooks. CACAO is a promising open standard. Subscribers should evaluate whether the CACAO standard eases maintaining the cybersecurity posture for their ICS. CACAO allows the definition of variables enabling a context-aware and asset-centric approach for quick and efficient remediation. For example, subscribers can define CACAO templates for action classes or even more specific operations, dynamically populate them with variables, and automatically generate context-aware playbooks. At the time of our research, security advisories are still premature, allowing only partial automation. However, implementing the CACAO standard is associated with great efforts. As long as there is no CACAO interpreter, subscribers must manually develop the CACAO playbook integration and execution. The main weak points of CACAO are the premature definitions, low adoption, and a small community.

Prioritize vulnerabilities. Organizations should think about prioritizing ICS vulnerabilities. In our small test environment, we faced several advisories and relevant CACAO playbooks leading to the questions of prioritization. Released in 2021, the Exploit Prediction Scoring System (EPSS) could be useful in ICS environments as it enables efficient prioritization of vulnerability remediation. EPSS considers exploit availability and likelihood, going beyond traditional vulnerability scores to identify critical vulnerabilities that require immediate attention. Given limited resources in ICS environments, EPSS allows operators to allocate resources more efficiently by focusing on vulnerabilities with higher exploitation likelihood, addressing the most critical risks first. Last, EPSS utilizes data from various sources, enhancing vulnerability prioritization accuracy, and providing reliable and actionable insights. In summary, EPSS empowers ICS operators to make informed vulnerability management decisions, safeguarding critical infrastructures effectively against potential cyber threats.

7 Conclusion

Security advisories for ICS vulnerabilities include alternative remediation measures when simply updating to the newest version is not an option. We have generated ICS vulnerability playbooks using open CSAF and CACAO standards. Our approach is the first to combine the fields of security advisories and playbooks addressing organizations' need to handle ICS vulnerabilities. While security advisories foster informing about vulnerabilities, playbooks are intended for workflow actions and eventually support automated execution. We have shown that crucial remediation advice can be included in CACAO playbooks by implementing a process model and experimenting with an industrial use case. ICS security advisories exist in various formats. Therefore, conversion to the CSAF standard is central to automated playbook generation. Toward the creation of individual workflow actions, we built upon matching terms to classify different remediation measures. In 79 security advisories, we identify a high prevalence of update advice and fewer practical remediation steps. Our results lead us to recommendations for security advisory publishers and automated vulnerability handling. Improving security advisories' structure and the content will help vulnerability playbook generation.

Future research can focus on further integration of open standards and their various features. While we use matching terms to extract workflow actions, artificial intelligence (e.g., large language models) might be able to build technical commands and add conditional workflow logic. Toward automated playbook execution, we also see the necessity to incorporate organization-specific factors as remediation measures could be deliberately kept vague to serve all architectures and systems equally well. Additionally, our work is based on available ICS data. As a result, our vulnerability playbooks are specific to ICS. It is worth investigating vulnerability playbook generation for IT assets. Future research could also compare more ICS advisories from plenty vendors to deepen the discussion on recommendations, but also measure the scalability of our vulnerability playbook generation process in larger environments. Another crucial aspect for future research is to target complex and dynamic documentation requirements in large scale ICS environments. Nevertheless, we see the two emerging open standards with increasing number of adopters shaping tomorrow's security operations.

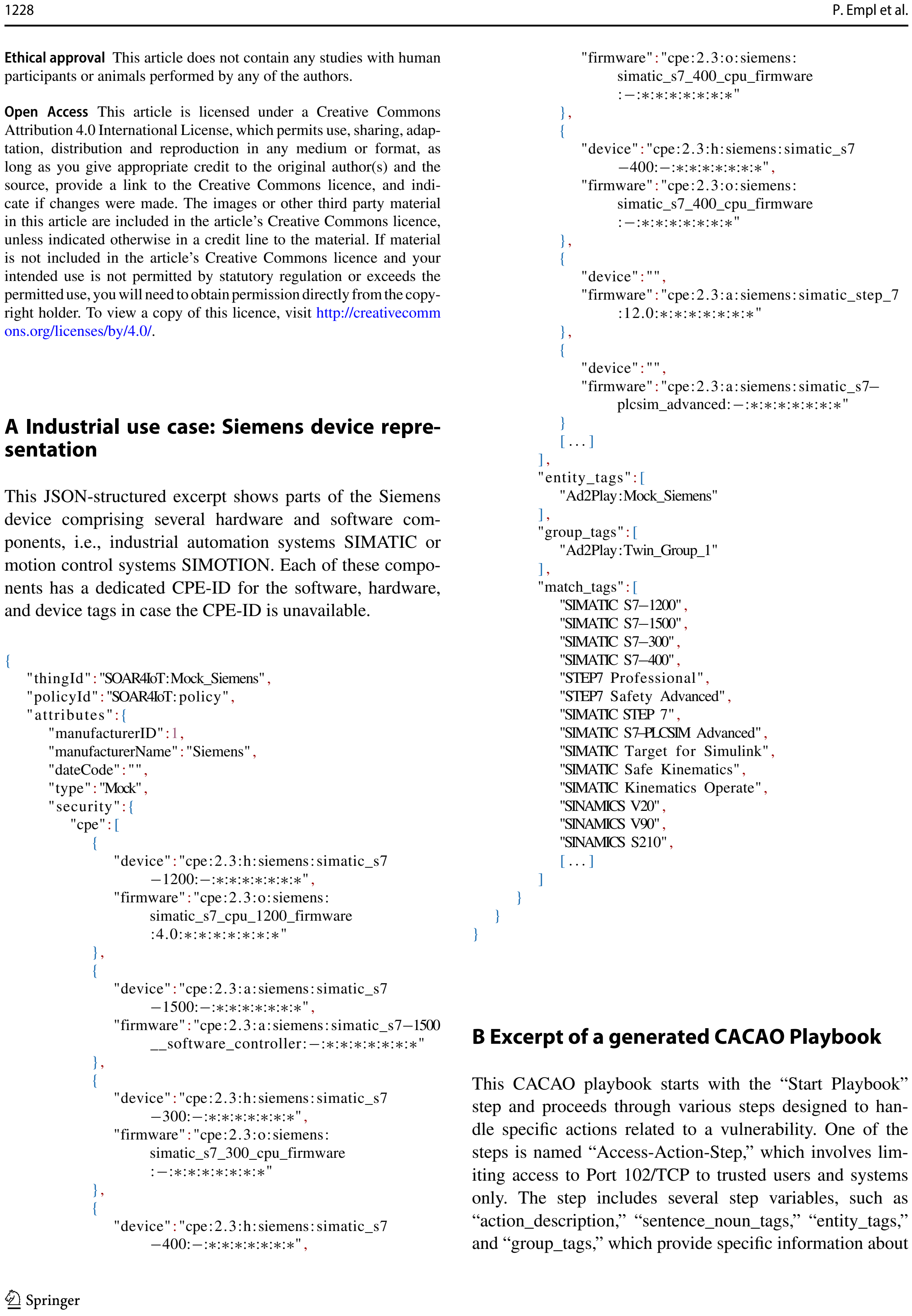
Acknowledgements This research was supported by the German Federal Ministry of Education and Research (BMBF) as part of the DEVISE project and by the Bavarian Ministry of Economic Affairs, Regional Development and Energy (DIK0338/01) as part of the INSIST project. We thank Dehn SE for the great cooperation and feedback.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The prototype and data supporting the findings of this research paper are openly available and accessible through a GitHub repository: https://github.com/ad2play/ad2play.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.



Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

A Industrial use case: Siemens device representation

This JSON-structured excerpt shows parts of the Siemens device comprising several hardware and software components, i.e., industrial automation systems SIMATIC or motion control systems SIMOTION. Each of these components has a dedicated CPE-ID for the software, hardware, and device tags in case the CPE-ID is unavailable.

```
{
  "thingId":"SOAR4IoT:Mock_Siemens",
  "policyId":"SOAR4IoT:policy",
  "attributes":{
    "manufacturerID":1,
    "manufacturerName":"Siemens",
    "dateCode":"",
    "type":"Mock",
    "security":{
      "cpe":[
        {
          "device":"cpe:2.3:h:siemens:simatic_s7
              −1200:−:∗:∗:∗:∗:∗:∗:∗",
          "firmware":"cpe:2.3:o:siemens:
              simatic_s7_cpu_1200_firmware
              :4.0:∗:∗:∗:∗:∗:∗:∗"
        },
        {
          "device":"cpe:2.3:a:siemens:simatic_s7
              −1500:−:∗:∗:∗:∗:∗:∗:∗",
          "firmware":"cpe:2.3:a:siemens:simatic_s7−1500
              __software_controller:−:∗:∗:∗:∗:∗:∗:∗"
        },
        {
          "device":"cpe:2.3:h:siemens:simatic_s7
              −300:−:∗:∗:∗:∗:∗:∗:∗",
          "firmware":"cpe:2.3:o:siemens:
              simatic_s7_300_cpu_firmware
              :−:∗:∗:∗:∗:∗:∗:∗"
        },
        {
          "device":"cpe:2.3:h:siemens:simatic_s7
              −400:−:∗:∗:∗:∗:∗:∗:∗",
          "firmware":"cpe:2.3:o:siemens:
              simatic_s7_400_cpu_firmware
              :−:∗:∗:∗:∗:∗:∗:∗"
        },
        {
          "device":"cpe:2.3:h:siemens:simatic_s7
              −400:−:∗:∗:∗:∗:∗:∗:∗",
          "firmware":"cpe:2.3:o:siemens:
              simatic_s7_400_cpu_firmware
              :−:∗:∗:∗:∗:∗:∗:∗"
        },
        {
          "device":"",
          "firmware":"cpe:2.3:a:siemens:simatic_step_7
              :12.0:∗:∗:∗:∗:∗:∗:∗"
        },
        {
          "device":"",
          "firmware":"cpe:2.3:a:siemens:simatic_s7−
              plcsim_advanced:−:∗:∗:∗:∗:∗:∗:∗"
        }
        [...]
      ],
      "entity_tags":[
        "Ad2Play:Mock_Siemens"
      ],
      "group_tags":[
        "Ad2Play:Twin_Group_1"
      ],
      "match_tags":[
        "SIMATIC S7−1200",
        "SIMATIC S7−1500",
        "SIMATIC S7−300",
        "SIMATIC S7−400",
        "STEP7 Professional",
        "STEP7 Safety Advanced",
        "SIMATIC STEP 7",
        "SIMATIC S7−PLCSIM Advanced",
        "SIMATIC Target for Simulink",
        "SIMATIC Safe Kinematics",
        "SIMATIC Kinematics Operate",
        "SINAMICS V20",
        "SINAMICS V90",
        "SINAMICS S210",
        [...]
      ]
    }
  }
}
```

B Excerpt of a generated CACAO Playbook

This CACAO playbook starts with the "Start Playbook" step and proceeds through various steps designed to handle specific actions related to a vulnerability. One of the steps is named "Access-Action-Step," which involves limiting access to Port 102/TCP to trusted users and systems only. The step includes several step variables, such as "action_description," "sentence_noun_tags," "entity_tags," and "group_tags," which provide specific information about the step. On successful completion of the “Access-Action-Step,” it proceeds to another step.

```
{
        "_id": "62c44b364466fa24127ad4e7",
        "type": "playbook",
        "spec_version": "1.0",
        "id": "playbook—4e105ae9–e4b7–53e0–935a–
              fed1125ca376",
        "name": "AUTOGENERATED Playbook from sourced CSAF
              file(s) 62c3f0a499cf2533865814eb",
        "created_by": "identity—a9becb6a–d006–518a–a0a1–66
              a7bc70675e",
        "created": "2023–07–05T14:31:18.672Z",
        "external_references": [
            {
                "name": "CSAF File",
                "description": "Id of CSAF file that was
                      used for the creation of the playbook",
                "external_id": "62c3f0a499cf2533865814eb",
                "_id": "62c44b364466fa24127ad4e8"
            }
        ],
        "workflow_start": "step—e7d19860–a84a–563d–8705–02
              beb8f03441",
        "workflow": {
            "step—e7d19860–a84a–563d–8705–02beb8f03441": {
                "type": "start",
                "name": "Start Playbook",
                "on_completion": "step—3f0d51ad–7500–5dca
                      –90e2–8ed00a4d9a4f",
                "_id": "62c44b364466fa24127ad4e9",
            },
            [...]
            "step—a5f44526–2a7a–5dfb–820d–74ce00db5cf1": {
                "type": "single",
                "name": "Access–Action–Step",
                "step_variables": {
                    "[$$action_description$$]": {
                        "type": "string",
                        "description": "This is the sentence
                              that triggered the Pattern and
                              includes the action",
                        "value": "Limit access to Port 102/
                            TCP to trusted users and
                            systems only.",
                        "constant": true,
                        "_id": "62c44b364466fa24127ad4db"
                    },
                    "[$$sentence_noun_tags$$]": {
                        "type": "string",
                        "description": "This is the
                            stringified array of the nouns
                            used in the sentence that
                            triggered the pattern",
                        "value": "Limit,access,Port,TCP,
                            users,systems",
                        "constant": true,
                        "_id": "62c44b364466fa24127ad4dc"
                    },
                    "[$$entity_tags$$]": {
                        "type": "string",
                        "description": "These are the
                            entity_tags related to the
                            Twins",
                        "value": "Ad2Play:Mock_Siemens",
                        "constant": true,
                        "_id": "62c44b364466fa24127ad4dd"
                    },
                    "[$$group_tags$$]": {
                        "type": "string",
                        "description": "These are the
                             group_tags related to the Twins
                             ",
                        "value": "Ad2Play:Twin_Group_1",
                        "constant": true,
                        "_id": "62c44b364466fa24127ad4de"
                    }
                }
                [...]
            }
        }
}
```

References

1. Lawson, C., Price, A.: market guide for security orchestration, automation and response solutions (2022)
2. Biden, J.R.J.: Executive order on improving the nation’s cybersecurity. https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/. Last accessed 2023-06-05 (2022)
3. Forum of Incident Response and Security Teams (FIRST). Automation sig. https://www.first.org/global/sigs/automation/. Last accessed 2023-06-05 (2023)
4. Stevens, R., Votipka, D., Dykstra, J., Tomlinson, F., Quartararo, E., Ahern, C., Mazurek, M.L.: How ready is your ready? Assessing the usability of incident response playbook frameworks. In: Proceedings of the 2022 SIGCHI Conference on Human Factors in Computing Systems (CHI ’22) (ACM), pp. 1–18. https://doi.org/10.1145/3491102.3517559 (2022)
5. OASIS. Cacao security playbooks version 1.0—committee specification 02. https://docs.oasis-open.org/cacao/security-playbooks/v1.0/security-playbooks-v1.0.html. Last accessed 2023-06-05 (2021)
6. Cybersecurity and Infrastructure Security Agency (CISA), Federal government cybersecurity incident and vulnerability response playbooks. Tech. rep., Cybersecurity and Infrastructure Security Agency (CISA) (2021)
7. Wang, B., Li, X., de Aguiar, L.P., Menasche, D.S., Shafiq, Z.: Characterizing and modeling patching practices of industrial control systems. In: Proceedings of the 2017 ACM on Measurement and Analysis of Computing Systems (POMACS ’17), vol. 1(1), p. 1. https://doi.org/10.1145/3078505.3078524 (2017)
8. Cybersecurity & Infrastructure Security Agency (CISA). Ics-cert advisories. https://www.cisa.gov/uscert/ics/advisories. Last accessed 2023-06-05 (2023)
9. OASIS. Common security advisory framework version 2.0—committee specification 03. https://docs.oasis-open.org/csaf/csaf/v2.0/csaf-v2.0.html. Last accessed 2023-06-05 (2022)
10. National Vulnerability Database (NVD). Cve-2022-34819 detail. https://nvd.nist.gov/vuln/detail/CVE-2022-34819. Last accessed 2023-06-05 (2022)
11. Common Vulnerabilities and Exposures (CVE). https://cve.mitre.org/. Accessed 20 July 2023 (2023)
12. National Vulnerability Database (NVD) Common Platform Enumeration (CPE). https://nvd.nist.gov/products/cpe. Accessed 20 July 2023 (2023)

13. Common Vulnerability Scoring System (CVSS). https://www.first.org/cvss/. Accessed 20 July 2023 (2023)
14. OASIS Common Security Advisory Framework (CSAF) Technical Committee. https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=csaf. Accessed 20 July 2023 (2023)
15. Skopik, F., Settanni, G., Fiedler, R.: A problem shared is a problem halved: a survey on the dimensions of collective cyber defense through security information sharing. Comput. Secur. **60**, 154 (2016). https://doi.org/10.1016/j.cose.2016.04.003
16. Gascon, H., Grobauer, B., Schreck, T., Rist, L., Arp, D., Rieck, K.: Mining attributed graphs for threat intelligence. In: Proceedings of the 7th ACM on Conference on Data and Application Security and Privacy (CODASPY '17) (ACM, 2017), pp. 15–22. https://doi.org/10.1145/3029806.3029811
17. Schlette, D., Caselli, M., Pernul, G.: A comparative study on cyber threat intelligence: the security incident response perspective. IEEE Commun. Surv. Tutor. **23**(4), 2525 (2021). https://doi.org/10.1109/COMST.2021.3117338
18. OASIS. Open command and control (OpenC2) language specification version 1.0—committee specification 02. https://docs.oasis-open.org/openc2/oc2ls/v1.0/oc2ls-v1.0.html. Last accessed 2023-06-05 (2019)
19. MITRE. Detection, denial, and disruption framework empowering network defense (D3FEND). https://d3fend.mitre.org/. Last accessed 2023-06-05 (2023)
20. ATC Project. RE&CT framework documentation. https://atc-project.github.io/atc-react/. Last accessed 2023-06-05 (2020)
21. West-Brown, M.J., Stikvoort, D., Kossakowski, K.P., Killcrece, G., Ruefle, R., Zajicek, M.: Handbook for computer security incident response teams (CSIRTs). Tech. rep., Defense Technical Information Center. https://doi.org/10.21236/ada413778 (2003)
22. Senthivel, S., Dhungana, S., Yoo, H., Ahmed, I., Roussev, V.: Denial of engineering operations attacks in industrial control systems. In: Proceedings of the 8th ACM Conference on Data and Application Security and Privacy (CODASPY '22) (ACM), CODASPY '18, pp. 319–329. https://doi.org/10.1145/3176258.3176319 (2018)
23. Ghena, B., Beyer, W., Hillaker, A., Pevarnek, J., Halderman, J.A.: Green lights forever: Analyzing the security of traffic infrastructure. In: Bratus, S., Lindner, F.F. (eds.), Proceedings of the 8th USENIX Workshop on Offensive Technologies (WOOT '14). USENIX Association (2014)
24. Li, F., Durumeric, Z., Czyz, J., Karami, M., Bailey, M., McCoy, D., Savage, S., Paxson, V.: You've got vulnerability: exploring effective vulnerability notifications. In: Holz, T., Savage, S. (eds.), Proceedings of the 25th USENIX Security Symposium (USENIX Security '16). USENIX Association, pp. 1033–1050 (2016)
25. Allodi, L., Banescu, S., Femmer, H., Beckers, K.: Identifying relevant information cues for vulnerability assessment using CVSS. In: Proceedings of the 8th ACM Conference on Data and Application Security and Privacy (CODASPY '18) (ACM), CODASPY '18, pp. 119–126 (2018). https://doi.org/10.1145/3176258.3176340
26. Fenz, S., Ekelhart, A., Weippl, E.: Semantic potential of existing security advisory standards. In: Proceedings of the FIRST 2008 Conference-Forum of Incident Response and Security Teams (FIRST '08) (2008)
27. Fenz, S., Ekelhart, A., Weippl, E.: Fortification of IT security by automatic security advisory processing. In: Proceedings of the 22nd International Conference on Advanced Information Networking and Applications (AINA '08) (IEEE), pp. 575–582. https://doi.org/10.1109/aina.2008.69 (2008)
28. Challande, A., David, R., Renault, G.: Building a commit-level dataset of real-world vulnerabilities. In: Proceedings of the 12th ACM Conference on Data and Application Security and Privacy (CODASPY '22) (ACM), CODASPY '22, pp. 101–106. https://doi.org/10.1145/3508398.3511495 (2022)
29. Mavroeidis, V., Eis, P., Zadnik, M., Caselli, M., Jordan, B.: On the integration of course of action playbooks into shareable cyber threat intelligence. In: Proceedings of the 2021 IEEE International Conference on Big Data (Big Data '21) (IEEE), pp. 2104–2108. https://doi.org/10.1109/bigdata52589.2021.9671893 (2021)
30. Akbari Gurabi, M., Mandal, A., Popanda, J., Rapp, R., Decker, S.: SASP: a semantic web-based approach for management of sharable cybersecurity playbooks. In: Proceedings of the 17th International Conference on Availability, Reliability and Security (ARES '22) (ACM), pp. 1–8. https://doi.org/10.1145/3538969.3544478 (2022)
31. Shaked, A., Cherdantseva, Y., Burnap, P.: Model-based incident response playbooks. In: Proceedings of the 17th International Conference on Availability, Reliability and Security (ARES '22) (ACM), pp. 1–7. https://doi.org/10.1145/3538969.3538976 (2022)
32. Islam, C., Babar, M.A., Nepal, S.: A multi-vocal review of security orchestration. ACM Comput. Surv. **52**(2), 1 (2019). https://doi.org/10.1145/3305268
33. Empl, P., Schlette, D., Zupfer, D., Pernul, G.: SOAR4IoT: securing IoT assets with digital twins. In: Proceedings of the 17th International Conference on Availability, Reliability and Security (ARES '22) (ACM), pp. 1–10. https://doi.org/10.1145/3538969.3538975 (2022)
34. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. J. Manag. Inf. Syst. **24**(3), 45 (2007). https://doi.org/10.2753/MIS0742-1222240302
35. Dietz, M., Pernul, G.: Digital twin: empowering enterprises towards a system-of-systems approach. Bus. Inf. Syst. Eng. **62**(2), 179 (2020). https://doi.org/10.1007/s12599-019-00624-0
36. Schlette, D., Menges, F., Baumer, T., Pernul, G.: Security enumerations for cyber-physical systems. In: Singhal, A., Vaidya, J. (eds.), Data and Applications Security and Privacy XXXIV—34th Annual IFIP WG 11.3 Conference, DBSec: Regensburg, Germany, June 25–26, 2020, Proceedings, Lecture Notes in Computer Science, vol. 12122, pp. 64–76. Springer (2020). https://doi.org/10.1007/978-3-030-49669-2
37. National vulnerability database. https://nvd.nist.gov/. Accessed on 20 Jul 2023

Appendix

Curriculum Vitae

Philip M. Empl

Chair of Information Systems – IFS
Faculty of Business, Economics, Management Information Systems
University of Regensburg, Germany

Education

2020 – 2024 **PhD Student**
University of Regensburg, Germany

2017 – 2020 **MSc Management Information Systems**
University of Regensburg, Germany

2014 – 2017 **BSc Management Information Systems**
University of Regensburg, Germany

Research Projects

SISSeC (2019-2022). Secure Industrial Semantic Sensor Cloud funded by the Federal Ministry of Economics and Climate Protection (16KN085725).

INSIST (2021-2024). Industrial IoT Security Operations Center funded by the Bavarian Ministry of Economic Affairs, Regional Development, and Energy (DIK0338/0).

Teaching

2020 – 2024 **Co-Lecturer** – Security of data-intensive Applications
Graduate lecture at University of Regensburg

2020 – 2024 **Tutor** – IT-Security I
Undergraduate lecture at University of Regensburg

2020 – 2024 **Tutor** – VAWI IT-Security
VAWi – Virtual continuing education in information systems

2019 – 2020 **Student Tutor** – Information Systems - Developments and Trends
Graduate lecture at University of Regensburg

2018 – 2020 **Student Tutor** – IT-Security I
Undergraduate lecture at University of Regensburg

2016 – 2017 **Student Tutor** – Theoretical Informatics
Undergraduate lecture at University of Regensburg

Reviewing Activities

Conferences. CAiSE 2023, ARES 2022, ESORICS 2022, ICICS 2022, CAiSE 2022, ICISSP 2022, DBSec 2022, ARES 2021, CCGrid 2021, ESORICS 2021, ICS 2021, ISC 2021, WI 2020.

Journals. Computers & Security, Journal of Computer Security.

Publications

[1] Wagner, G., Empl, P., & Schryen, G. (2020). Designing a Novel Strategy for Exploring Literature Corpora. In *Proceedings of the 28th European Conference on Information Systems (pp. 44:1–44:17). AIS.*.

[2] Putz, B., Dietz, M., Empl, P., & Pernul, G. (2021). Ethertwin: Blockchain-Based Secure Digital Twin Information Management. *Information Processing & Management, 58*(1), 102425. Elsevier.

[3] Empl, P., & Pernul, G. (2021). A Flexible Security Analytics Service for the Industrial IoT. In *Proceedings of the 2021 ACM Workshop on Secure and Trustworthy Cyber-physical Systems* (pp. 23–32). SciTePress.

[4] Empl, P., Schlette, D., Zupfer, D., & Pernul, G. (2022). SOAR4IoT: Securing IoT Assets with Digital Twins. In *Proceedings of the 17th International Conference on Availability, Reliability and Security* (pp. 4:1–4:10). ACM.

[5] Empl, P., & Pernul, G. (2023). Digital-Twin-Based Security Analytics for the Internet of Things. *Information, 14*(2), Article 95. MDPI.

[6] Empl, P., Hager, H., & Pernul, G. (2023). Digital Twins for IoT Security Management. In *Proceedings of the 37th Annual IFIP WG 11.3 Conference on Data and Applications Security and Privacy* (pp. 141—149). Springer Nature.

[7] Schlette, D., Empl, P., Caselli, M., Schreck, T., & Pernul, G. (2024). Do You Play It by the Books? A Study on Incident Response Playbooks and Influencing Factors. In *Proceedings of the 45th IEEE Symposium on Security and Privacy* (pp. 59:1–59:19). IEEE Computer Society.

[8] Empl, P., Schlette, D., Stöger, L., & Pernul, G. (2024). Generating ICS Vulnerability Playbooks with Open Standards. *International Journal of Information Security 23*(2), 1215–1230. Springer.

[9] Empl, P., Böhm, F., & Pernul, G. (2024). Process-Aware Intrusion Detection in MQTT Networks. In *Proceedings of the 14th ACM Conference on Data and Application Security and Privacy* (pp. 91–102). ACM.

[10] Empl, P., Koch, D., Dietz, M., & Pernul, G. (2024). Digital Twins in Security Operations: State of the Art and Future Perspectives. Submitted to *ACM Computing Surveys.* ACM.

[11] Hornsteiner, M., Empl, P., Bunghardt, T., & Schönig, S. (2024). Reading between the Lines: Process Mining in OPC UA Networks. *Sensors, 24*(14), 4497. MDPI.

[12] Ackermann, L., Käppel, M., Marcus, L., Moder, L., Dunzer, S., Hornsteiner, M., Liessmann, A., Zisgen, Y., Empl, P., Herm, L.-V., Neis, N., Neuberger, J., Poss, L., Schaschek, M., Weinzierl, S., Wördehoff, N., Jablonski, S., Koschmider, A., Kratsch, W., Matzner, M., Rinderle-Ma, S., Röglinger, M., Schönig, S., Winkelmann, A. (2024). Recent Advances in Data-Driven Business Process Management.

Talks and Presentations

2024	Process-Aware Intrusion Detection in MQTT Networks *CODASPY Conference 2024*
2024	Do You Play It by the Books? *S&P Conference 2024*
2023	Digital Twins for IoT Security Management *DBSec Conference 2023*
2022	SOAR4IoT: Securing IoT Assets with Digital Twins *ARES Conference 2022*
2020	Secure Industrial Semantic Sensor Cloud *TRIOKON 2020*