Intrusion Detection System in Digital Twins for Industrial Control Systems

1st Sahar Zamanian
Department of Electrical and Information Technology
Lund University, EIT
Lund, Sweden
sahar.zamanian@eit.lth.se

2nd Maria Kihl
Department of Electrical and Information Technology
Lund University, EIT
Lund, Sweden
maria.kihl@eit.lth.se

***Abstract*—Digital twins open up new possibilities in terms of monitoring, optimizing, and predicting the state of cyber-physical systems (CPSs). Furthermore, we argue that a fully functional virtual replica of a CPS can also play an important role in securing the system. This paper describes ongoing work; we present a host-based intrusion detection system that allows users to create and execute digital twins, closely matching their physical counterparts. We describe a novel approach to automatically detect the manipulated replica from the specification, taking advantage of engineering data-exchange formats. From a security perspective, an identical simulated environment (in terms of the system specification) can be freely explored and tested by security professionals, without risking negative impacts on live systems. Going one step further, security modules on top of the framework support security analysts in monitoring the current state of CPSs. We plan to demonstrate the viability of the IDS in a proof-of-concept, including the automated detection of digital twins and the monitoring of security and safety rules.**

***Index Terms*—Intrusion Detection System, Industrial Control System, Deep Learning, Network Flow, Cyber Security, Network Attack**

I. Introduction

The concept of Digital Twins (DTs) has emerged as a cornerstone in the development of next-generation cyber-physical systems, particularly within the context of Industry 4.0 and Industrial Automation and Control Systems (IACS). A Digital Twin is a virtual representation of a physical entity that mirrors its real-time state, behavior, and history through continuous data synchronization [1], [2]. This bidirectional link between the physical and digital domains enables predictive maintenance, process optimization, and real-time monitoring, offering significant economic and operational benefits in industries ranging from manufacturing to energy and healthcare care [3], [4]. However, increasing reliance on DTs introduces new cyber-security risks, particularly due to the need for persistent connectivity, high-frequency data exchange, and integration with external systems such as cloud infrastructures and IoT devices. These emerging challenges underscore the urgent need for robust security architectures and intrusion detection frameworks specifically tailored for DT environments [5].

This article presents ongoing research work. We plan to tackle the security issue by looking into a scenario in which the digital twin system is vulnerable to potential compromise through a Man-in-the-Middle (MITM) attack with the interception of synchronization data by an adversary. Specifically, the attacker uses IP spoofing by intercepting network packets. If an attacker successfully intercepts the intermediate network, gains access to the synchronization data, and subsequently manipulates and alters the synchronization states, the integrity of the digital twin system becomes compromised. Our planned work will focus on identifying anomalies in the synchronization data in a digital twin-based IACS security architecture.

II. Related work

In this section, we present other papers that have performed related research.

A. Digital twin

Eckhart et al. in [6] conceptualized the notion of a Digital Twin as a digital counterpart of a physical entity, which can be either living or non-living. There are two primary approaches for Digital twins aimed at Industrial Control Systems (ICSs), as identified in [7]: the information/knowledge-driven approach and the data-driven approach. The former relies on the physical system's specifications to develop virtual system prototypes, whereas the latter leverages real-time data sourced from physical environment devices to construct the system model.

The challenge of securing Industrial Automation and Control Systems (IACS) within the scope of digital twins has been extensively reviewed in various comprehensive surveys, including in [8] and [9]. Specifically, the study in [10] introduced a security-aware Cyber-Physical System (CPS) Twinning framework. This innovative framework is designed to create a digital twin of an ICS based on its specifications, offering two modes of operation: a simulation mode, in which the digital twin functions as an independent emulation, and a replication mode, in which synchronization between the digital and physical twins is enabled.

B. Intrusion Detection Systems

Intrusion detection systems (IDS) are cyber-security tools designed to monitor network traffic, find suspicious activities and deviations from established security policies, and then alert the system administrator when such activities or violations are detected. The primary operational objective of an IDS is to analyze and systematically evaluate data related to anomalies within network traffic patterns. In [11], the authors propose an Intrusion Detection System (IDS) aimed at Digital Twins (DT). They leverage an open-source digital twin framework designed for industrial control systems, extending its capabilities to encompass attack simulation and defense mechanisms. However, their implementation is based on a local machine environment and does not take advantage of a cloud-based infrastructure. In the follow-up work [12], anomaly detection and anomaly classifications based on supervised machine learning algorithms in manufacturing systems are considered. The proposed IDSs for DTs in [13], [14] with different scenarios are evaluated by measuring signals.

Inspired by the work in [12] and [10], we will investigate the area of state synchronization as a security facilitator. However, in comparison with the work in [12], we plan to investigate IDSs that can protect DTs from external attacks, instead of attacks in the factory domain. Therefore, we propose an IDS model that allows us to identify attacks in the virtual domain and prevent them from reaching the physical domain.

C. Man-In-The-Middle attack (MITM)

Adversary activities within the Digital Twin (DT) processes can disrupt the entire system, presenting significant challenges. A DT can be attacked in several ways, as described in [8]. An overview of security threats and their corresponding countermeasures is provided in [15].

One notable security attack in DT is the Man-In-The-Middle attack (MITM). This attack involves intercepting communication between two entities, without them suspecting that the communication channel between the physical and digital twin has been compromised [9]. Dietz et al. [16] introduced a framework designed specifically for DT. This framework shows how an MITM attack can be executed in a simulated industry environment and analyzes its impact on the involved systems. Several other research papers, such as [17], [18] and [16], explore Intrusion Detection Systems (IDSs) specifically designed to counteract Man-in-the-Middle (MITM) attacks.

However, as far as we know, cloud-based IDSs for DTs with the aim of controlling state replications and classifying attacks on state synchronization have not been explored.

III. Targeted system

In our work, we plan to use the system architecture proposed by [10], as illustrated in Figure 1. The system has three fundamental properties: a DT, a Physical Twin (PT), an IDS, and a network that connects them. The physical counterpart represents an industrial system comprising multiple physical states that correspond to an automation system that includes various types of machinery and robotic components. The physical twin is deployed in an isolated factory network, where it can only interact with the DT using synchronization data sent between the internal synchronization gateways (Synch GWs). The critical component in this architecture is the network connecting the physical and digital twins' synchronization gateways. This network introduces a vulnerability to potential cyber-attacks due to the data transmissions between the physical and digital twins.

The DT consists of several states that should represent the PT states. These states are continuously updated via the network. For our work, we have assumed that both the DT and the IDS can be deployed in a trusted execution environment, using containers or virtual machines (VMs) on suitable cloud resources. Only one external network connection is allowed, and all synchronization occurs via the Synch GWs. The IDS is deployed in the DT and connected to the synchronization gateway.

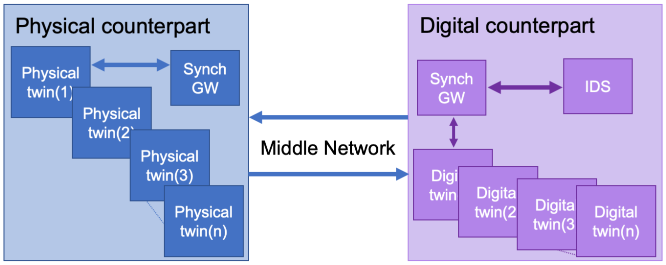

Fig. 1. The proposed Digital Twin scenario

IV. Methodology

A. Attack Model

In our work, we will focus on the Man-In-The-Middle attack in the network, since this attack model is specifically designed for systems with cloud-based data sharing and control in IACS. We assume that the physical twin component of the system is well-protected and isolated from external network attacks. Adversaries will attempt to compromise the system's integrity by manipulating the network traffic and change the communication between the physical and digital twins.

The attacker is launching the MITM attack on the targeted system through the following steps:

1) Intercepting, modifying, and replaying communication between the physical and digital counterparts, and vice versa by IP spoofing.
2) Attack the synchronization data by sending arbitrary states to the digital twin. This means that the attacker can send arbitrary input from the physical states to the digital twin states.

Figure 2 shows a visual representation of the potential substitution of states within the DT when an attacker accesses synchronization information within the synchronization gateway. It is evident from Figure 2 that if an adversary successfully manipulates the states in the DT, the system is compromised. Consequently, it is crucial to develop robust mechanisms for timely detection and mitigation of malicious activities.

V. Discussion

This paper presents ongoing research aimed at addressing the security vulnerabilities of DT systems within IACS.

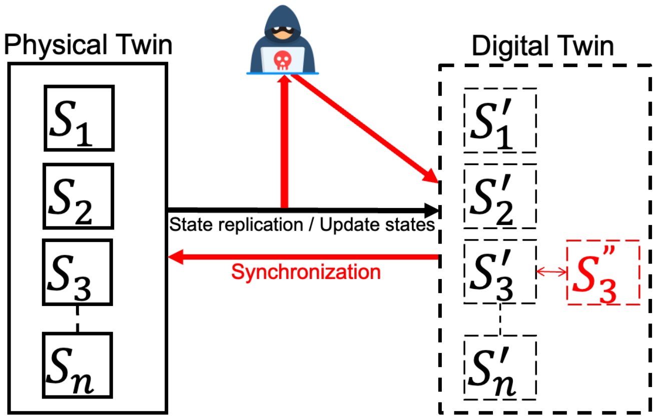

Fig. 2. Attack Model

We focus particularly on safeguarding the synchronization mechanism between the physical and digital twins, which we identified as a critical point of vulnerability in cyber-physical architectures. Through a cloud-based IDS, our approach is designed to detect anomalies in state synchronization, with a special emphasis on identifying malicious activities such as MITM attacks. The overarching goal is to ensure that digital twins operate securely and faithfully mirror their physical counterparts without being compromised by adversarial manipulations.

A key challenge that emerged during our exploration is the complexity inherent in securing state synchronization between distributed systems operating in separate domains. Synchronization, while enabling real-time mirroring and control, also opens up potential attack vectors if not properly monitored. Specifically, the interception and manipulation of synchronization data can lead to a complete compromise of the digital twin, allowing attackers to introduce arbitrary states, replay outdated information, or spoof identities.

The existing literature offers numerous insights into IDS solutions for industrial systems, but very few studies have focused on the security of state synchronization within DT systems, especially those exploiting cloud-based infrastructures. Many previous approaches have concentrated on securing the physical system or relied on local environments for implementation, overlooking the growing trend toward cloud-hosted digital twin environments. Moreover, while some work has addressed MITM attacks in general DT contexts, there is a lack of targeted models and tools that focus on detecting and classifying synchronization-level anomalies in cloud-based DT architectures.

This gap formed the basis for our motivation. We recognized the need for an IDS model that not only understands the semantics of DT synchronization, but is also capable of operating in scalable, cloud-based infrastructures. Unlike traditional IDSs that monitor network flow or static system behavior, our system monitors dynamic, time-sensitive synchronization exchanges and leverages state modeling to detect inconsistencies. This approach provides early detection of malicious manipulations before they can propagate into the physical domain.

To explore this, we designed a targeted system architecture that includes a physical twin, a digital twin and a synchronization gateway, all connected via a network. We consider the network to be the primary threat vector. The IDS component, deployed within a trusted execution environment in the digital twin cloud infrastructure, continuously monitors synchronization patterns and validates state integrity using finite state machine (FSM) modeling. Both the physical and digital twins are modelled as FSMs, and we apply defined synchronization functions and transition operations to track deviations that may indicate potential attacks.

Our next steps involve implementing a proof-of-concept (PoC) for this architecture using real-world industrial data. We will investigate various attack scenarios including spoofed synchronization messages and replay attacks. We plan to evaluate the system's performance based on multiple metrics, such as detection accuracy, latency, and false positive rate. Additionally, we aim to explore the integration of machine learning models to enhance the classification of detected anomalies and extend the framework's adaptability to other forms of cyber-physical systems beyond manufacturing.

This research aims to contribute to a more resilient digital twin ecosystem by proactively addressing emerging security challenges, particularly those associated with synchronization integrity. Our ultimate goal is to ensure that digital twins enhance operational efficiency without compromising the trust and safety of industrial systems.

VI. Open Questions

This research presents a conceptual and methodological foundation for securing digital twin systems against synchronization-based attacks, but there are several open questions that require further investigations:

- **How can real-time intrusion detection be optimized at scale?** The current framework assumes a trusted execution environment and isolated physical twins. However, large-scale deployments include multiple digital twins and physical assets that may introduce latency, scalability issues, and performance bottlenecks. More research is needed to explore efficient real-time detection algorithms that scale with system complexity.
- **What are the implications of adversarial machine learning on IDS for Digital Twins?** Given the increasing integration of machine learning techniques in IDS systems, the susceptibility of such models to adversarial inputs poses a significant threat. Investigating how adversaries could exploit these vulnerabilities to bypass detection or disrupt state synchronization remains an important question.
- **How can trust be established and maintained across heterogeneous synchronization gateways?** The proposed synchronization model assumes secure communication between physical and digital twins. However, in real-world scenarios where multiple vendors and diverse

hardware/software platforms exist, establishing trust between heterogeneous gateways is non-trivial and still an open challenge.

- **How to evaluate the system's robustness against sophisticated threat models?** The current approach addresses state manipulation via MITM attacks. However, future work must consider hybrid threats, such as combining insider threats, firmware-level attacks, and cloud-side vulnerabilities, to test the robustness of the proposed architecture.
- **What regulatory or standardization efforts are required for DT security in IACS environments?** As digital twins become integral to industrial systems, it is essential to investigate how existing cyber-security standards can be adapted or extended to cover DT-based architectures and what new guidelines may be necessary.

Addressing these questions will deepen the understanding of secure digital twin architectures and explore how more resilient, adaptive, and intelligent intrusion detection systems tailored for next-generation industrial environments can be developed.

Acknowledgments

This work was supported by the SSF project SEC4FACTORY under grant no. SSF RIT17-0032. The authors are part of the Excellence Center at Linköping-Lund on Information Technology (ELLIIT) strategic research area, the LTH profile area for AI and Digitalization, and the LU Profile area Natural and Artifical Cognition. Maria Kihl is partially funded by the competence centre NextG2Com in the VINNOVA program for Advanced Digitalisation with grant number 2023-00541. Also she is partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

References

[1] Michael Grieves and John Vickers. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems*, pages 85–113. Springer, 2014. Conceptual origin of Digital Twins.

[2] Fei Tao, Meng Zhang, Ang Liu, and A. Y. C. Nee. Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, 15(4):2405–2415, 2019.

[3] Aidan Fuller, Zane Fan, Charles Day, and Colin Barlow. Digital twin: Enabling technologies, challenges and open research. *IEEE Access*, 8:108952–108971, 2020.

[4] Fei Tao, Bin Xiao, Qinglin Qi, Jiangfeng Cheng, and Ping Ji. Digital twin modeling. *Journal of Manufacturing Systems*, 64:372–389, 2022.

[5] Fei Tao, Qinglin Qi, Lihui Wang, and AYC Nee. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering*, 5(4):653–661, 2019.

[6] Matthias Eckhart and Andreas Ekelhart. Securing cyber-physical systems through digital twins. *Ercim News*, 115:22–23, 2018.

[7] Abiodun Ayodeji, Yong-kuo Liu, Nan Chao, and Li-qun Yang. A new perspective towards the development of robust data-driven intrusion detection for industrial control systems. *Nuclear engineering and technology*, 52(12):2687–2698, 2020.

[8] Cristina Alcaraz and Javier Lopez. Digital twin: A comprehensive survey of security threats. *IEEE Communications Surveys & Tutorials*, 24(3):1475–1503, 2022.

[9] Salwa Alem, David Espes, Laurent Nana, Eric Martin, and Florent De Lamotte. A novel bi-anomaly-based intrusion detection system approach for industry 4.0. *Future Generation Computer Systems*, 145:267–283, 2023.

[10] Christian Gehrmann and Martin Gunnarsson. A digital twin based industrial automation and control system security architecture. *IEEE Transactions on Industrial Informatics*, 16(1):669–680, 2019.

[11] Seba Anna Varghese, Alireza Dehlaghi Ghadim, Ali Balador, Zahra Alimadadi, and Panos Papadimitratos. Digital twin-based intrusion detection for industrial control systems. In *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pages 611–617. IEEE, 2022.

[12] Tijana Markovic, Miguel Leon, Bjorn Leander, and Sasikumar Punnekkat. A modular ice cream factory dataset on anomalies in sensors to support machine learning research in manufacturing systems. *IEEE Access*, 11:29744–29758, 2023.

[13] Fatemeh Akbarian, Emma Fitzgerald, and Maria Kihl. Intrusion detection in digital twins for industrial control systems. In *2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pages 1–6. IEEE, 2020.

[14] William Tärneberg, Per Skarin, Christian Gehrmann, and Maria Kihl. Prototyping intrusion detection in an industrial cloud-native digital twin. In *2021 22nd IEEE International Conference on Industrial Technology (ICIT)*, volume 1, pages 749–755. IEEE, 2021.

[15] Enis Karaarslan and Mohammed Babiker. Digital twin security threats and countermeasures: An introduction. In *2021 International Conference on Information Security and Cryptology (ISCTURKEY)*, pages 7–11. IEEE, 2021.

[16] Marietheres Dietz, Manfred Vielberth, and Günther Pernul. Integrating digital twin security simulations in the security operations center. In *Proceedings of the 15th International Conference on Availability, Reliability and Security*, pages 1–9, 2020.

[17] Matthias Eckhart and Andreas Ekelhart. A specification-based state replication approach for digital twins. In *Proceedings of the 2018 workshop on cyber-physical systems security and privacy*, pages 36–47, 2018.

[18] Matthias Eckhart and Andreas Ekelhart. Towards security-aware virtual environments for digital twins. In *Proceedings of the 4th ACM workshop on cyber-physical system security*, pages 61–72, 2018.